Understanding Exoplanet Atmospheres

Thesis by
Zhaoxi Michael Zhang

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended February 8, 2022

Zhaoxi Michael Zhang
ORCID: 0000-0002-0659-1783

ACKNOWLEDGEMENTS

I would like to thank my advisor Heather Knutson, who gave me complete intellectual freedom to explore my hare-brained ideas. Most of them did not pan out, but the ones that did make up 5 out of the 6 papers in this thesis. She also gave me great scientific advice, provided me with all the equipment I needed, and promptly answered emails and looked over drafts.

I would like to thank Jim Fuller, not only for serving on my candidacy and thesis committees, but also for advising me on a supernova research project that did not make it into the thesis. Jim was a great adviser, and the project gave me valuable experience in a completely different area of astrophysics.

I would like to thank my colleagues, without whom my work would not have been possible. In particular, I'd like to thank my former SURF mentee Happy Chang for being a delight to work with, and my close collaborator Fei Dai, who made my current research project possible by pointing out observationally favorable TESS candidates.

I would like to thank the Caltech Sovereignty Club, where we talk about international affairs, history, sociology, and everything else unrelated to STEM. It is a much needed break from science. In particular, Eduardo Beltrame deserves thanks for founding the club, and all active members and guest speakers deserve thanks for keeping it a vibrant forum for discussion.

I would like to thank my friends, including my childhood friend Ziyi Cai, and the ones I made at Caltech. In particular, I'd like to thank Nikita Kamraj for spending a lot of time with me during the pandemic, and Zhihui Li for sharing screens with me every weekday during the pandemic to keep both of us productive.

ABSTRACT

The study of exoplanet atmospheres is a blossoming field. Over the past two decades, dozens of hot gas giant atmospheres have been observed using a variety of techniques with both space and ground telescopes, revealing the presence of water, the ubiquity of clouds, the presence of equatorial jets, the existence of photoevaporation, and much more. For the far more abundant planets smaller than $\sim 3R_{\oplus}$, which potentially have a wide variety of exotic atmospheric compositions, observations are more challenging. We are just beginning to characterize their atmospheres.

This thesis consists of 6 papers on the topic of exoplanet atmospheres. In paper 1, we use *Spitzer* observations to make rudimentary 1D maps of two hot giant planets, allowing us to infer atmospheric circulation properties and compare them to models and to similar observations of other giant planets. In paper 2, we present PLATON, a fast, open source, easy to use, and easy to understand Python package that calculates transmission spectra for exoplanets and retrieves atmospheric characteristics based on observed spectra. PLATON supports the most common atmospheric parameters, in addition to less commonly included features such as a Mie scattering cloud model and unocculted starspot corrections. In paper 3, we add significant improvements to PLATON, including updated molecular opacities and emission spectra capability. In addition, we perform the most comprehensive retrieval on published *HST* and *Spitzer* transmission and emission spectra of the archetypal hot Jupiter HD 189733b, finding that they are well-matched by a moderately metal-enhanced atmosphere with a solar C/O ratio where the terminator is dominated by extended nm-sized hazes.

Papers 4-6 cover mass loss from sub-Neptunes. In paper 4, we present a tight upper limit on the amount of escaping helium from the the archetypal super Earth 55 Cnc e, suggesting that it has no primordial (H/He) atmosphere. In paper 5, we obtain the first detection of an escaping atmosphere from a young mini Neptune by measuring Lyα absorption from HD 63433c. We do not detect absorption from the inner planet, suggesting that the inner planet may have lost its primordial atmosphere while the outer one has not. In paper 6, we detect escaping helium from a young mini Neptune for the first time. The inferred mass loss rate is high enough to strip a significant portion of the atmosphere within the planet's lifetime; combined with the previous paper, these observations support the canonical explanation of mini Neptunes as rocky planets with a substantial primordial H/He atmosphere and validate models predicting that mini Neptunes can transform into super Earths.

PUBLISHED CONTENT AND CONTRIBUTIONS

Michael Zhang, Heather A. Knutson, Tiffany Kataria, Joel C. Schwartz, Nicolas B. Cowan, Adam P. Showman, Adam Burrows, Jonathan J. Fortney, Kamen Todorov, Jean-Michel Desert, Eric Agol, and Drake Deming. Phase Curves of WASP-33b and HD 149026b and a New Correlation between Phase Curve Offset and Irradiation Temperature. *AJ*, 155(2):83, February 2018. doi: 10.3847/1538-3881/aaa458. M.Z. analyzed the Spitzer data, interpreted the global circulation model results, found the statistical trends, and wrote the paper.

Michael Zhang, Yayaati Chachan, Eliza M. R. Kempton, and Heather A. Knutson. Forward Modeling and Retrievals with PLATON, a Fast Open-source Tool. *PASP*, 131(997):034501, March 2019. doi: 10.1088/1538-3873/aaf5ad. M.Z. conceived the project, designed the package, wrote the vast majority of the code, and wrote the paper.

Michael Zhang, Yayaati Chachan, Eliza M. R. Kempton, Heather A. Knutson, and Wenjun (Happy) Chang. PLATON II: New Capabilities and a Comprehensive Retrieval on HD 189733b Transit and Eclipse Data. *ApJ*, 899(1):27, August 2020. doi: 10.3847/1538-4357/aba1e6. M.Z. conceived the project, wrote the vast majority of the code, and wrote the paper.

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, Antonija Oklopcic, and Renyu Hu. No Escaping Helium from 55 Cnc e. *AJ*, 161(4):181, April 2021. doi: 10.3847/1538-3881/abe382. M.Z. conceived the project, wrote the observing proposal, obtained the data, carried out the data analysis and the 1D simulations, and wrote the paper.

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, and Oscar Barragán. Escaping Helium from TOI 560.01, a Young Mini-Neptune. *AJ*, 163(2):67, February 2022a. doi: 10.3847/1538-3881/ac3fa7. M.Z. conceived the project, wrote the observing proposals, obtained the data, carried out the data analysis and the 1D simulations, and wrote the paper.

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, Leonardo A. dos Santos, Luca Fossati, Gregory W. Henry, David Ehrenreich, Yann Alibert, Sergio Hoyer, Thomas G. Wilson, and Andrea Bonfanti. Detection of Ongoing Mass Loss from HD 63433c, a Young Mini-Neptune. *AJ*, 163(2):68, February 2022b. doi: 10.3847/1538-3881/ac3f3b. M.Z. conceived the project, wrote the observing proposals, obtained the data, carried out the fiducial data analysis, and wrote the paper.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Chapter 1

INTRODUCTION

In 1995, astronomers discovered the first planet around another main sequence star: 51 Peg b (Mayor & Queloz, 1995). This planet is an exotic world unlike any in the solar system. It is a tidally locked, Jupiter-sized gas giant with an orbital period of 4.2 days and equilibrium temperature of 1300 K, virtually guaranteeing that its atmospheric composition, clouds, circulation patterns, and formation history are very different from that of the solar system gas giants. Subsequent discoveries showed that 51 Peg b was the first example of a previously unknown class of planets, called hot Jupiters.

Due to their size, high thermal emission, and puffy atmospheres, hot Jupiters are the most observationally favorable exoplanets. The Hubble Space Telescope and ground-based telescopes routinely measure their compositions, cloud properties, and circulation patterns. We now know (Madhusudhan, 2019) that clouds are extremely common on these planets, as is water vapor and eastward winds carrying substantial heat fluxes. Several other molecules have also been detected. In addition, an extensive menagerie of atoms and ions have been detected in the upper atmospheres of some ultra-hot Jupiters (e.g. Hoeijmakers et al. 2019).

The advent of ground and space-based transit surveys like WASP, HATnet, and Kepler has made it possible to discover exoplanets smaller than hot Jupiters en masse. Thousands of exoplanets are now known, ranging from planets smaller than Mars to giants almost heavy enough to fuse deuterium in their cores (Batalha, 2014). The field of exoplanets has seen an explosion in interest, data, and progress. Exoplanets have continued to surprise astronomers. For example, the most common type of planet among those discovered is one that does not exist in the solar system: the sub-Neptunes, with radii between 1 and 4 $R_\oplus$. These planets are fundamentally mysterious. What are they like? How did they form? Do they have atmospheres, and if so, what are they made out of?

Adding to the mystery, the sub-Neptune population is divided into two sub-populations (Figure 1.1): the smaller but denser super Earths, and the bigger but less dense mini Neptunes (Fulton et al., 2017; Fulton & Petigura, 2018). These sub-populations are divided by a radius gap around 1.9 $R_\oplus$, where few planets reside. The prevailing

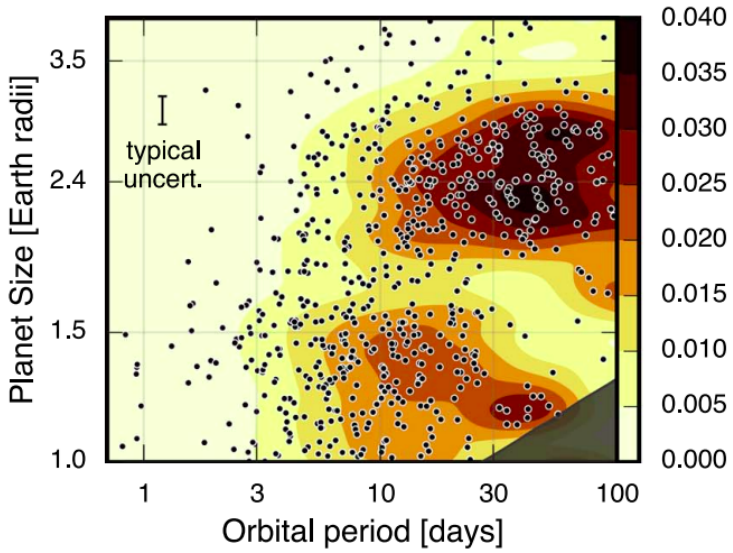


Figure 1.1: Distribution of known transiting exoplanets and planet candidates discovered by the Kepler survey, by radius and period (Fig. 6 of Fulton & Petigura 2018, reproduced with permission). Notice the two distinct populations: the bigger mini Neptunes and the smaller super Earths.

theory is that the mini Neptunes have a primordial hydrogen/helium atmosphere that make up a few percent of their mass, while the super Earths do not because high-energy radiation from the star has evaporated it away. However, the validity of this theory, and the precise nature of the two sub-populations, are subjects of ongoing research. In particular, it is the subject of the last three papers of this thesis.

1.1 How exoplanet atmospheres are studied

There are three main methods to study exoplanet atmospheres: transmission spectra, emission spectra, and phase curves. Typically, astronomers study the atmospheres of transiting exoplanets, which periodically pass in front of their stars, because transmission spectra and low-resolution emission spectra are only possible for transiting planets.

Transmission spectroscopy

In transmission spectroscopy, the star is monitored spectroscopically before, during, and after the planetary transit, with the change in the spectrum during the transit being attributed to absorption from the planet (Figure 1.2). If the planet has an atmosphere, the amount of starlight absorbed will be wavelength dependent because the cross sections of absorbers in the atmosphere are wavelength dependent, and this wavelength dependence can be used to detect molecules. For example, there

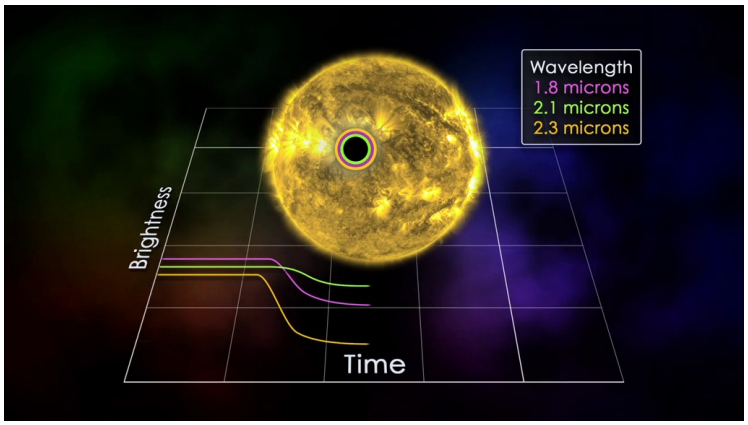


Figure 1.2: A planet transiting its host star. Source: NASA, Goddard Space Flight Center (`https://svs.gsfc.nasa.gov/11428`).

is a strong water absorption feature at 1.4 μm, in the middle of the *HST*/WFC3 G141 bandpass. This combination of telescope, instrument, and grism is a favorite of exoplanet atmosphere observers, and has detected water features on more than a dozen planets (Madhusudhan, 2019). Transmission spectroscopy is easiest for large, hot planets with low surface gravity and low mean molecular weight, because these planets have large and puffy atmospheres.

The ubiquity of clouds is one of the earliest and most robust discoveries of transmission spectroscopy, but also the most frustrating. Clouds block the line of sight to the lower atmosphere, dramatically weakening spectral features, often to the point of undetectability. In the tens of transmission spectra observed to date, every single one has a water absorption feature smaller than 2 scale heights, whereas 5-10 scale heights is expected for a clear atmosphere (Madhusudhan, 2019). Cloud particles typically have no spectral features of their own which can be detected with current technology. Among the many mini Neptunes observed in transmission, only K2-18b shows a definitive absorption feature (Tsiaras et al., 2019; Benneke et al., 2019), while the others are either flat or ambiguous. One particularly nearby mini Neptune orbiting a small star (an *a priori* favorable target) was observed with *HST*/WFC3 across 15 transits—more than enough to detect even a clear H_2O or CO_2 atmosphere, which have signals several times lower than a clear H_2/He atmosphere—yet the transit spectrum was completely flat (Kreidberg et al., 2014).

For cloudy atmospheres, all hope is not lost. One can observe the transit at spectral resolutions high enough to resolve the individual molecular lines (e.g. Snellen et al. 2010), of which there are hundreds to billions, depending on the molecule and the wavelength range. While the observations are much too noisy to detect

individual lines, cross correlating with a template could yield a strong detection. This is expected to mitigate the problem of clouds because molecules absorb many orders of magnitude more strongly in the cores of molecular lines than at typical wavelengths, potentially raising their photospheres far above the cloud deck (Gandhi et al., 2020). This technique has detected many atoms, molecules, and ions in several planets. For example, 5 molecules were detected in the hot Jupiter HD 209458b (Giacobbe et al., 2021).

Transmission spectroscopy can also be used to measure mass loss from escaping atmospheres, even when clouds exist (see Owen & Campos Estrada 2020 for a review). Gaseous planets around young and active stars experience intense stellar X-ray and extreme ultraviolet irradiation, which heats their upper atmospheres to thousands of Kelvin and drives a supersonic outflow. For mini Neptunes, whose atmospheres comprise only a few percent of their mass, this photoevaporation can strip the entire atmosphere over the course of the planet's life, shrinking its radius by tens of percent while barely affecting its mass. The planet turns into a super Earth. Mass loss is thus expected to play a key role in the formation and evolution of super Earths and mini Neptunes. It is the leading explanation for why these two populations exist in the first place, separated by a radius gap where few planets exist (Fulton et al., 2017).

We can directly observe mass loss by looking for absorption in either the Lyα line (1216 Å) from ground state hydrogen, or the line (10,833 Å) from triplet state helium. The two methods are complementary: aside from detecting different atoms, the former probes high-velocity gas in the tenuous outer regions of the outflow (because interstellar absorption wipes out the line core), while the latter probes lower-velocity gas closer to the planet. Escaping hydrogen and helium have been detected in Jupiter-sized and Neptune-sized planets (starting with Vidal-Madjar et al. 2003 for hydrogen and Spake et al. 2018 for helium), but until recently, attempts to detect photoevaporation in the regime where it dramatically affects exoplanet demographics—namely young mini Neptunes—have not succeeded. In the last two papers of this thesis, we present a detection of escaping hydrogen from one planet in this regime, and escaping helium from another. Aside from shedding light on planet evolution, these mass loss measurements provide additional evidence that most mini Neptunes have thick hydrogen/helium atmospheres—which has been theorized, but not proven, as there are other potential explanations for their low density.

Emission spectroscopy

In emission spectroscopy, a spectrograph monitors the star as the planet passes behind it, and the spectrum of the planet is deduced from the difference between the in-eclipse and out-of-eclipse spectra. At infrared wavelengths, this spectrum captures thermal emission from the planet. An isothermal atmosphere would have a blackbody spectrum. An atmosphere where temperature increases with depth shows molecular absorption features, i.e. less emission at wavelengths of high opacity, because the photosphere is higher and colder where opacity is higher. Conversely, an "inverted" atmosphere whose temperature decreases with depth shows molecular emission features. Inverted atmospheres are possible in a limited altitude range in the presence of molecules which absorb strongly in the optical, but not in the infrared. Just like with transmission spectra, clouds could wipe out molecular features, but the larger range of temperatures on the dayside compared to the terminator and the more favorable geometry make clouds less of a problem. Just like with transmission spectra, high-resolution spectroscopy can be used to mitigate the problem of clouds and measure relative molecular abundances.

Emission spectra are easiest to obtain for large, hot planets orbiting small stars. Unlike transmission spectra, emission spectral features are not much smaller for high mean molecular weight atmospheres than for hydrogen/helium atmospheres. Emission spectra also depend strongly on the temperature-pressure profile, which itself is governed by a complex combination of stellar radiation, interior heat, convection, radiative transfer, clouds, and horizontal circulation. This is a mixed blessing, for it means that emission spectra can constrain the temperature-pressure profile, but that attempts to measure atmospheric composition or other properties are complicated by the unknown profile.

Favorite instruments for measuring emission spectra include *HST*/WFC3 (the G141 grism, 1.1–1.7 μm, covers the water band), and *Spitzer*/IRAC (3.6 and 4.5 μm) before it was decommissioned. From the ground, high resolution emission spectra have been taken by CARMENES, VLT/CRIRES, Keck/NIRSPEC, and many others (see Birkby 2018 for a review). Emission spectroscopy has detected molecular absorption from a handful of planets in addition to constraining their temperature-pressure profile (Madhusudhan, 2019). The large number of planets that *Spitzer* has observed makes statistical studies possible (e.g. Wallack et al. 2021), shedding light on heat recirculation and composition trends.

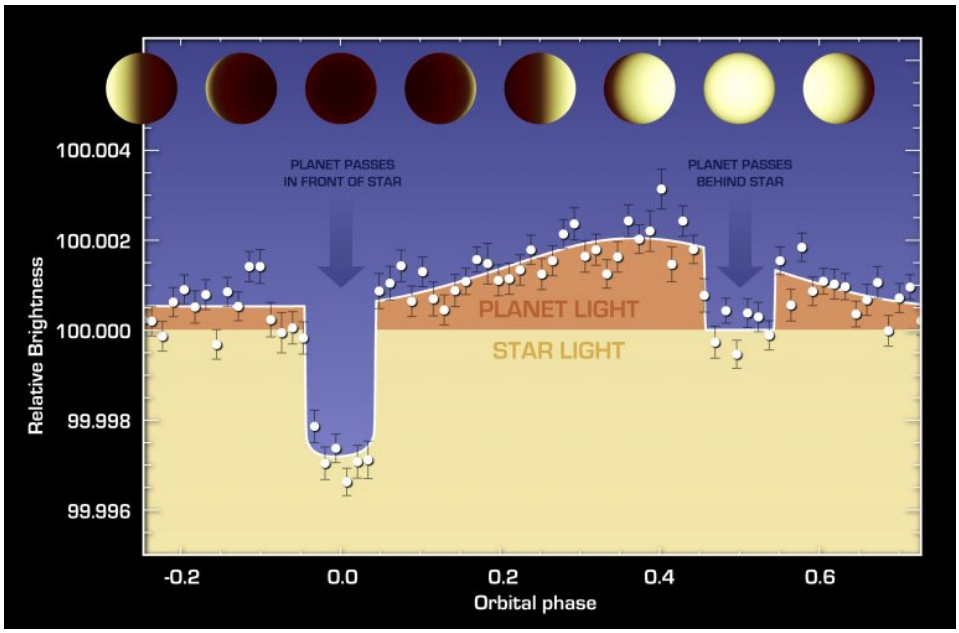


Figure 1.3: The phase curve of a transiting planet. Credit: NASA/JPL-Caltech/Univ. of Cambridge.

Phase curves

A phase curve (Figure 1.3) is a record of the brightness of the star + planet system over an entire orbit. A phase curve contains a transit, an eclipse, and flux variations in between the transit and eclipse. The flux varies because the planet presents different phases as it orbits; for tidally locked planets, a phase always corresponds to the same hemisphere. A phase curve in the infrared measures thermal emission, enabling the reconstruction of a crude one-dimensional thermal map of the planet.

Many useful properties can be measured from a phase curve. The dayside and nightside temperatures can be deduced from the dayside and nightside emission. These temperatures can be used to compute the albedo and heat recirculation efficiency, as parameterized by a toy model (Cowan & Agol, 2011). Additionally, the phase at which emission peaks is usually slightly before the secondary eclipse, corresponding to a hot spot eastward of the substellar position. Since the hot spot is displaced eastward due to eastward equatorial jets which advect significant amounts of heat, the phase offset sheds light on the nature of atmospheric circulation. Lastly, a spectroscopic phase curve can measure how atmospheric composition changes as a function of longitude.

Spitzer has observed the phase curves of ~20 planets in the 3.6 and/or 4.5 μm bands (see Bell et al. 2021 for a comprehensive reanalysis of 4.5 μm phase curves), and a few

at higher wavelengths (e.g. the first phase curve, Knutson et al. 2007). The results broadly confirm some predictions from 3D general circulation models, such as the higher day-night temperature contrast for hotter planets. However, many predicted trends are surprisingly absent. No correlation has been observed between phase offset and temperature, nor between temperature and heat recirculation efficiency (Bell et al., 2021). On an individual planet level, the match is even worse. For example, the cloud-free general circulation models that we use tend to drastically overpredict the phase offset (Zhang et al., 2018), although active modelling of clouds is expected to reduce the discrepancy (Roman & Rauscher, 2019). Much remains to be understood about heat recirculation on exoplanets.

Finding favorable planets

The study of exoplanet atmospheres would not be possible without the discovery of nearby transiting exoplanets to study. While many nearby hot Jupiters were discovered by ground-based transit surveys such as HATnet and WASP, discovering sub-Neptunes from the ground has been much more challenging because of the exacting photometric precision required. The spacecraft Kepler has discovered thousands of smaller transiting planets by staring at a 10° x 10° patch of sky for more than 3 years. However, there are only so many nearby stars in this small patch of sky. For this reason, the vast majority of Kepler discoveries are too faint for a mass measurement, and for atmospheric characterization with current telescopes.

In contrast to Kepler, TESS is a spacecraft which hunts for transiting exoplanets by scanning the entire sky, with each sector observed continuously for 30 days. TESS can detect almost every short-period planet that would be favorable for atmospheric followup. For this reason, TESS has dramatically increased the number of sub-Neptunes favorable for atmospheric characterization. It is no coincidence that the last two papers of my thesis, in addition to my ongoing work, are entirely dedicated to studying planets discovered by TESS.

1.2 Theoretical understanding

To extract useful information from a transmission spectrum, emission spectrum, or phase curve, one needs a model. The extracted information should ideally shed light not only on the properties of the planet in question, but on broader questions like planet demographics and formation history.

Models come in various levels of complexity. The most complex are 3D general circulation models which include chemistry, clouds, radiative transfer, and advection

in a self-consistent way (e.g. Lines et al. 2018). As mentioned in the previous section, these models do not generally match observations, and it is usually not obvious why, and time consuming to find out how changes in the parameters translate to changes in the result. At the other extreme are parameterized 1D models which do not strive for self-consistency, but are instead very fast. The parameters are inferred from the data via a Bayesian method, such as MCMC or nested sampling. This process is called a "retrieval." Until recently, there were very few open source retrieval codes available (TauREx being an exception, Waldmann et al. 2015) and none that were easy to use, inspiring us to write Planetary Atmospheric Tool for Observer Noobs (PLATON) (Zhang et al., 2018, 2019).

Retrievals come with their own set of problems. It is easily possible to both fit the data very well and infer completely incorrect parameters, due to a combination of missing physics, the lack of consistency constraints, difficult-to-model instrumental systematics, and theoretical uncertainties in molecular opacities. The last has been significantly mitigated in recent years due to a concerted effort to improve molecular line lists and make them publicly available, especially on the part of the ExoMol collaboration (Tennyson & Yurchenko, 2012).

Retrievals can be used to infer either individual molecular abundances, or the overall metallicity and C/O ratio of the atmosphere, depending on the chosen parametrization. The metallicity and C/O ratio parameterization requires an assumption of chemical equilibrium, potentially with parameterized perturbations. The parameterization is common partially because researchers have long sought to use exoplanet atmospheric compositions to understand planet formation (Madhusudhan, 2019). Generally speaking, grains form early in the history of a protoplanetary disk, and are then aggregated into pebbles, planetesimals, and planets. The solids are metal-rich, while the remaining gas has reduced metallicity because the heavy elements had been preferentially taken out of it by the solids. Thus, the metallicity of the exoplanet's atmosphere could potentially reveal the amount of solids that it accreted relative to the amount of gas. In addition, the protoplanetary disk is expected to have a series of "ice lines" where H_2O, CO_2, and then CO condense as one moves farther from the star. The C/O ratio of the gas and of the solids both change as these species condense (Öberg et al., 2011), so the C/O ratio of the atmosphere should shed light on the planet's formation location. However, a wide variety of processes complicate this simple story. As a small sample of the problems, planets migrate during the formation process; there are vertical temperature gradients; many thermal

and photochemical processes change the disk composition, not just condensation; in mature planets, condensates can rain out of the atmosphere, removing them from view; measuring metallicity and C/O ratio requires assumptions about equilibrium chemistry which may be inaccurate (e.g. Moses et al. 2013); and as we showed in Zhang et al. (2019), even the most exquisite data for the best measured exoplanet in existence (HD 189733b) do not constrain the C/O ratio very well. There is still much work to do before atmospheric composition can be used to make definitive statements about planet formation.

1.3 Structure of this thesis

This thesis comprises 8 chapters. Chapter 1 is this introduction. Chapters 2–7 contain the six papers listed in Published Content and Contributions, in chronological order. Chapter 8 is a conclusion which explores potential future directions for the field of exoplanet atmospheric science.

References

Batalha, N. M. 2014, Proceedings of the National Academy of Sciences, 111, 12647

Bell, T. J., Dang, L., Cowan, N. B., et al. 2021, MNRAS, 504, 3316

Benneke, B., Wong, I., Piaulet, C., et al. 2019, ApJ, 887, L14

Birkby, J. L. 2018, arXiv e-prints, arXiv:1806.04617

Cowan, N. B., & Agol, E. 2011, ApJ, 729, 54

Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109

Gandhi, S., Brogi, M., & Webb, R. K. 2020, MNRAS, 498, 194

Giacobbe, P., Brogi, M., Gandhi, S., et al. 2021, Nature, 592, 205

Hoeijmakers, H. J., Ehrenreich, D., Kitzmann, D., et al. 2019, A&A, 627, A165

Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, Nature, 447, 183

Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, Nature, 505, 69

Lines, S., Mayne, N. J., Boutle, I. A., et al. 2018, A&A, 615, A97

Madhusudhan, N. 2019, ARA&A, 57, 617

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, ApJ, 763, 25

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJ, 743, L16

Owen, J. E., & Campos Estrada, B. 2020, MNRAS, 491, 5287

Roman, M., & Rauscher, E. 2019, ApJ, 872, 1

Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049

Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, Nature, 557, 68

Tennyson, J., & Yurchenko, S. N. 2012, MNRAS, 425, 21

Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., & Yurchenko, S. N. 2019, Nature Astronomy, 3, 1086

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J. M., et al. 2003, Nature, 422, 143

Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107

Wallack, N. L., Knutson, H. A., & Deming, D. 2021, AJ, 162, 36

Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2019, PASP, 131, 034501

Zhang, M., Knutson, H. A., Kataria, T., et al. 2018, AJ, 155, 83

Chapter 2

PHASE CURVES OF WASP-33B AND HD 149026B AND A NEW CORRELATION BETWEEN PHASE CURVE OFFSET AND IRRADIATION TEMPERATURE

Michael Zhang, Heather A. Knutson, Tiffany Kataria, Joel C. Schwartz, Nicolas B. Cowan, Adam P. Showman, Adam Burrows, Jonathan J. Fortney, Kamen Todorov, Jean-Michel Desert, Eric Agol, and Drake Deming. Phase Curves of WASP-33b and HD 149026b and a New Correlation between Phase Curve Offset and Irradiation Temperature. *AJ*, 155(2):83, February 2018. doi: 10.3847/1538-3881/aaa458. M.Z. analyzed the Spitzer data, interpreted the global circulation model results, found the statistical trends, and wrote the paper.

2.1 Abstract

We present new 3.6 and 4.5 μm *Spitzer* phase curves for the highly irradiated hot Jupiter WASP-33b and the unusually dense Saturn-mass planet HD 149026b. As part of this analysis, we develop a new variant of pixel level decorrelation that is effective at removing intrapixel sensitivity variations for long observations (>10 hours) where the position of the star can vary by a significant fraction of a pixel. Using this algorithm, we measure eclipse depths, phase amplitudes, and phase offsets for both planets at 3.6 μm and 4.5 μm. We use a simple toy model to show that WASP-33b's phase offset, albedo, and heat recirculation efficiency are largely similar to those of other hot Jupiters despite its very high irradiation. On the other hand, our fits for HD 149026b prefer a very high albedo and an unusually high recirculation efficiency. We also compare our results to predictions from general circulation models, and find that while neither planet matches the models well, the discrepancies for HD 149026b are especially large. We speculate that this may be related to its high bulk metallicity, which could lead to enhanced atmospheric opacities and the formation of reflective cloud layers in localized regions of the atmosphere. We then place these two planets in a broader context by exploring relationships between the temperatures, albedos, heat transport efficiencies, and phase offsets of all planets with published thermal phase curves. We find a striking relationship between phase offset and irradiation temperature—the former drops with increasing temperature until around 3400 K, and rises thereafter. Although

some aspects of this trend are mirrored in the circulation models, there are notable differences that provide important clues for future modeling efforts.

2.2 Introduction

The *Spitzer Space Telescope* was designed and constructed prior to the discovery of the first transiting exoplanet, but it has nevertheless become an important tool in the study of exoplanet atmospheres. In particular, the development of techniques to correct instrumental systematics and derive precise timeseries photometry from *Spitzer* data has enabled the first measurements of thermal emission from a diverse array of exoplanets (Seager & Deming, 2010). These measurements, in the form of secondary eclipses and phase curves, allow us to characterize the temperatures, albedos, heat transport efficiencies, and phase offsets of these planets (e.g., Cowan & Agol, 2011). For planets with observations at multiple wavelengths we can also constrain their atmospheric compositions, investigate their vertical pressure-temperature profiles, and probe the presence of clouds (Burrows et al., 2010). These *Spitzer* phase curves provide invaluable information about the fundamental physical processes that drive the atmospheric circulation patterns of these tidally-locked planets, and can be compared to predictions from general circulation models (GCMs; e.g., Heng & Showman, 2015). Although both models and observations are generally in good agreement on the dayside emission spectra of hot Jupiters, there are significant discrepancies in the measured night-side spectra (e.g. Zellem et al., 2014) and models that provide a good match to the measured phase curve amplitude and phase offset in a single bandpass often have difficulties matching phase curve data for the same planet at additional wavelengths (i.e. Knutson et al., 2012). This suggests that there are aspects of the atmospheric circulation, cloud properties, magnetic fields, and chemistry of these planets that are not adequately captured in current GCMs. The GCMs we use in this paper, for example, neglect clouds, magnetohydrodynamics, and disequilibrium chemistry, although some of these topics have been investigated in more focused modeling studies (e.g., Cooper & Showman 2005; Parmentier et al. 2016; Rogers 2017).

In this paper, we examine multi-wavelength phase curve observations for two planets with unusual characteristics as compared to the broader sample of transiting hot Jupiters. WASP-33b is a 2.2 M_J planet with a radius of 1.5 R_J orbiting a 1.5 M_{sun} δ Scuti star with a period of 1.22 days (Lehmann et al., 2015). With an irradiation temperature ($T_0 = T_{eff}/\sqrt{a_*}$) of 3890 K, this planet is one of the most highly irradiated hot Jupiters currently known. The star itself has pulsations at a

variety of frequencies, with the dominant mode at 21 d^{-1}. These pulsations have an amplitude of roughly 1 mmag, about 1/4 that of the secondary eclipse depth (von Essen et al., 2014). Previous authors have measured broadband thermal emission from WASP-33b's dayside at a variety of wavelengths, including 0.91 μm (Smith et al., 2011), 1.05 μm (von Essen et al., 2015), 2.14 μm (Deming et al., 2012; de Mooij et al., 2013), and *Spitzer*'s 3.6 and 4.5 μm bands (Deming et al., 2012). von Essen et al. (2015) summarize these results and combine them to obtain an average dayside brightness temperature of 3358 ± 165 K. More recently, Haynes et al. (2015) reported evidence for a temperature inversion in the 1.1–1.6 μm dayside spectrum of this planet, and Nugroho et al. (2017) used a cross-correlation technique to detect TiO in the 0.62–0.88 μm dayside spectrum with the High-Dispersion Spectrograph on Subaru. It has long been suggested (Hubeny et al., 2003; Fortney et al., 2008; Burrows et al., 2008) that additional opacity from molecules such as gas-phase TiO and VO could lead to the formation of temperature inversions in the most highly irradiated atmospheres, and this indeed appears to be the case for WASP-33b.

Atmospheric circulation models generally predict that more highly irradiated planets should have larger day-night temperature contrasts. According to Showman & Guillot (2002), the day-night temperature difference can be thought of as resulting from a competition between the radiative cooling timescale τ_{rad} and the timescale of advection by wind, τ_{adv}. Because the radiative timescale decreases much faster with increasing temperature than the advective timescale, more highly irradiated planets should have steeper day-night temperature gradients. Perez-Becker & Showman (2013) and Komacek & Showman (2016) show that the full picture is more complicated, but the general idea is still that radiation outcompetes other heat transport mechanisms for the most highly irradiated planets, causing a larger day-night temperature difference. The typical pressure at whic incident starlight is absorbed is also important for atmospheric circulation, and the presence of a dayside temperature inversion will therefore also affect the redistribution of energy to the planet's night side (Showman et al., 2009; Lewis et al., 2014).

HD 149026b is a 0.36 M_J planet with a radius of 0.65 R_J, and orbits a subgiant G0 IV star of metallicity [Fe/H] = 0.36 with a period of 2.9 days (Sato et al., 2005). Its small radius and correspondingly large density suggest the presence of a large heavy-element core. Dodson-Robinson & Bodenheimer (2009) summarize the many attempts to estimate the mass of this core, concluding that plausible estimates range from 45-110 $M_\oplus$, corresponding to 39-96% of total mass. Given the high

metallicity of the star and the planet's large core-mass fraction, this planet seems likely to have a high atmopsheric metallicity. Although Carter et al. (2009) analyzed four spectroscopic transit observations with *Hubble*'s NICMOS instrument (1.1–2.0 μm), the uncertainties from these data were too high to provide useful constraints on the planet's transmission spectrum. Stevenson et al. (2012) subsequently obtained *Spitzer* secondary eclipse observations at 3.6, 4.5, 5.8, 8.0, and 16 μm and found a brightness temperature of 2000 ± 60 K at 3.6 μm and 1600 − 1800 K at longer wavelengths. When they fit these data with chemical equilibrium models, they found that they preferred models with large amounts of CO and CO_2, 30× solar metallicity, no temperature inversion, and moderate heat redistribution.

Lewis et al. (2010) studied the effect of metallicity on the warm Neptune GJ436b and found that high metallicity models had equatorial jets and strong day-night temperature variations, while lower metallicity models had weak temperature variations and high latitude jets. By contributing opacity, metals raise the photosphere to a higher altitude, where atmospheric dynamics are less important and radiative cooling is more efficient. For planets with condensate cloud layers, increasing the atmospheric metallicity also increases the amount of cloud-forming material and the corresponding cloud opacity. Although most circulation models do not currently include clouds, the presence of spatially inhomogeneous cloud layers can significantly alter the shape of both optical and infrared phase curves (Parmentier et al., 2016; Heng & Demory, 2013; Shporer & Hu, 2015).

We describe our new 3.6 and 4.5 μm phase curve observations for WASP-33b and HD 149026b in Section 2.3 and our analysis of this data in Section 2.4. In Section 2.5, we combine a simple toy model and more sophisticated GCM simulations for each planet to interpret these observations and search for patterns in the full sample of published thermal phase curve observations. Finally, we make concluding remarks in Section 2.6.

2.3 Observations

All observations were taken with the 3.6 and 4.5 μm arrays of the IRAC instrument on *Spitzer* (Fazio et al., 2004) during the post-cryogenic (warm) mission. Start dates, total durations (including downlink time), AORs, and other information about the observations used in this paper are presented in Table 2.1. Observations were timed to begin before a secondary eclipse and end after the following secondary eclipse, and all frames were taken in subarray mode without the now-standard

Table 2.1: *Spitzer* Observation Details

Planet	$\lambda(\mu m)$	Date (UTC)	Dur. (h)	Frames	Exp (s)	Order	Ap. (pix)
W	3.6	2012/06/04	37.2	311,552	0.36	2	2.5
W	4.5	2012/04/11	37.2	311,680	0.36	1	2.8
HD	3.6	2011/04/08	81.2*	663,104	0.36	2	2.8
HD	4.5	2011/04/08	81.6*	663,104	0.36	3	2.6

*Including 2.3 h of downlink time

peak-up pointing optimization (Ingalls et al., 2012), which was implemented after these observations were executed. Due to data volume constraints the HD 149026b observations required a single downlink break near the middle of each phase curve, resulting in a 2-3 hour gap in coverage. The WASP-33b observations were executed without any breaks for downlinks.

2.4 Analysis

Overview

We extract a photometric timeseries for each phase curve observation using aperture photometry, then fit the data with a combined astrophysical and noise model as described in the sections below.

Photometry

Subarray images are 32x32 pixels. We estimate and subtract the sky background from each image by excluding all pixels within a radius of 12 pixels from the star, rejecting outliers using sigma clipping, and then calculating the biweight location of the remaining pixels. The biweight location is robust and efficient statistic implemented in `astropy` (Astropy Collaboration et al., 2013), and we find that it gives results comparable to methods used in previous studies (Ingalls et al., 2016). For both planets, the sky background contributes less than 1% of the total flux at 3.6 μm and less than 0.5% of the total flux at 4.5 μm for our preferred apertures.

We estimate the position of the star in each image using an iterative flux-weighted centroiding method with a circular aperture of radius 3 pixels, and perform aperture photometry using the `photutils` module (Bradley et al., 2016) from `astropy`. We consider apertures with fixed radii ranging from 1.5 to 5.0 pixels in steps of either 0.1 pixels (1.5 – 3.0 pixels) or 0.5 pixels (3.0 – 5.0 pixels).

We omit the first 0.1 days of data for each data set, which is normal procedure for *Spitzer* analyses (Deming et al., 2015a) and removes an obvious ramp at the

beginning of the observations. The HD 149026b 3.6 observations have a downlink gap in the middle, so in addition to removing 0.1 days of data from the very beginning, we also remove 0.1 days from the post-downlink segment.

Instrumental noise model

The largest flux variations in our raw *Spitzer* light curves are not astrophysical, but instead result from well-known intrapixel sensitivity variations combined with telscope pointing jitter (e.g., Charbonneau et al. 2005; Grillmair et al. 2012). Although there are several different approaches to correcting for these effects, pixel level decorrelation (PLD; Deming et al. 2015a) has been among the most successful to date in fits to shorter (< 10 h) observations (Ingalls et al., 2016). Following the updated definition of PLD in Benneke et al. (2017), we model the light curve as:

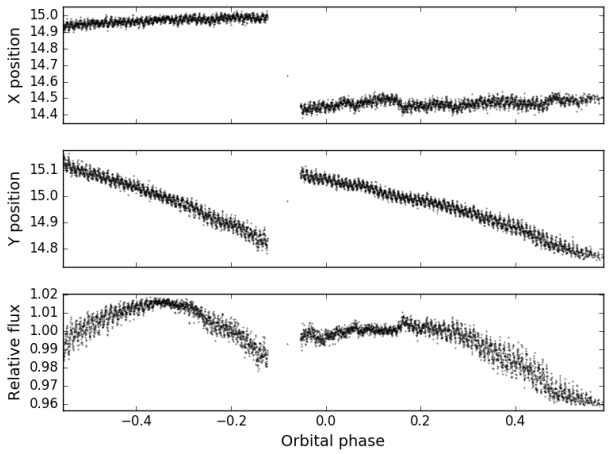
$$L(t) = f(t)(1 + m(t - t_0)) \left(\sum_{i=1}^{9} c_i P_i(t) \right), \tag{2.1}$$

where m is the slope, $f(t)$ is the true brightness, $P_i(t)$ are the normalized fluxes in a 3×3 pixel box centered on the position of the star, c_i are nine coefficients giving the relative weight of each pixel, and the final two terms are meant to model temporal variations in sensitivity. In each image, we remove astrophysical flux variations by dividing the individual pixel values by the sum of the flux across all nine pixels.

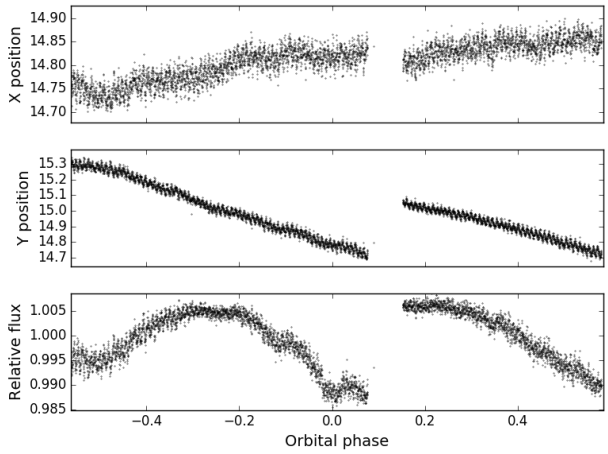
We do not necessarily expect a linear relationship between individual pixel values and the total flux across the aperture. PLD was originally formulated as the first term of a Taylor series expansion and therefore works best when applied to data where the star moves over a relatively small range (typically on the order of $1/10^{th}$ of a pixel) of pixel positions (Wong et al., 2015a). In our observations the star drifted by as much as half a pixel (see Figure 2.1), and we found that the standard linear PLD produced correspondingly poor fits. We account for this increased drift by developing a new variant of PLD:

$$L(t) = f(t)(1 + m(t - t_0)) \left(\sum_{i=1}^{n} \sum_{j=1}^{9} c_{ij} P_j(t)^i \right), \tag{2.2}$$

where the linear slope m is a free parameter and n is the highest order used in the model. This is similar to Luger et al. (2016), except that we neglect cross terms.



(a) 3.6 µm

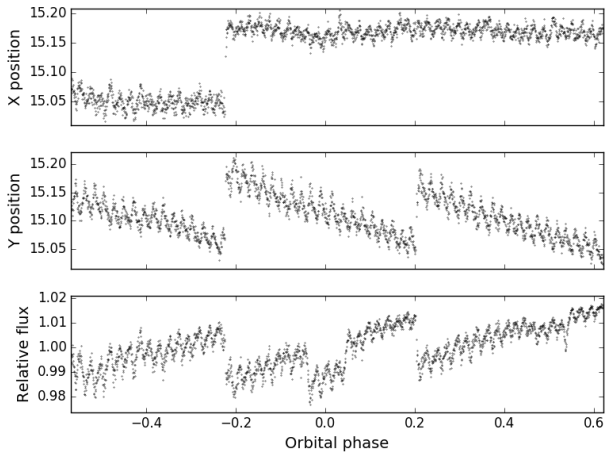


(b) 4.5 µm

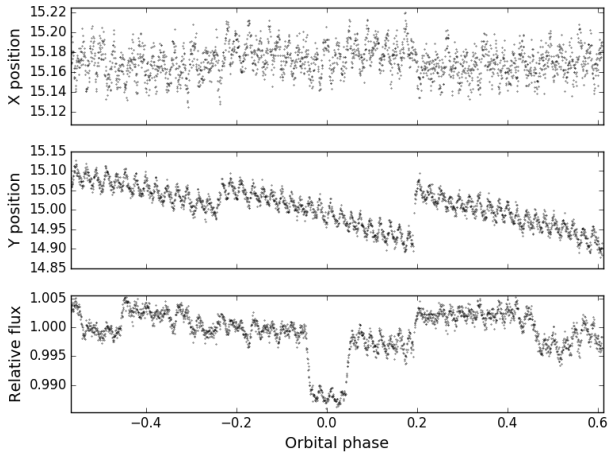
Figure 2.1: Raw photometry and x and y position as a function of orbital phase for HD 149026b. The top panel shows the 3.6 µm data, and the bottom panel shows the 4.5 µm data. Fluxes have been divided by the median value, and all measurements are shown binned into sets of 128 points, corresponding to a time step of 51 seconds.

We experimented with cross terms but found that they did not improve the quality of the fits. This, combined with the combinatorial explosion in the number of cross terms as the order is increased, convinced us to drop the cross terms.

Following Benneke et al. (2017), we have opted to include the linear term in all our fits. The inclusion of a linear term is standard in many analyses (e.g., Stevenson et al. 2012; Deming et al. 2015b) and can account for a variety of instrumental and astrophysical noise sources that are not adequately corrected by the basic instrumental noise model. We find that adding the slope decreases the value of the BIC substantially for WASP-33b's 3.6 µm phase curve ($\Delta BIC = -34.5$) and HD-



(a) 3.6 µm



(b) 4.5 µm

Figure 2.2: Raw photometry and x and y position as a function of orbital phase for WASP-33b. See Fig. 2.1 caption for additional information.

149026b's 3.6 µm phase curve ($\Delta BIC = -15.3$), while having little effect on the value for HD 149026b 4.5 µm ($\Delta BIC = 0.46$) and increasing it for WASP-33b 4.5 µm ($\Delta BIC = 7.7$). Nevertheless, we include the linear term for all light curves for uniformity. We also considered a quadratic term but found that it resulted in an increased BIC for all four visits.

We tried fits in which n ranged as high as sixth order, but found that going beyond third order terms never led to lower BIC. Although we include the $m(t - t_0)$ term in our instrumental noise model, following Deming et al. (2015a), it could also represent an astrophysical drift in the stellar brightness.

Since we use Monte Carlo Markov Chain (MCMC) to fit all parameters, one challenge is the signficant degree of degeneracy between individual pixel light curves, which can result in long convergence times for MCMC fits. We reduce these degeneracies and improve convergence times by carrying out a principle component analysis (PCA) on the $N \times 9n$ matrix of central pixel data where N is the total number of images in each phase curve observation, resulting in a $N \times 9n$ matrix of reprojected central pixel data.

Astrophysical model

Our astrophysical model consists of a transit, a secondary eclipse, and a phase curve. To model the transit and eclipse, we developed a GPU-accelerated version of `batman` (Kreidberg, 2015), which is roughly 10 times faster than the CPU version. This code has since been merged with the main repository[1]. When calculating the transit and eclipse shapes we take the period from Smith et al. (2011) (WASP-33b) and Carter et al. (2009) (HD 149026b), and allow the transit timing, inclination, transit depth, eclipse depth, eclipse phase (common to both eclipses), and a/R_* to vary as free parameters in our fits. We model the transit using a four-parameter nonlinear limb-darkening law, with coefficients derived via linear interpolation from Sing (2010). For WASP-33b, we assumed $T_{eff} = 7400K$, log g = 4.3, and [M/H] = 0.1 (Collier Cameron et al., 2010). For HD 149026b, we assumed $T_{eff} = 6160K$, log(g)=4.278, and [M/H]=0.36 (Torres et al., 2008). `batman` calculates the eclipse shape from geometry alone, thus neglecting limb darkening and all other sources of planetary brightness variation. Published radial velocity measurements and secondary eclipse times for both WASP-33b and HD 149026b indicate that the orbital eccentricities for both of these planets are consistent with zero (Kovács et al., 2013; von Essen et al., 2015), and we therefore fix the eccentricities of both planets to zero in our fits.

Following Cowan & Agol (2008), we model the planet's phase variation as a series expansion in sine and cosine, where we only consider first-order sinusoidal terms:

$$L_p = C + c_1 \cos(2\pi t/P) + c_2 \sin(2\pi t/P), \tag{2.3}$$

where P is the orbital period. Although we also explored fits with second order harmonic terms, we found that these did not improve the quality of the fit for either planet.

[1] `https://github.com/lkreidberg/batman`

Noise model

For HD 149206, which has a relatively quiet host star, we assume the noise is Gaussian and uncorrelated (i.e., white) and allow the value of the per-point uncertainty in each bandpass to vary as a free parameter in our fits. However, as discussed in §2.2, WASP-33 has quasi-periodic stellar oscillations on the order of 0.1% that need to be accounted for in order to achieve a good fit.

It is possible to model these oscillations using sinusoidal functions or wavelets, as done by other authors (i.e., Kovács et al. 2013; Deming et al. 2012; von Essen et al. 2014). However, we decided to use the Gaussian process code `celerite` (Foreman-Mackey et al., 2017) to fit these pulsations non-parametrically. Gaussian Processes treats the pulsations as a form of correlated noise whose properties are described by a parameterized covariance matrix fitted to our data. This avoids the need to impose a functional form on the oscillations, and allows the oscillation modes to depart from perfect periodicity over the course of the observation. We also tried using a combination of three sinusoids to fit the stellar pulsations, but found that for the WASP-33b 3.6 μm light curve, this decreased the standard deviation of the residuals by only 10%, compared to 54% for the GP code. Three sinusoids require six free parameters, as compared to the five parameters of our GP model.

`celerite` models the covariance matrix with a function that depends only on $\tau = |t_i - t_j|$, the time difference between two observations. We define the covariance function as the sum of two radial kernels patterned after a simple harmonic oscillator along with a diagonal kernel to represent the white noise (the latter term is functionally the same as the white noise parameter in our HD 149026 fits). The kernel $k(\tau)$ representing the correlated noise component is then:

$$k(\tau) = S_0 \omega_0 Q e^{-\frac{\omega_0 \tau}{2Q}} (\cos(\eta \omega_0 \tau) + \frac{1}{2\eta Q} \sin(\eta \omega_0 \tau)), \tag{2.4}$$

where $\eta = |1 - (4Q^2)^{-1}|^{1/2}$. Following the recommendation of Foreman-Mackey et al. (2017), we combine two of these kernels to model the stellar variations. In the first element of our combined kernel, Q is fixed to $\frac{1}{\sqrt{2}}$ while S_0 and ω_0 are allowed to vary in our fits; this represents a non-oscillatory component that decays rapidly with τ. In the second component, all three parameters are allowed to vary in order to model the oscillatory component of the stellar noise, which shows a dominant frequency of approximately 21 d^{-1} in a Lomb-Scargle periodogram. We give the second ω_0 an appropriate name: $\omega_{stellar}$. This is the parameter presented in Table

2.2. Our noise model thus consists of 6 parameters: S_0, ω_0 for the first kernel; S_0, Q, ω_0 for the second kernel; and a white noise term.

Markov Chain Monte Carlo Fits

We explore the parameter space for our model and determine best-fit parameters using a Markov chain Monte Carlo (MCMC) analysis. We carry out our fits using the `emcee` package (Foreman-Mackey et al., 2013), which is a Python implementation of an affine-invariant ensemble sampler. This approach allows for more efficient exploration of highly correlated parameters spaces, as proposed steps are generated using an ensemble of walkers whose positions are distributed along the regions of highest probability. Our models have 24–48 free parameters, depending on planet and wavelength, so we carry out our fits using 250 walkers in order to ensure sufficient sampling of the parameter space.

We obtain a starting point for our fits by using the published transit and eclipse parameters to get a model light curve, dividing the data by the model, and obtaining PLD parameters by fitting the residuals using linear regression. For WASP-33, the Gaussian process parameters were estimated by plotting the autocorrelation and manually tweaking the parameters until we achieved a good match. We then generated initial positions for each of the 250 walkers by taking the reference values calculated above and randomly perturbing each dimension. Each dimension is first perturbed by a number drawn from a normal distribution with a mean of 0 and standard deviation of 1% of the nominal value. We then do an absolute perturbation, with each dimension being perturbed by a number drawn from a normal distribution with a mean of 0 and standard deviation of 0.01. This two-step perturbation ensures that dimensions whose initial values are 0, as well as dimensions whose initial values are far from 0, are both sufficiently spread out. Although this is a very broad distribution relative to the final uncertainties in these parameters, it ensures that our fits are able to reliably identify the global maximum in the likelihood function.

We first run `emcee` for 100,000 steps, resulting in a total of 25 million steps in the combined chain. We then take the single step with the highest likelihood from this chain, which should be very close to the global maximum, and initialize a new set of 250 walkers to a Gaussian ball centered around that point. The standard deviation of this ball for a given dimension is 10^{-4} of the initial value. We then burn in this new set of walkers and run `emcee` for an additional 100,000 steps. We determine our final posterior probability distributions using the last half of this chain, and confirm

that the acceptance rate for this section of the chain is between 20 and 50 percent. We visually inspect the progress of randomly chosen walkers to check that there is no overall trend, and that the number of steps is appreciably larger than the period of quasiperiodic oscillations, if any. Finally, we check for convergence by calculating the average autocorrelation length for all walkers and ensure that the number of steps for each walker is at least ten times the length for each model parameter.

Model Selection and Optimization of Photometry

For each planet and each wavelength, we need to choose the optimal order (n in Equation 2.2) for our PLD model, the photometric aperture used to generate our light curve, and the size of the bins used in the fits. As discussed in Deming et al. (2015a) and Kammer et al. (2015), the PLD method performs better when it is fit using binned light curves.

We first determine the optimal PLD order by fitting light curves generated using an aperture of 2.5 and a bin size of 128, which are representative of the optimal values in previous fits to *Spitzer* data sets (e.g., Wong et al. 2015b). We run MCMC fits for models with PLD orders ranging between 1 and 5. In each fit, we calculate the Bayesian Information Criterion (BIC) value (Kass & Raftery, 1995) for every position in the chain, and then calculate the median BIC over the entire chain. We then select the order with the lowest median BIC for our final version of the model. A spot check reveals that using maximum BIC instead of median BIC does not change the result.

We next choose the optimal photometric aperture by repeating our MCMC fits to photometry generated using all 20 apertures, where we fix the order of our PLD model to the optimized value and keep the same bin size as in our previous fits. In this case all of our models have the same number of free parameters, and we therefore select the aperture that produces the highest median likelihood over our MCMC chain. We also consider an alternative aperture selection metric where we compare the median (best-fit) white noise parameters for each aperture on the assumption that the best aperture should have the smallest white noise value. We find that this gives very similar results to our previously adopted likelihood metric. We list our final choice of aperture for each observation in Table 2.1.

The bin size is more complicated to optimize, as it represents a trade-off between minimizing the noise on short versus long timescales. If it is too big, we average over pointing variations and this can degrade the quality of the PLD model and increase

the uncertainties in our model parameters. If it is too small, the PLD parameters adjust themselves in such a way as to minimize residuals on the shortest timescales (~seconds) at the expense of large timescale residuals—even though the latter is closer to the timescale of the astrophysical variations and can bias our estimates for the astrophysical model parameters. We consider bin sizes ranging from 1 to 4096 and find that a bin size of 128, corresponding to a time interval of 51 seconds, is a good compromise. 51 seconds is much longer than the shortest pointing jitter variations, which have a timescale of seconds, but much shorter than the timescale of any astrophysical signal. We therefore use 128 as the bin size for our final fits to all four phase curves.

The optimal results are presented in Table 2.1.

2.5 Discussion

Figures 2.3 and 2.4 each show the phase curve model, the systematics-corrected observations, and the fit residuals for the highest likelihood iteration of the MCMC chain. Figure 2.5 shows the posterior probability distributions for our HD 149026b 4.5 μm model. The triangle plot for our HD 149026b 3.6 μm model has similar correlations, while the model parameters for WASP-33b are much more Gaussian and less correlated.

In Table 2.2, we present the results of our fits, including the secondary eclipse depth, phase amplitude, and phase offset at 3.6 and 4.5 μm for each planet. Here, the secondary eclipse depth (F_p) is defined as the planetary flux at the center of eclipse divided by the stellar flux. Unlike the most commonly used secondary eclipse model, which assumes that the brightness from the planet is constant over the duration of the eclipse, this model allows for variations in the planet's brightness during this interval (see Lewis et al. 2013 for a discussion of the importance of this approach when fitting planets on highly eccentric orbits). The phase amplitude is defined as the amplitude of the sinusoidal phase curve, $A = \sqrt{c_1^2 + c_2^2}$. The phase offset is the difference in degrees between the secondary eclipse center and the phase curve maximum, with a negative offset meaning the maximum occurs before the center of the eclipse.

In Table 2.2, we also present noise properties for each light curve. These include the measured white noise (σ_{white}), the theoretical photon noise (σ_{photon}), the standard deviation of the residuals of the best fit model (σ_{tot}), and the lag-1 autocorrelation of said residuals—the last being a measure of correlated noise.

Table 2.2: Best-Fit Parameters

Parameter	WASP-33b 3.6 μm	WASP-33b 4.5 μm	HD 149026b 3.6 μm[1]	HD 149026b 4.5 μm
Eclipse depth (ppm)	3506 ± 173	4250 ± 160	430 ± 19	385 ± 23
Amplitude (ppm)	936 ± 105	1792 ± 94	189^{+27}_{-39}	164^{+22}_{-24}
Phase offset (°)	−12.8 ± 5.8	−19.8 ± 3.0	32.2^{+17}_{-15}	$-24.3^{+5.5}_{-4.7}$
Transit center (BJD_{UTC})	2456029.62604 ± 0.00016	2456024.74659 ± 0.00014	2455661.78488 ± 0.00021	2455673.28848 ± 0.00022
R_p/R_s	0.108 ± 0.001	0.103 ± 0.0011	0.0519 ± 0.0004	0.0503 ± 0.0004
a/R_*	$3.65^{+0.03}_{-0.05}$	$3.65^{+0.04}_{-0.05}$	$6.38^{+0.5}_{-0.4}$	6.67 ± 0.4
b	$0.150^{+0.072}_{-0.089}$	$0.16^{+0.08}_{-0.10}$	$0.48^{+0.09}_{-0.15}$	$0.38^{+0.12}_{-0.20}$
i	$87.6^{+1.4}_{-1.2}$	$87.6^{+1.5}_{-1.3}$	$85.6^{+1.6}_{-1.1}$	$86.7^{+1.7}_{-1.4}$
$\phi_{eclipse}$	0.50023 ± 0.00028	0.50045 ± 0.00024	0.4989 ± 0.00033	$0.50039^{+0.00050}_{-0.00044}$
slope (ppm/day)	1590 ± 203	60 ± 170	510 ± 130, 140 ± 100	−81 ± 28
σ_{white}(ppm)	356 ± 5.6	451 ± 7	270 ± 4.4, 305 ± 4.2	348 ± 3.6
σ_{photon}(ppm)	306	411	227	313
σ_{tot}(ppm)	628[2]	708[3]	290	347
$\omega_{stellar}$ (rad/day)	129 ± 1.3	130 ± 1.5	–	–
Lag-1 autocorrelation	0.69[4]	0.60[5]	0.08	−0.02

[1] The HD 149026b phase amplitude and offset should be treated with skepticism due to data quality issues; see Subsection 2.5

[2] After subtracting the Gaussian Process stellar pulsation model, the standard deviation decreases to 343 ppm for the best-fit model.

[3] Post-GP: 438 ppm

[4] Post-GP: -0.027

[5] Post-GP: -0.047

Overall quality of fits and problems with the HD 149026b 3.6 μm phase curve

For every observation except HD 149026b at 3.6 μm, our higher order PLD model appears to provide a satisfactory fit to the data; as shown in Table 2.2, these observations have a measured white noise only 10-20% higher than the photon noise limit. As shown in Figure 2.3, the HD 149026b 4.5 μm observations have reasonable residuals, with no prominent unremoved systematics. WASP-33b residuals are harder to evaluate visually, but the small measured white noise indicates that most sources of error other than photon noise have been accounted for.

In contrast to the good general quality of the other light curves, the HD 149026b 3.6 μm observation should be treated with skepticism. We find that the data strongly prefer a large positive phase offset, which is inconsistent with the negative offset at 4.5 μm and is difficult to reproduce with thermal emission from standard GCMs assuming synchronous rotation (Heng & Showman, 2015).

The data themselves are of unusually low quality. This observation is divided into two segments with a 2.4 hour gap, corresponding to a telescope downlink break. In each segment the star's position varies over an approximately oval region with a dimension of 0.5 pixels in the x direction and 0.2 pixels in the y direction, and the two oval regions are themselves separated by 0.5 pixels. During the first segment the star is relatively close to the center of the central pixel, but after the repointing required for the data downlink the star's position in the second segment falls on an adjacent pixel. It is very close to the edge, with that pixel receiving 40% of the light and the second brightest pixel receiving 20%.

All of this bodes poorly for PLD correction, or for any other correction algorithm. Not surprisingly, we find that the RMS of the fit residuals for the second segment is 17% higher than in the first segment, providing tangible evidence for the persistence of these edge effects. It should be noted that none of these problems appear in the other three data sets. Both WASP-33 observations are continuous, and the star's position shifts over an area no bigger than 0.2 by 0.2 pixels. Although the 4.5 μm observations for HD 149026 also include a downlink break in the middle, the telescope was able to return the star to approximately the same position at the end of the downlink and as a result the data from both segments span a single 0.7 x 0.2 pixel oval centered near the middle of the pixel.

We experimented with many different models for the data. We initially tried fitting the 3.6 μm HD 149026 data with a single systematics model utilizing the same 3×3 pixel postage stamp centered on the middle of the array that we used for our other

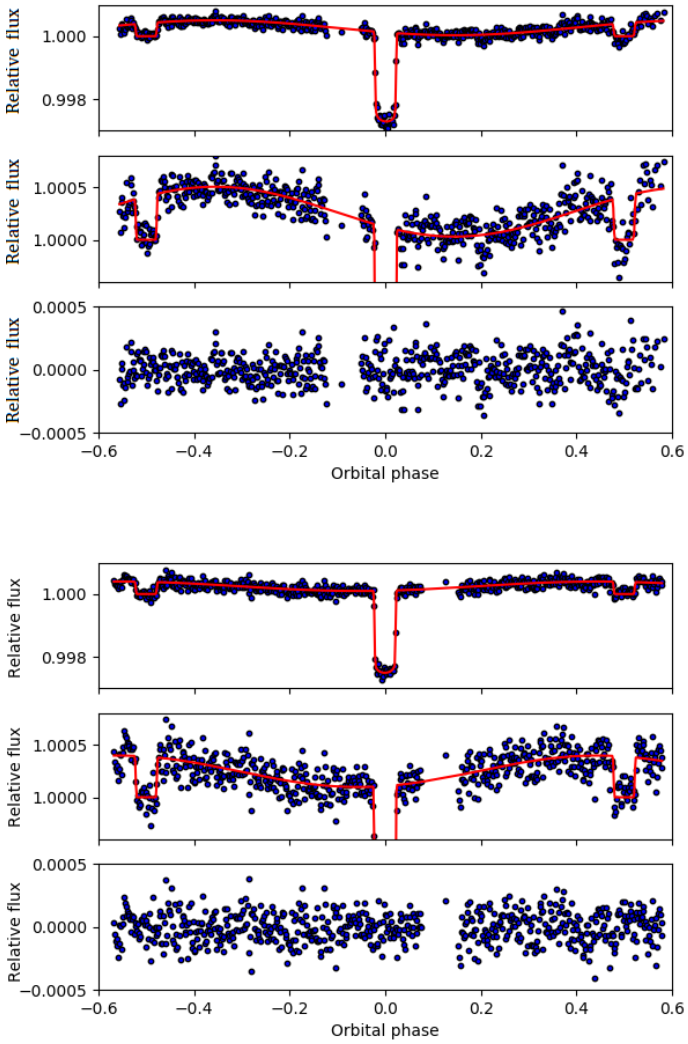


Figure 2.3: Normalized light curve for our 3.6 µm (top) and 4.5 µm (bottom) observations of HD 149026b with our instrumental noise model divided out (blue filled circles) and a representative model fit overplotted for comparison (red line). The upper two panels show the same light curve and model with a different y axis range. We show the residuals from this solution in the lower panel. All three panels use a bin size of 1024 points (6.8 minutes).

data sets. However, because the star is offset relative to this postage stamp during the second segment of data, this fit resulted in prominent systematics in the residuals for the second segment and a best-fit phase offset of approximately 60°. We then considered a separate Gaussian Process noise model for the two segments, where the second segment was represented by a simple harmonic oscillator kernel. The

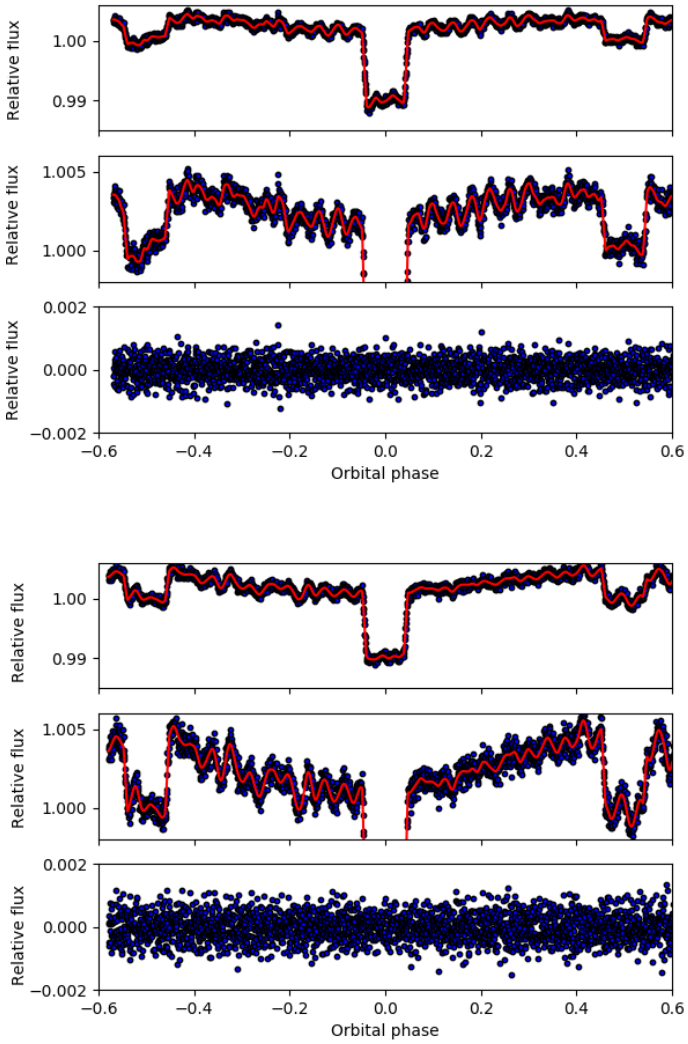


Figure 2.4: Normalized light curve for our 3.6 µm (left) and 4.5 µm (right) observations of WASP-33b with the instrumental noise model divided out (blue filled circles) and a representative model, including the Gaussian Process stellar pulsation model, overplotted for comparison (red line). All three panels use a bin size of 128 points (51 seconds). See Figure 2.3 caption for more details.

results did not change. Other models we tried including fitting only the first segment and introducing a separate linear slope for both segments.

In the end, we settled upon a separate noise and systematics model for each segment. Each segment therefore has its own PLD parameters, error parameter, and linear slope. This drastically reduced the residual systematics in the second segment. Compared to the model where both segments had the same noise and systematics model, this segmented model has 9 additional free parameters and $\Delta BIC = -926$.

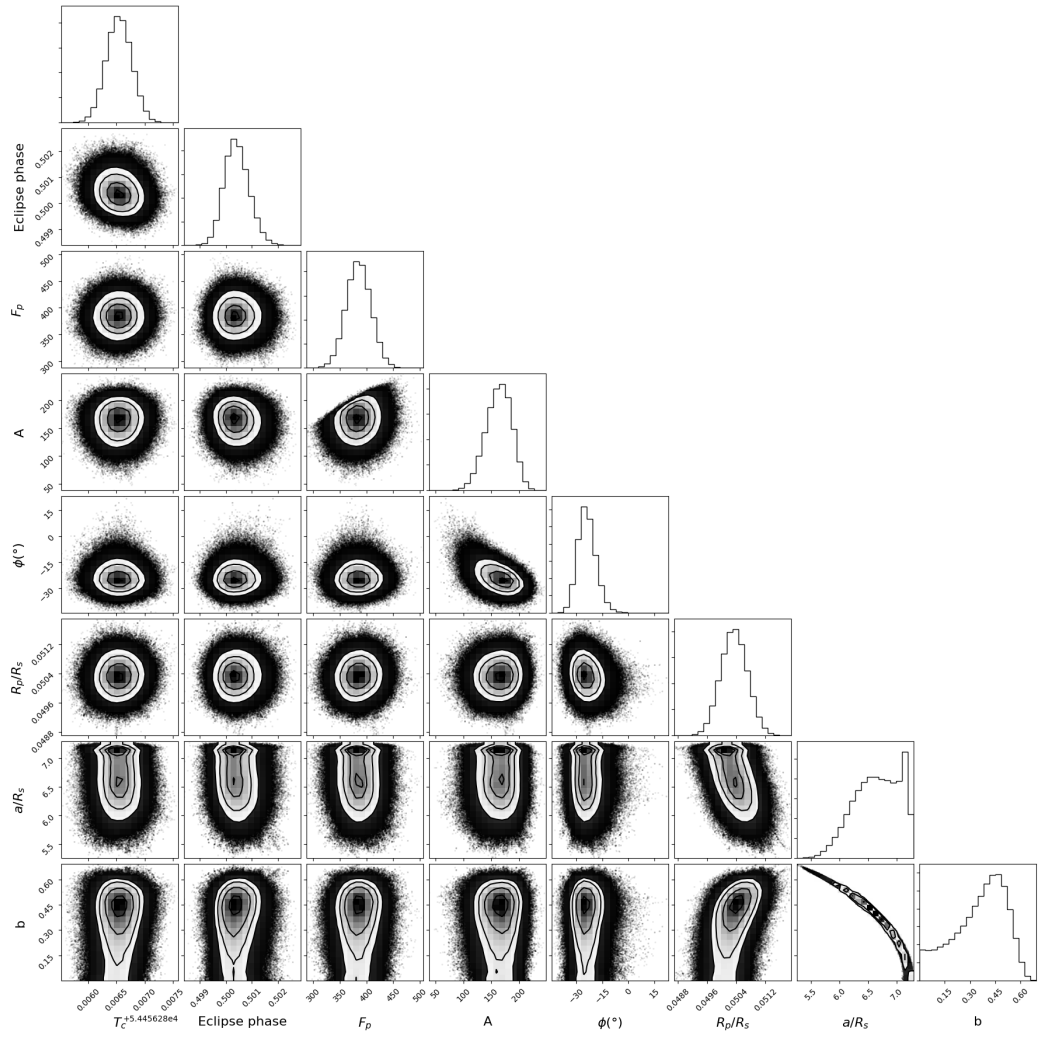


Figure 2.5: Posterior probability distributions for our fit to HD 149026b's 4.5 µm phase curve; also known as a triangle plot. The triangle plot for HD 149026b 3.6 µm is similar, while those for WASP-33b are more Gaussian and less correlated.

We found a best-fit phase offset of approximately 30°, compared to 60° with the simpler model.

We stress that the phase offset and amplitude are likely unreliable even in this improved version of the fits. We adjusted the bin size to see what effect it has on the phase offset, and found that it monotonically decreases from 80 degrees west to 80 degrees east as the bin size increases from 1 to 4096. Similarly, the phase amplitude ranges from 190 ppm to 950 ppm, although it does not change monotonically with bin size. Although we remain concerned about the reliability of the phase curve fit in this bandpass, we conclude that the measured transit should be relatively unaffected by these structures due to its short timescale and large amplitudes. The secondary

eclipse depth is somewhat more problematic. As can be seen in Figure 2.3, the light curve has visible systematics after the downlink break, with an upward fluctuation before the eclipse and a downward fluctuation during the eclipse. These likely bias the estimated eclipse depth. Despite this bias, we find that our 3.6 μm secondary eclipse depth is in agreement with the published value in this band from Stevenson et al. (2012).

Transit parameters and updated ephemerides

We recovered the transit time, a/R_*, i, and R_p/R_s from the chain. The results are shown in Table 2.2. Notably, we do not see any of the transit anomalies seen by Kovács et al. (2013) in the WASP-33b light curves. These anomalies included a 8 mmag rise in brightness across the transit (Figure 8 of Kovács et al. 2013) and a 1.5 mmag mid-transit bump (their Figure 10), both of which were seen by multiple observers.

Table 2.3: Updated Ephemerides for both planets

Parameter	HD 149026b	WASP-33b
Period (d)	2.87588874	1.21987089
T_0(BJD$_{\rm UTC}$)	2454456.78760	2454163.22367
Error in period (d)	5.9×10^{-7}	1.5×10^{-7}
Error in T_0 (d)	0.00016	0.00022

We combined our best-fit transit times for both planets with previously published transit times to calculate updated ephemerides. For HD 149026b, we used Charbonneau et al. (2006), Winn et al. (2008), Nutzman et al. (2009), Knutson et al. (2009), Carter et al. (2009) and Stevenson et al. (2012).

For WASP-33b, we used Collier Cameron et al. (2010), Smith et al. (2011), Kovács et al. (2013), von Essen et al. (2014), Johnson et al. (2015), and Turner et al. (2016). The updated ephemerides are shown in Table 2.3, and O-C plots for all transits are shown in Figure 2.6 and 2.7. We tested the goodness of fit with χ^2, finding that both planets are consistent with a linear ephemeris—WASP-33b to well within 1σ, and HD-149026b to within 2σ (p=0.07). For HD 149026b, both the period and transit timing are fully consistent with Carter et al. (2009). For WASP-33b, both the period and transit timing are fully consistent with Kovács et al. (2013).

It is notable that for both planets, the radius ratio is inconsistent between the two channels, differing by 3% for HD 149026b and 5% for WASP-33b. This could be due to imperfect modeling of stellar oscillations for WASP-33b, or to uncor-

rected systematics for both planets. The difference corresponds to roughly five atmospheric scale heights for HD 149026b and twelve scale heights for WASP-33b. Similarly large discrepancies have been reported in ground-based transit observations of other planets (e.g., Mancini et al., 2016), but these appear to be inconsistent with most model predictions as well as space-based transmission spectroscopy of similar planets (e.g., Sing et al. 2016).

Kovács et al. (2013) carried out a comprehensive analysis of ground-based WASP-33b light curves, consisting of amateur and professional data in the optical and near infrared bands. They found $R_p/R_s = 0.1143 \pm 0.0002$, which is a remarkable 6-10σ higher than our *Spitzer* values. However, the authors note anomalies in many of their data sets, including a mid-transit hump, a skewed transit shape, and discrepancies in transit depth measurements that are much larger than the formal errors.

HD 149026b has more consistent transit depths in the literature. Winn et al. (2008) found $R_p/R_s = 0.0491^{+0.0018}_{-0.0005}$ in Stromgren (b+y)/2 photometry, Nutzman et al. (2009) found $R_p/R_s = 0.05158 \pm 0.00077$ at 8 μm, and Carter et al. (2009) found $R_p/R_s = 0.05416^{+0.00091}_{-0.00070}$ with NICMOS (1.1-2.0 μm). The last measurement is $2-3\sigma$ higher than our *Spitzer* values, but our results are consistent with Winn et al. (2008) and Nutzman et al. (2009).

For HD 149026b, our measured secondary eclipse depths are fully consistent with those measured by Stevenson et al. (2012) in the same *Spitzer* bands using BLISS mapping. For WASP-33b, the eclipse depth is in good agreement with Deming et al. (2012) at 4.5 μm, although it is 1.7σ higher at 3.6 μm. This might be because of imperfect modeling of stellar pulsations, leading to underestimated error bars in both papers. Since we have a longer observational baseline over which to characterize stellar pulsations and measured two eclipses instead of one, our measurement of the eclipse depth should be less sensitive to the effects of stellar pulsations than Deming et al. (2012).

Constraints on atmospheric circulation

General circulation models

We present cloud-free GCMs for HD 149026b and WASP-33b calculated using the Substellar and Planetary Atmospheric Radiation and Circulation (SPARC) model (Showman et al., 2009), which couples the MITgcm (Adcroft et al., 2004) to a two-stream implementation of the multi-stream, plane-parallel radiative transfer code of Marley & McKay (1999). The MITgcm is an atmospheric and oceanic

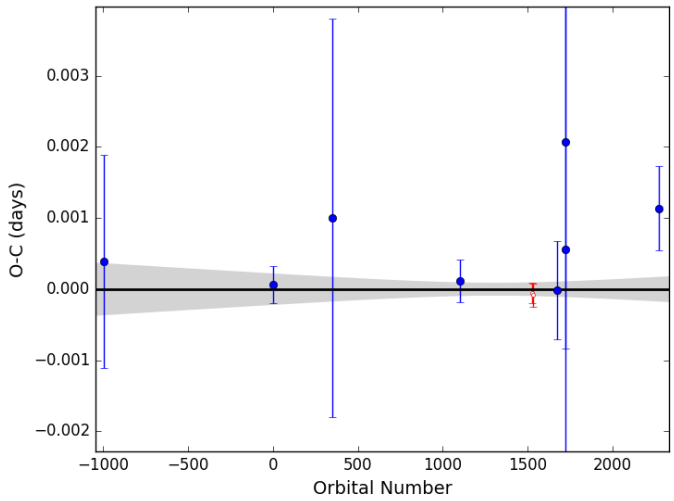


Figure 2.6: Observed minus calculated transit times for WASP-33b calculated using our updated ephemeris. Previously published results are shown as blue filled circles, and our results are shown as red open circles. The black line indicates the predicted transit times at each epoch assuming a constant ephemeris, and the gray region indicates the 1σ confidence interval.

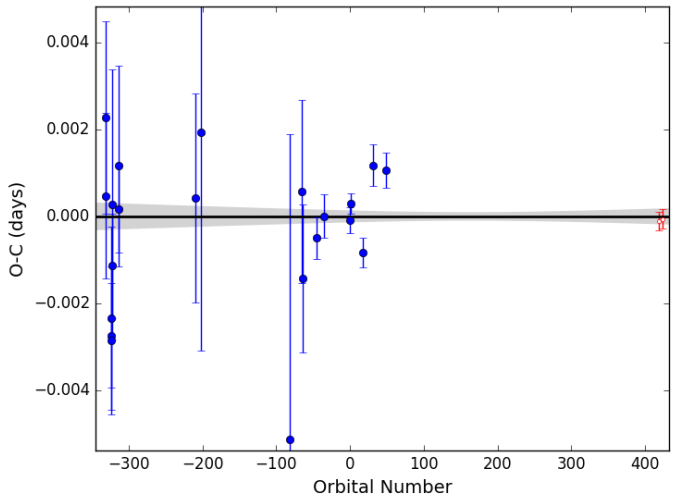


Figure 2.7: Observed minus calculated transit times for HD 149026b calculated using our updated ephemeris; see Fig. 7 caption for more details.

circulation model that solves the primitive equations, which are relevant for stably stratified atmospheres with large horizontal/vertical aspect ratios (generally true for hot Jupiters). The equations are solved using a finite-volume discretization on a cubed-sphere grid, which allows longer time steps and increases the accuracy near the poles as compared to a traditional longitude-latitude grid. The radiative transfer

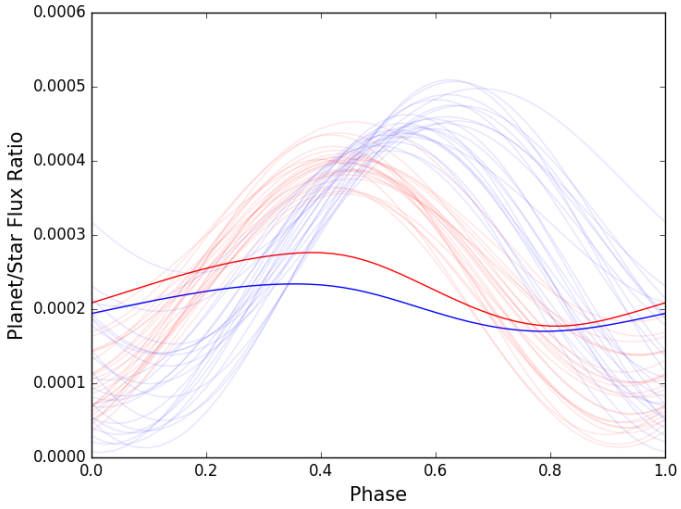
Table 2.4: Phase curve parameters from GCMs

Planet	TiO?	Metallicity	Band (μm)	F_p(ppm)	A (ppm)	$\phi(^\circ)$
WASP-33b	No	1×	3.6	4086	1664	−9.9
WASP-33b	No	1×	4.5	4597	1903	−9.6
WASP-33b	Yes	1×	3.6	4151	1688	−8.8
WASP-33b	Yes	1×	4.5	4779	1981	−7.6
HD 149026b	No	30×	3.6	430	167	−11
HD 149026b	No	30×	4.5	616	252	−7
HD 149026b	No	1×	3.6	219	31.9	−55
HD 149026b	No	1×	4.5	260	49.6	−45

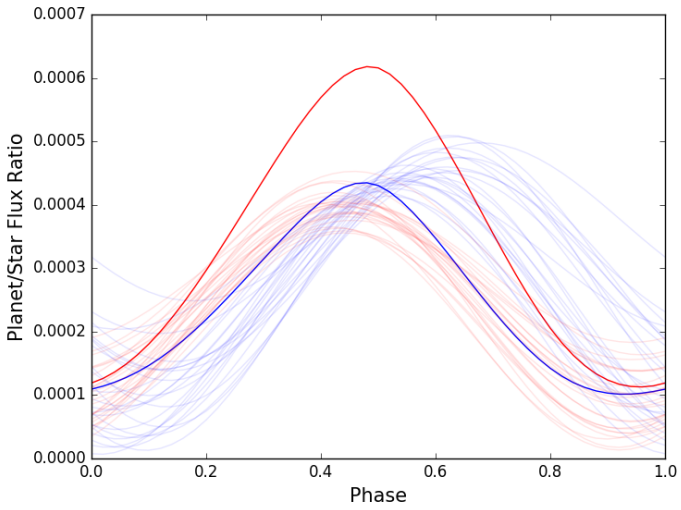
code employs the correlated-k method with 11 bands optimized for accuracy and computational efficiency. The opacities are calculated assuming local thermodynamic and chemical equilibrium. This code has been used extensively to model the atmospheric circulation of exoplanets over a wide range of planetary properties (e.g., Lewis et al. 2010; Kataria et al. 2015, 2016; Wakeford et al. 2017). After running the GCM, we extract light curves following the method of Fortney et al. (2006).

We list the predicted eclipse depths, amplitudes, and phase offsets at 3.6 and 4.5 μm for each model in Table 2.4; these can be compared directly to the measured values in Table 2.2.

For HD 149026b, we consider models with solar and 30× metallicity and compare the resulting phase curves to our best-fit phase curve model in Figure 2.8. We only consider models without TiO, as Stevenson et al. (2012) found that this planet's dayside emission spectrum was best described by a model without a temperature inversion. We find that the solar metallicity GCM predicts a relatively small phase curve amplitude in both bandpasses, in sharp disagreement with our data. The 30× solar metallicity model has a higher opacity in both *Spitzer* bandpasses and therefore probes lower pressures (higher altitudes) than the solar metallicity model, leading to larger predicted phase curve amplitudes. This model comes closer to matching the data, although it underestimates the amplitude at 3.6 μm and overestimates it at 4.5 μm. We note that neither the 1D models shown in Stevenson et al. (2012) nor the 3D GCMs are able to match the measured secondary eclipse depths at 3.6 and 4.5 μm, and speculate that these discrepancies in both secondary eclipse depths and phase curve amplitudes might be resolved by increasing the amount of CO and/or CO_2 in the atmosphere. Both CO and CO_2 have absorption bands in the 4.5 μm *Spitzer* bandpass, increasing their abundance will accordingly decrease the planet's



(a) Solar metallicity

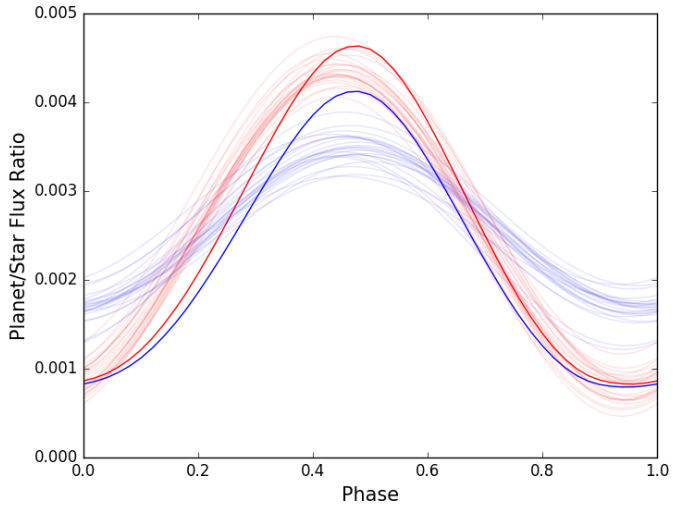


(b) 30× solar metallicity

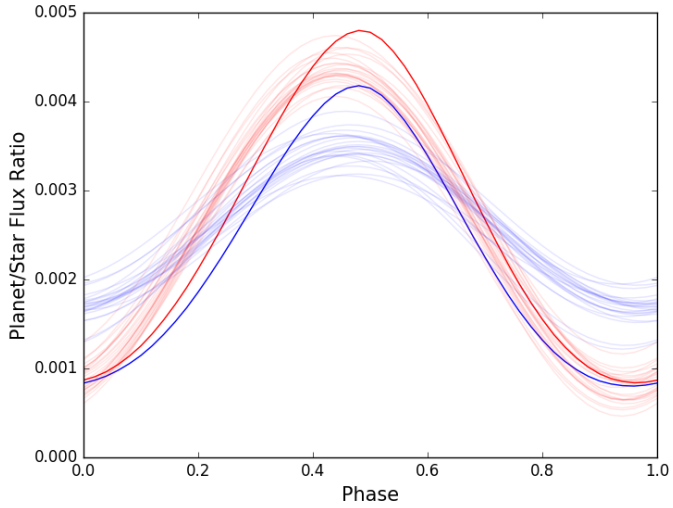
Figure 2.8: Comparison of the GCM-generated phase curves (thick lines) for HD 149026b with our measured phase curves (thin lines). The 24 thin lines each represent one randomly selected MCMC step, and the dispersion in these lines is therefore representative of the uncertainties in the measured phase curve shape. No TiO is included. 3.6 μm results are plotted as blue curves while 4.5 μm results are in red.

brightness in this band relative to the 3.6 μm band. Previous models for GJ 436b (Moses et al., 2013; Morley et al., 2017) serve as a useful demonstration of the effect of very high atmospheric metallicities ($> 200 - 300\times$ solar) on the strength of the CO absorption in the 4.5 μm band. On the planet's night side, which is cool

enough to fall near the transition from CO to methane-dominated carbon chemistry, disequilibrium chemistry due to quenching and horizontal transport could increase the relative amount of CO and CO_2 (Cooper & Showman, 2006).



(a) Without TiO



(b) With TiO

Figure 2.9: Comparison of GCM-generated phase curves (thick lines) for WASP-33b with our measured phase curves (thin lines). The 24 thin lines each represent a randomly selected MCMC step, and the dispersion in these lines is therefore representative of the uncertainties in the measured phase curve shape. Both models assume solar metallicity. 3.6 μm results are plotted as blue curves and 4.5 μm results are in red.

For WASP-33b, which has a much lower bulk density than HD 149026b, we consider only the solar metallicity case for our GCMs. As discussed in Section 2.2, this planet is one of the most highly irradiated hot Jupiters discovered to date; its dayside emission spectrum is best matched by models with a temperature inversion and appears to hint at the presence of gas-phase TiO. We show predictions for two models in Figure 2.9, including one with and the other without TiO, in order to evaluate the effect of this molecule on its dayside emission spectrum and day-night circulation. We find that the differences in the phase curves for these two models are fairly subtle, and although the data are somewhat better matched by the model without TiO, both models disagree with the observations at the 2σ level. In this case the models predict a larger phase curve amplitude and secondary eclipse depth in both bands.

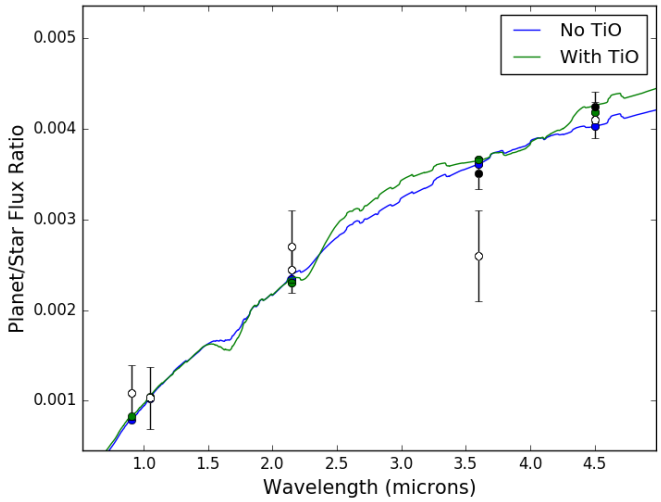


Figure 2.10: Model emission spectrum for WASP-33b, compared with observations (black filled circles for our values, black open circles for literature values from Smith et al. 2011; von Essen et al. 2015; Deming et al. 2012; de Mooij et al. 2013). Each blue or green point represents the band-averaged flux ratio corresponding to the observation at the same wavelength.

In Figure 2.10 and 2.11, we show the GCM-derived emission spectra for the two planets. Our WASP-33b observations are consistent with both GCMs, which have very similar emission spectra at *Spitzer* wavelengths. HD 149026b, on the other hand, is highly inconsistent with both GCMs in the *Spitzer* 3.6 and 4.5 μm bands. Interestingly, eclipse observations in the 5.8, 8, and 16 μm *Spitzer* bands seem to favor the solar metallicity model over the 30× solar metallicity model. We speculate that this could be because the GCMs assume gas-only opacity with no hazes or

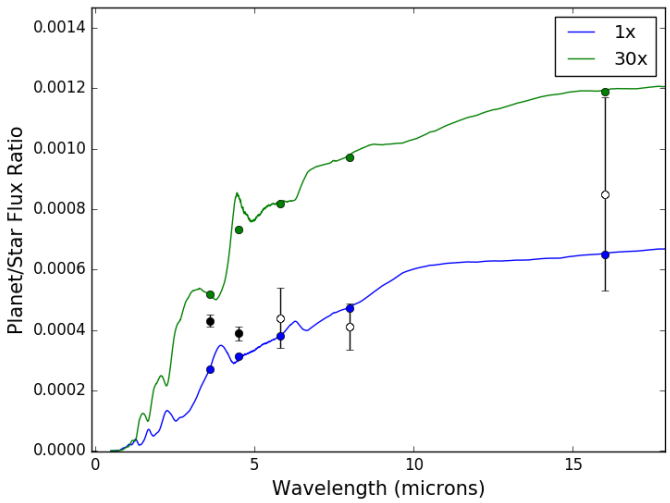


Figure 2.11: Model emission spectrum for HD 149026b, compared with observations (black filled circles for our values, black open circles for literature values from Stevenson et al. (2012)). Each blue or green point represents the band-averaged flux ratio corresponding to the observation at the same wavelength.

clouds, resulting in a very low albedo; in contrast, our phase curve data appear to favor a high albedo (see discussion in Subsection 2.5). A high albedo would cool the planet's dayside, bringing the 30x model into bet- ter agreement with the data. However, even with a reduced amplitude the 30x solar model is a poor match for the observed spectral slope across the 3.6-4.5 um bands; this could be explored with additional models in the future. We note that our preference for the high metallicity model is primarily driven by the need to match the large observed phase curve amplitudes in both bands rather than the shape of the planet's dayside emission spectrum. As an alternative to changing the planet's dayside albedo, the addition of localized clouds on the planet's night side could increase the phase curve amplitude for the solar metallicity model (e.g., Stevenson 2016), bringing it into better agreement with our observations.

A simple toy model for albedo and recirculation efficiency

In this section, we use a simple toy model first presented in Cowan & Agol (2011) to calculate average brightness temperatures, albedos, and circulation efficiencies for the two planets under the assumption that their thermal emission is well-approximated by a blackbody and that these observations probe a similar range in pressures across all wavelengths and longitudes. In this model, a planet's thermal

phase curve is described by two parameters: a Bond albedo (A_B) and a heat redistribution efficiency (ε). Planets absorb a fraction $1 - A_B$ of the stellar flux on their daysides, redistribute this energy to the nightside with the stated efficiency, and then emit as a blackbody. In the $\varepsilon = 1$ case, the entire planet has the same temperature, and energy balance gives $T_p = \frac{T_0}{\sqrt{2}}(1 - A_B)^{1/4}$ where $T_0 = \frac{T_s}{\sqrt{a_*}}$ and $a_* = \frac{a}{R_s}$. For the $\varepsilon = 0$ case (i.e., no heat redistribution to the night side), the corresponding dayside temperature is given by $T_d = (2/3)^{1/4}(1 - A_B)^{1/4}T_0$ and the nightside temperature is zero. If we define ε so that it linearly interpolates between these two extremes we arrive at an analytic description for ε and A_B as a function of the effective blackbody temperatures for the planet's dayside (T_d) and nightside (T_n):

$$\varepsilon = \frac{8}{5 + 3(T_d/T_n)^4} \tag{2.5}$$

$$A_B = 1 - \frac{5T_n^4 + 3T_d^4}{2T_0^4}. \tag{2.6}$$

Here, and throughout the paper, we define the nightside and dayside to mean an orbital phase of 0 and 0.5, respectively. Previous studies have alternated between this definition and one in which day and night correspond to the hottest and coldest hemispheres on the planet.

The hemisphere-averaged planet brightness temperature is given by Equation 6 in Cowan & Agol (2011), and depends on two things: the brightness temperature of the star at the observed wavelength, and the ratio $\psi(\lambda)$ between normalized planetary flux and transit depth at the observed wavelength. We first calculate the stellar brightness temperature in each band using the closest model in the BT-NextGen (AGSS2009) spectral grid, as provided by the Spanish Virtual Observatory[2]. The spectral grid spacing is fine enough that choosing the adjacent model changes the brightness temperature by less than 1%. As a check on BT-NextGen, we also calculated brightness temperatures using Phoenix models (Husser et al., 2013), and found that the results differed on average by only 0.4%. We next calculate the planet-star brightness ratio $\psi(\lambda)$ for the day side by dividing the eclipse depth F_p by the transit depth, and for the night side by dividing $F_p - 2A\cos\phi$ by the transit depth, where A is the phase amplitude and ϕ is the phase offset. We then convert this to a brightness temperature for the planet using the stellar brightness temperature calculated earlier.

[2] `http://svo2.cab.inta-csic.es/theory/newov2/index.php`

Table 2.5: Dayside and nightside brightness temperatures, Bond albedo, and recirculation efficiency for each channel

Planet	$\lambda(\mu m)$	$T_{b,day}(K)$	$T_{b,night}(K)$	A_B	ε
WASP-33b	3.6	3082 ± 92	1952^{+125}_{-134}	$0.25^{+0.09}_{-0.10}$	0.34 ± 0.06
WASP-33b	4.5	3209^{+89}_{-87}	1498^{+114}_{-118}	$0.25^{+0.08}_{-0.09}$	0.12 ± 0.03
HD 149026b	3.6	1941 ± 46	1133^{+290}_{-270}	$0.36^{+0.10}_{-0.16}$	$0.26^{+0.26}_{-0.16}$?
HD 149026b	4.5	1649 ± 49	1018^{+115}_{-116}	$0.66^{+0.05}_{-0.06}$	$0.31^{+0.11}_{-0.10}$

Table 2.6: Averaged brightness temperatures, Bond albedo, and recirculation efficiencies

Parameter	WASP-33b	HD 149026b
$T_{\mathrm{d}ay}$	3144 ± 114	1804 ± 98
$T_{\mathrm{n}ight}$	1757 ± 88	1032 ± 120
A_B	$0.25^{+0.09}_{-0.10}$	$0.53^{+0.09}_{-0.11}$
ε	$0.22^{+0.05}_{-0.04}$	$0.24^{+0.11}_{-0.09}$
$T_{\varepsilon=0}$	3514 ± 30	2276 ± 37
T_d/T_0	0.81 ± 0.04	0.72 ± 0.04

We obtain uncertainties on these brightness temperature estimates using the posterior probability distributions from our MCMC fit. For each step in our MCMC chain, we calculate $\psi(\lambda)$ for the day and night sides from the chain itself. With these parameters, we then calculated $\mathrm{T_d}$ and $\mathrm{T_n}$ in each bandpass, using the error-weighted average of both bandpasses to calculate the planet's corresponding albedo and recirculation efficiency. Finally, we compare to Figure 7 in Cowan & Agol (2011) by calculating the quantities $T_{\varepsilon=0} = (2/3)^{1/4}T_0$ and $T_d/T_0 = (1 - A_b)^{1/4}(\frac{2}{3} - \frac{5}{12}\varepsilon)$ for each planet. Although WASP-33b has a higher $T_{\varepsilon=0}$ than all of the planets in Figure 7 of Cowan & Agol (2011), we find that its temperature ratio is fully consistent with that of other highly irradiated ($T_{\varepsilon=0} > 2500K$) planets. Similarly, T_d/T_0 for HD 149026b is in good agreement with the values for other planets with similar irradiation levels despite lingering questions about the reliability of the 3.6 μm results. We list the relevant values for each planet in Tables 2.5 and 2.6.

Comparison with other planets

WASP-33b is very unusual among the more than two hundred hot Jupiters discovered to date, being the second most irradiated hot Jupiter currently known (KELT-9b being the first). Despite this peculiarity, its albedo and recirculation efficiency appear largely similar to those of other hot Jupiters observed to date. In order to compare our planets to other hot Jupiters, we produce an updated version of Figure

3 from Schwartz et al. (2017), which plots contours corresponding to the albedo and efficiency values estimated in 2.5. Our version of the plot is shown in Figure 2.12a. Although we largely follow the method described in this paper, our approach differs in several aspects:

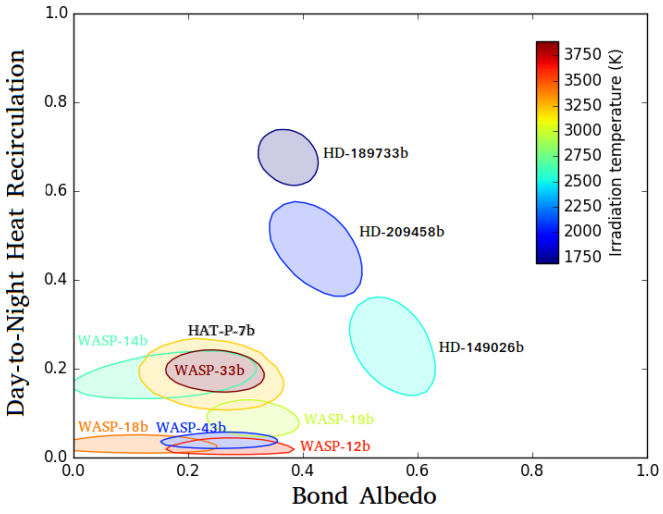
1. Uncertainties are propagated using a Monte Carlo method, instead of dividing up the albedo-efficiency parameter space into cells and computing χ^2 for each cell.

2. If the dayside or nightside flux in an iteration is negative, we exclude the entire iteration, while Schwartz et al. (2017) set the corresponding temperature to zero. This tends to slightly lower the recirculation efficiency.

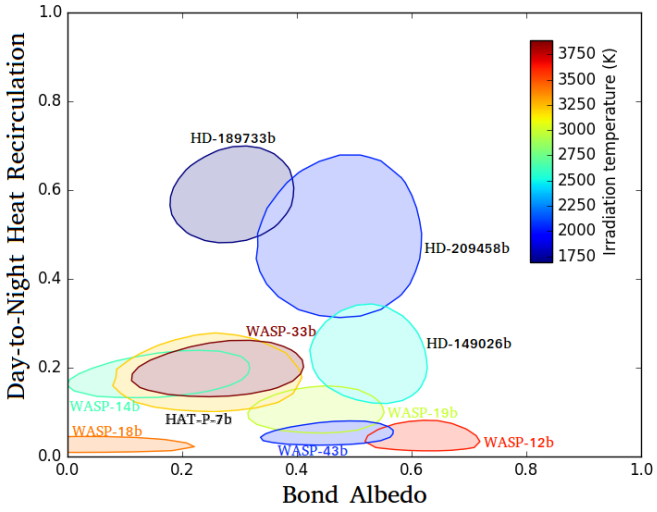
3. Schwartz et al. (2017) assume a geometric albedo of 7% and subtract the reflected light eclipse depth from the measured eclipse depth. The actual geometric albedos of these planets are poorly constrained by current observations and an assumed albedo of 7% has a negligible effect on our results, so we instead assume an albedo of zero.

Another complication is in the treatment of WASP-12b. The WASP-12b phase curve paper (Cowan et al., 2012) included results from two analysis methods: polynomial fitting and point-by-point decorrelation. We used the latter set of phase curve and eclipse depth parameters in our paper as it results in a more consistent phase curve offset between the two bands. Although the two methods produced similar phase curve parameters for WASP-12b at 4.5 μm, they were very different for 3.6 μm, and our results differ substantially depending on which version we choose. We downloaded the data in each bandpass ourselves, as well as an additional pair of phase curve observations taken in 2013. All four data sets were analyzed with the same higher-order PLD approach used for WASP-33b and HD 149026b. For 4.5 μm, our two results were consistent with each other and with both methods in Cowan et al. (2012). For 3.6 μm, our two results were consistent neither with each other nor with either method in Cowan et al. (2012). We therefore conclude that the properties of WASP-12b are not well-constrained by the current observations, although we still show it in our plots.

We show two versions of the albedo-efficiency plot in Figure 2.12 , including one with all published thermal secondary eclipse and phase curve data and another



(a) All wavelengths included



(b) Only *Spitzer* 3.6 and 4.5 μm data included

Figure 2.12: Albedo and recirculation efficiency for all exoplanets with published infrared phase curves, calculated following Schwartz et al. (2017), assuming the planet radiates as a blackbody.

which only considers 3.6 and 4.5 μm *Spitzer* data in order to ensure a more uniform analysis. The only planet that moved significantly was WASP-12b.

This figure shows that WASP-33b has a Bond albedo and recirculation efficiency that appear largely similar to those of other hot Jupiters despite its high irradiation level. HD 149026b, however, appears to have an unusually high albedo in our toy model. Its best-fit albedo is higher than that of all other planets with thermal phase curves, and it

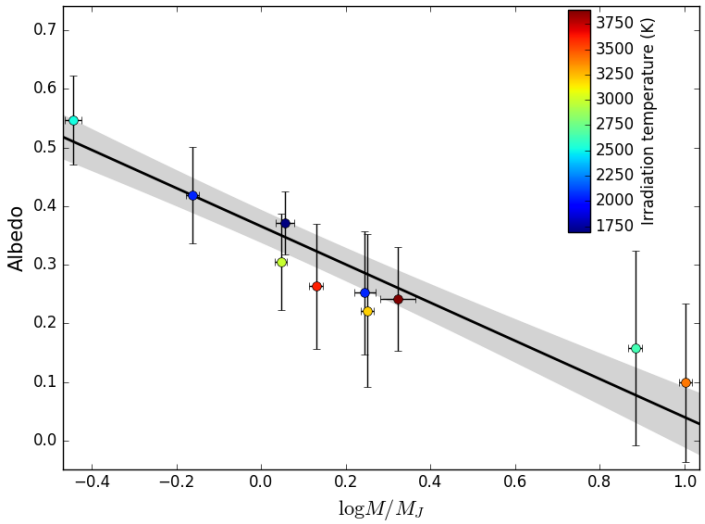


Figure 2.13: Logarithm of mass vs. albedo for exoplanets with phase curves. Albedo is calculated using the simple toy model described in Schwartz et al. (2017), in which the planet is assumed to radiate as a blackbody. We also overplot the best-fit linear function as a black line, with the 1σ confidence interval shown in gray.

is also higher than any of the optical geometric albedos measured by Kepler as shown by Figure 7 of Schwartz & Cowan (2015). This might reasonably be explained by the presence of a reflective cloud layer in this planet's upper atmosphere; the presence of such a cloud layer could be confirmed with future transmission spectroscopy.

Parmentier et al. (2016) calculated the effective cloud coverage for a range of equilibrium temperatures, three cloud top pressures, and a number of cloud compositions. Although they focused on investigating the role of clouds at optical wavelengths rather than in the *Spitzer* bands, we can nonetheless utilize their results to explore the potential cloud species that might be present in HD 149026b's atmosphere. They report evidence for the presence of silicate clouds for $T_{eq} > 1600K$, presence of MnS clouds for $T_{eq} < 1600K$, and absence of silicate clouds for $T_{eq} < 1600K$. HD 149026b has a zero-albedo equilibrium temperature of 1700 K, very close to the 1600 K dividing line. If there are silicate clouds, their Figure 13 shows that the dayside effective cloud coverage is expected to be 30-80%, depending on the cloud top pressure, while the nightside coverage is 50-100%. If there are MnS clouds but no silicate clouds, the dayside cloud coverage would be 0-20%, while the nightside cloud coverage would be 20-100%.

Mahapatra et al. (2017) used a kinetic, non-equilibrium cloud formation model to study cloud structures and compositions. For HD 149026b, they found that clouds

are likely composed of many different species, with TiO_2 dominant at the cloud top ($10^{2.5}$ bar) while species such as Fe, SiO, and $MgSiO_3$ are common deeper down.

One of the most interesting aspects of Figure 2.12 is that it shows no obvious correlations between irradiation temperature, albedo, and efficiency. In fact, it appears that planets with very different irradiation temperatures can have similar albedos and recirculation efficiencies. We plotted mass versus efficiency, irradiation temperature versus albedo, and irradiation temperature versus efficiency, finding that the two least massive planets in our sample—HD 149026b and HD 209458b—are outliers in both albedo and efficiency. In Wong et al. (2015b) we previously suggested a possible correlation between mass and albedo, but after additional data was collected, we concluded in Wong et al. (2016) that a simple mass-albedo correlation was no longer tenable. However, we see a strong correlation when we plot these two parameters in Figure 2.13. The linear model has a lower BIC than the constant-albedo model ($\Delta BIC = -12.5$), indicating a strong preference for the linear model. In addition, the error bars on the albedo seem overestimated, possibly due to the large systematic error we deliberately introduce (in accordance with Cowan & Agol 2011) in converting from brightness temperature to physical temperature. After subtracting the best fit linear model, we find $\chi^2 = 1.86$ for the residuals; with 8 degrees of freedom, there is only a 1.5% probability of obtaining a χ^2 this low. If the errors were correctly estimated, ΔBIC would be even more negative, preferring the linear model even more strongly. Our best fit line has slope $m = -0.326 \pm 0.047$ and intercept $b = 0.366 \pm 0.016$. Coincidentally, 0.366 is very close to the Bond albedo of Jupiter itself.

The physical explanation for the decrease in albedo with mass is unclear. One possibility is that increased surface gravity makes it harder for cloud particles to be kept aloft, as shown in Equation 10 of Heng & Demory (2013). The increased cloudiness at low surface gravity has been observed on brown dwarfs (Faherty et al., 2016). The main difficulty with this explanation is that we have also plotted the relation between surface gravity and albedo, and although an anticorrelation is seen, it is much less statistically significant than the mass-albedo correlation ($\Delta BIC = -2.4$, compared to $\Delta BIC = -12.5$).

Phase curve offsets

The toy model discussed above derives recirculation efficiency from the observed nightside flux. However, the nightside flux can only be measured by reference to

the secondary eclipse, which is hours or days away. This makes the measurement particularly sensitive to instrumental noise sources on long timescales, including the long-term pointing drift present in many phase curve observations. As an example, the *Spitzer* phase curves for WASP-43b imply a negative nightside flux, which is unphysical (Keating & Cowan, 2017).

In this section we explore correlations between phase offset and other planetary parameters. In GCMs, both the size of the phase offset and the relative temperature gradient between the day and night sides increase with increasing depth (pressure) in the atmosphere (e.g., Showman et al. 2009). We therefore consider whether or not the measured phase offset might be useful as a proxy for recirculation efficiency. We plotted recirculation efficiency against phase offset and found that although planets with very large phase offsets have somewhat high efficiencies and planets with very small phase offsets have somewhat low efficiencies, the correlation is by no means exact. We conclude that either phase offset is an imperfect proxy for recirculation efficiency in practice, or the recirculation efficiency calculated using our simple toy model is simply not accurate enough for the correlation to be obvious.

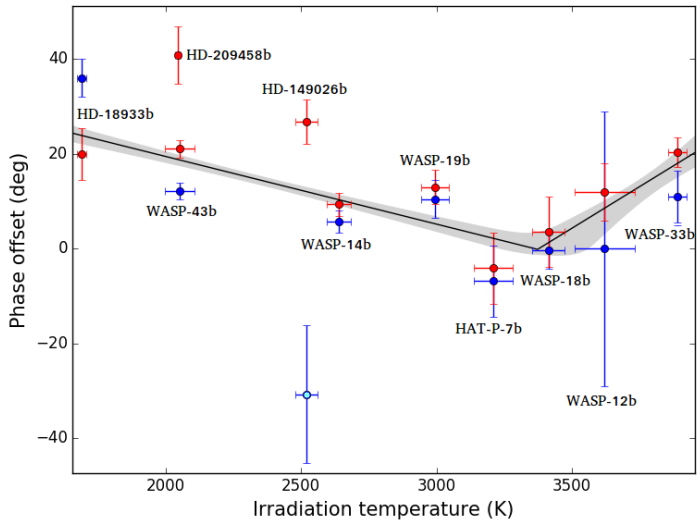


Figure 2.14: Phase offset vs. irradiation temperature for all hot Jupiters on circular orbits with thermal phase curves. Blue represents 3.6 μm while red represents 4.5 μm. The light blue point is the 3.6 μm observation for HD 149026b, which we discuss in §2.5. The black lines represent the best-fit bilinear model, while the gray region indicates the 1σ confidence interval.

Figure 2.14 shows a strong correlation between a planet's irradiation temperature $T_0 = T_{eff}/\sqrt{a_*}$ and its phase offset. There is a clear downward trend until $3400K$,

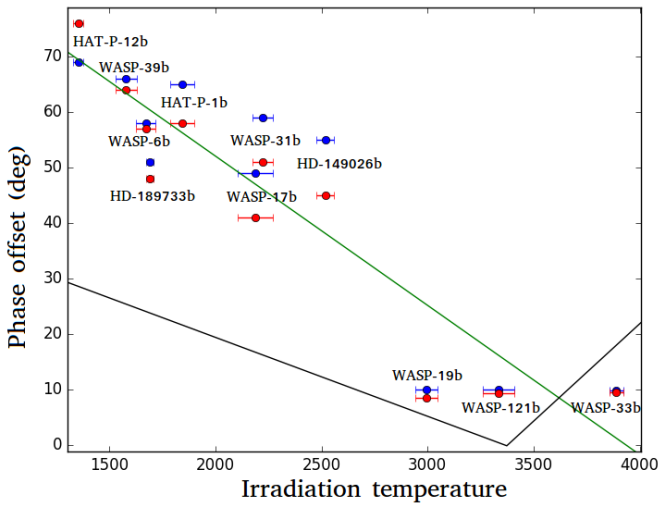


Figure 2.15: Phase offset vs. irradiation temperature for GCM-modeled planets. The green line is the best fit linear model to the GCM data, while the black lines represent the best fit bilinear model to the observations. The black lines are identical to the ones in Figure 2.14.

after which the trend reverses direction. We tested the significance of the trend by fitting the data with five models: a constant phase model, a linear model, a bilinear model, a bilinear model with the slope of the second line segment fixed to 0, and a bilinear model with both the slope and intercept of the second line segment fixed to 0. We obtain a BIC of 185, 168, 136, 145, and 186, respectively. Thus, the reversal at $3400K$, despite being based on only three data points, is significant from a purely statistical perspective. The fit results for the first line segment are $b_1 = 47.7° \pm 4.7°$, $m_1 = -0.014 \pm 0.002$ deg/K; for the crossover point, $T_c = 3410 \pm 110$; and for the second line segment, $m_2 = 0.039^{+0.017}_{-0.011}$ deg/K.

Despite the statistical significance, the rise is still based only three planets, and may not be real. The test assumes Gaussian errors, while the actual errors are in reality both asymetric and non-Gaussian. Even more importantly, *Spitzer* light curves are notorious for having bizarre and unexplained instrumental systematics which can affect the fitted parameters in ways that are subtle and difficult to diagnose. In this paper alone, we have seen this for HD 149026b in the 3.6 µm band, and WASP-12b for two different observations in the 3.6 µm band. Other examples of problematic behavior include Stevenson et al. (2017), where two separate visits in the same band resulted in very different nightside fluxes. A sprinkling of unmodelled

systematics, plus a smattering of bad luck, could be sufficient to destroy the final rise in temperature.

On the other hand, there are physical reasons to be less skeptical. First, although the reversal at $3400K$ has never been predicted or previously noted, the initial drop is unsurprising—phase offsets are expected to decrease with increasing temperature because the radiative timescale drops steeply with temperature (Perez-Becker & Showman, 2013; Komacek & Showman, 2016). Second, it is also clear that some kind of break must occur at $3400K$—if the downward trend continued, the phase offset would become westward at higher temperatures, which is physically implausible.

The only significant outlier in this trend is the 3.6 μm observation for HD 149026b, which we discuss in §2.5. Additionally, as discussed in §2.5, Cowan et al. (2012) presented two contradictory sets of results for WASP-12b's phase curve, based on two distinct analysis methods. Had we used the other version, the 4.5 μm phase curve offset would have been nearly identical, but the 3.6 μm phase curve offset would be at a physically implausible 53°—another clear outlier. The fact that both potential outliers in this plot are based on problematic data sets gives confidence to the reality of the trend. The trend is even more striking when one considers that it is between two relatively reliably measured quantities. The irradiation temperature is dependent only on the stellar effective temperature and a/R_*, both of which are easily measured. The phase offset is harder to measure (e.g., §2.5), but unlike albedos and efficiencies, it is a purely empirical quantity.

To understand this trend, we took previously published SPARC GCM simulations (Kataria et al., 2016) and plotted phase offset against irradiation temperature in Figure 2.15. The sample of planets in these GCMs is different from those presented in the observations. All models are at solar metallicity and have no TiO, in order to ensure a more uniform comparison. As expected, the predicted phase offsets from these models decrease with increasing irradiation temperature until approximately 3000 K. The offsets decrease at a rate of -0.017 deg/K. However, instead of rising towards the highest temperatures they instead plateau around a minimum phase offset of approximately 10°. It is also worth noting that in these models the 3.6 μm bandpass has a larger phase shift (by an average of 4.1°) for all but the coolest planet, indicating that this wavelength probes deeper into the atmosphere. In our observations, the 4.5 μm phase curves have larger phase shifts (and thus deeper photospheres) than the 3.6 μm phase curves for every planet except the coldest,

the average difference being 6.2 degrees. The universality of this trend among our relatively diverse sample of hot Jupiters is suggestive, and should help guide future modeling efforts in this area.

There are a combination of factors that may or may not explain the discrepancy between observations and GCMs. For example, the addition of high altitude clouds to the GCMs could help by decreasing the dayside photospheric pressure, which would systematically reduce the size of the predicted phase offsets and provide a better match to the observational data. Such clouds appear to provide a good match to the optical phase curve offsets measured for the hot Jupiters located in the Kepler field (Parmentier et al., 2016; Demory et al., 2013; Shporer & Hu, 2015; Angerhausen et al., 2015) and have also been postulated to explain other infrared phase-curve observations (e.g. Kataria et al., 2015; Stevenson et al., 2017). However, Roman & Rauscher (2017) complicate this explanation by showing that clouds do not always lead to lower phase offsets. They find that in the case of Kepler-7b, inhomogenous clouds distributed along the western terminator result in a higher phase offset, while global clouds result in a marginally lower offset, compared to the clear atmosphere case. A super-solar metallicity atmosphere could also provide a similar effect; enhanced metallicity results in enhanced opacities, such that the photosphere is higher in the atmosphere where the day-night contrast is larger and the phase offset is smaller (e.g., Kataria et al. 2015). At high temperatures, the presence of dayside temperature inversions produced by gas-phase TiO/VO might also affect the predicted phase offsets, as this will change the opacity of the atmosphere and hence what altitudes are probed (e.g., Showman et al. 2009).

MHD effects such as Lorentz drag and Ohmic dissipation are also likely to be important (e.g., Perna et al. 2010; Batygin & Stevenson 2010; Menou 2012; Rauscher & Menou 2012; Ginzburg & Sari 2016). Because Lorentz drag and Ohmic dissipation are two facets of the same underlying processes, we can turn to the literature on Ohmic dissipation and radius inflation in hot Jupiters to determine the regime in which these effects become important. For hot Jupiters with an appreciable magnetic field, the effect of this magnetic field on the atmospheric circulation will depend on the ionization fraction of the planet's upper atmosphere. Previous models have concluded that alkali metals such as Na and K will provide the dominant source of ions in these atmospheres (e.g., Batygin & Stevenson 2010). As the ionization fraction increases the strength of the Lorentz drag and also the amount of radius inflation due to Ohmic dissipation increase as well, resulting in a peak radius inflation

at equilibrium temperatures of around 1500 K. At higher temperatures, the atmospheric circulation is effectively suppressed by the magnetic drag and less energy is deposited in the planet's interior. Thorngren & Fortney (2017) find compelling evidence for a peak in the radius inflation of hot Jupiters around 1500 K and a decline thereafter, in good agreement with these models. With this picture in mind, the addition of Lorentz drag around 1500 K is likely the explanation for why the observed phase offsets at higher temperatures decrease to values consistent with zero, while the GCMs predict a minimum phase offset around 10°. However, this simple picture appears to conflict with the observed increase in phase offset for the most highly irradiated planets, as the amount of Lorentz drag should remain constant over this highly irradiated regime. Although there has been some recent work on atmospheric circulation in the MHD-dominated regime (e.g., Rogers 2017), it is not yet clear whether more careful modeling can reproduce the observed trend in phase offsets at these temperatures.

As a last point, we consider possible explanations for the relative offsets observed between the two *Spitzer* bands. As noted earlier, the measured 4.5 μm offset is consistent larger than the measured 3.6 μm offset for every planet except for HD 189733b, but the opposite is consistently true in the model predictions for these planets. Atmospheric chemistry would seem to be the obvious explanation: CH_4 is a major absorber within the 3.6 μm bandpass, while CO is a strong absorber within the 4.5 μm bandpass, so the relative abundances of these two molecules could easily shift the relative photospheric pressures in these two bands. Increasing the amount of CH_4 via vertical mixing or other disequilibrium chemistry processes would increase the opacity and decrease the photospheric pressure in the 3.6 μm band, resulting in a larger day-night contrast and smaller phase offset. Similarly, decreasing the amount of CO in the atmosphere would decrease the opacity in the 4.5 μm band, shifting the photosphere to higher pressures with a smaller day-night temperature contrast and a larger phase offset. Still, an enhanced methane abundance would require a drastic departure from equilibrium chemistry, as CO is expected to be the major carbon-bearing molecule at temperatures relevant to hot Jupiters. Madhusudhan et al. (2014), for example, show that the delineation between CO and CH_4 dominance at 1500 K is at 10 bars, rising to 100 bars at 2000 K. For hot planets the abundance of CH_4 at photospheric pressures (approx. 100 mbar) should be tiny. Even for HD 189733b, the coldest planet in Figure 14, the CH_4 abundance is nearly 3 orders of magnitude below the CO abundance for any reasonable photospheric pressure, as shown in Figure 3 of Madhusudhan et al. (2014).

2.6 Conclusions

In this paper, we present new phase curve observations for WASP-33b and HD 149026b at 3.6 μm and 4.5 μm. Our measured parameters are in good agreement with previously published transit and secondary eclipse observations of these two planets, and we use our new phase curve observations to investigate the atmospheric circulation patterns of these two planets. We use a simple toy model to estimate the brightness temperatures, albedo, and recirculation efficiency of both planets under the assumption that they emit as blackbodies and find that WASP-33b appears generally similar to other hot Jupiters despite its unusually high irradiation level. On the other hand, HD 149026b has a typical recirculation efficiency but an albedo of 0.6—the highest ever measured. This albedo strongly suggests the presence of clouds, which could easily be confirmed with *HST* transmission spectroscopy. Intriguingly, we find strong evidence for a correlation between the masses of planets with published thermal phase curves and their inferred albedos; this may be indicative of the role that surface gravity plays in the settling of cloud particles.

We also compared our measured phase curves for these two planets to predictions from GCMs. For HD 149026b, we considered models with 1× and 30× solar metallicity, both of which provided an unusually poor match to the observed phase curve shapes. Based on this planet's high inferred albedo and enhanced bulk metallicity, it seems likely that even higher metallicity GCMs incorporating clouds could provide a better match to these data. For WASP-33b, we considered models with and without TiO; although there were still some discrepancies, these models were overall a much better fit than in the case of HD 149026b. We note that MHD effects likely dominate the atmospheric circulation for highly irradiated planets like WASP-33b, and present an obvious avenue for future investigations.

Lastly, we placed these two planets in context by comparing their observed phase offsets in each band to those of other planets. We find a strong correlation between measured phase offset and irradiation temperature, where the observed offset decreases with increasing irradiation temperature to a minimum around 3400 K, and then rises again for the most highly irradiated planets. Although this decreasing trend with increasing irradiation is predicted by GCM simulations of these planets, in practice the size of the observed phase offsets for the coolest planets appear to be lower than predicted by the GCMs. We propose that this can be explained by the presence of high altitude cloud layers in these atmospheres, which decrease the photospheric pressure probed in these two bands. At higher temperatures, we find

that the observed phase offsets decrease to zero for irradiation temperatures near 3400 K, while the GCMs predict a minimum phase offset of 10° for planets in this temperature regime. We propose that this discrepancy can be resolved by the inclusion of MHD effects such as Lorentz drag, which would serve to further reduce the speed of atmospheric winds and decrease the size of the observed phase offset. We note that the trend of increasing phase offset with increasing temperature for the most highly irradiated planets is not well-matched by this simple picture, but perhaps could be explained with more sophisticated circulation models incorporating the full range of MHD effects. Finally, we propose that the relative phase offsets at 3.6 and 4.5 μm, which are consistently the opposite of those predicted in the GCMs, might be explained by a change in the assumed atmospheric compositions and/or chemistries of these planets.

2.7 Acknowledgements

This work is based on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Support for this work was provided by NASA through an award issued by JPL/Caltech. H.A.K. acknowledges support from the Sloan Foundation.

References

Adcroft, A., Campin, J.-M., Hill, C., & Marshall, J. 2004, Monthly Weather Review, 132, 2845

Angerhausen, D., DeLarme, E., & Morse, J. A. 2015, PASP, 127, 1113

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33

Batygin, K., & Stevenson, D. J. 2010, ApJ, 714, L238

Benneke, B., Werner, M., Petigura, E., et al. 2017, ApJ, 834, 187

Bradley, L., Sipocz, B., Robitaille, T., et al. 2016, astropy/photutils: v0.3

Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436

Burrows, A., Rauscher, E., Spiegel, D. S., & Menou, K. 2010, ApJ, 719, 341

Carter, J. A., Winn, J. N., Gilliland, R., & Holman, M. J. 2009, ApJ, 696, 241

Charbonneau, D., Allen, L. E., Megeath, S. T., et al. 2005, ApJ, 626, 523

Charbonneau, D., Winn, J. N., Latham, D. W., et al. 2006, ApJ, 636, 445

Collier Cameron, A., Guenther, E., Smalley, B., et al. 2010, MNRAS, 407, 507

Cooper, C. S., & Showman, A. P. 2005, ApJ, 629, L45

—. 2006, ApJ, 649, 1048

Cowan, N. B., & Agol, E. 2008, ApJ, 678, L129

Cowan, N. B., & Agol, E. 2011, ApJ, 729, 54

Cowan, N. B., Machalek, P., Croll, B., et al. 2012, ApJ, 747, 82

de Mooij, E. J. W., Brogi, M., de Kok, R. J., et al. 2013, A&A, 550, A54

Deming, D., Fraine, J. D., Sada, P. V., et al. 2012, ApJ, 754, 106

Deming, D., Knutson, H., Kammer, J., et al. 2015a, ApJ, 805, 132

—. 2015b, ApJ, 805, 132

Demory, B.-O., de Wit, J., Lewis, N., et al. 2013, ApJ, 776, L25

Dodson-Robinson, S. E., & Bodenheimer, P. 2009, ApJ, 695, L159

Faherty, J. K., Riedel, A. R., Cruz, K. L., et al. 2016, ApJS, 225, 10

Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10

Foreman-Mackey, D., Agol, E., Angus, R., & Ambikasaran, S. 2017, ArXiv

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fortney, J. J., Cooper, C. S., Showman, A. P., Marley, M. S., & Freedman, R. S. 2006, ApJ, 652, 746

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419

Ginzburg, S., & Sari, R. 2016, ApJ, 819, 116

Grillmair, C. J., Carey, S. J., Stauffer, J. R., et al. 2012, in Proc. SPIE, Vol. 8448, Observatory Operations: Strategies, Processes, and Systems IV, 84481I

Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, ApJ, 806, 146

Heng, K., & Demory, B.-O. 2013, ApJ, 777, 100

Heng, K., & Showman, A. P. 2015, Annual Review of Earth and Planetary Sciences, 43, 509

Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011

Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6

Ingalls, J. G., Krick, J. E., Carey, S. J., et al. 2012, in Proc. SPIE, Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, 84421Y

Ingalls, J. G., Krick, J. E., Carey, S. J., et al. 2016, AJ, 152, 44

Johnson, M. C., Cochran, W. D., Collier Cameron, A., & Bayliss, D. 2015, ApJ, 810, L23

Kammer, J. A., Knutson, H. A., Line, M. R., et al. 2015, ApJ, 810, 118

Kass, R. E., & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773

Kataria, T., Showman, A. P., Fortney, J. J., et al. 2015, ApJ, 801, 86

Kataria, T., Sing, D. K., Lewis, N. K., et al. 2016, ApJ, 821, 9

Keating, D., & Cowan, N. B. 2017, ArXiv e-prints, 1709.03502

Knutson, H. A., Charbonneau, D., Cowan, N. B., et al. 2009, ApJ, 703, 769

Knutson, H. A., Lewis, N., Fortney, J. J., et al. 2012, ApJ, 754, 22

Komacek, T. D., & Showman, A. P. 2016, ApJ, 821, 16

Kovács, G., Kovács, T., Hartman, J. D., et al. 2013, A&A, 553, A44

Kreidberg, L. 2015, PASP, 127, 1161

Lehmann, H., Guenther, E., Sebastian, D., et al. 2015, A&A, 578, L4

Lewis, N. K., Showman, A. P., Fortney, J. J., Knutson, H. A., & Marley, M. S. 2014, ApJ, 795, 150

Lewis, N. K., Showman, A. P., Fortney, J. J., et al. 2010, ApJ, 720, 344

Lewis, N. K., Knutson, H. A., Showman, A. P., et al. 2013, ApJ, 766, 95

Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100

Madhusudhan, N., Knutson, H., Fortney, J. J., & Barman, T. 2014, Protostars and Planets VI, 739

Mahapatra, G., Helling, C., & Miguel, Y. 2017, MNRAS, 472, 447

Mancini, L., Kemmer, J., Southworth, J., et al. 2016, MNRAS, 459, 1393

Marley, M. S., & McKay, C. P. 1999, Icarus, 138, 268

Menou, K. 2012, ApJ, 745, 138

Morley, C. V., Knutson, H., Line, M., et al. 2017, AJ, 153, 86

Moses, J. I., Line, M. R., Visscher, C., et al. 2013, ApJ, 777, 34

Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, ArXiv e-prints, 1710.05276

Nutzman, P., Charbonneau, D., Winn, J. N., et al. 2009, ApJ, 692, 229

Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, ApJ, 828, 22

Perez-Becker, D., & Showman, A. P. 2013, ApJ, 776, 134

Perna, R., Menou, K., & Rauscher, E. 2010, ApJ, 719, 1421

Rauscher, E., & Menou, K. 2012, ApJ, 750, 96

Rogers, T. M. 2017, Nature Astronomy, 1, 0131

Roman, M., & Rauscher, E. 2017, ApJ, 850, 17

Sato, B., Fischer, D. A., Henry, G. W., et al. 2005, ApJ, 633, 465

Schwartz, J. C., & Cowan, N. B. 2015, MNRAS, 449, 4192

Schwartz, J. C., Kashner, Z., Jovmir, D., & Cowan, N. B. 2017, ArXiv e-prints, 1707.05790

Seager, S., & Deming, D. 2010, ARA&A, 48, 631

Showman, A. P., Fortney, J. J., Lian, Y., et al. 2009, ApJ, 699, 564

Showman, A. P., & Guillot, T. 2002, A&A, 385, 166

Shporer, A., & Hu, R. 2015, AJ, 150, 112

Sing, D. 2010, Limb Darkening, `http://www.astro.ex.ac.uk/people/sing/David_Sing/Limb_Darkening.html`, accessed: 2017-03-29

Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Nature, 529, 59

Smith, A. M. S., Anderson, D. R., Skillen, I., Collier Cameron, A., & Smalley, B. 2011, MNRAS, 416, 2096

Stevenson, K. B. 2016, ApJ, 817, L16

Stevenson, K. B., Harrington, J., Fortney, J. J., et al. 2012, ApJ, 754, 136

Stevenson, K. B., Line, M. R., Bean, J. L., et al. 2017, AJ, 153, 68

Thorngren, D. P., & Fortney, J. J. 2017, ArXiv e-prints, 1709.04539

Torres, G., Winn, J. N., & Holman, M. J. 2008, ApJ, 677, 1324

Turner, J. D., Pearson, K. A., Biddle, L. I., et al. 2016, MNRAS, 459, 789

von Essen, C., Mallonn, M., Albrecht, S., et al. 2015, A&A, 584, A75

von Essen, C., Czesla, S., Wolter, U., et al. 2014, A&A, 561, A48

Wakeford, H. R., Sing, D. K., Kataria, T., et al. 2017, Science, 356, 628

Winn, J. N., Henry, G. W., Torres, G., & Holman, M. J. 2008, ApJ, 675, 1531

Wong, I., Knutson, H. A., Lewis, N. K., et al. 2015a, ApJ, 811, 122

—. 2015b, ApJ, 811, 122

Wong, I., Knutson, H. A., Kataria, T., et al. 2016, ApJ, 823, 122

Zellem, R., Griffith, C. A., Lewis, N. K., Swain, M. R., & Knutson, H. A. 2014, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 46, AAS/Division for Planetary Sciences Meeting Abstracts, 104.04

Chapter 3

FORWARD MODELING AND RETRIEVALS WITH PLATON, A FAST OPEN-SOURCE TOOL

Michael Zhang, Yayaati Chachan, Eliza M. R. Kempton, and Heather A. Knutson. Forward Modeling and Retrievals with PLATON, a Fast Open-source Tool. *PASP*, 131(997):034501, March 2019. doi: 10.1088/1538-3873/aaf5ad. M.Z. conceived the project, designed the package, wrote the vast majority of the code, and wrote the paper.

3.1 Abstract

We introduce PLanetary Atmospheric Transmission for Observer Noobs (PLATON), a Python package that calculates transmission spectra for exoplanets and retrieves atmospheric characteristics based on observed spectra. PLATON is easy to install and use, with common use cases taking no more than a few lines of code. It is also fast, with the forward model taking much less than one second and a typical retrieval finishing in minutes on an ordinary desktop. PLATON supports the most common atmospheric parameters, such as temperature, metallicity, C/O ratio, cloud-top pressure, and scattering slope. It also has less commonly included features, such as a Mie scattering cloud model and unocculted starspot corrections. The code is available online at https://github.com/ideasrule/platon under the open source GPL-3.0 license.

3.2 Introduction

Transmission spectroscopy is an essential tool for understanding exoplanet atmospheres. By measuring the transit depth at different wavelengths, astronomers have detected a variety of molecules, including many detections of water and a few of carbon monoxide (Madhusudhan et al., 2014; Deming & Seager, 2017). With sufficiently wide wavelength coverage, it is possible to constrain not only compositions, but temperatures and cloud properties as well. The inverse problem of retrieving atmospheric parameters given observed transit depths is crucial to the interpretation of observations. The *HST* and *Spitzer* archives alone already hold observations of dozens of transiting exoplanets, and the volume of data will only increase as TESS discovers more planets around bright stars and ground-based transit spectroscopy

matures. A fast, easy to use, and open source retrieval code is needed in order to help the community quickly extract atmospheric parameters from these observations.

The exoplanet community has developed many retrieval codes over the years, as reviewed in Madhusudhan (2018). The first retrieval code was published in Madhusudhan & Seager (2009), but many more have been developed since then, including Benneke & Seager (2012), Line et al. (2013), and Waldmann et al. (2015). These codes use a wide variety of techniques. Most use either MCMC or nested sampling to map the posterior distribution, but some use self-consistent temperature-pressure profiles while others use parametric profiles; some enforce chemical equilibrium while others allow individual molecular abundances to vary freely; some compute eclipse depths while others only compute transit depths; some use opacity sampling while others use the correlated-k method. Despite the large number of codes, we were only able to find one documented, publicly available retrieval code with a dedicated paper (Tau-REX; Waldmann et al., 2015). The Bayesian Atmospheric Radiative Transfer (BART) code is partially described in one subsection of a dissertation (Blecic, 2016) and one subsection of a paper (Blecic et al., 2017). CHIMERA (Line et al., 2013) appears to be publicly available[1], but has no documentation or examples.

Although the papers describing these codes rarely have benchmarks, it appears that a typical retrieval takes days to weeks on a typical desktop computer with 4 cores (Zingales & Waldmann, 2018). We aimed to create a publicly available fast retrieval code that can be run on a typical desktop, and takes minutes to hours instead of days to weeks.

Recently, Kempton et al. (2017) introduced Exo-Transmit. Exo-Transmit calculates transmission spectra for exoplanets of arbitrary size and composition, with pre-packaged data files that make it easy to specify common compositions and pressure-temperature profiles. However, with a run time on the order of tens of seconds, Exo-Transmit is too slow to incorporate into a retrieval code. Exo-Transmit also relies on pre-packaged data files to specify the composition of the atmosphere, and does not allow the user to specify arbitrary metallicities, C/O ratios, or abundances. In addition, Exo-Transmit is written in C and therefore requires additional steps to interface with Python-based codes.

Inspired by Exo-Transmit, we developed PLATON to address these issues. PLATON was written from scratch in pure Python, and supports both Python 2.7 and

[1] `https://github.com/ExoCTK/chimera`

3.5 on both Linux and Mac machines. It uses the same opacity data and a few of the same algorithms as Exo-Transmit. Its forward model contains all the functionality of Exo-Transmit, but is 100-1000x faster—fast enough to incorporate into a Markov chain Monte Carlo (MCMC) or nested sampling retrieval. These speed improvements come from three main sources. First, it loads all data files once and holds them in memory afterwards. Second, it performs transit depth calculations only for the wavelength ranges requested by the user, and only for atmospheric layers above the cloud pressure. Third, the radiative transfer is formulated as a matrix-vector multiplication, which `numpy` can compute very efficiently by calling a highly optimized Basic Linear Algebra Subsystems (BLAS) routine. This decreases the computation time from 1-2 seconds to 30 milliseconds. Aside from these optimizations, PLATON also incorporates several new features, including the ability to model Mie scattering, account for unocculted starspots, compute equilibrium abundances at user-specified metallicities and C/O ratios, and account for varying gravitational acceleration as a function of height for extended atmospheres. This forward modelling code was then incorporated into a retrieval module within PLATON.

PLATON is in many ways complementary to Tau-REX. Whereas Tau-REX considers a small number of molecules (which could be automatically chosen by a machine learning algorithm), PLATON considers a large number of pre-determined molecules. Whereas Tau-REX performs free retrievals, allowing for the abundances of individual molecules to be retrieved, PLATON assumes equilibrium chemistry throughout the atmosphere. In ease of installation and use, however, PLATON has a clear advantage. Installing Tau-REX is a complex multi-step process which involves signing up for an account, installing CMake, OpenMPI, and MultiNest manually, soft linking libraries (on Macs), installing some Python packages, compiling Tau-REX's C++ and Fortran libraries, and downloading a separate set of input data. After installation, the first author ran into numerous problems running the tests, mostly related to problems with the dependencies. In contrast, PLATON's installation process is a single line, which also installs all dependencies:

```
pip install platon
```

Or alternatively, after cloning the Git repository:

```
python setup.py install
```

We describe PLATON's underlying model structure and validation in Section 2, and its incorporation into a retrieval framework in Section 3.

3.3 Forward model

Transmission spectroscopy has the benefit of being relatively simple to model. By its nature, it probes higher altitudes (P < 1 bar), where the atmosphere can be approximated as isothermal (Burrows et al., 2008). In addition, although nightside pollution of transit depths will be significant for *JWST* (Kipping & Tinetti, 2010), planetary emission is negligible compared to transmission of starlight. This drastically simplifies the radiative transfer calculation and limits the amount of model uncertainty. Our forward model uses the opacity files from Exo-Transmit, in addition to some of its algorithms. This section will discuss the algorithms and opacities briefly, but more detail can be found in Kempton et al. (2017).

Overview

To calculate a transmission spectrum, we assume an isothermal atmosphere with equilibrium chemistry.[2] We first divide the atmosphere into 500 layers, equally spaced in log P from 10^{-4} to 10^{8} Pa. The composition of each layer is calculated assuming equilibrium chemistry, from which follows the mean molecular weight. The physical depth of each layer is then obtained by solving the hydrostatic equation:

$$\frac{dP}{dr} = -\frac{GM}{r^2}\frac{\mu m_{amu} P}{kT}. \tag{3.1}$$

The integration constant is set by P(R) = 1 bar: that is, the radius which the user specifies for the planet is taken to be the radius at a reference pressure of 1 bar. The reference pressure is arbitrary, but we picked 1 bar because it is easy to remember and close to standard atmospheric pressure on Earth. After solving the hydrostatic equation, the opacity of each layer is then calculated from its composition. We include gas absorption, collisional absorption, and Rayleigh scattering, the last of which can be modified with a user-specified slope and amplitude.

Finally, we carry out the radiative transfer calculation. Consider a ray passing through the atmosphere with impact parameter r, where r is the distance from planet center to one of the layers. This ray passes through all layers of the atmosphere above r, but lingers in each layer for different distances. By calculating the distance it takes to traverse each layer and multiplying by that layer's absorption coefficient, we calculate the optical depth experienced by this ray:

[2] Arbitrary P-T profiles and compositions are also supported, but this is not the recommended use case. We refer the user to the documentation for information on these features.

$$\tau_\lambda(r_i) = \sum_{j=i}^{N} \alpha_\lambda(j)\Delta l(i,j). \tag{3.2}$$

The transit depth can then be calculated by adding up the cross sectional areas of each layer, weighted by the amount of light let through:

$$D_\lambda = (R_{bot}/R_s)^2 + 2\sum_{i=1}^{N} \frac{r dr}{R_s^2}(1 - e^{-\tau_\lambda}) \tag{3.3}$$

where R_{bot} is the radius at the bottom-most layer.

As a final step, we correct the transit depth for unocculted star spots. This step is necessary for active stars because the hot unspotted regions have a different spectrum from the colder star spots, and a planet that blocks only the former will appear bigger than it actually is. Worse, this radius inflation depends on wavelength, giving rise to a spurious transit spectrum shape (McCullough et al., 2014; Rackham et al., 2018). To correct for this effect, we use:

$$D_{\lambda,c} = D_\lambda \frac{S(\lambda, T_{clear})}{fS(\lambda, T_{spot}) + (1-f)S(\lambda, T_{clear})} \tag{3.4}$$

where f is the spot fraction, T_{spot} is the spot temperature, T_{clear} is the temperature of unspotted regions, and S is the spectrum of the stellar surface. We obtain both $S(\lambda, T_{clear})$ and $S(\lambda, T_{spot})$ by interpolating the BT-NextGen (AGSS2009) stellar spectral grid (Allard et al., 2012), as provided by the Spanish Virtual Observatory[3]. We neglect any contributions to the spot spectrum that are not purely due to temperature.

Absorption coefficients

We use the same data files as Exo-Transmit for gas absorption and collisional absorption coefficients, although these will be updated in version 3. Per-species gas absorption coefficients were computed on a wavelength-temperature-pressure grid using line lists, with a resolution of $\lambda/\Delta\lambda = 1000$ over the range 0.3-30 μm. The sources of these line lists are shown in Table 2 of Lupu et al. (2014). For most molecules, the line list comes from HITRAN, but they use miscellaneous other sources, such as Freedman et al. (2008, 2014). All collisional absorption coefficients are from HITRAN. We refer the reader to Kempton et al. (2017) for more information.

[3] `http://svo2.cab.inta-csic.es/theory/newov2/index.php`

For Rayleigh scattering, we compute the absorption coefficient using the equation:

$$\alpha_\lambda = \frac{128}{3\pi^5}\frac{P}{kT}p^2\lambda^{-4} \tag{3.5}$$

where p is the polarizability of a species. For both Exo-Transmit and PLATON, polarizabilities are obtained from the CRC Handbook of Chemistry and Physics. We use an updated list of polarizability values, including several species assumed to have zero polarizability in Exo-Transmit: H, O, C, N, C_2H_4, H_2CO, OCS, OH, SO_2, Na, and K. The polarizability of OH is not in the CRC handbook, and was instead taken from Pluta et al. (1988). A few species have no published polarizability values that we can find: MgH, SH, SiH, SiO, TiO, and VO. We adopt 0 for their polarizability.

We compute total absorption coefficients at temperature and pressure grid points bordering the atmospheric temperature-pressure profile. 2D bilinear interpolation is then used to get the coefficients at each point along the temperature profile. This is the same algorithm used by Kempton et al. (2017).

Atmospheric composition

In the sphere of atmospheric composition, we depart from Exo-Transmit by providing an equilibrium chemistry model computed with `GGchem` (Woitke et al., 2018). The user can provide two parameters: metallicity and C/O ratio. If the user does not specify these parameters, we assume solar metallicity and solar C/O ratio (0.53). From these parameters, we compute the abundances of 34 atomic and molecular species in every layer of the atmosphere. These 34 species are the same as those included in Kempton et al. (2017), except that we leave out C_2H_6 and SH because these species are not included in the default `GGchem` databases.

While many public codes exist for equilibrium chemistry (Blecic et al., 2016; Woitke et al., 2018; Stock et al., 2018), they are currently too slow for use in retrievals. This is why we opted to use `GGchem` to generate a grid of abundances. `GGchem` computes equilibrium species abundances from atomic abundances by minimizing the total Gibbs free energy of the mixture. We generate separate grids, one with and the other without condensation.

The abundance grids have 5 dimensions: species name, temperature, pressure, metallicity, and C/O ratio. The temperature and pressure grid points are chosen to match the opacity grid, with temperature ranging from 300 to 3000 K in 100 K intervals and pressure from 10^{-4} to 10^8 Pa in decade intervals. The pressure

range is determined by our opacity data, which are only available for this range. The temperature range is set mostly by our opacity data, which span 100-3000 K at 100 K intervals. We did not generate abundances down to 100 K because `GGchem` becomes unstable at very low temperatures. Metallicity ranges from $\log_{10}(Z) = -1$ to $\log_{10}(Z) = 3$ in steps of 0.05, while C/O ratio ranges from 0.2 to 2.0 in steps of 0.2. We then perform 4D linear interpolation over the grid, obtaining log(Abundance) as a function of T, log(P), log(Z), and log(C/O). The relative error introduced by interpolation in metallicity is on the order of 10^{-4}. This is significantly smaller than the error introduced by interpolation in C/O, which is on the order of a few percent. Some molecules (such as water and HCN) have a sharp abundance transition at C/O $\sim$ 1, giving rise to interpolation errors of tens of percent in this regime. However, the abundances change so rapidly around C/O $\sim$ 1 that our assumption of uniform equilibrium abundances across the whole planet also becomes invalid, as it is unlikely that the entire planet has exactly the same C/O ratio. This makes the interpolation errors less significant. Nevertheless, our development branch contains an abundance grid with twice the resolution in C/O, increasing to 4x the resolution around C/O $\sim$ 1. Interpolation errors are less than 1% in almost all cases with this new grid. The new grid will be incorporated in our next release.

For retrievals, the current data is almost never good enough to constrain any parameter to better than one grid spacing. For forward models, we recommend using parameters that correspond exactly to a grid point in temperature, metallicity, and C/O to avoid interpolation errors.

Clouds and hazes: parametric

By default, PLATON accounts for clouds and hazes by allowing the cloud-top pressure, scattering amplitude, and scattering slope to vary as free parameters in the fit. The cloud-top pressure defines the height in the atmosphere below which no light can penetrate. The user specifies this parameter, or can set it to infinity for a clear atmosphere. The scattering amplitude and slope are a simple way to parameterize scattering properties without invoking a microphysical model. The default value—an amplitude (A) of 1 and a slope of 4—corresponds to pure Rayleigh scattering. If A is changed while the slope is fixed at 4, the absorption coefficient for scattering is simply multiplied by A at all wavelengths. If the slope s is not 4, we set the scattering absorption coefficient such that it is A times the Rayleigh absorption at the reference wavelength of 1 μm, and is proportional to λ^{-s} at all wavelengths. This parameterization makes no assumptions about the underlying physics while

allowing users to see at a glance how strong the scattering is compared to Rayleigh at the reference wavelength (by comparing A to 1), and how the scattering strength behaves with wavelength (by looking at the slope). Since the reference wavelength is only a matter of definition and does not change the underlying model, the user can set it to any value.

Clouds and hazes: Mie scattering

In addition to the parametric approach, PLATON supports Mie scattering (Benneke et al., subm.). Clouds and hazes often have particles in the micron range, comparable to the wavelengths of most transit spectra observations. This makes Mie scattering an important component of atmospheric physics (Marley et al., 2013). For example, Benneke et al., subm. use Mie scattering to explain the anomalously low 3.6 and 4.5 μm *Spitzer* transit depths for GJ 3470b as compared to its WFC3 (1.1–1.6 μm) transit depths.

In the Mie scattering mode, the user specifies a cloud-top pressure (P_{cloud}), a complex refractive index ($m = n - ik$), a mean particle size (r_m), a geometric standard deviation for particle size (σ_g), a maximum number density (n_0), and a fractional scale height (f). The atmosphere is assumed to contain particles of the same refractive index at all altitudes above the "cloud top pressure." Below this pressure, the atmosphere is assumed to be perfectly opaque, corresponding to perfectly opaque clouds. The number density is given by:

$$n = n_0 e^{-\frac{h}{f H_{gas}}}, \tag{3.6}$$

where h is the height above the cloud top, f is specified by the user, and H_{gas} is the gas scale height.

The extinction cross section of a single particle of radius r is given by:

$$\sigma = \pi r^2 Q_{ext}(m, 2\pi r/\lambda). \tag{3.7}$$

To calculate Q_{ext}, we use the same algorithm as LX-MIE (Kitzmann & Heng, 2018), except implemented in Python instead of C++. This algorithm is fast, simple, stable, and does not lead to overflows. Since the calculation of Q_{ext} is time intensive (often taking hundreds of milliseconds), we cache the results of every calculation for the lifetime of the transit depth calculator object. Every time the value of $Q_{ext}(m, x)$ is required, the cache is first consulted. If at least one value in the cache has the same

m and an x within 5% of the requested x, we perform linear interpolation on all cache values with the same m and return the interpolated value. If no cache value satisfies these criteria, we consider this a cache miss. Q_{ext} is then calculated for all cache misses and added to the cache.

We assume that the particles follow a log-normal radius distribution:

$$P(r) = \frac{1}{\sigma r \sqrt{2\pi}} \exp\Big(- \frac{(\ln r - \ln r_m)^2}{2\sigma_g^2} \Big), \tag{3.8}$$

in which case the effective cross section of one particle becomes:

$$\bar{\sigma}(\lambda) = \int_0^\infty \pi r^2 Q_{ext}(m, \frac{2\pi r}{\lambda}) \frac{1}{\sigma\sqrt{2\pi} r} e^{-\frac{(\ln r - \ln r_m)^2}{2\sigma_g^2}} dr. \tag{3.9}$$

After a change of variables to $z = \frac{\ln r - \ln r_m}{\sigma}$, we have:

$$\bar{\sigma}(\lambda) = \int_{-\infty}^\infty e^{-z^2/2} \frac{\sqrt{\pi}}{2} r_m^2 e^{2\sigma z} Q_{ext}(m, \frac{2\pi r_m e^{\sigma z}}{\lambda}) dz. \tag{3.10}$$

We integrate this equation by computing the integrand at 100 different values of z ranging from -5 to 5, then using the trapezoid rule. These values were chosen to be densely spaced near 0 and less densely spaced at high or low z. The number of points, the range of z values, and the spacing of z values were chosen so that the integral is as accurate as possible over a wide range of refractive indicies and mean particle sizes without being too computationally expensive.

Wavelength binning

The user can request the transit depth in specific wavelength bins. In these situations, PLATON uses the stellar spectrum to compute a properly binned transit depth. This is important for broad-band instruments like *Spitzer*/IRAC, where the stellar spectrum changes significantly from one side of the bandpass to the other.

Using the user specified stellar temperature, PLATON interpolates the BT-NextGen (AGSS2009) stellar spectral grid to get a stellar spectrum, assuming log(g) = 4.5 (typical for main sequence stars) and solar metallicity. If the user also specifies a spot coverage fraction and spot temperature, PLATON computes the spot spectrum by interpolation, then computes the weighted average of the unspotted and spotted spectrum. The spectrum, which is in units of erg/cm^2/s/A, is converted to a photon flux and multiplied by the spacing between adjacent wavelengths in our wavelength

grid. The multiplication is necessary because our wavelength grid is uniform in log space, but not in linear space, as each grid spacing is 0.1% larger than the previous. Finally, the unbinned transit depths are converted to a binned transit depth by a weighted average, the weight at each wavelength being the photon flux multiplied by the grid spacing.

Validation

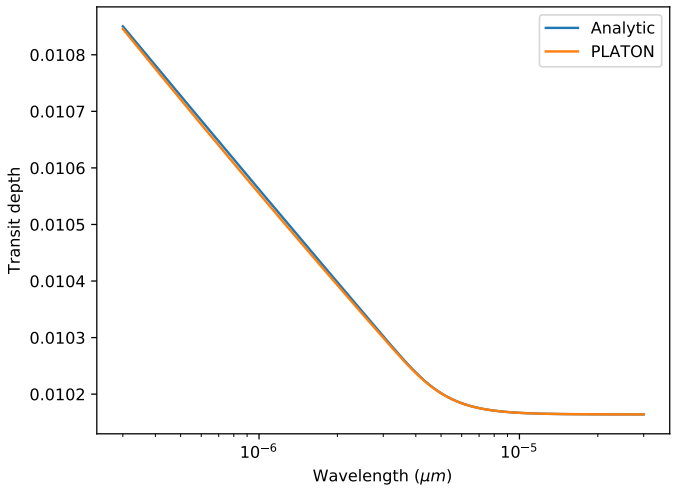


Figure 3.1: Transit depths with all opacity except Rayleigh scattering turned off. The transit depth obeys a power law to within 0.5% blueward of 3 μm, as derived by Shabram et al. (2011). Redward of 3 μm, transit depths flatten out because the atmosphere becomes so transparent that the photospheric pressure becomes greater than 10^8 Pa, the maximum pressure we consider.

We validate our code with unit tests covering every component of the package. In cases where the expected answer can be computed analytically with simplified inputs, such as the interpolation routines or the radiative transfer, we check against these analytic results. An example is shown in Figure 3.1, where we reproduce the theoretical transit depths for a hydrogen atmosphere with only Rayleigh scattering. We make use of Equation 27 of Bétrémieux & Swain (2017) to derive:

$$\kappa = \frac{128\pi^5}{3}\frac{p^2}{\lambda^4 \mu m_H} \tag{3.11}$$

$$\tau_s = \frac{P_s \kappa}{g}\sqrt{\frac{2\pi R_s}{H}} \tag{3.12}$$

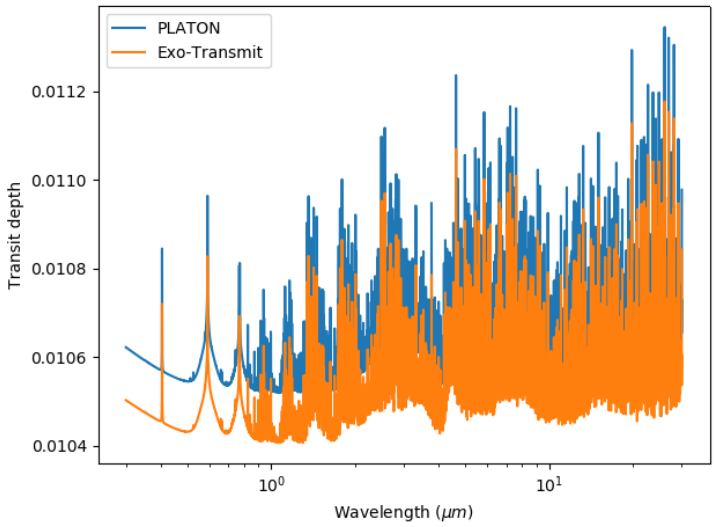
$$R = R_s + H(\gamma + \ln(\tau_s) + E_1(\tau_s)), \tag{3.13}$$

where H is the scale height, $\gamma = 0.57721$ is a constant, p=0.8059×10^{-24} cm^3 is the polarizability of the hydrogen molecule, R_s is the radius at the surface, and τ_s is the slant optical depth at the surface, and E_1 is the exponential integral. PLATON does not model a real surface, but we do not model radiative transfer at pressures greater than 10^8 Pa, creating an effective surface at that pressure.

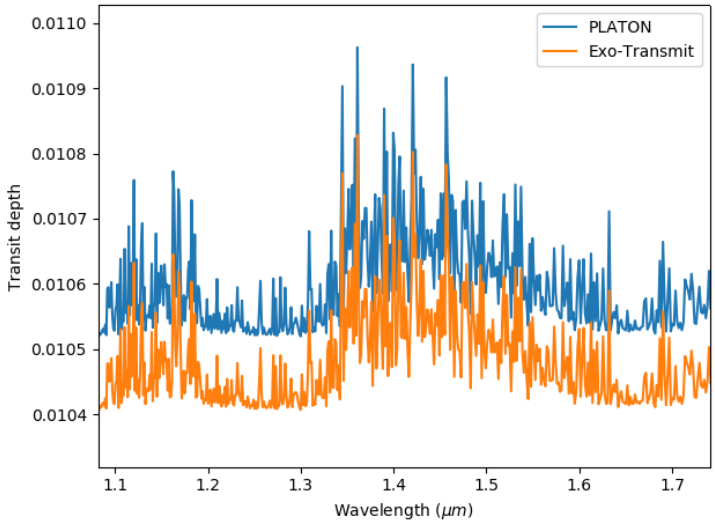
In cases where it is possible to compare against published results, such as the $Q_{ext}(m, x)$ calculations for selected m and x, we verify that we reproduce the published results in Kitzmann & Heng (2018). For methods where such a check would be impractical or take too long, such as the retrieval, we simply ensure the method runs without error and returns a result in the expected format. We use Travis CI to ensure that all unit tests are run after every commit. Travis CI creates pristine virtual machines after every commit, installs PLATON on each machine, and runs the unit tests to make sure they pass. We use four virtual machines to test PLATON in four configurations: Ubuntu with Python 2.7, Ubuntu with Python 3, OS X with Python 2.7, and OS X with Python 3. This way, we ensure that PLATON can be installed and successfully run on the most common configurations.

We also compare our transmission spectrum to the output of Exo-Transmit. In an old version of the code, which was intended to match Exo-Transmit output exactly, the transmission spectrum matched to within machine precision for all wavelengths. In the current version, we have made many design decisions that increase accuracy or usability at the cost of breaking the exact match between the two codes. Figure 3.2 shows a comparison between PLATON and Exo-Transmit spectra for a typical hot Jupiter. The main difference between the two is a uniform shift, with PLATON reporting 115 ppm deeper transit depths. Figure 3.3 shows that if this shift is removed, the two spectra agree almost perfectly.

There are two reasons why PLATON reports higher transit depths. First, PLATON correctly takes into account the decline of gravity with height, while Exo-Transmit assumes constant gravity. This makes PLATON's atmospheres larger—an effect that is most pronounced for super-Earths, but discernible even for Jupiters. Second, Exo-Transmit truncates the atmosphere, taking into account only the region between 0.1 Pa and 10^5 Pa. We consider the entire pressure range for which we have absorption data, spanning 10^{-4} Pa to 10^8 Pa. In practice, we have found that increasing the upper pressure limit has no effect because the photosphere is well above 10^5 Pa, but decreasing the lower pressure limit does increase the transit depth substantially.



(a) Full wavelength coverage

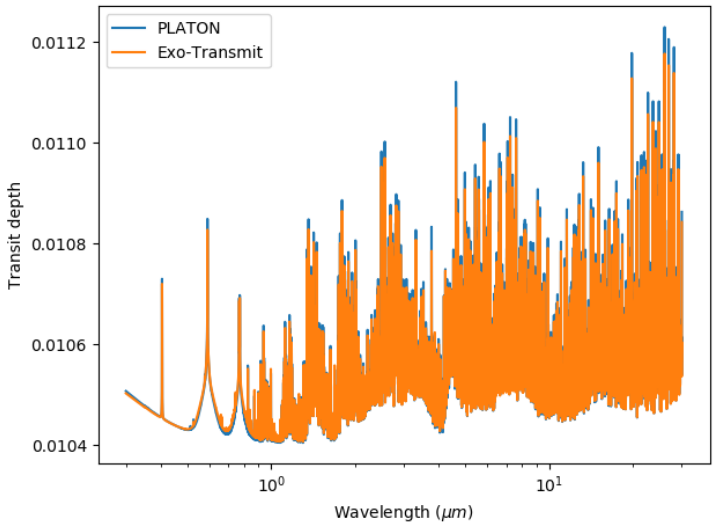


(b) Zoomed into WFC3 band

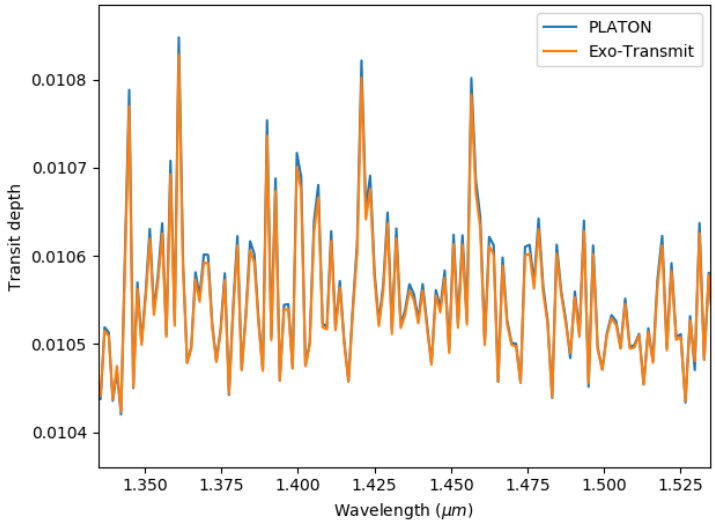
Figure 3.2: Transmission spectrum of a 1200 K Jupiter mass, Jupiter radius planet orbiting a Sun-like star, as computed by PLATON and Exo-Transmit. The difference is mainly due to PLATON truncating the atmosphere at a much lower pressure.

This is especially the case at high metallicities, where strong lines saturate and the transit depths of strong lines correspond to the 0.1 Pa height limit in Exo-Transmit.

As a final validation step, we compare our forward model for HAT-P-26b against that of another atmospheric code, ATMO (Goyal et al., 2018a). The result is shown in Figure 3.4. Due to our opacity sampling algorithm, our transit depths are much



(a) Full wavelength coverage



(b) Zoomed to show individual lines

Figure 3.3: Same as Figure 3.2, but with PLATON transit depths shifted downwards by 115 ppm.

spikier than ATMO's, but the transit depths match decently once they are binned to the resolution of *HST*.

3.4 Retrieval

We support retrievals of atmospheric parameters with either MCMC or multimodal nested sampling. MCMC is implemented by `emcee` (Foreman-Mackey et al., 2013), which uses an affine invariant ensemble sampler. Multimodal nested sampling was first described by Shaw et al. (2007), but we use the pure Python implementation `nestle`, available on Github at https://github.com/kbarbary/nestle. With the exception of the complex refractive index, any parameter that can be passed to the forward model can be included in the retrieval. These parameters are shown in Table 3.1. The last parameter in Table 3.1, error multiple, multiplies every observational error. It accounts for over- or under- estimation of error bars by the user, assuming all error bars are off by the same factor.

The user can choose to fit or freeze any parameter. However, if the user chooses to use the parametric method to account for clouds and hazes, the last 4 parameters in Table 3.1 are irrelevant. PLATON requires that the log particle number density be set to $-\infty$ to avoid confusion. If the user chooses to use Mie scattering, the complex refractive index must be specified, along with the last four parameters in the table. PLATON requires that the log scattering amplitude be set to the default value of 0 and the slope be set to the default value of 4, again to avoid confusion. If a parameter is included in the fit, the user can choose a uniform prior or a Gaussian prior. Gaussian priors are most appropriate for parameters like the planet mass, where a published value exists, but with error bars substantial enough to broaden the distribution of the atmospheric parameters.

Benchmarks

We benchmark PLATON on a typical desktop computer to illustrate its performance. The computer runs Ubuntu 16.04 LTS with a Core i7 7700k CPU and 16 GB of RAM. PLATON never uses more than 500 MB of memory, since that is the total size of its data files, and should perform just as well on computers with less than 16 GB of RAM.

When the forward model is first initialized, PLATON loads all relevant data files into memory. This takes 390 ms, but is only done once. Table 3.2 shows the amount of time taken to compute transit depths within the most commonly used bands once PLATON is initialized. The time taken most directly depends on the number of wavelength grid points within the band. Since grid points are spaced uniformly in logarithmic space with R=1000, the number of grid points is proportional to

Parameter name	Supported range	Log space?	Default
Stellar radius	Any	No	None
Stellar temperature	Any	No	None
Stellar spot fraction	Any	No	0
Stellar spot temperature	Any	No	None
Planet mass	Any	No	None
Planet radius (at 1 bar)	Any	No	None
Temperature	300-3000 K	No	None
Metallicity	0.1-1000x solar	Yes	Solar
C/O ratio	0.2-2.0	No	0.53
Cloud-top pressure	∞, or 10^{-4} - 10^{8} Pa	Yes	∞
Scattering amplitude	Any	Yes	1
Scattering slope	Any	No	4
Error multiple	Any	No	1
Particle number density	Any	Yes	1 m^{-3}
Particle radius	Any	Yes	1 μm
Particle radius standard deviation	Any	No	0.5
Particle fractional scale height	Any	No	1

Table 3.1: Parameters that can be retrieved. PLATON supports a limited range for some parameters, as shown in the second column. The third column indicates whether PLATON takes the base-10 log of the parameter during retrieval.

Table 3.2: Benchmarks for desktop computation of forward model

Band	Wavelength range (μm)	Time (ms)
All wavelengths	0.3-30	393
STIS	0.293-1.019	57
WFC3	1.119-1.628	29
Spitzer 3.6 μm	3.2-4.0	16
Spitzer 4.5 μm	4.0-5.0	16

the ratio between the maximum and minimum wavelengths. In the future, if we increase the resolution we support, run times will increase proportionally. It should be noted that all values in the table are computed for HD 209458b, assuming a clear atmosphere and no Mie scattering. If there are clouds, PLATON neglects the part of the atmosphere below the cloud-top pressure, substantially improving its performance. If Mie scattering is turned on, the calculation of Q_{ext} can consume hundreds of milliseconds on the first run of the forward model. However, every run populates the Q_{ext} cache, and a future run with similar parameters will have a high cache hit rate. Mie calculations will then consume negligible time.

Table 3.3: Benchmarks for retrieval using real HD 209458b data

Band(s)	Algorithm	Time (min)	Likelihood evaluations
WFC3	nested sampling	6.3	20,999
STIS + WFC3 + *Spitzer*	nested sampling	29	17,205
WFC3	MCMC	13	50,051
STIS + WFC3 + *Spitzer*	MCMC	57	50,051

Table 3.3 shows the typical performance of retrievals. Users can run the second and fourth benchmarks themselves by running `examples/retrieve_multinest.py` and `examples/retrieve_emcee.py`, respectively. These benchmarks use published HD 209458b data from Knutson et al. (2007), Deming et al. (2013), and Evans et al. (2015). The nested sampling runs were performed with 100 live points, and the MCMC runs had 50 walkers and 1000 steps. These correspond to 17,000 and 50,000 likelihood evaluations, respectively. The number of likelihood evaluations, and hence the run time, is proportional to the number of live points or steps.

For MCMC, we verified convergence in two ways. First, we ran another retrieval with 10,000 steps and checked that it gave similar posteriors. Second, we estimated the autocorrelation length using a variant of the "new" method recommended by the author of emcee, Dan Foreman-Mackey, on his webpage[4]. Namely, for each dimension corresponding to a parameter, we compute the autocorrelation length of the chain, flattened along the walker dimension. The estimates are averaged together to obtain an average autocorrelation length. If the chain is less than 50 times longer than this average, we consider the estimate unreliable and run more iterations. If the chain is more than 50 times longer, we consider the estimate reliable, and the chain to have converged. The factor of 50 comes from the recommendation of the same webpage. For our benchmark retrieval, we find an autocorrelation length of 7 steps—corresponding to 350 samples, as we have 50 walkers. For the chain with 10,000 steps, we compute an autocorrelation length of 24 steps. Both numbers are far smaller than the total number of steps, but the fact that the estimates are discrepant shows the disturbing fact that the autocorrelation length estimate increases with the number of samples. This is a property of the algorithm, not of our chains. It is one of the reasons why monitoring convergence is not trivial, and is why we have decided to refer the user to the extensive online and published literature on this topic.

[4] `https://emcee.readthedocs.io/en/latest/tutorials/autocorr`

Comparison with retrieval results in literature

To evaluate the performance of our retrieval tool, we focus on planets with large datasets and published retrieved results in the literature. WASP-39b and HAT-P-26b are excellent candidates for such an exercise as their transmission spectra extend all the way from the optical (*STIS*) to near-infrared (*Spitzer*). Their atmospheric properties have been inferred from ATMO Retrieval Code (ARC), which couples the ATMO models to a L-M least-squares minimizer and a Differential Evolution Chain Monte Carlo analysis (Tremblin et al., 2015; Wakeford et al., 2017). ATMO is used to compute 1-D T-P profiles for atmospheres in hydrostatic and radiative-convective equilibrium and calculate the transmission and emission spectra for a given atmospheric profile. We choose to compare with ARC because ATMO has been benchmarked against the Met Office SOCRATES code (used for the Earth) and because ATMO has a public scalable grid of transmission spectra (Goyal et al., 2018a) that can be used for retrievals. More importantly, the data for these planets provide relatively tight constraints on their atmospheric water abundance (WASP-39b) and heavy element abundance (HAT-P-26b) in the ATMO retrieval (Wakeford et al., 2017, 2018). These features provide useful tests for PLATON's retrieval capabilities and allow direct comparison of well constrained quantities using different retrieval models.

We run retrievals on these two planets with both PLATON and the ATMO grid. We use the ATMO grid which accounts for rainout, which has 5 dimensions: temperature, gravity, metallicity, C/O ratio, scattering factor, and cloud parameter (representing a constant additional opacity across the atmosphere at all wavelengths). Since there is no parameter in PLATON that corresponds exactly to the cloud parameter, we set the cloud parameter to zero in the ATMO models and the cloudtop pressure to infinity in the PLATON models. This restricts our retrievals to cases without an optically thick cloud layer while still allowing for enhanced scattering due to optically thin hazes. We use uniform priors for temperature, log(Z), C/O ratio, and log(scattering factor), with the minimum and maximum set to the minimum and maximum of the ATMO grid. 5D linear interpolation is used to compute the transmission spectrum for a certain set of parameters from the transmission spectra at the grid points, then scaled to the planetary and stellar radii. For consistency, we use the same priors for PLATON, even though PLATON supports a wider parameter space. Figures 3.6 and 3.7 show the resulting posterior probability distributions for HAT-P-26b and WASP-39b, respectively, while the best fits are shown in Figure 3.5. HAT-P-26b shows strikingly good agreement between PLATON and ATMO, with

Table 3.4: Best Fit Parameters for HAT-P-26b

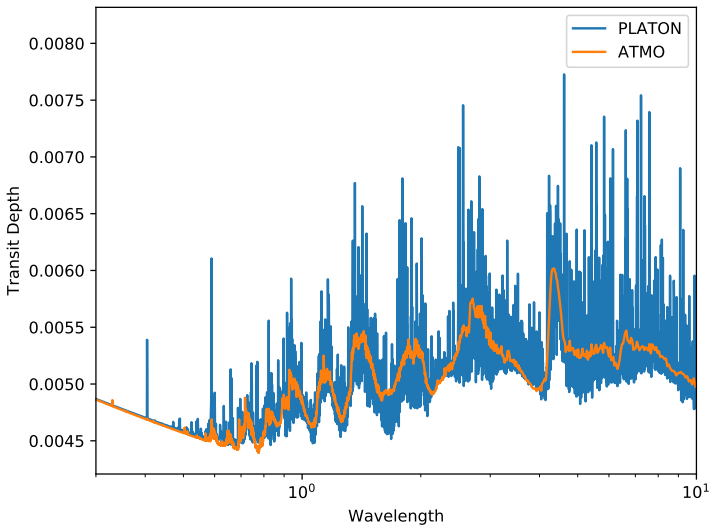
Property	PLATON	ATMO grid	Wakeford et al 2017
Temperature (K)	903 ± 96	878^{+101}_{-87}	
C/O ratio	$0.46^{+0.12}_{-0.08}$	$0.45^{+0.09}_{-0.07}$	
$\log_{10}(Z / Z_\odot)$	$1.96^{+0.23}_{-0.39}$	$1.87^{+0.27}_{-0.45}$	1.566 ± 1.7034
$\log_{10}$(Scattering Factor)	$2.37^{+0.38}_{-0.71}$	$2.21^{+0.46}_{-0.62}$	
error multiple	$1.68^{+0.29}_{-0.22}$	$1.60^{+0.28}_{-0.21}$	

Table 3.5: Best Fit Parameters for WASP-39b

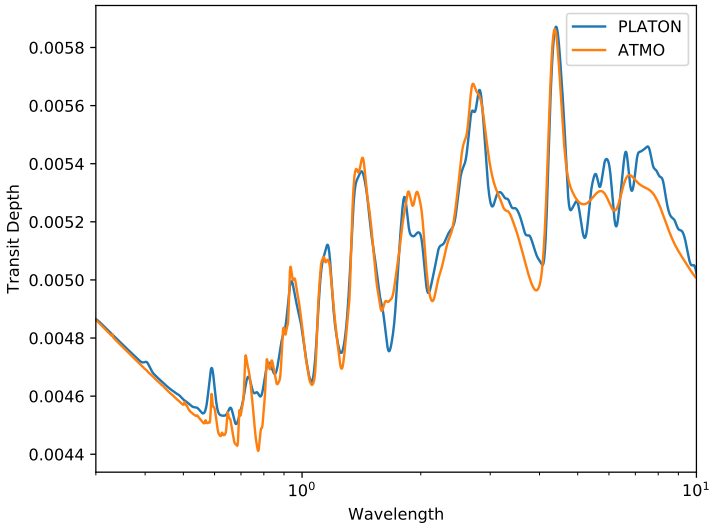
Property	PLATON	ATMO grid	Wakeford et al 2018
Temperature (K)	1031^{+58}_{-69}	1129^{+73}_{-38}	1030^{+30}_{-20}
C/O ratio	$0.60^{+0.08}_{-0.12}$	$0.59^{+0.1}_{-0.15}$	$0.31^{+0.08}_{-0.05}$
$\log_{10}(Z / Z_\odot)$	$2.23^{+0.05}_{-0.11}$	$2.23^{+0.05}_{-0.08}$	$2.18^{+0.12}_{-0.16}$
$\log_{10}$(Scattering Factor)	$1.17^{+0.33}_{-0.34}$	$1.30^{+0.49}_{-0.65}$	

all posteriors being nearly identical. For WASP-39b, on the other hand, there are slight discrepancies. PLATON favors a temperature ~100 K lower, a discrepancy of roughly 1σ. In addition, ATMO exhibits an odd multimodal distribution in scattering factor and C/O ratio that is not seen in PLATON. This multimodal distribution is not seen in the retrieval from Wakeford et al. (2018), which uses the ATMO code directly instead of using a generic grid. The temperature retrieved by PLATON is also more consistent with Wakeford et al. (2018) than with the temperature we retrieved using the generic ATMO grid. Additionally, we do not see a multimodal distribution if we use the ATMO grid which does not include rainout, nor do we see it if the observations of the two Na and K spectral features are excluded. We therefore conclude that the odd behavior of our ATMO retrievals using the generic grid is likely due to a rare coincidence of factors.

Tables 3.4 and 3.5 show the median and error of each retrieved parameter for HAT-P-26b and WASP-39b, respectively. The error bars are derived by comparing the 84th and 16th percentiles to the median. We have also included published retrieval results, when available. For HAT-P-26b, all values are consistent. For WASP-39b, PLATON favors lower temperature, but similar C/O ratio, metallicity, and scattering factor compared to the ATMO generic grid. The published metallicity and temperature are both very similar to PLATON's results, while the published C/O ratio is lower than that of PLATON. It is unclear why the published C/O ratio,



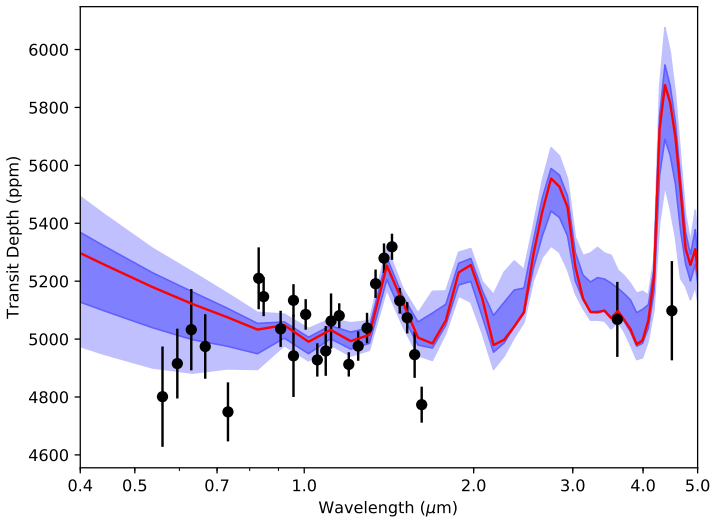
(a) Full wavelength resolution



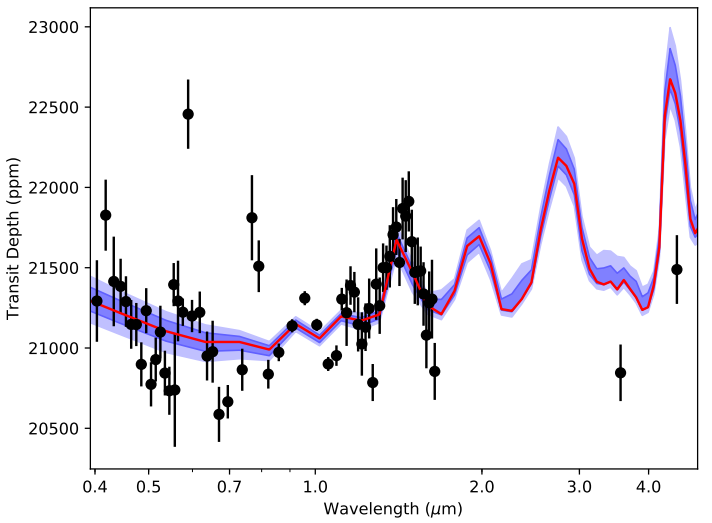
(b) Rough *HST* resolution

Figure 3.4: Transmission spectrum for HAT-P-26b generated from ATMO (obtained from Goyal et al., 2018b) and PLATON for an isothermal atmosphere with temperature 851 K, 10x solar metallicity, C/O ratio of 0.35, and haze factor/scattering factor of 1. In the right panel, we show the transmission spectrum smoothed by a Gaussian filter of $\sigma = 15$ to roughly match the resolution of *HST*.

which was retrieved by directly using the ATMO code, is lower than that retrieved by both PLATON and the ATMO grid.



(a) HAT-P-26b



(b) WASP-39b

Figure 3.5: The best fit transmission spectrum for HAT-P-26b and WASP-39b (red) and the $\pm 1\sigma$ as well as the $\pm 2\sigma$ uncertainties shown in blue.

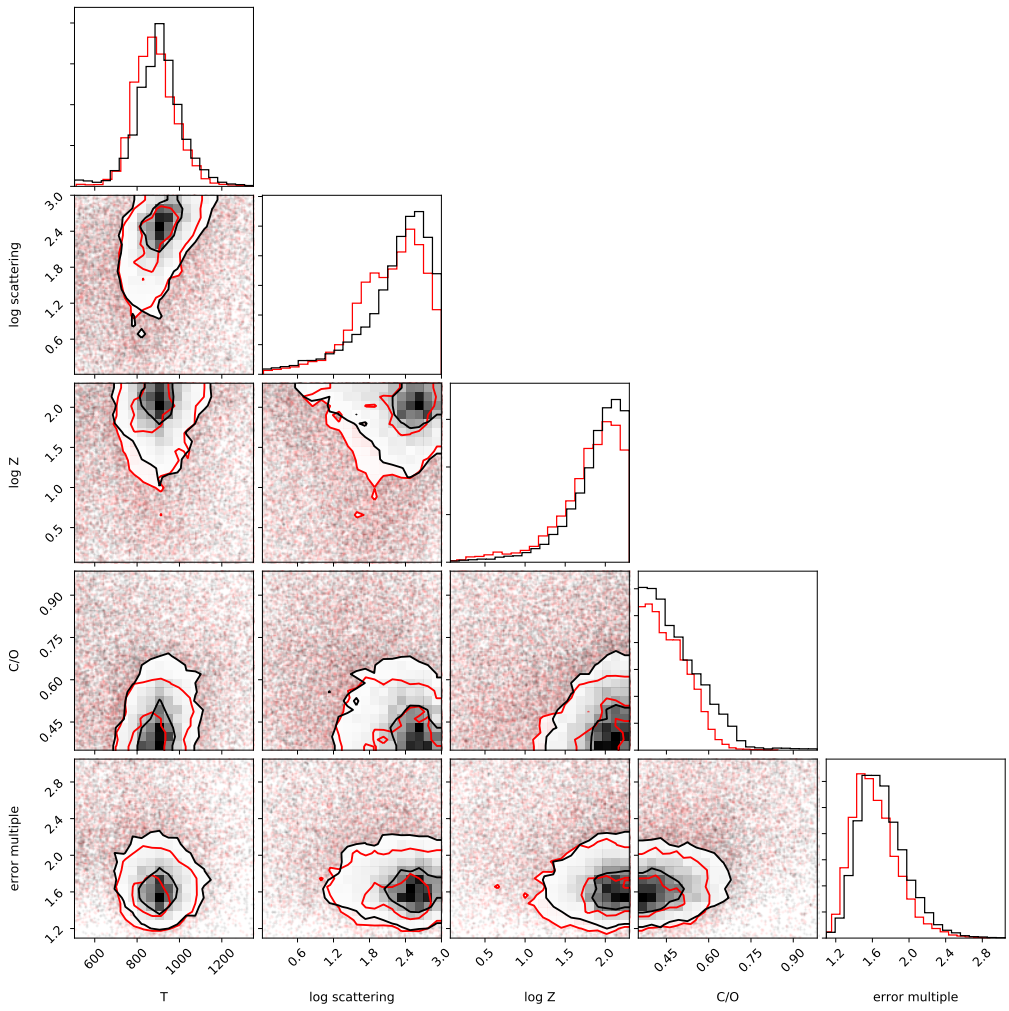


Figure 3.6: Posterior probability distribution of planetary and atmospheric properties for HAT-P-26b. Red: retrieved with PLATON; black: retrieved with ATMO generic grid.

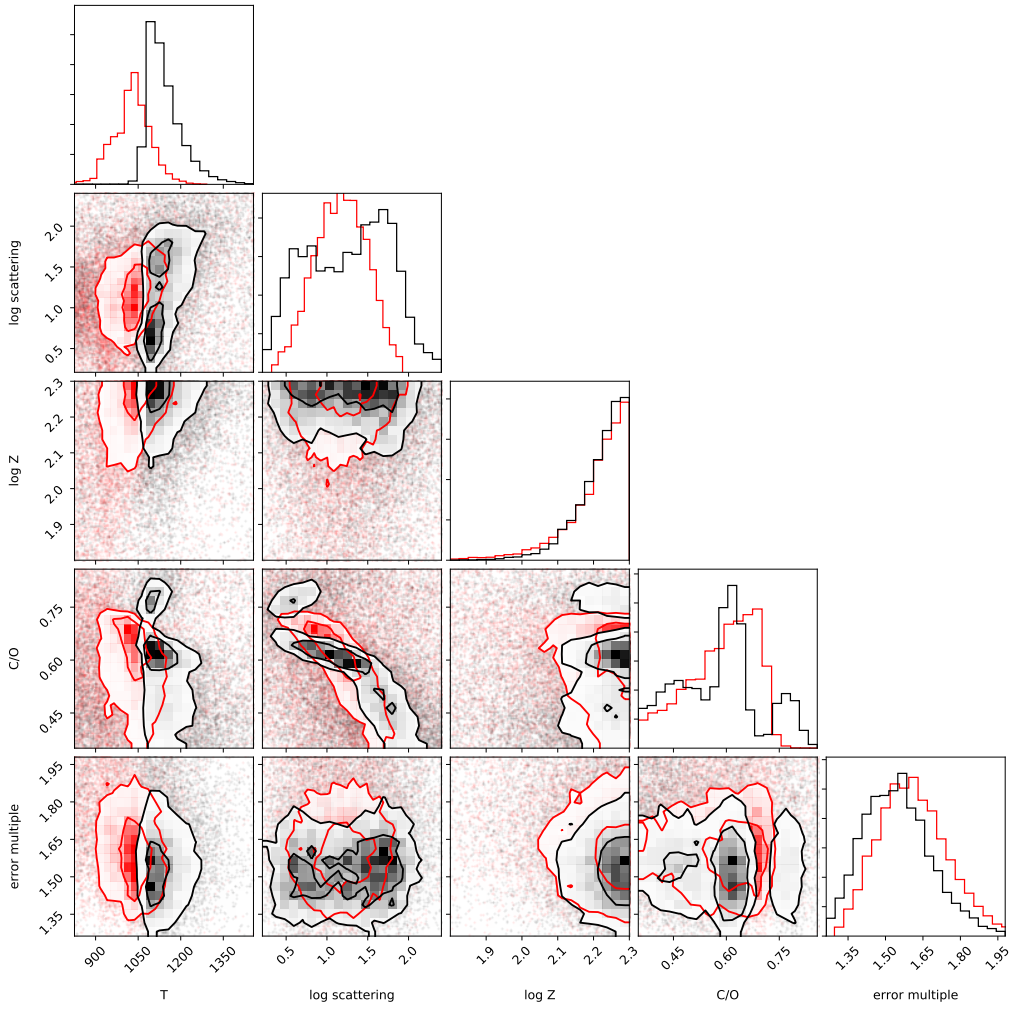


Figure 3.7: Posterior probability distribution of planetary and atmospheric properties for WASP-39b. Red: retrieved with PLATON; black: retrieved with ATMO generic grid.

3.5 Precautions, limitations, and best practices

PLATON is a powerful tool, but we were required to make some compromises in order to achieve our goal of a fast, easy to use package written in pure Python. In this section, we address some of the limitations of PLATON that users should be aware of. The biggest source of error in PLATON comes from its relatively low R=1000 spectral resolution, which does not allow us to resolve individual lines at typical atmospheric pressures (P < 1 bar). When we sample these unresolved lines using our relatively coarse wavelength grid (this method is known as "opacity sampling") it leads to spikiness in the cross sections and corresponding transmission/emission spectra, resulting in the errors visible in Figure 3.8. The idea behind opacity sampling is that even though the sampling resolution is much lower than that needed to resolve individual lines, it is still much higher than the instrumental resolution, and the spikiness in transit/eclipse depths can be smoothed out to an acceptable level by binning to instrumental resolution.

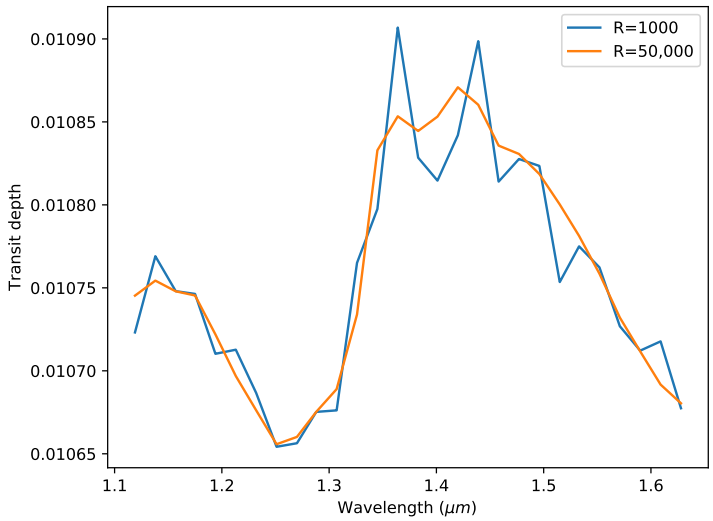


Figure 3.8: Transit depths across the WFC3 band for a 1200 K Jupiter orbiting a Sun-like star, computed assuming water-dominated opacity at two resolutions: R=1000 (native PLATON), and R=50,000 (assumed to be the truth). The depths are binned with a bin size of 19 nm. PLATON deviates from the true value by up to 50 ppm, although the deviation is much smaller at most wavelengths.

In practice this means that users will obtain the best results when they utilize wavelength bins that are large relative to the intrinsic model resolution. The number of wavelength grid points in a bin must be large enough to effectively average out the spikiness in transit depths caused by opacity sampling within that bin. For example, a hot Jupiter at T=1200 K with solar metallicity has transit depths that

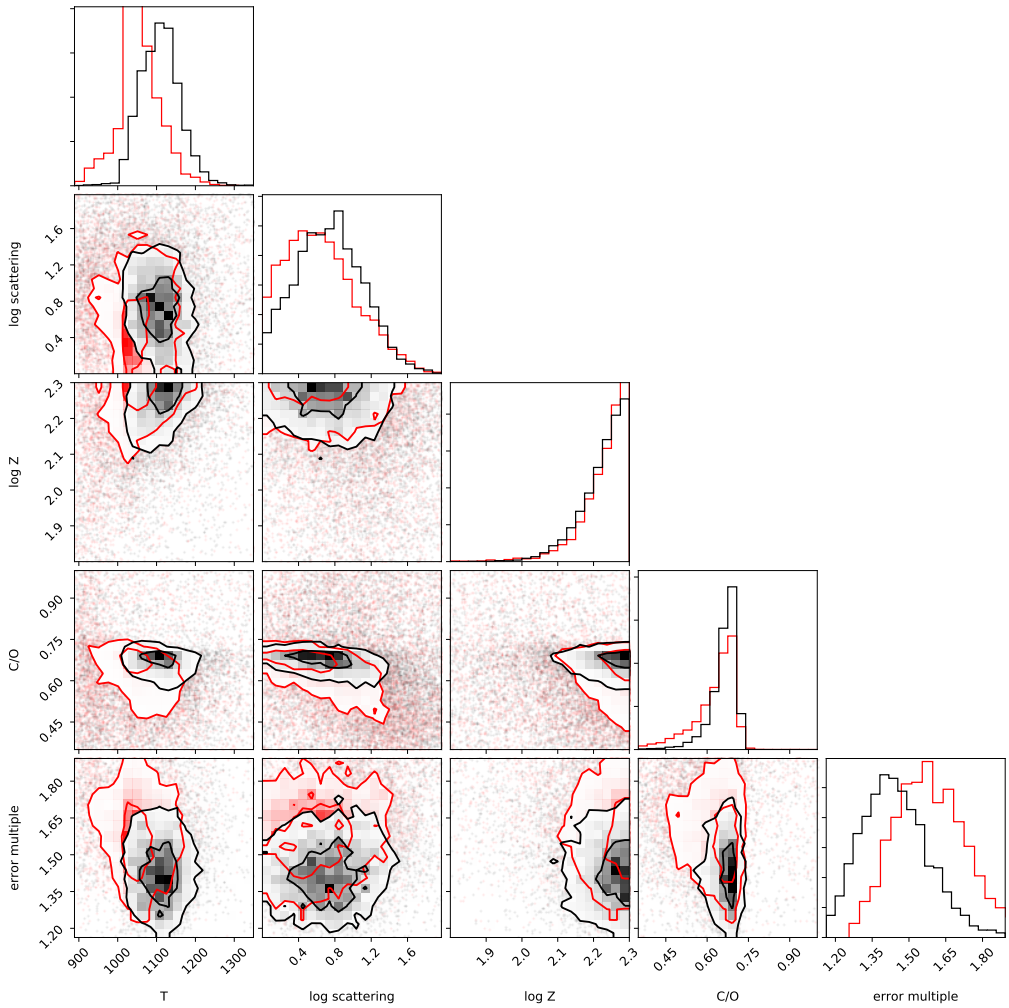


Figure 3.9: Comparison of retrievals on WASP-39b transit depths above 500 nm, using two opacity data resolutions: R=1000 (red) and R=10,000 (black). As expected, the R=1000 retrieval has a higher error_multiple and wider posteriors, reflecting the errors introduced by opacity sampling. The low-resolution retrieval favors slightly lower temperatures (by 60 K), but the posteriors are otherwise very similar.

fluctuate roughly 100 ppm from grid point to grid point at 1.4 μm, the middle of the *HST* WFC3 band. At this wavelength, the wavelength grid points are separated by 1.4 nm, so a 30 nm wide wavelength bin would include 21 points. The error generated by opacity sampling is then ~100/sqrt(21) = 22 ppm, which is appreciably smaller than the typical published uncertainties on WFC3 data sets (i.e. Wakeford et al. 2017, 2018). This sampling error is effectively "white" in wavelength space—i.e., it does not systematically bias the transit depth in one direction, but instead introduces random scatter around the true value. If the observational data have errors

comparable to or smaller than 22 ppm, retrievals with R=1000 will likely give only slightly wider posteriors than if we used line-by-line radiative transfer. However, it is worth remembering that our model also makes other simplifying assumptions, such as an isothermal atmosphere, cloud opacity parameterization, and equilibrium chemistry, that may cause deviations from the measured transmission spectrum that are much greater than either the observational errors or the sampling errors.

If the default R=1000 opacities are too low resolution for a given retrieval, we also provide optional opacity files at R=2000 and R=10,000, available at `http://www.astro.caltech.edu/~mz/absorption.html`. These only span 0.5-12 μm, not 0.3-30 μm, but the full wavelength coverage will be added in version 3. Users can replace the existing opacity files in PLATON's data directory with a higher resolution version and then run PLATON as usual. No code changes are necessary, but run time and memory usage will grow linearly with resolution.

Our tests indicate that these high-resolution data are not usually necessary for STIS, WFC3, or *Spitzer* data. Figure 3.9 compares a retrieval performed at R=10,000 to one performed at R=1000. The higher error_multiple and wider posteriors caused by opacity sampling can clearly be seen, but the posteriors are otherwise very similar. The high-resolution retrieval was 15 times slower, taking 215 minutes with 100 live points compared to 14 minutes at R=1000. Fisher & Heng (2018) performed a similar test—they tested resolutions of 1, 2, 5, and 10 cm^{-1} for WFC3 retrievals and found that "for all of these values, the posterior distributions of T, X_{H_2O}, and k_{cloud} are somewhat similar." The difference between the best and worst resolutions tested amounted to 0.5 sigma in temperature and water mixing ratio. In addition, the 2D posterior distributions in their Figure 6 are very similar. Our resolution of R=1000 corresponds to a wavenumber resolution of 6-9 cm^{-1}, so we are between their 5 and 10 cm^{-1} results in terms of the error introduced by opacity sampling.

We next provide advice on how to ensure an accurate posterior when using nested sampling or MCMC. For nested sampling, we have found that 100 live points is usually sufficient to obtain accurate 1D posteriors unless the posterior space is highly multimodal. However, it is difficult to produce a publication-quality plot of the 2D posteriors with only 100 live points, as the low density of points in parameter space makes the contours look ragged and broken up. We therefore recommend 100 live points for exploratory data analysis, and 1000 live points to generate publication-quality corner plots. For MCMC, we find that 1000 steps with 50 walkers (resulting in 50,000 samples) is typically enough to sample more than

50 times the autocorrelation length. We recommend 1000 steps for exploratory data analysis, and 10,000 steps to produce publication-quality posteriors. For cases with widely separated multimodal distributions, we note that the MCMC method will perform poorly as walkers will have a hard time moving between peaks. This is a well known limitation of MCMC, and we therefore recommend that users switch to nested sampling in these cases.

Other best practices

As its name implies, PLATON is designed to be accessible to users with minimal experience in modelling atmospheres. To that end, we offer several suggestions to help newcomers do retrievals with minimal pain. The first step in a retrieval is to decide which parameters to fix and which to retrieve, and following that, whether to use uniform or Gaussian priors. The optimal choice varies on a case by case basis according to the quality of the user's data in comparison to the published data. Nevertheless, there are a few guidelines that nearly always apply. The stellar temperature, for example, has almost no impact on the transmission spectrum, and published values are always much more accurate than what one can derive from a retrieval on an exoplanet atmosphere. The stellar temperature should be fixed. Similarly, the spot fraction should be fixed to 0 for inactive stars, and restricted to an appropriately small range for active stars. On the other hand, the metallicity and cloud-top pressure are typically not known in advance and should therefore be included as free parameters in the fit with log-uniform priors. The stellar radius and planet mass usually have published values with an intermediate accuracy—not low enough that they should be ignored, but not high enough that they can be safely fixed to a single value. This is a good use case for Gaussian priors, which take into account the published value's mean and standard deviation, while still allowing the parameter to vary within the published uncertainties as part of the fit.

We recommend that users begin with the parametric cloud and haze model and only switch to Mie scattering if the parametric model does not result in a good fit. The parametric model is less physically motivated, but it has fewer free parameters and is less time consuming. If the user does decide to try Mie scattering, we recommend leaving the particle size geometric standard deviation at the default value of 0.5. This parameter has been measured for Earth aerosols (e.g. Shen et al. 2015; Pinnick et al. 1978; Elias et al. 2009) and ranges from 0.25 to 0.7, with a typical value of 0.5. It is worth noting that the aerosol literature usually quotes e^{σ_g} instead of σ_g.

When we refer to standard deviation, we always mean the σ_g parameter in Equation 3.8.

After choosing which parameters to fit, the user must pick default values and corresponding ranges for all model parameters used in the fit. There is no point including a region of parameter space that is clearly unphysical, such as a temperature of 300 K for the most irradiated planet ever discovered, nor is there a point in restricting a parameter to a narrower range than justified by the current state of knowledge.

Next, the user might wonder whether to use MCMC or nested sampling. In our experience it is best to start with nested sampling, because it is faster (usually finishing within minutes) and has a natural stopping point automatically determined by `nestle`. We have, however, found pathological cases where `nestle` samples the parameter space extremely inefficiently; in these case, MCMC is necessary to have a reasonably fast retrieval. With MCMC, it is necessary to set the number of walkers and the number of steps by hand, and checking for convergence is not trivial. However, one advantage of MCMC is that it gives an unbiased sample from the posterior distribution, making it easy to plot the posterior probability distributions (also known as "corner plots"). We refer to Foreman-Mackey et al. (2013) for tips on choosing `emcee` parameters. With `nestle`, the best one can do is get biased samples along with their weights, which can then be resampled into equal-weight samples. Overall, the results from MCMC and nestled sampling are very similar, and we encourage users to try both to see which one they prefer.

3.6 Conclusion

We have developed a transmission spectrum calculator and retrieval tool in pure Python. We release it on GitHub, and encourage the community to use it, contribute to it, and incorporate it in whole or part into other software. This paper describes version 2.0 of the package, but we intend to keep it under continuous development—adding features, writing more unit tests, increasing user friendliness, and fixing bugs. There is already a beta version of an eclipse depth calculator and retriever, and we intend to allow users to specify wavelength-dependent refractive indicies for Mie scattering in the near future.

PLATON is not designed to model every physical phenomenon on exoplanets, or even to keep up with the cutting edge in theory. Its niche is to be a fast, simple, and easily modifiable tool: a PLanetary Atmospheric Transmission tool for Observer Noobs.

3.7 Acknowledgements

M.Z. acknowledges Plato (Greek:ΠΛΑΤΩΝ) for his insightful philosophy and his contribution to the package's name.

We thank Carlos E. Munoz for generating high-resolution opacity data in a format suitable for PLATON. We thank Jayesh Goyal for assistance in comparing PLATON to ATMO.

Support for this work was provided by *HST* GO programs 13431, 13665, and 14260.

Software: numpy, scipy, matplotlib, emcee, nestle, corner, nose, Travis-CI

References

Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765

Benneke, B., & Seager, S. 2012, ApJ, 753, 100

Bétrémieux, Y., & Swain, M. R. 2017, MNRAS, 467, 2834

Blecic, J. 2016, ArXiv e-prints, 1604.02692

Blecic, J., Dobbs-Dixon, I., & Greene, T. 2017, ApJ, 848, 127

Blecic, J., Harrington, J., & Bowman, M. O. 2016, ApJS, 225, 4

Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436

Deming, D., Wilkins, A., McCullough, P., et al. 2013, ApJ, 774, 95

Deming, L. D., & Seager, S. 2017, Journal of Geophysical Research (Planets), 122, 53

Elias, T., Haeffelin, M., Drobinski, P., et al. 2009, Atmospheric Research, 92, 443

Evans, T. M., Aigrain, S., Gibson, N., et al. 2015, MNRAS, 451, 680

Fisher, C., & Heng, K. 2018, MNRAS, 481, 4698

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, ApJS, 214, 25

Freedman, R. S., Marley, M. S., & Lodders, K. 2008, ApJS, 174, 504

Goyal, J. M., Wakeford, H. R., Mayne, N. J., et al. 2018a, MNRAS

Goyal, J. M., Mayne, N., Sing, D. K., et al. 2018b, MNRAS, 474, 5158

Kempton, E. M.-R., Lupu, R., Owusu-Asare, A., Slough, P., & Cale, B. 2017, PASP, 129, 044402

Kipping, D. M., & Tinetti, G. 2010, MNRAS, 407, 2589

Kitzmann, D., & Heng, K. 2018, MNRAS, 475, 94

Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007, ApJ, 655, 564

Line, M. R., Wolf, A., Zhang, X., & Yung, Y. 2013, in American Astronomical Society Meeting Abstracts, Vol. 221, American Astronomical Society Meeting Abstracts #221, 224.02

Lupu, R. E., Zahnle, K., Marley, M. S., et al. 2014, ApJ, 784, 27

Madhusudhan, N. 2018, ArXiv e-prints, 1808.04824

Madhusudhan, N., Knutson, H., Fortney, J. J., & Barman, T. 2014, Protostars and Planets VI, 739

Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24

Marley, M. S., Ackerman, A. S., Cuzzi, J. N., & Kitzmann, D. 2013, Clouds and Hazes in Exoplanet Atmospheres, ed. S. J. Mackwell, A. A. Simon-Miller, J. W. Harder, & M. A. Bullock, 367–391

McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, ApJ, 791, 55

Pinnick, R. G., Hoihjelle, D. L., Fernandez, G., et al. 1978, Journal of Atmospheric Sciences, 35, 2020

Pluta, T., Sadlej, A. J., & Bartlett, R. J. 1988, Chemical Physics Letters, 143, 91

Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122

Shabram, M., Fortney, J. J., Greene, T. P., & Freedman, R. S. 2011, ApJ, 727, 65

Shaw, J. R., Bridges, M., & Hobson, M. P. 2007, MNRAS, 378, 1365

Shen, X. J., Sun, J. Y., Zhang, X. Y., et al. 2015, Atmospheric Environment, 120, 307

Stock, J. W., Kitzmann, D., Patzer, A. B. C., & Sedlmayr, E. 2018, MNRAS, 479, 865

Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, ApJ, 804, L17

Wakeford, H. R., Sing, D. K., Kataria, T., et al. 2017, Science, 356, 628

Wakeford, H. R., Sing, D. K., Deming, D., et al. 2018, AJ, 155, 29

Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107

Woitke, P., Helling, C., Hunter, G. H., et al. 2018, A&A, 614, A1

Zingales, T., & Waldmann, I. P. 2018, ArXiv e-prints, 1806.02906

Chapter 4

PLATON II: ECLIPSE DEPTHS AND UPDATED OPACITIES

Michael Zhang, Yayaati Chachan, Eliza M. R. Kempton, Heather A. Knutson, and Wenjun (Happy) Chang. PLATON II: New Capabilities and a Comprehensive Retrieval on HD 189733b Transit and Eclipse Data. *ApJ*, 899(1):27, August 2020. doi: 10.3847/1538-4357/aba1e6. M.Z. conceived the project, wrote the vast majority of the code, and wrote the paper.

4.1 Abstract

Recently, we introduced PLanetary Atmospheric Tool for Observer Noobs (`PLATON`), a Python package that calculates model transmission spectra for exoplanets and retrieves atmospheric characteristics based on observed spectra. We now expand its capabilities to include the ability to compute secondary eclipse depths. We have also added the option to calculate models using the correlated-k method for radiative transfer, which improves accuracy without sacrificing speed. Additionally, we update the opacities in `PLATON`—many of which were generated using old or proprietary line lists—using the most recent and complete public line lists. These opacities are made available at R=1000 and R=10,000 over the $0.3 - 30$ μm range, and at R=375,000 in select near IR bands, making it possible to utilize `PLATON` for ground-based high resolution cross correlation studies. To demonstrate `PLATON`'s new capabilities, we perform a retrieval on published *HST* and *Spitzer* transmission and emission spectra of the archetypal hot Jupiter HD 189733b. This is the first joint transit and secondary eclipse retrieval for this planet in the literature, as well as the most comprehensive set of both transit and eclipse data assembled for a retrieval to date. We find that these high signal-to-noise data are well-matched by atmosphere models with a C/O ratio of $0.66^{+0.05}_{-0.09}$ and a metallicity of 12^{+8}_{-5} times solar where the terminator is dominated by extended nanometer-sized haze particles at optical wavelengths. These are among the smallest uncertainties reported to date for an exoplanet, demonstrating both the power and the limitations of *HST* and *Spitzer* exoplanet observations.

4.2 Introduction

The emission spectra of exoplanets provide unique insights into their atmospheric properties (e.g., Madhusudhan, 2018). By measuring the difference between the in-

eclipse and out-of-eclipse flux when the planet passes behind the star ('secondary eclipse'), one can measure the flux emitted by the planet as a function of wavelength. Secondary eclipse observations probe the compositions and temperature-pressure profiles of their dayside atmospheres. This technique was used to derive the first atmospheric composition measurement for a Neptune-mass planet (GJ 436b; Stevenson et al. 2010), the first definitive detection of a thermal inversion in the atmosphere of an ultra-hot Jupiter (Haynes et al., 2015), and dayside water abundance measurements for several hot Jupiters (e.g. Kreidberg et al. 2014; Line et al. 2016; Pinhas et al. 2019). Emission spectroscopy at high spectral resolution ($R > 20,000$) has led to the detection of CO, H_2O, and HCN in exoplanet atmospheres (i.e. Snellen et al. 2010; Birkby 2018), including in non-transiting planets like 51 Pegasi b (Brogi et al., 2013). By measuring the atmospheric compositions of these planets, we can obtain new insights into present-day atmospheric processes such as disequilibrium chemistry (e.g., Moses et al. 2013), as well as their past formation and migration histories (e.g., Öberg et al. 2011; Madhusudhan et al. 2014; Ali-Dib 2017; Cridland et al. 2019; Booth & Ilee 2019).

Ideally, we would extract atmospheric parameters from observed emission spectra using a Bayesian retrieval code. However, there is an overall lack of open source retrieval codes that can handle emission spectra. A similar lack of retrieval codes for transit spectra motivated us to write `PLATON` (Zhang et al., 2019), a fast, open source, easy to use, and easy to understand forward modeling and retrieval code that traces its lineage back to Exo-Transmit (Kempton et al., 2017). `PLATON` has since been used in several papers: a few exploring the atmospheric properties of observed planets (Chachan et al., 2019; Kirk et al., 2019; Guo et al., 2020), and one demonstrating the possibility of using K-means clustering to speed up retrievals by 40% (Hayes et al., 2020). In the latter study, the speed of `PLATON` was especially useful due to the necessity of running many retrievals. We now expand `PLATON`'s capabilities to include thermal emission and compare the resulting models and atmospheric retrievals to that of another retrieval code in order to validate this new functionality.

We utilize the open source TauREx code (Waldmann et al., 2015; Al-Refaie et al., 2019), which has been used in multiple published studies (i.e. Komacek et al. 2019; Shulyak et al. 2019), for this comparison. Specially, we compare to TauREx 3, the latest release. `PLATON` and TauREx were developed independently of each other, and there are several key differences between their functionalities. TauREx is a sophisticated code which supports free retrieval of chemical abundances in

addition to equilibrium chemistry retrievals. PLATON only allows for retrievals using equilibrium chemistry, with atmospheric metallicity and C/O ratio as free parameters. When calculating this equilibrium chemistry, PLATON uses GGchem, which can account for losses due to condensation (Woitke et al., 2018), whereas TauREx assumes that everything stays in the gas phase. PLATON also uses opacities generated from the latest line lists for water (Polyansky et al., 2018), methane (Rey et al., 2017), and ammonia (Coles et al., 2019), which are significantly more complete and accurate than the line lists that were available at the time TauREx was released.

The two codes also differ in their treatment of aerosols. Both support Rayleigh scattering, although PLATON also supports Rayleigh-like scattering by allowing for a variable scattering strength and slope ($\sigma(\lambda) = A\sigma_{\mathrm{R}ayleigh}(\lambda)(1\mu m/\lambda)^s$). Both also support Mie scattering, but with different parameterizations. TauREx has three ways of approximating Mie opacity: a gray opacity, the parameterization of Lee et al. (2013) ($Q_{ext} = \frac{5}{Q_0 x^{-4} + x^0.2}$), and an analytical calculation of the Mie opacity of spherical particles with a size distribution given by Equation 36 or Equation 37 of Sharp & Burrows (2007). PLATON analytically calculates the Mie opacity of spherical particles with a lognormal size distribution, as explained in Zhang et al. (2019). In this study, we explore the impact of these differences on our models using the benchmark hot Jupiter HD 189733b (Bouchy et al., 2005) as our test case.

Aside from TauREx, other open source retrieval codes include the recently released Helios-r2 (Kitzmann et al., 2020) and the Bayesian Atmospheric Radiative Transfer (BART) code (Blecic, 2016; Blecic et al., 2017). Helios-r2 is primarily intended for brown dwarfs and supports both free retrievals and gas-only equilibrium retrievals without clouds. It does not use a parameterized T/P profile, but retrieves the temperatures of individual layers in the atmosphere, with constraints on how much the temperature can vary from layer to layer. This approach is suitable for the high signal-to-noise regime of brown dwarfs, but not ideal for exoplanets.

BART is partially described in one subsection of a dissertation (Blecic, 2016) and one subsection of a paper (Blecic et al., 2017), but has not been described in detail in a peer reviewed paper. BART uses a custom MCMC code, MC^3, to perform retrievals using either free abundances or gas-only equilibrium chemistry, the latter of which is computed by TEA (Blecic et al., 2016). It supports two T/P profile parameterizations: Madhusudhan & Seager (2009) and Guillot (2010). However, BART does not provide opacity data, and the opacity calculator it provides does not

support ExoMol line lists. In addition, TEA is slow, taking 2-3 s of CPU time per temperature/pressure point, and has convergence problems below 400 K (Woitke et al., 2018).

Last in our roundup of retrieval codes is petitRADTRANS, a forward modelling code that does not support retrievals, but is fast enough to be wrapped in a Bayesian retrieval framework (Mollière et al., 2019). Mollière et al. (2019) benchmark against `PLATON` and find that although `PLATON` is much faster overall, the two codes are comparable in speed for the same number of wavelength points. petitRADTRANS only supports free abundances, and uses either correlated-k (R=1000) or line-by-line (R=10^6) radiative transfer, with a T/P profile parameterization that is a variant of Guillot (2010). To our knowledge, `PLATON` is unique among open source retrieval codes in supporting equilibrium condensation in a Bayesian retrieval framework, and our opacities are based on the most up-to-date line lists.

We utilize `PLATON` to carry out the first joint retrieval on published emission and transmission spectroscopy for HD 189733b, resulting in improved constraints on its atmospheric composition. `PLATON` is also capable of calculating transit and eclipse spectra at R=375,000, and we compare our model to published high-resolution ($R \sim 100,000$) CRIRES emission spectroscopy for this planet to search for previously reported signatures of H_2O and HCN (de Kok et al., 2013; Birkby et al., 2013; Brogi et al., 2016; Cabot et al., 2019) in addition to an unreported molecule: CH_4.

In Section 4.3, we describe the emission spectrum and secondary eclipse depth calculator. Section 4.4 describes the opacity update, while Section 4.5 describes other new features and improvements in `PLATON`, including the new correlated-k capability and optional model parameters, including a chemical quench pressure and a *HST*/WFC3 offset. In Section 4.6, we perform a joint retrieval of the transit and eclipse spectra of HD 189733b to infer its atmospheric properties and confirm published high-resolution detections of H_2O while calling into question the reported high-resolution detection of HCN. We summarize our conclusions in Section 4.7.

4.3 Emission Spectrum and Secondary Eclipse depth calculator

Algorithm

Our emission spectrum and secondary eclipse depth calculator utilizes much of the same code as the transit depth calculator presented in our first paper (Zhang et al., 2019). Given a planetary mass, radius, metallicity, C/O ratio, and temperature-pressure profile, we compute equilibrium molecular abundances for 250 pressures

uniformly distributed in log(P). We then solve the hydrostatic equation to determine the height corresponding to each pressure. With molecular abundances, temperatures, and pressures as a function of height, the emergent flux is given by:

$$F_{p\lambda} = 2\pi \int_0^\infty \int_0^1 B_\lambda(\tau_\lambda) e^{-\tau_\lambda/\mu} d\mu d\tau_\lambda \tag{4.1}$$

where $B_\lambda(\tau_\lambda)$ is the Planck function at an optical depth of τ_λ, and μ is the cosine of the viewing angle with respect to the vertical. Here, we are making the assumption that the source function is the Planck function, which in turn requires that scattering contributes negligibly to the emission. Adding scattering as an emission source would make the problem much more complex because it would be non-local: the source function at a certain location would depend on scattered photons from other locations.

Directly integrating Equation 4.1 would require evaluating the integrand hundreds of millions of times—once for every combination of wavelength, τ_λ, and μ. We instead rewrite the double integral as a single integral with a special function:

$$F_{p\lambda} = 2\pi \int_0^\infty B_\lambda(\tau_\lambda) E_2(\tau_\lambda) d\tau_\lambda \tag{4.2}$$

where E_2 is the exponential integral, defined as:

$$E_2(x) = \int_1^\infty \frac{e^{-xt}}{t^2} dt. \tag{4.3}$$

We discretize and reformulate the equation as follows:

$$F_{p\lambda} \approx 2\pi \sum_0^N B(\tau_{\lambda,i}) E_2(\tau_{\lambda,i}) \Delta\tau_{\lambda,i} \tag{4.4}$$

$$\approx 2\pi \sum_0^{N-1} -B(\tau_{\lambda,i})(E_3(\tau_{\lambda,i+1}) - E_3(\tau_{\lambda,i})), \tag{4.5}$$

$$\tag{4.6}$$

where the second equation follows from the first because the integral of $E_2(x)$ is $-E_3(x) + C$. Reformulating the equation in this way has the virtue of guaranteeing

that the result is exactly correct for an isothermal atmosphere. For an isothermal atmosphere, B is constant; therefore, for every layer i except the two boundaries, the $E_3(\tau_{i+1})$ term is cancelled by the $-E_3(\tau_i)$ term in layer i+1. The numerical integration therefore gives $2\pi B(E_3(\tau_{\lambda,0}) - E_3(\tau_{\lambda,N}))$. Assuming the top of the atmosphere has an optical depth of 0 and the bottom has an optical depth of infinity, we obtain $F = \pi B$—exactly the correct result for an isothermal atmosphere. In practice, we find that this trick reduces the error for non-isothermal atmospheres as well, especially when the temperature gradient is small. This method of replacing $g(x)\Delta x$ with $G(x_2) - G(x_1)$ (where g(x) is the derivative of G(x)) before integrating was inspired by the TauREx source code, which uses the same technique in combination with Gaussian quadrature.

The E_3 exponential integral is a special function defined by scipy, eliminating the need to perform integrals to evaluate it. It is also continuous, infinitely differentiable, and approaches 0.5 as $x \to 0$ and 0 as $x \to \infty$. These properties mean that E_3 poses no problems for numerical integration. To further speed up the code, we evaluate $E_3(x)$ on a logarithmic grid spanning $x = 10^{-6}$ to $x = 10^2$ during initialization of the eclipse depth calculator, and interpolate from this grid thereafter. The interpolation is accurate to 1.2×10^{-5}, and is therefore a negligible source of error. The more common approach to integrating Equation 4.1, adopted by TauREx and HELIOS-r2, is to use Gaussian quadrature. Compared to our approach, Gaussian quadrature is 4 times slower and introduces errors of ~0.2% for the test planet in Subsection 4.3 when 4 points are used (Figure 4.1). This error is utterly negligible compared to the other sources of error we explore in Subsection 4.3.

To derive a monochromatic eclipse depth from the emergent flux, we multiply by the square of the planet-to-star radius ratio and divide by the stellar emergent flux:

$$D_\lambda = \left(\frac{R_{p\lambda}}{R_s}\right)^2 \frac{F_{p\lambda}}{F_{s\lambda}}. \tag{4.7}$$

The stellar spectrum is calculated by interpolating the BT-Settl (AGSS2009) stellar spectral grid (Allard et al., 2012), as provided by the Spanish Virtual Observatory[1]. Here, the wavelength-dependent planet radius $R_{p\lambda}$ is defined as the radius at which the radial optical depth reaches one, assuming that the limb has the same temperature-pressure profile as the dayside. (This is not to be confused with the

[1] `http://svo2.cab.inta-csic.es/theory/newov2/index.php`

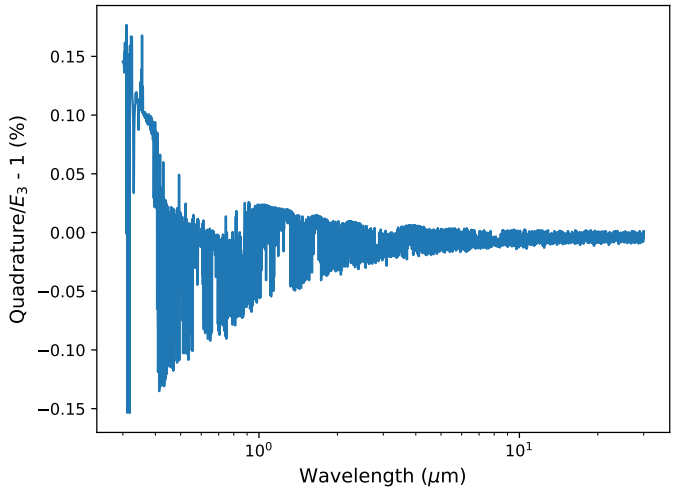


Figure 4.1: Fractional error introduced by using Gaussian quadrature with 4 points instead of E_3, for the test planet in Subsection 4.3.

white light planet radius the users pass into `PLATON`, which is the radius at a pressure of 1 bar.) This approach differs from those of previous studies (i.e. Waldmann et al. 2015), which typically fix the radius to the value measured from the optical or near-infrared transit depth. This is larger than the effective planet size for emission spectroscopy, as starlight transmitted through the planet's limb should reach an optical depth of one at lower pressures than for direct emission. Our approach, although an improvement, is still only an approximation. The planetary limb is usually colder than the day side, and even if it were not, Equation 4.7 is only strictly true if the atmospheric scale height is a negligible fraction of the planetary radius. These inaccuracies are expected to cause errors on the order of 1% for hot Jupiters, since the scale height of a hot Jupiter is of order 1% of the radius. For further discussion of the difficulty in choosing a photospheric radius and the error this introduces, we refer the reader to Fortney et al. (2019).

When fitting observational data, we typically want band-integrated fluxes and secondary eclipse depths. The band-integrated flux is equal to the number of photons emitted from the planet within the band, and the secondary eclipse depth is that value divided by the number of photons emitted from the star within the band. The eclipse depth for the band λ_1–λ_2 is then:

$$D = \frac{\int_{\lambda_1}^{\lambda_2} R_{p\lambda}^2 F_{p\lambda} \lambda d\lambda}{\int_{\lambda_1}^{\lambda_2} R_s^2 F_{s\lambda} \lambda d\lambda}. \tag{4.8}$$

We utilize Equation 4.8 when the user defines custom wavelength bins. Otherwise, we compute monochromatic fluxes and secondary eclipse depths at full spectral resolution (R=1000) using Equation 4.7.

Temperature-pressure profile

The temperature-pressure profile is a crucial component of any atmospheric emission model. This profile determines whether molecular features will be seen in absorption or emission and their relative strengths. Although it is possible to predict the T/P profile theoretically using energy balance arguments, such a self-consistent calculation is computationally expensive. This is because the radiative intensity, opacity, chemical abundance, and temperature at each pressure level all depend on one another, requiring an iterative procedure to solve for all components simultaneously. These iterative procedures are too slow for our purposes here. We instead forgo self-consistency and retrieve a parameterized T/P profile along with the atmospheric composition. We support three parametric forms for the T/P profile, listed in order of increasing complexity: isothermal, Line et al. (2013), and Madhusudhan & Seager (2009).

Isothermal

An isothermal profile always gives a blackbody planetary spectrum. This can be derived theoretically from Equation 4.1.

Line et al. (2013)

This physically motivated parameterization was invented by Guillot (2010) to shed light on exoplanet atmospheres, and then subsequently extended by Line et al. (2013). Guillot (2010) used the two-stream approximation with radiation partitioned into two distinct wavelength channels: thermal and visible. Starlight is considered purely visible, while planetary emission is considered purely thermal. The planetary atmosphere is assumed to have a single opacity κ_{th} that applies to thermal radiation everywhere, and a single opacity κ_v that applies to visible radiation everywhere. Under these simplifying assumptions, the averaged dayside T/P profile can be derived analytically:

$$\tau = \frac{\kappa_{th} P}{g} \tag{4.9}$$

$$m(\gamma) \equiv 1 + \frac{1}{\gamma}[1 + (\frac{\gamma\tau}{2} - 1)e^{-\gamma\tau}] + \gamma(1 - \frac{\tau^2}{2})E_2(\gamma\tau)] \tag{4.10}$$

$$T^4 = \frac{3}{4}T_{int}^4(\frac{2}{3} + \tau) + \frac{1}{2}T_{eq}^4 m(\gamma) \tag{4.11}$$

where T is the temperature at a certain height in the atmosphere, τ is the optical depth of thermal radiation from the top of the atmosphere corresponding to that height, $\gamma \equiv \kappa_v/\kappa_{th}$, $T_{eq} \equiv T_*\sqrt{\frac{R_*}{2a}}$ is the equilibrium temperature assuming zero albedo, E_2 is the exponential integral with n=2, and T_{int} is a temperature reflecting the amount of internal heat.

Line et al. (2013) introduce an albedo into this formulation, so that the equilibrium temperature is now $T_{eq} \equiv \beta T_*\sqrt{\frac{R_*}{2a}}$. They also introduce a second visible channel with its own opacity to allow for temperature inversions, so that the temperature is now:

$$T^4 = \frac{3}{4}T_{int}^4(\frac{2}{3} + \tau) + (1 - \alpha)\eta(\gamma_1) + \alpha\eta(\gamma_2) \tag{4.12}$$

where $\gamma_1 \equiv \kappa_{v1}/\kappa_{th}$, $\gamma_2 \equiv \kappa_{v2}/\kappa_{th}$, and α partitions the visible radiation between the two channels. Unlike Line et al. (2013), we impose the constraint that $\alpha <= 0.5$: that is, the second visible stream is by definition the minor one. Without this constraint, the two streams are interchangeable.

In total, this parameterization has six free parameters: κ_{th}, γ_1, γ_2, α, β, and T_{int}. Following Line et al. (2013), we recommend fixing T_{int} to 100 K in most cases because internal heat usually contributes negligibly to the short-period transiting planets that are most amenable to atmospheric characterization using the secondary eclipse technique. Exceptions to this rule may include eccentric planets, planets whose cooling is delayed, or planets whose radii are inflated.

Madhusudhan & Seager (2009)

Madhusudhan & Seager (2009) introduced a purely empirical T/P profile for exoplanet atmosphere modeling, which was designed to be flexible enough to approximate most published theoretical T/P profiles from forward models without having an excessive number of free parameters. This model divides the atmosphere into three

layers: a deep isothermal layer (caused by the limited interior flux compared to the stellar flux, as $dT/dr \propto L(r)$ in a radiative zone where the diffusion approximation holds), an intermediate layer that can support a thermal inversion, and an outer layer, intended to represent the optically thin region. The parametric profile is agnostic about physical assumptions (e.g. convection, optical depths) that set the temperature and pressure structure of the atmosphere. The temperatures of the three layers are then:

$$T_{outer} = T_0 + \frac{ln(P/P_0)^2}{\alpha_1^2}$$
$$T_{mid} = T_2 + \frac{ln(P/P_2)^2}{\alpha_2^2}$$
$$T_{inner} = T_3.$$

There are six free parameters in this model: T_0, P_1, α_1, α_2, P_3, T_3. Following Madhusudhan & Seager (2009), we set P_0 to the pressure at the top of the atmosphere, which for us is 10^{-4} Pa. T_2 and P_2 are set by the requirement that temperature must be continuous across region boundaries:

$$ln(P_2) = \frac{\alpha_2^2(T_0 + ln(P_1/P_0)^2/\alpha^2 - T_3) - ln(P_1)^2 + ln(P_3)^2}{2ln(P_3/P_1)}$$
$$T_2 = T_3 - \frac{ln(P_3/P_2)^2}{\alpha_2^2}.$$

Benchmarking, speed advice

One of the goals of `PLATON` is to be fast. To illustrate typical speeds, we benchmark `PLATON` on a typical desktop computer to illustrate its performance. The computer runs Ubuntu 16.04 LTS with a Core i7 7700k CPU and 16 GB of RAM.

Table 4.1: Benchmarks for desktop computation of emission forward model

Band	λ (μm)	Time (ktables/R=1k/R=10k)
All wavelengths	0.3–30	0.26/0.24/4.8 s
WFC3	1.119–1.628	0.029/0.029/0.26 s
Spitzer 3.6 μm	3.2–4.0	0.021/0.021/0.11 s
Spitzer 4.5 μm	4.0–5.0	0.020/0.021/0.11 s

When the forward model is first initialized, `PLATON` loads all relevant data files into memory. This takes 0.34 s (1.6 s for R=10,000), but is only done once. Table 4.1 shows the amount of time taken to compute eclipse depths within the most commonly used bands once `PLATON` is initialized. The time taken depends linearly on the number of wavelength grid points within the band. Since grid points are spaced uniformly in logarithmic space, the number of grid points is proportional to the ratio between the maximum and minimum wavelengths. The time taken also depends approximately linearly on the resolution, for the same reason. Because our correlated k algorithm runs at R=100 with 10 Gaussian quadrature points, it performs the same number of radiative transfer computations as the R=1000 opacity sampling method, explaining the very similar running times.

It is difficult to give a representative running time for nested sampling retrievals, because this is highly dependent on the problem at hand. The running time is proportional to the total logarithmic wavelength range, the number of live points used, and the log of the ratio between the prior parameter hypervolume and the posterior hypervolume. The hypervolume ratio depends on the width of the priors, and on the quality of the data: an exquisite dataset takes longer to retrieve on.

Despite these variations, some rough numbers are possible. With 1000 live points, generously wide priors, and the exquisite HD 189733b dataset, `dynesty` required 400,000 likelihood evaluations for the eclipse-only retrieval. A typical retrieval with a lower signal-to-noise data set and a less conservative prior range would require fewer likelihood evaluations; using 200 live points instead of 1000 would cut the number of evaluations by a factor of 5. Taking 200,000 evaluations as a typical value for 1000 live points, we see that retrieving on a dataset of WFC3, *Spitzer* 3.6 μm, and *Spitzer* 4.5 μm observations will take 0.8 hours with 200 live points and R=1000 opacities (or correlated k coefficients); 4 hours with 1000 live points and R=1000 opacities; 5.3 hours with 200 live points and R=10,000 opacities; and 27 hours with 1000 live points and R=10,000 opacities.

We recommend a staged approach to retrievals. Exploratory data analysis can be done with R=1000 opacities and 200 live points. In the process, intermittent spot checks should be performed with R=10,000 opacities and 200 live points to check the effect of resolution, and with R=1000 opacities and 1000 live points to check the effect of sparse sampling. When one is satisfied with the exploratory data analysis and is ready to finalize the results, one should run a final retrieval with R=10,000 opacities and 1000 live points. This is the approach we followed for HD 189733b,

although had we stuck with the low-resolution, sparsely sampled retrieval, none of our conclusions would have changed.

If these running times are still too slow, there is one trivial way to speed up the code by a factor of a few: by going to PLATON_DIR/data/Absorption and removing the absorption files of all molecules that have a negligible effect on the spectrum, which prevents `PLATON` from taking their opacities into account. For a hot Jupiter, for example, the vast majority of the molecules in `PLATON` (see Table 4.4 for a list) are unimportant. One might reasonably include CO, CO_2, CH_4, H_2O, NH_3, H_2S, and HCN in an emission retrieval, but neglect the other 23. This decreases the R=10k running time for the entire 0.3–30 μm range from 4.8 s to 2.1 s. In the near future, we will implement an opacity zeroer in `PLATON` that implements this functionality without having to touch the data files.

Validation

To validate `PLATON`'s new emission spectroscopy mode, we test two cases: an isothermal atmosphere and a non-isothermal atmosphere modelled on that of HD 189733b. In the case of an isothermal atmosphere, the planetary flux should be equal to that of a blackbody:

$$F_\lambda = \pi B_\lambda(T). \tag{4.13}$$

We find that this is indeed the case. The numerically evaluated flux differs from the theoretical expectation by an amount consistent with machine precision ($\sim 2^{-52} \approx 2 \times 10^{-16}$). This is not a surprise, as our numerical integration algorithm (Equation 4.6) gives exactly the correct answer for the special case of an isothermal atmosphere.

To validate the non-isothermal atmosphere case, we compared the output of `PLATON` to that of TauREx 3 (Al-Refaie et al., 2019). Using both codes, we simulated a planet meant to represent HD 189733b. This test planet has the parameters given in Table 4.2. These parameters are, respectively, the stellar temperature T_s, stellar radius R_s, planetary mass M_p, planetary radius R_p, planetary atmospheric metallicity Z relative to the Sun, planetary C/O ratio, planetary equilibrium temperature T_{eq}, and the five parameters (β, κ_{th}, κ_{v1}, κ_{v2}, and α) specifying the T/P profile following the formulation of Line et al. (2013). Because the chemical equilibrium model of TauREx does not include condensation, we passed include_condensation=False to `PLATON`, causing `PLATON` to also use a gas-only chemical equilibrium model.

TauREx is distributed with opacities for CO_2, NH_3, CH_4, CO, and H_2O; we therefore zero out the abundances of all other active gases in `PLATON` for the purposes of this test. In addition, we used a blackbody as the stellar spectrum in both codes, as the PHOENIX spectra that TauREx supports do not extend redward of 5 μm. We generated an emission spectrum from both codes, both binned to R=100 from the native resolution of 1000 for `PLATON` and 15,000 for TauREx, and compared the resulting wavelength-dependent eclipse depths. As shown in Figure 4.2, the median absolute difference between the two is 2.1%, with a 95th percentile of 17% and a maximum of 39%.

Table 4.2: Parameters of test planet

Parameter	Value
T_s	5052 K
R_s	0.751 $R_\odot$
M_p	1.129 M_J
R_p	1.144 R_J
$Z_p/Z_\odot$	20
C/O	0.7
T_{eq}	1189 K
β	1
κ_{th}	3.8×10^{-3} m^2 kg^{-1}
κ_{v1}	1.9×10^{-3} m^2 kg^{-1}
κ_{v2}	5.2×10^{-4} m^2 kg^{-1}
α	0.331

This test planet is chosen to have properties broadly similar to those of HD 189733b. T_s, R_s, and M_p are taken from Stassun et al. 2017, and we selected values for the other parameters that approximately reproduce HD 189733b's transit and eclipse spectra.

There are a number of differences between `PLATON` and Tau-REx that could explain the discrepancy in the predicted eclipse depths. First, `PLATON` performs radiative transfer at a spectral resolution of R=1000, while we ran Tau-REx with R=15,000 opacity files. Second, the two codes also handle equilibrium chemistry differently. Tau-REx uses ACE (Agúndez et al., 2012) and only considers the elements H, He, C, O, and N and 106 molecules composed of those elements, while `PLATON` also includes F, Na, Mg, Si, P, S, Cl, K, Ti, and V, along with 300 molecules composed of those elements. This results in abundance differences in the tens of percent, as seen in Figure 4.4. Third, `PLATON` also uses a wavelength-dependent radius to convert emergent flux to luminosity (Equation 4.7), but Tau-REx does not. Fourth, `PLATON`'s opacities are generated from newer line lists than the opacities currently available

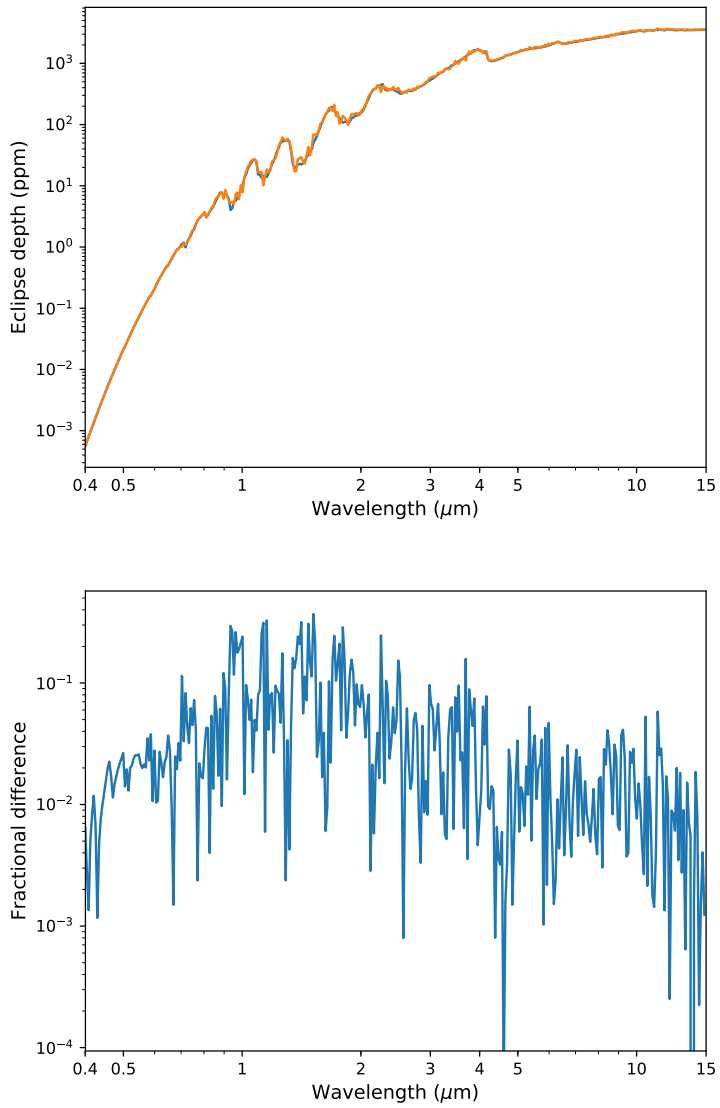


Figure 4.2: PLATON vs. Tau-REx comparison. Top: Comparison of eclipse depths computed by `PLATON` and Tau-REx for a planet meant to approximate HD 189733b. Bottom: fractional differences between the two models.

from Tau-REx's website. Finally, there are slight differences in the algorithm used to calculate the planet-star flux ratio. For example, in calculating the emergent flux (Equation 4.1), Tau-REx evaluates the integral over viewing angles by sampling four viewing angles and using Gaussian quadrature, whereas `PLATON` evaluates the integral analytically, which is equivalent to using an infinite number of viewing angles.

To disentangle which factors cause most of the differences, we modified `PLATON` step by step to more closely approximate Tau-REx's algorithm, re-measuring the

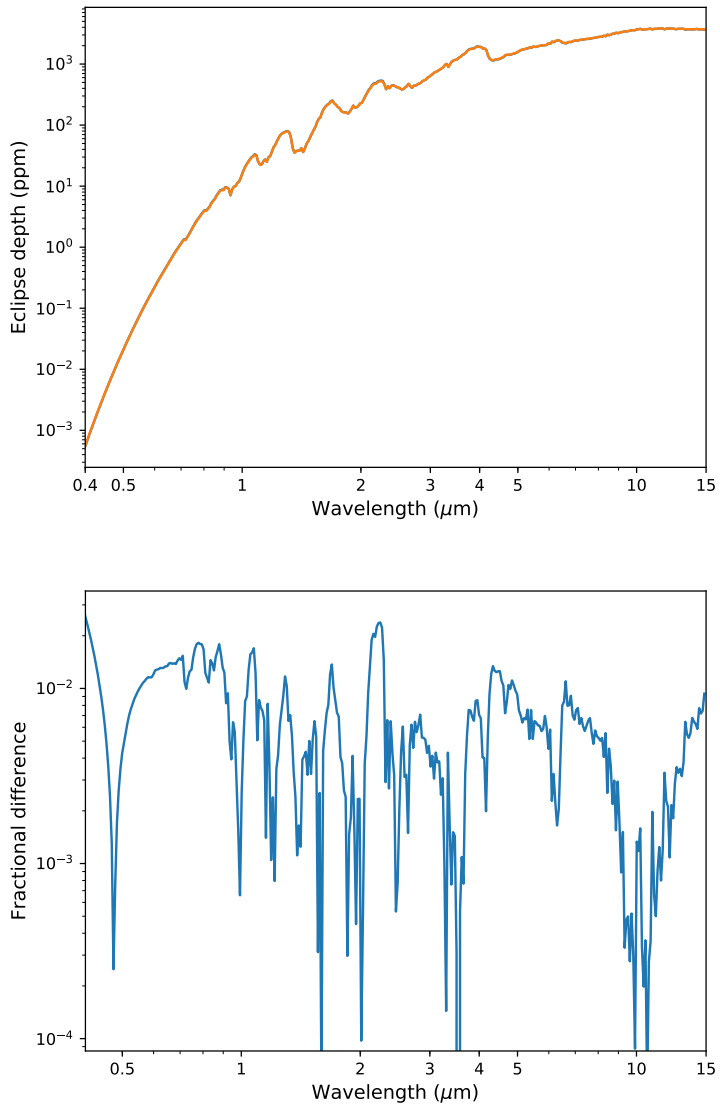


Figure 4.3: Comparison of eclipse depths computed by `PLATON` and Tau-REx (top) and corresponding residuals (bottom), where both codes are fed the same mixing ratios and the same opacities at the same resolution of R=10,000.

discrepancies at each step. The results are summarized in Table 4.3, and described in detail below. First, we replaced the default R=1000 opacities with R=10,000 opacities, which are also publicly available. This reduced the median absolute difference between Tau-REx and `PLATON` from 2.1% to 1.6%, with the 95th percentile at 7.4% and some wavelengths having a discrepancy of up to 17%. We conclude that a higher resolution leads to significant improvement in the agreement between `PLATON` and Tau-REx at some, but not most, wavelengths. Next, we disabled equilibrium chemistry and set constant abundances with altitude for CH_4, CO_2, CO,

Table 4.3: Effect of removing differences between PLATON and TauREx

Removed differences	Median diff. (%)	95th percentile diff. (%)	Max diff. (%)
None	2.2	19	37
Resolution	1.2	7.6	22
Resolution, Chemistry	1.2	4.6	11
Resolution, Chemistry, Opacities	0.6	1.7	2.7
Chemistry, Opacities	1.9	17	30
Chemistry, Opacities with ktables	2.1	8.1	20

This table shows the discrepancies between eclipse depths calculated by PLATON and TauREx after binning to R=100, for a wavelength range of 0.4–15 μm. As differences between the two codes are eliminated, their eclipse depths become more and more similar, as expected.

H_2O, NH_3, H_2, and He. This results in a median difference of 1.1% with a 95th percentile of 5.1% and maximum discrepancies of up to 10%. Following this, we replaced the line absorption cross sections in Tau-REx with those used in PLATON. This decreased the median difference to 0.8%, the 95th percentile to 2.1%, and the maximum difference from 10% to 2.7%. The spectra produced by the two codes are compared in Figure 4.3. The remaining discrepancies are likely due to slight differences in the radiative transfer code, especially the number of viewing angles (4 in Tau-REx vs. infinite in PLATON) and the precise interpolation method for absorption cross sections. As a final step, we replaced the R=10,000 opacities with the default R=1000 opacities and redid the comparison. This time, we obtained a median difference of 1.9%, a 95th percentile of 15%, and a maximum of 31%. If we do the test with correlated k radiative transfer instead of opacity sampling, these numbers are 2.2%, 7.4%, and 19%.

Based on these tests, we concluded that differences in resolution, chemistry, and opacities are all significant contributors to the discrepancies between the two codes. As shown in Table 4.3, the median error caused by these differences are on the order of 2%, with the 95th percentile being 10-20%, and the maximum difference being a few tens of percent.

Retrieval comparison

To test the validity of PLATON retrievals, we performed an equilibrium chemistry retrieval comparison between PLATON and TauREx using a synthetic spectrum. We

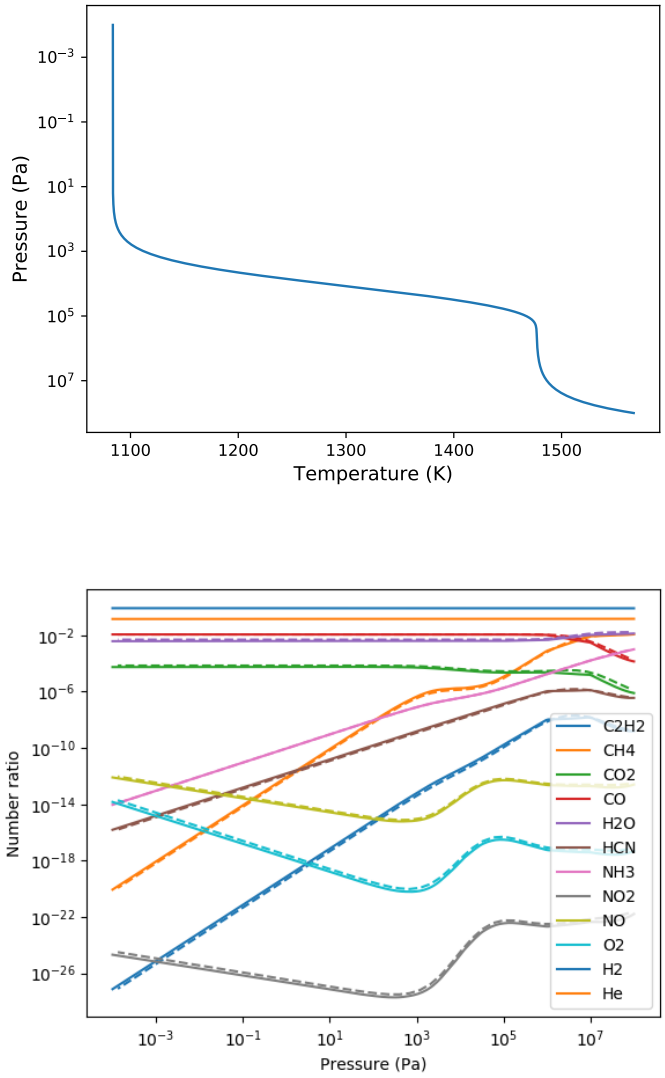


Figure 4.4: Differences in molecular abundances under equilibrium chemistry conditions between GGchem (used by `PLATON`) and ACE (used by Tau-REx) for the test planet, with the T/P profile shown in the upper panel. These differences are typically tens of percent.

used TauREx to generate the 0.4–6 μm emission spectrum using the stellar and planetary parameters in Table 4.2. The spectrum is binned down to a resolution of R=100, and 100 ppm of white noise is added to every binned eclipse depth. We then ran a retrieval on the synthetic spectrum using both TauREx and `PLATON`, with 6 free parameters (corresponding uniform priors, all generously wide, in brackets): $\log Z$ (-1–3), C/O (0.2–2.0), $\log \kappa_{th}$ (-3.4 – -1.4), $\log \gamma_{v_1}$ (-1 – 1), $\log \gamma_{v_2}$ (-1 – 1), and α (0–0.5). The `PLATON` retrieval used gas-only equilibrium abundances rather

than the default condensation equilibrium abundances, while the stellar spectrum was set to a blackbody in both codes in order to ensure that any differences in the results were due to differences in the planet model and not the stellar model. We utilized nested sampling with 1000 live points for all retrievals. For `PLATON`, the package we used was `dynesty` (Speagle, 2020); for TauREx 3, it was `nestle`.

We ran two comparisons. In the first comparison, we generated the emission spectrum using `PLATON` R=10,000 opacities, and included the same set of molecular opacities in both the `PLATON` and TauREx retrievals: NH_3, CH_4, CO, CO_2, and H_2O. In `PLATON`, the opacities of all other molecules were set to 0. Figure 4.5 shows the results of this first retrieval comparison. In general, the two codes give very similar posteriors. The 1D posteriors of $\log_{\gamma_{v_1}}$, $\log_{\gamma_{v_2}}$, and α are indistinguishable. The log(Z) and C/O posteriors show discrepancies at the 0.7σ level because the equilibrium abundances of active gases differ by a few tens of percent between `PLATON` and TauREx, which in turn is because the former includes many times more atoms and molecules in its calculations than the latter (see Subsection 4.3). As a result of these differences, `PLATON` prefers slightly higher temperatures at a given pressure (ΔT ~30 K at 100 mbar), which is reflected in the slightly higher κ_{th}.

In the second comparison, we generated the emission spectrum using TauREx R=15,000 opacities and used those same opacities in the TauREx retrieval. For the `PLATON` retrieval, we used the the full list of molecules with the same R=10,000 opacities as in the previous comparison. The differences between the two retrievals therefore reflect both the effect of including different numbers of molecules, and differences in the line lists used for those molecules. Figure 4.6 shows the results of this second retrieval comparison. Even though the `PLATON` 1D posteriors are still consistent with the input planet parameters at the 1.7σ level, they are more discrepant than in Figure 4.6, with `PLATON` obtaining a metallicity 4× lower and a C/O ratio 0.18 lower than the TauREx retrieval.

Having obtained this result, the natural question to ask is what causes the discrepancy: `PLATON`'s newer line lists, or its inclusion of more molecules? The answer is the latter. We examined `PLATON`'s best fit model and found that it underestimated the planetary emission around 3.8 μm, where an opacity window caused a spike in planetary emission. Removing molecules one by one from the atmosphere, we find that H_2S is the cause of the discrepancy: removing it alone from the atmosphere makes the best fit spectrum line up perfectly with TauREx's simulated data. Indeed, when we repeat the `PLATON` retrieval while including only the molecules

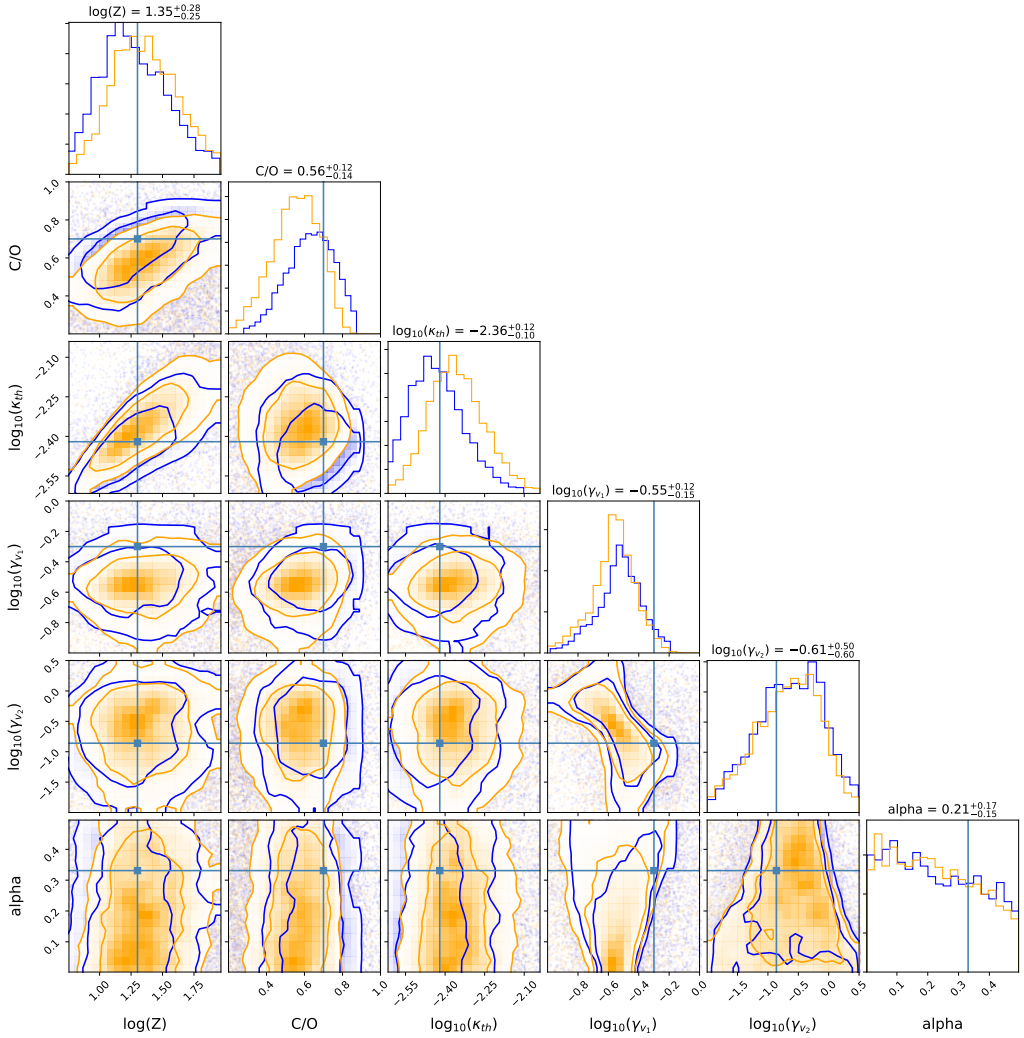


Figure 4.5: Posterior distributions retrieved by `PLATON` (orange) and TauREx (blue) when both codes use the same opacities and include the same list of molecules. The contours contain 68% and 95% of posterior mass. The numbers on top of each column show the values inferred by `PLATON`. The teal horizontal and vertical lines show the truth values.

that TauREx includes, the resulting posteriors are almost identical to those of Figure 4.5. This underscores the importance of erring on the side of caution when choosing which active gases to include in a model. Emission spectroscopy has the inconvenient property that it is the *lack* of absorption that causes the most easily detectable emission—and so even a trace gas with a relatively low opacity can have a significant impact on the spectrum at wavelengths where other gases also have less absorption.

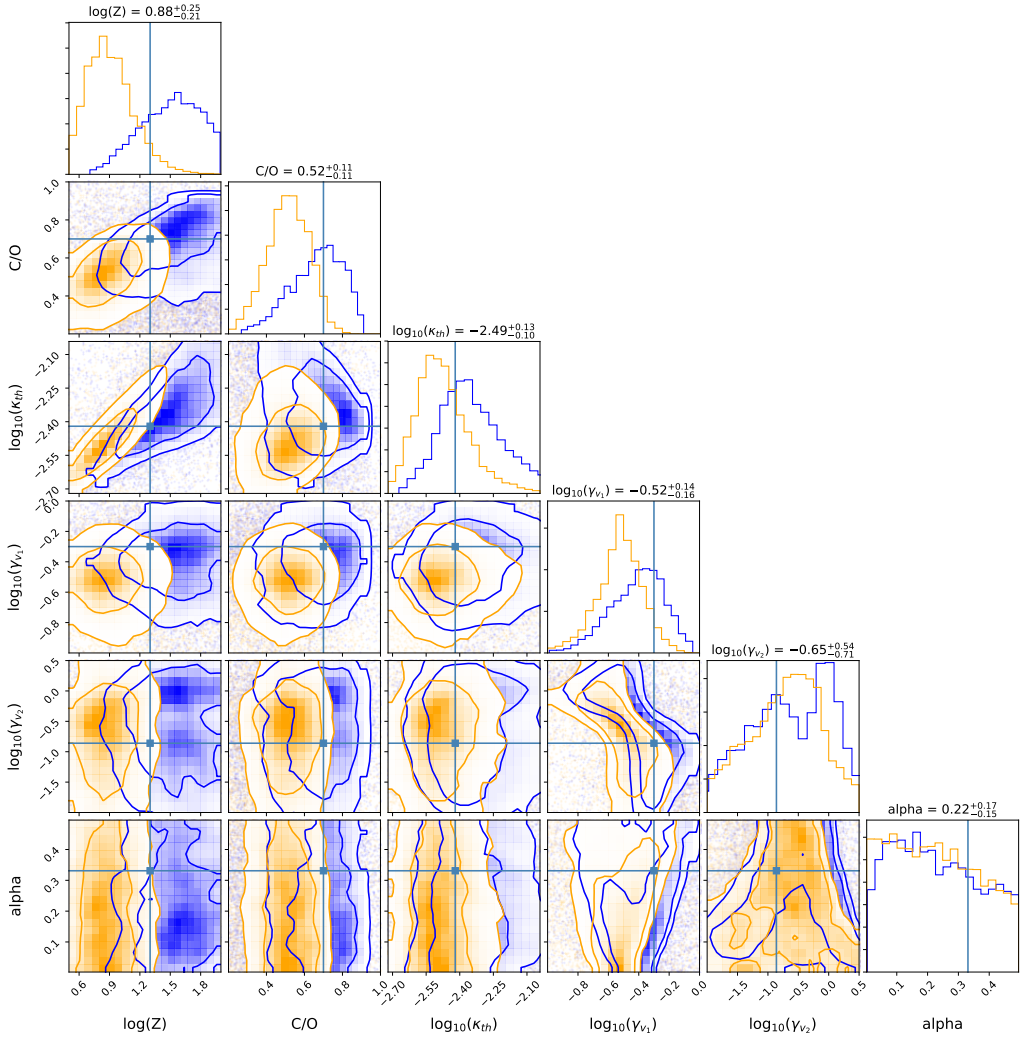


Figure 4.6: Same as Figure 4.5, but for the second comparision. Here, TauREx opacities are used to generate the spectrum and for TauREx's retrieval. The `PLATON` retrieval uses its own opacities, with all molecules included. The discrepancy is due to the inclusion of H_2S in `PLATON`, but not in TauREx.

4.4 Opacity update

One fundamental building block of any atmospheric code is the calculation of opacities. There are three types of opacities we consider: scattering, line, and collisional. As discussed in Zhang et al. (2019), scattering opacities are calculated by `PLATON` itself, and collisional opacities are calculated using the limited data available from HITRAN (Richard et al., 2012; Karman et al., 2019). Line opacities are calculated from lists of transitions from one quantum state to another, giving the position, intensity, and broadening parameters of the transitions. For this update to `PLATON`, we focused on line opacities.

In the original PLATON release, our opacity data were taken directly from Exo-Transmit (Kempton et al., 2017). Exo-Transmit, in turn, calculated its opacities from line lists generated from a large number of sources, listed in Table 2 of Lupu et al. (2014). These include HITRAN, HITEMP, private communications, Freedman et al. (2008), and Freedman et al. (2014), among many others. Many of these line lists are outdated, proprietary, or both. In addition, the program used to generate opacity data from the line lists is not public, making it difficult to reproduce our opacity calculations.

Table 4.4: Sources of line lists

Source	Molecules
ExoMol	C_2H_4, CO, H_2CO, H_2S, H_2O, HCl, HCN, MgH, NH_3, NO, OH, PH_3, SH, SiH, SiO, SO_2, TiO, VO
HITRAN 2016	C_2H_2, C_2H_6, HF, N2, NO_2, O_2, O_3, OCS
CDSD-4000	CO_2
Rey et al 2017	CH_4
NIST	Na, K

We address these shortcomings by regenerating PLATON's opacity data using the public line lists in Table 4.4. For each molecule, we generate absorption cross sections from line lists using the method outlined in ExoCross (Yurchenko et al., 2018). The cross sections are generated for 30 temperatures (100 – 3000 K in 100 K increments), 13 pressures (10^{-4} – 10^{8} Pa in decade increments), and 4616 wavelengths (0.3 – 30 μm, with uniform spacing in logarithmic space). The resolution of our wavelength grid is not high enough to resolve individual lines at typical atmospheric pressures (P < 1 bar), leading to spikiness in the wavelength-dependent cross sections, and therefore in the final transit and secondary eclipse depths. This approach to radiative transfer is called "opacity sampling." The idea behind opacity sampling is that even though the sampling resolution is much lower than that needed to resolve individual lines, it is still much higher than the instrumental resolution, and the spikiness in the resulting models can be smoothed out by binning to instrumental resolution.

We generate cross sections by assuming a Voigt profile for every line, with the Gaussian standard deviation set by the temperature of the gas, and the Lorentzian portion set by the pressure broadening coefficients γ_{ref} and n:

$$\gamma(P,T) = \gamma_{ref}\frac{P}{P_{ref}}\left(\frac{T_{ref}}{T}\right)^{n} \tag{4.14}$$

where γ_{ref} and n are expected to vary depending on the line considered and the species responsible for the broadening.

ExoMol

ExoMol (Tennyson & Yurchenko, 2018) is a database of molecular line lists intended for modeling the atmospheres of exoplanets and cool stars. The lists are generated using a combination of ab initio calculations and empirical data. Many of the line lists represent significant improvements in completeness over the previous state of the art. For example, POKAZATEL, the water line list, has 6 billion transitions. This is an order of magnitude more than previous lists, and covers every possible transition between states below the dissociation energy of water (Polyansky et al., 2018). Compared to the Freedman et al. (2008) line lists used by previous versions of `PLATON`, the ExoMol line list has many times the number of transitions for water (6 billion vs. 200 million), NH_3 (10 billion vs. 34,000), H_2S (115 million vs. 188,000), PH_3 (50 billion vs. 20,000) and VO (377 million vs. 3.1 million).

The specific line lists we used are listed in Table 4.11, together with the number of transitions they contain, their maximum temperature of validity, and citations to the associated papers. For some molecules, ExoMol provides line lists for multiple isotopologues; in those cases we only include the most common isotopologue. The exception to this rule is TiO, an important molecule where multiple isotopologues have comparable abundances. For this molecule, we compute the absorption due to each isotopologue and add them in proportion to each isotopologue's abundance.

Although these ExoMol line lists are an improvement over what was available before, many are still incomplete. The calculations only include states below a certain J quantum number, and hence miss transitions between higher-energy states that become important at higher temperatures. Thus, many line lists are not valid for the full range of temperatures supported by `PLATON`, which ranges up to 3000 K. For example, the C_2H_4 line list, despite having 50 billion lines, is only valid below 700 K. For these molecules, we still generated cross sections for all temperatures. The cross sections are likely to be under-estimated at high temperatures due to missing lines, but as of this writing, there is no better alternative.

ExoMol reports line broadening coefficients for hydrogen and helium whenever available. In practice, however, there are no calculations or experimental data available for any broadening agent for the vast majority of lines. For example, ExoMol reports no broadening coefficients at all for H_2S, CO, MgH, NO, OH, SiH, SiO, or VO. For NO, it only reports air broadening coefficients, which we adopt for lack of a better alternative. For the other molecules, ExoMol reports broadening coefficients for H_2 and He in a consistent format (Barton et al., 2017), relying on the handful of studies that have reported coefficients for a small number of lines while resorting to default values for the rest. We used ExoMol broadening data when available and assumed that the broadening agent is a mixture of 85% H_2 and 15% He. For all molecules where ExoMol broadening data is not available, we assumed that $\gamma_{ref} = 0.07$ and n=0.5 at a $T_{ref} = 296K$ and $P_{ref} = 1$ bar. n=0.5 is the theoretically expected value from classical calculations, while $\gamma_{ref} = 0.07$ is a typical value adopted by ExoMol as the default. The only exception is C_2H_4, where we used broadening parameters measured by Bouanich et al. (2003) (H_2) and Reuter & Sirota (1993) (He) for 34 and 3 lines respectively by generating ExoMol-formatted broadening files from the measurements.

HITRAN 2016

HITRAN (Gordon et al., 2017) is a database of line lists sourced from a combination of observations, theory, and semi-empirical calculations. It is intended for use at terrestrial temperatures, and has a line intensity cutoff that makes it inaccurate for higher temperatures. Nevertheless, HITRAN is a valuable resource because it is the only source of line lists for many molecules. HITRAN specifies line broadening parameters by including γ and n in the description of every line. Although γ_{H2} and n_{H2} are included in HITRAN, very few lines have hydrogen broadening data. Therefore, we chose γ_{air} and n_{air} as the broadening parameters for every line.

For most molecules not included in ExoMol, Exo-Transmit used (and `PLATON` inherited) absorption data from HITRAN 2008. We regenerate the absorption data using HITRAN 2016 (Gordon et al., 2017), which has expanded wavelength coverage and improved accuracy. This update also fixed some errors in the old data that resulted from incorrect generation of absorption data from line lists. The HAPI API makes it easy to retrieve line lists for all isotopologues at once, with intensities appropriately scaled to the isotopologue abundance. Therefore, we considered all isotopologues for the molecules we took from HITRAN.

CDSD-4000

The Carbon Dioxide Spectroscopic Databank 4000 (CDSD-4000) is a line list meant for high temperatures, provided in a format similar to HITRAN. It has significantly more lines than HITRAN or HITEMP, and to our knowledge it is the most complete publicly available line list for carbon dioxide. CDSD-4000 has pressure broadening coefficients for air, which we adopt due to the absence of broadening coefficients for hydrogen or helium.

Rey et al 2017

The line list presented by Rey et al. (2017), which we name Rey for convenience, is the first theoretical methane line list suitable for high temperature applications. It is complete in the infrared range (0–13,400 cm^{-1}) up to a temperature of 3000 K, whereas the ExoMol line list "10to10" is only accurate to 1500 K (Yurchenko & Tennyson, 2014). Rey also claims to be the first theoretical methane line list with line positions accurate enough for high resolution cross-correlation studies. We have confirmed this claim by cross correlating brown dwarf models generated using both methane line lists to observational high-resolution spectra of a T brown dwarf, with all other opacity sources excluded. The cross correlation peak is 4.5σ with "10to10," but 15.3σ with Rey, indicating far superior line positions. In addition, with 150 billion lines, this line list is far more complete than either ExoMol's "10to10" line list (10 billion transitions) or the Freedman et al. (2008) line list that we used previously (200 million transitions).

Voigt cutoff

When generating absorption cross sections from line lists, the cutoff—namely, how far away from the line center the line is considered to end— is an important source of error. One could in principle omit the cutoff, but computational speed would suffer greatly. In addition, omitting the cutoff does not necessarily lead to better results, as the Voigt profile is only an approximation to the true line profile (Ngo et al., 2012) and is not accurate more than several Voigt widths away from the center. However, truncating the lines too soon would result in an underestimate of the true opacity due to the omission of millions of line wings.

We adopted a cutoff of 25 cm^{-1} for all molecules for pressures less than or equal to 1 bar. For pressures of 10 and 100 bar, we adopted a cutoff of 100 cm^{-1}. For a pressure of 1000 bar, we adopted a cutoff of 1000 cm^{-1}. This prescription was inspired by Sharp & Burrows (2007), who adopt a cutoff of min($25P$, 100) cm^{-1} where P is

in bars. Hedges & Madhusudhan (2016) studied the effect of different cutoffs and concluded that the Sharp & Burrows prescription significantly underestimates absorption at low pressures (P < 0.01 bar), but is accurate from 0.01 to 100 bar. We therefore modified the prescription to use 25 cm^{-1} for all pressures below or equal to 1 bar. Pressure broadening coefficients are almost never measured or calculated at very high pressures (100 – 1000 bar), so our opacity data in this regime should be regarded as highly speculative.

4.5 Other improvements

Aside from the opacity update and the eclipse depth calculator, many improvements have been made to `PLATON` since the publication of the last paper. These improvements introduce features, fix bugs, increase speed, and improve usability. A comprehensive list can be found in our release notes, but we list a few of the most noteworthy updates below.

- We now include H- opacity, calculated using the algorithm in John (1988). Because H- opacity is insignificant for most planets, we disable it by default.
- Nested sampling is now done by `dynesty` (Speagle, 2020) rather than `nestle`. Among other improvements, `dynesty` prints out the number of likelihood evaluations, log evidence (logz), and remaining evidence (dlogz) after each iteration. dlogz is an indicator of how far the algorithm is from completion. `nestle` did not have this indicator, which made waiting a frustrating experience.
- The eclipse depth calculator now evaluates the single integral in Equation 4.2 instead of the double integral in Equation 4.1, making it much faster, many times more memory efficient, and more accurate.
- Data arrays are transposed so that the wavelength index increases the fastest, followed by the pressure index, followed by the temperature index. This improves cache locality, which speeds up the code by a factor of 1/3.
- The number of atmospheric layers is decreased from 500 to 250, improving speed by 40% while increasing numerical error by only ~1%. In addition, we now use improved interpolation methods to further decrease numerical errors.

High resolution opacities

PLATON has a clean separation between data and code. As a result, all that is needed to operate PLATON at an arbitrary wavelength range and resolution is the appropriate opacity data files. No code changes are required. Since we published Zhang et al. (2019), we have generated high-resolution opacity data files for a variety of applications. This includes studying the atmospheres of cold brown dwarfs, of the ultra-hot super Earth 55 Cnc e, and of HD 189733b (see Subsection 4.6).

We now make these opacities public to enable anyone to perform line-by-line calculations with PLATON[2]. All opacities have a resolution of R=375,000, and are calculated at the wavelengths indicated by wavelengths.npy. These opacities can be used by deleting all files from the "Absorption" folder in PLATON's data directory, putting the downloaded absorb_coeffs files into the directory, and replacing the wavelengths.npy file in the data directory with the one in the downloaded zip file. The user must use a blackbody stellar spectrum (by passing stellar_blackbody=True to compute_depths) when using high resolution opacities. The user can also generate their own high-resolution opacity files using publicly available codes such as ExoCross (Yurchenko et al., 2018) or HELIOS-K (Grimm & Heng, 2015). As long as they are in the same format as the PLATON data files, PLATON will accept them with no code changes.

Table 4.5: High resolution absorption data

Filename	Wavelengths (μm)	Molecules
hispec	0.94 – 2.43	CH_4, CO, H_2O, H_2S, HCl, HCN, MgH, NH_3, NO_2, NO, O_2, O_3, OH, SH, SiH, SiO, SO_2, TiO, VO
*Y*_band	1.020 – 1.086	CH_4, H_2O, NH_3
*K*_band	1.89 – 2.40	C_2H_2, CH_4, CO, H_2O, HCN, Na, NH_3, SiO
*L*_band	2.86 – 3.70	C_2H_2, CH_4, CO, H_2O, H_2S, HCN, Na, SiO
*L*_band2	3.51 – 4.08	CH_4, CO_2, H_2O, HCN

We note that in principle other codes, such as TauREx, can also be used to calculate high resolution spectra if they are provided with custom user-generated high resolu-

[2] https://www.astro.caltech.edu/platon

tion opacity files. To the best of our knowledge, however, PLATON is the only code to make such opacity files publicly available.

Correlated-k

The gold standard of radiative transfer is the line-by-line method (Marley & Robinson, 2015): calculating the transit or eclipse depth on a wavelength grid fine enough to resolve individual molecular lines, which typically requires $R >> c/\sqrt{\frac{kT}{\mu m_H}} \approx 200,000$. The results are then binned to instrumental resolution. For many applications, including PLATON, this is computationally prohibitive. Opacity sampling sacrifices accuracy for speed by performing radiative transfer at a much lower resolution—R=1000, in the case of PLATON—and binning the depths thus obtained to instrumental resolution. In order for the survey to approximate the results of the census, a large sample size is required, meaning opacity sampling at R=1000 is only accurate if the user is binning to resolutions much below 1000. To take a concrete example, suppose PLATON is used to calculate the eclipse depth of HD 189733b over the 1.40–1.42 μm band using the R=10,000 opacities. PLATON would sample 140 wavelengths within this band and calculate the eclipse depth at each wavelength. These 140 eclipse depths would have a mean of 53 ppm and a standard deviation of 25 ppm. Therefore, the error caused by opacity sampling is $25/\sqrt{140}$=2 ppm. If the R=1000 opacities were used instead, this error would be $25/\sqrt{14} = 7$ ppm. Fortunately, 7 ppm is several times lower than the error of the WFC3 observations, but this will not be the case for every combination of planet, instrument, and wavelength band.

The correlated-k method (Lacis & Oinas, 1991) improves upon opacity sampling by taking into account the distribution of opacities within the passband. For example, it calculates the 10th percentile eclipse depth by using the pre-calculated 10th percentile of all molecular opacities, and likewise for the 20th percentile, 30th percentile, etc. Since the eclipse depth varies smoothly with percentile, we can use numerical integration to find the average eclipse depth. In effect, the correlated-k method converts an integration over wavelength into an integration over percentile; it replaces the integration of a highly non-smooth function $\int_{\lambda_1}^{\lambda_2} f(\lambda)d\lambda$ with the integration of a smooth function $\int_0^1 f'(g)dg$, which is in turn evaluated by Gaussian quadrature ($\sum_{i=1}^{N} w_i f'(g_i)$), where g is the percentile divided by 100. $f(\lambda)$ can be $R_p(\lambda)$, the emergent flux $F_p(\lambda)$, or any other radiative quantity, so long as its only dependence on wavelength is through the opacity, and so long as it varies smoothly with opacity.

Earlier versions of PLATON performed radiative transfer via opacity sampling at a default resolution of 1000 with an optional R=10,000 mode. We now give the user the option to choose between correlated-k (R=100) and opacity sampling. The correlated-k method provides the accuracy of R=50,000 opacity sampling for typical exoplanet applications, but runs at the same speed as the old default R=1000 PLATON mode. Our implementation of correlated-k rests on two approximations:

1. At any given wavelength λ within the band, $g(\lambda)$ is the same for every layer. That is, if a layer is more opaque at a certain wavelength than at x% of other wavelengths within the band, all other layers must also be more opaque at that wavelength than at x% of other wavelengths.

2. For each layer, if the individual gases were to be separated out, $g(\lambda)$ would be equal for all gases with significant opacity. That is, if one molecule absorbs more strongly at a certain wavelength than at x% of other wavelengths within the band, the same must be true for all other molecules.

The first assumption is the defining assumption of the correlated-k method, and explains its name: the opacity κ is assumed to be correlated throughout the atmosphere under consideration. For an atmosphere with one species with exactly one absorption or emission line within the band, it is exactly true. For an atmosphere with two layers, each of which is exclusively composed of a different gas, it is very inaccurate. A real atmosphere is in between these two extremes: the region dominating the features in a transmission or emission spectrum typically spans ~200 K in temperature (Figure 4.10) and 1-2 orders of magnitude in pressure (see Figures 4.10 and 4.11), which generally means that there is no change in the dominant gas absorber.

The second assumption, however, is only true when one molecule dominates the opacity. If two molecules contribute equally to the opacity, the assumption is no longer valid, as the absorption lines of different molecules will not in general overlap. In fact, the opposite assumption is more accurate: namely, that the opacities are completely uncorrelated between different gases. Lacis & Oinas (1991) take this approach, but adopting this assumption naively for PLATON would require O(n^N) radiative transfer calculations, where n is the number of discrete g values adopted (for us, 10) and N is the number of gases (for us, 30), making these computations intractable. There are methods of merging the opacity distributions of multiple gases

that do not scale exponentially—including the 'random overlap with resorting and rebinning' method introduced by Lacis & Oinas (1991) and named by Amundsen et al. (2017). These are more complicated to implement, and we may incorporate them into a future release of `PLATON`. The partially correlated approach attempts to take into account the correlations between gases (Zhang et al., 2003), but these sophisticated schemes are beyond the scope of `PLATON`.

We note that the simplicity of our approach comes at a cost: it systematically overestimates transmittance in most cases (Zhang et al., 2003). This overestimation is easy to understand with a toy scenario. Consider a gas with binary absorption properties: at 50% of wavelengths it has infinite absorption, while at the other wavelengths it has zero absorption. This gas would have a transmittance of 50%. Now consider adding a second gas, also with binary absorption properties. If the two gases have perfectly correlated absorption (which is our assumption), their absorption peaks fall on top of each other, and transmittance is still 50%. If their absorption is not perfectly correlated, the absorption peaks of the second gas block some of the light that would have went through the first gas, and total transmittance is less than 50%. If their absorption is perfectly anti-correlated, the total transmittance would be 0%.

In practice, `PLATON` uses the following correlated-k algorithm:

1. (Pre-computed) Compute the absorption coefficients of each atom/molecule at each temperature, pressure, and wavelength grid point, with a spectral resolution of R=50,000. Correlated-k coefficients are generated from the absorption coefficients with a resolution of R=100.

2. Divide the wavelength range under consideration into bands, with each band having a width of $\lambda/100$.

3. For each band, compute the transit/eclipse depth at 10 different opacity percentiles, and combine them via Gaussian quadrature. The transit depth at the 16th opacity percentile (for example) is calculated by assuming every gas, at every temperature and pressure, has an absorption coefficient equal to the pre-calculated 16th percentile absorption coefficient for that band at that temperature and pressure. We use 10 Gaussian quadrature points, which is sufficient to keep the integration error below 1% in most cases (Lacis & Oinas, 1991; Goody et al., 1989).

4. The R=100 transit or eclipse depths are then binned to the user-specified wavelength bins using the methods described in our first paper (Zhang et al., 2019).

We performed an experiment to deduce the accuracy of the correlated-k algorithm compared to a line-by-line calculation. The transit spectrum of a hot Jupiter (modelled after HD 209458b) was computed from $0.95 - 2.4$ μm using two methods: a line-by-line calculation at R=375,000 binned to R=100, and a correlated-k calculation using R=100 opacities with 10 Gaussian quadrature points. For the nominal model, which is dominated by water opacity, correlated-k performs extremely well and has a maximum error of only 8 ppm. For a model engineered to include three molecules with comparably significant absorption correlated-k still performs well, with a maximum error of 300 ppm. The emission spectrum tells a similar story. Correlated-k is accurate to 0.1% for the nominal model, and to 3.5% for the pessimistic model.

These tests also demonstrate that although correlated-k is very accurate, its errors are not random. Correlated-k almost always underestimates the transit depths and overestimates the eclipse depths. This is a consequence of overestimating the transmittance, which in turn is because (contrary to the second assumption above) the absorption properties of two molecules are in general not strongly correlated.

Beta features

Since we do not currently plan to write a third `PLATON` paper, we include a list of beta features that will likely become part of the official `PLATON`. This list also serves to illustrate what is possible with minimal hacking. All of these features were created as a result of requests from `PLATON` users other than the authors. Users are highly encouraged to contact the authors to suggest new features or improvements to existing features.

`PLATON` does not calculate disequilibrium chemistry from first principles, nor does it compute self-consistent temperature-pressure profiles. It is often useful to take abundance and temperature profiles from elsewhere and plug them into `PLATON`, using it as a radiative transfer engine to predict transit and eclipse depths. This is currently easy to do for vertically-constant abundances, but not for vertically-variable abundances. To make the latter possible, we created the branch custom_abundances on the GitHub repository. Examples of how to use it are found in examples/plot_transit_custom_abunds.py and examples/plot_eclipse_custom_abunds.py.

Some metallic species are not included in PLATON by default, but may become important in the optical for ultra-hot Jupiters. These include Ca, Fe, Ni, and Ti. We make these opacities available at https://www.astro.caltech.edu/platon/metal_opacities/. These can be used by placing them in PLATON_DIR/data/Absorption and adding the atoms to PLATON_DIR/data/species_info. These atoms are not incorporated into the equilibrium chemistry calculation, but the user can easily specify vertically constant abundances for them. We describe the procedure in PLATON's online documentation.

Lastly, Na and K each have two very strong lines in the optical, where the atmosphere is transparent enough that their far wings may become significant. Unfortunately, while the lines cores are accurately described by a Voigt profile, the Voigt profile can underestimate far wing absorption by orders of magnitude (Allard, N. F. et al., 2016; Allard et al., 2019). More accurate line profiles for these atoms were recently published by Allard, N. F. et al. (2016) (K) and Allard et al. (2019) (Na) using a semi-classical theory and assuming broadening by molecular hydrogen only. We use these line profiles to generate PLATON-friendly absorption coefficients at R=1000 and R=10,000, found at https://www.astro.caltech.edu/platon/metal_opacities/. By overwriting PLATON_DIR/ data/ Absorption/ absorb_coeffs_Na.npy and PLATON_DIR/ data/ Absorption/ absorb_coeffs_K.npy with these coefficients, the user can generate much more accurate hydrogen-broadened alkali line profiles with PLATON.

4.6 Retrieval on HD189733b

Published data sets

HD 189733b is one of the most favorable exoplanets for atmospheric characterization. It is a transiting hot Jupiter orbiting an exceptionally close (20 pc) K star with a *H* band magnitude of 5.6, and was one of the earliest transiting planets discovered (Bouchy et al., 2005). To demonstrate PLATON's new abilities, we perform a joint retrieval on the best available optical and near infrared transit and secondary eclipse data for HD 189733b from *HST* and *Spitzer*. To our knowledge, this is the first joint transit and secondary eclipse retrieval for this planet in the literature, as well as the most comprehensive set of both transit and secondary eclipse data assembled for a retrieval to date. The fixed stellar and planetary parameters are listed in Table 4.7. The transit depths we adopt are listed in Table 4.8, while the eclipse depths are listed in Table 4.10.

HD 189733b has been observed in transmission with *HST*/STIS (Sing et al., 2011) and WFC3 (Gibson et al., 2012; McCullough et al., 2014), *Spitzer* in all five IRAC bands (Tinetti et al., 2007; Beaulieu et al., 2008; Agol et al., 2010; Désert et al., 2011; Morello et al., 2014), and *Spitzer*/MIPS at 24 μm (Knutson et al., 2009). Pont et al. (2013) carried out a uniform re-analysis of all transit data obtained to date including corrections for stellar activity; we utilize their transmission spectral data in our analysis. This planet was also observed in transit by *HST*/NICMOS in spectroscopic (Swain et al., 2008) and photometric (Sing et al., 2009) modes, but this instrument was less stable than WFC3, and the reliability of the spectroscopic NICMOS observations was questioned in a subsequent study (Gibson et al., 2011; Deming & Seager, 2017). We therefore exclude these older NICMOS observations from our analysis. We also exclude the higher resolution observations of the sodium line published in (Huitson et al., 2012), as `PLATON` is not designed to model absorption at the very low pressures probed by the core of this line.

In emission, HD 189733b has been observed with *HST*/STIS (Evans et al., 2013), *HST*/NICMOS (Swain et al., 2009), *HST*/WFC3 (Crouzet et al., 2014), *Spitzer*/IRAC in all four bands (Knutson et al., 2007; Charbonneau et al., 2008; Agol et al., 2010; Knutson et al., 2012), *Spitzer*/IRS at 5 – 14 μm (Deming et al., 2006; Grillmair et al., 2008; Todorov et al., 2014), and *Spitzer*/MIPS at 24 μm (Knutson et al., 2009). Because `PLATON` does not model reflected light, we limit our retrieval to the infrared data only, which are expected to be dominated by thermal emission. As with the transmission spectrum, we exclude the NICMOS observations from our retrieval. We also exclude the *Spitzer*/IRS emission spectrum, as over the years different groups have obtained contradictory results. Most recently, Todorov et al. (2014) found that the overall amplitude of the IRS eclipse depth can shift up and down depending on the method used to correct for systematics. Indeed, a comparison to our best fit model spectrum reveals that these data are consistent with the broadband observations if they are shifted upwards by 30%. In addition to transit and secondary eclipse observations, HD 189733b's phase curve has also been measured in the 3.6, 4.5, 8.0, and 24 μm *Spitzer* bands (Knutson et al., 2007, 2009, 2012).

Based on the observations listed above, HD 189733b is one of the most extensively observed transiting planets to date. Previous studies of HD 189733b's optical transmission spectrum found that it appears to have a strong scattering slope and attenuated absorption features due to the presence of high-altitude scattering particles

(Pont et al., 2008; Sing et al., 2011; Pont et al., 2013). These scattering particles are possibly some form of silicate condensate (Lecavelier Des Etangs et al., 2008; Lee et al., 2015; Helling et al., 2016). Because this planet is expected to be tidally locked, it should develop a super-rotating equatorial band of wind that transports heat from the day side to the night side (e.g., Showman & Polvani 2011). It is observed to have a relatively modest day-night temperature gradient (Knutson et al., 2007, 2009, 2012) and models predict that it may also have spatially inhomogeneous cloud coverage (Lee et al., 2015; Lines et al., 2018).

In the infrared, the effect of the scattering particles on HD 189733b's transmission spectrum is reduced. McCullough et al. (2014) report the detection of a spectroscopically resolved water feature at 1.4 μm that is consistent with the model of scattering aerosol reported by Pont et al. (2013), although they also argue that the optical slope can be explained by stellar activity alone (see Section 4.6). In emission, previous studies have detected spectroscopically resolved water absorption at 1.4 μm (Crouzet et al., 2014) and (debatably) in the mid-infrared (Grillmair et al., 2008; Todorov et al., 2014), and have placed additional constraints on the abundances of carbon monoxide, carbon dioxide, and methane based on the relative depths of the broadband *Spitzer* secondary eclipse data (e.g., Line et al. 2010; Lee et al. 2012). We discuss the results of these retrievals in more detail in Section 4.6.

`PLATON` retrieval

When modeling HD 189733b's transmission spectrum, our retrieval uses a complex refractive index with a real component of 1.7, a value in between that of $MgSiO_3$ and SiO_2 (which have $n \sim 1.5$ at optical wavelengths) and TiO_2 (which has $n \sim 2.4$). Since the true composition of condensates in the atmosphere is unknown, and multiple condensates may well be important, we allow the imaginary component of the refractive index to vary as a free parameter in our fit. We do not include clouds in our dayside models; even if the clouds observed on the terminator extended over the entire dayside, we would expect them to have lower optical depths when viewed in emission at infrared wavelengths (e.g., Fortney 2005). Indeed, previous retrieval studies of HD 189733b's dayside atmosphere have found that cloud-free models provide a good fit to the available data (e.g. Barstow et al. 2014).

We carry out our retrievals using the `dynesty` nested sampling package with static sampling and R=10,000 opacities. We used 1000 live points, a convergence criteria of $\Delta\log(z) = 1$, multi-ellipsoidal bounds, and a random-walk sampling method. The

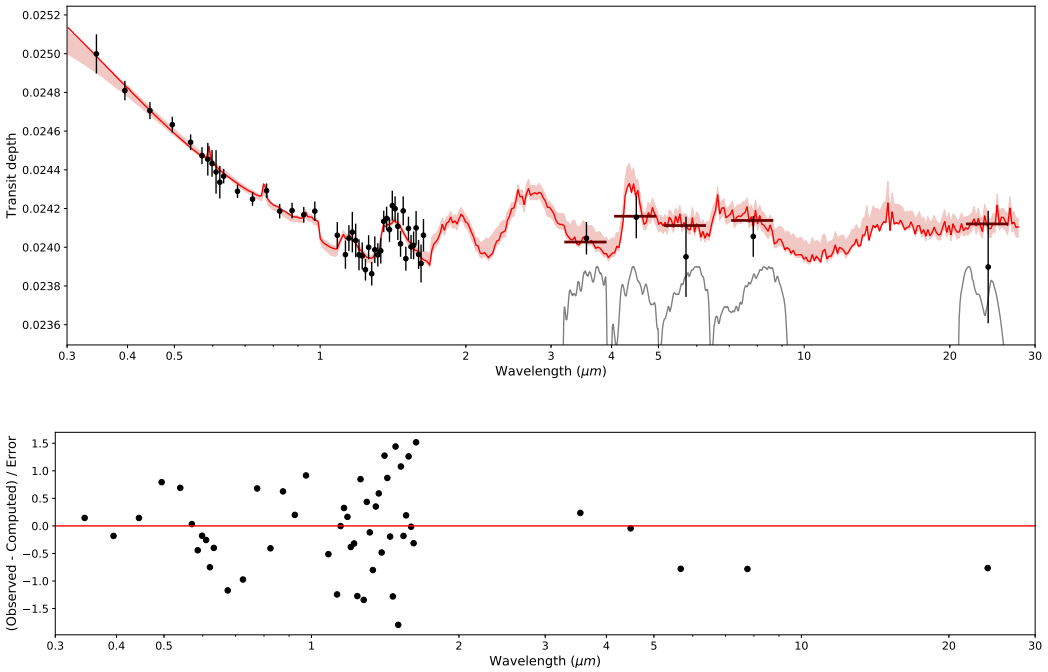


Figure 4.7: Best fit and residuals. Top: Best fit transit spectra from retrieval on HD 189733b. The best-fit `PLATON` model and corresponding 1σ uncertainty window are shown as a red line and red shaded region, respectively. The activity-corrected observations drawn from Pont et al. (2013) and listed in Table 4.8 are shown as black filled circles. Bottom: difference between observed and computed transit depths, in units of measurement error σ. Adopting the nightside emission pollution correction of Kipping & Tinetti (2010) would reduce the errors on the 8μm and 24μm points by 0.6σ and 0.4σ, respectively, bringing the observations into nearly perfect agreement with the model.

fixed parameters, listed in Table 4.7, are the stellar radius, stellar temperature, and planetary mass. The free parameters, listed in Table 4.6, are the planetary radius R_p at 1 bar, the metallicity Z relative to solar, the C/O ratio, the limb temperature T_{limb}, the Line et al. (2013) T/P profile parameters (thermal opacity κ_{th}, visible-to-thermal opacity ratio of first visible stream γ, visible-to-thermal opacity ratio of second visible stream γ_2, percentage apportioned to the second visible stream α, effective albedo β), mean haze particle radius r_m, haze particle number density n, ratio of haze scale height to gas scale height h_{frac}, WFC3 instrumental offsets ($\Delta_{wfc3,t}$ for transit and $\Delta_{wfc3,e}$ for eclipse), and the imaginary portion of the haze refractive index k. In Figure 4.7 and 4.8, we show the best fit transit and eclipse spectra from our retrieval. Our best fit model is a good fit overall to the data, with a χ^2 of 30.7 for the transit spectrum and 38.3 for the eclipse spectrum for a total $\chi^2 = 69$. With 52 transit depths, 34 eclipse depths, and 15 free parameters, the p value is 0.55. The

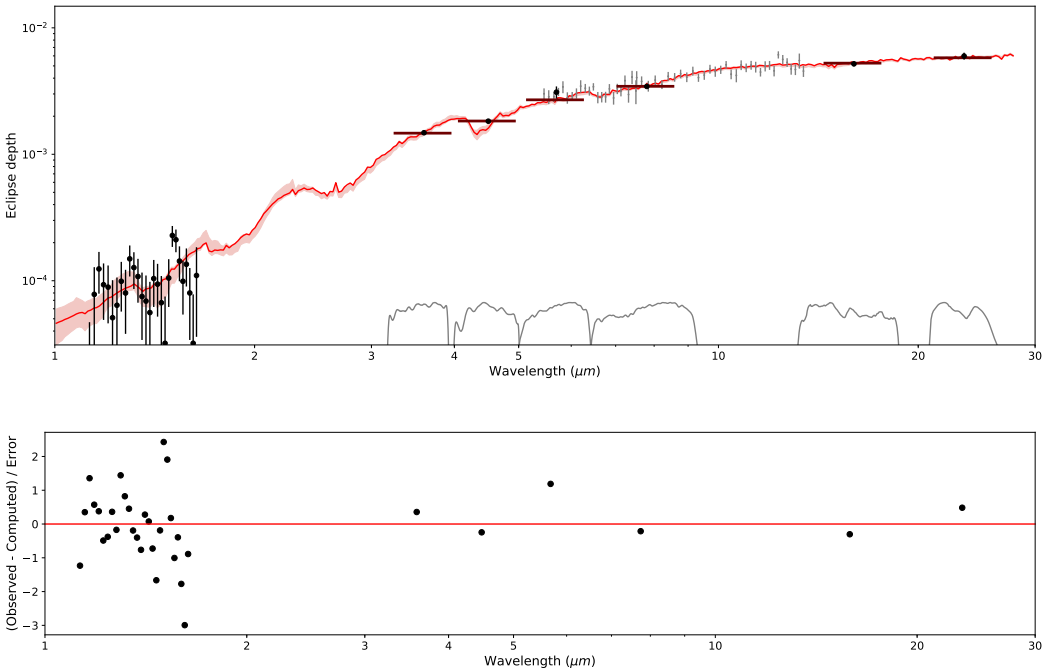


Figure 4.8: **Top:** Best fit eclipse spectra from retrieval on HD 189733b. The best-fit `PLATON` model and corresonding 1σ uncertainty window are shown as a red line and red shaded region, respectively. Data used in the fit (see list in Table 4.10) are shown as black filled circles. The IRS eclipse depths from Todorov et al. (2014) are shown in grey and were not included in the retrieval. These eclipse depths have been shifted up by 30% to match the model—a plausible shift, since Todorov et al. (2014) mentions that its results are 20% below those of Grillmair et al. (2008). Bottom: difference between observed and computed secondary eclipse depths, in units of measurement error σ.

single largest point of disagreement (3.1σ) between model and data occurs at the very red end of the WFC3 emission spectrum, where edge effects may impact the reliability of data.

In Table 4.6, we tabulate the 1D posterior distributions from the retrieval. In Figure 4.9, we show the 2D posterior distributions. We find that the data prefer supersolar metallicities ($Z = 7 - 20$) and a C/O ratio between $0.47 - 0.69$. This C/O ratio is consistent with the solar value, but is somewhat low compared to the stellar C/O ratio of 0.90 ± 0.15 (Teske et al., 2014). However, Teske et al. (2014) also report that their estimate for the stellar C/O ratio depends on what data they include and how the non-LTE correction is performed. They report C/O ratios ranging from 0.69 to 1.2 for different data analysis choices, and conclude that although the C/O ratio could be below 0.75, it is very likely above 0.80. If so, the planetary atmospheric C/O ratio

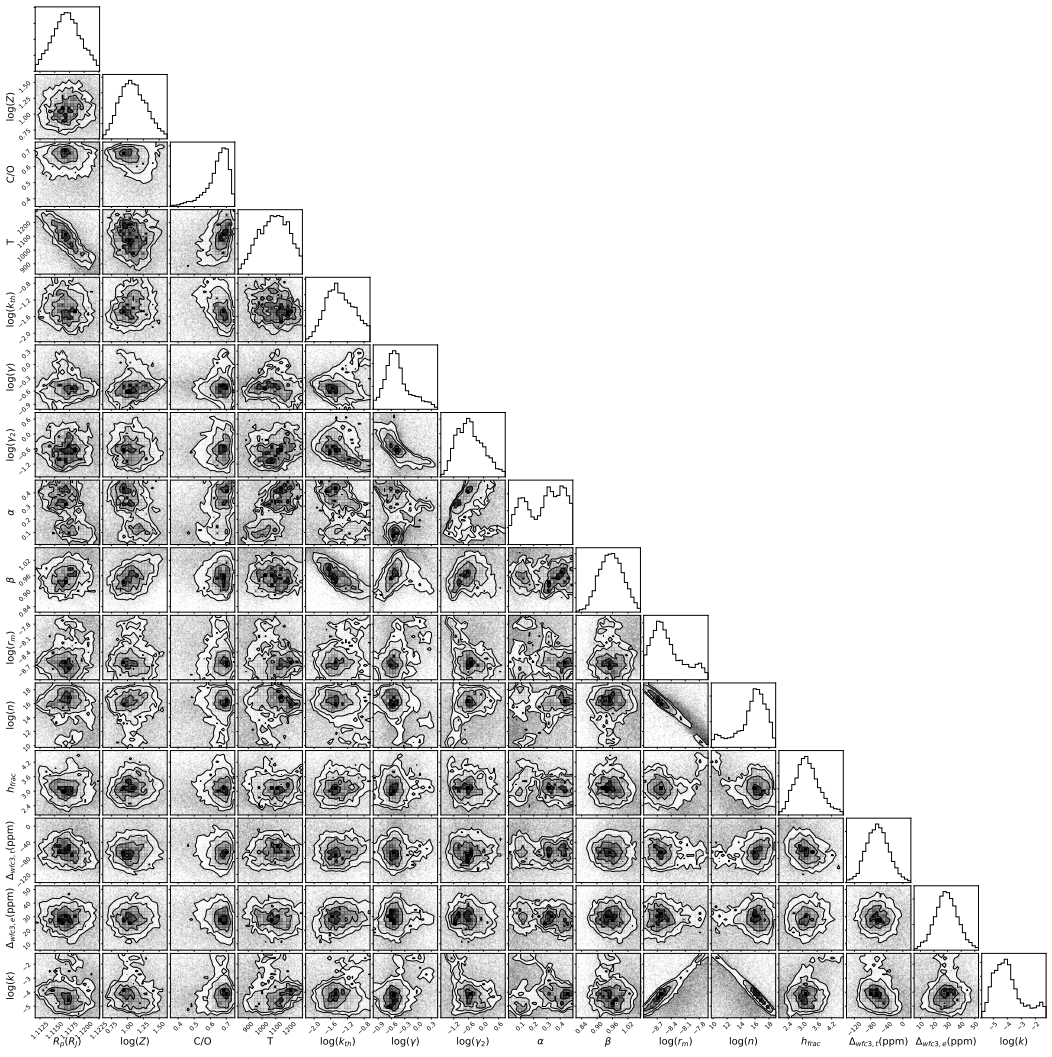


Figure 4.9: Posterior distribution of the fiducial retrieval on HD 189733b.

would be slightly suppressed relative to the stellar value, in good agreement with theoretical predictions for gas giant planets with atmospheric metallicities enhanced by the accretion of solids (Espinoza et al., 2017).

Our observations also constrain HD 189733b's dayside pressure-temperature profile. We find no evidence for a dayside temperature inversion, as shown in Figure 4.10. The parameter α, which partitions the visible radiation into two separate channels and corresponds to the flux in the channel with opacity γ_2, is consistent with 0. This means that there is no need for a second visible wavelength channel and that the simpler double-gray parameterization of Guillot (2010) is sufficient. We also find that the overall shape of the pressure-temperature profile is consistent with a low albedo and efficient day-night redistribution of heat, as β is consistent with 1.

Table 4.6: Retrieved parameters

Parameter	Posterior	Best	Prior
R_p (R_J)	1.117 ± 0.002	1.114	$[1.11, 1.13]$
$\log_{10} Z$	$1.08^{+0.23}_{-0.20}$	0.956	$[-1, 3]$
C/O	$0.66^{+0.05}_{-0.09}$	0.69	$[0.2, 2]$
T_{limb} (K)	1089^{+110}_{-120}	1203	$[500, 1300]$
$\log_{10} \kappa_{th}$ (m^2 kg^{-1})	$-1.40^{+0.40}_{-0.32}$	-1.44	$[-5, 0]$
$\log_{10} \gamma$	$-0.51^{+0.36}_{-0.21}$	-0.66	$[-4, 1]$
$\log_{10} \gamma_2$	$-0.58^{+0.62}_{-0.50}$	-0.47	$[-4, 1]$
α^*	< 0.47	0.441	$[0, 0.5]$
β	0.95 ± 0.05	0.938	$[0.5, 2]$
$\log_{10} r_m$(m)*	< -7.8	-8.59	$[-9, -5]$
$\log_{10} n(m^{-3})$	$16.0^{+1.5}_{-2.7}$	14.4	$[8, 21]$
h_{frac}	$3.2^{+0.6}_{-0.5}$	3.19	$[0.5, 5]$
$\Delta_{wfc3,t}$(ppm)	-63 ± 28	-72	0 ± 100
$\Delta_{wfc3,e}$(ppm)	29 ± 9	33	0 ± 39
$\log_{10} k^*$	$* < -1.7$	-3.86	$[-6, 0]$

For these parameters, the 95th percentile upper bound is reported.
The 'best' column reports the parameters of the best fit model.

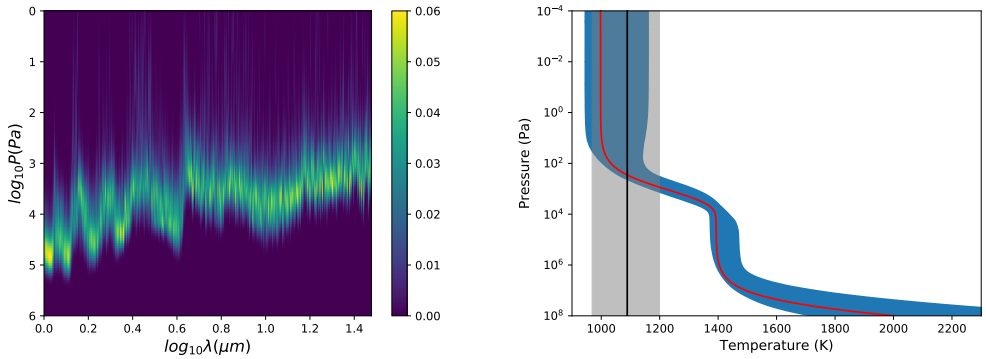


Figure 4.10: Inferences from the retrieval. Left: emission contribution function of HD 189733b, as indicated by the best fit solution. Right: the T/P profile as indicated by the best fit solution (red), along with the 2σ uncertainties on the T/P profile (blue). The median limb temperature is indicated in black, while the 1σ uncertainty in limb temperature is indicated in gray.

We place constraints on the sizes and locations of the scattering particles near HD 189733b's terminator. The mean particle size is constrained to be less than 14 nm, and is consistent with arbitrarily small values. However, this size constraint is dependent on the assumed value of the imaginary refractive index, with more absorbent particles requiring a larger mean particle size (see Subsection 4.6 for

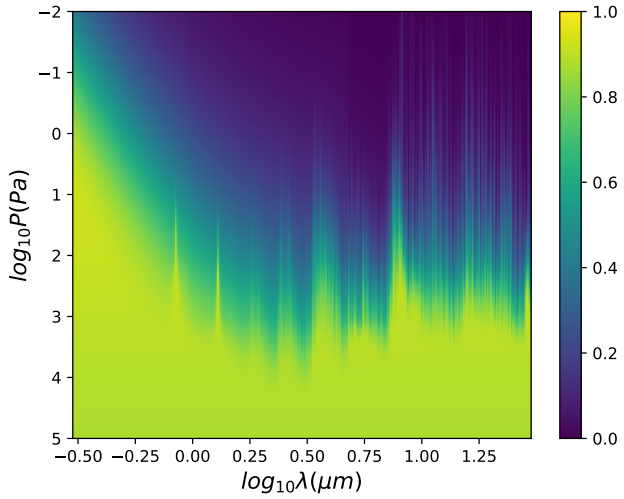


Figure 4.11: Contribution function of the transmission spectrum. The absolute scale is in arbitrary units.

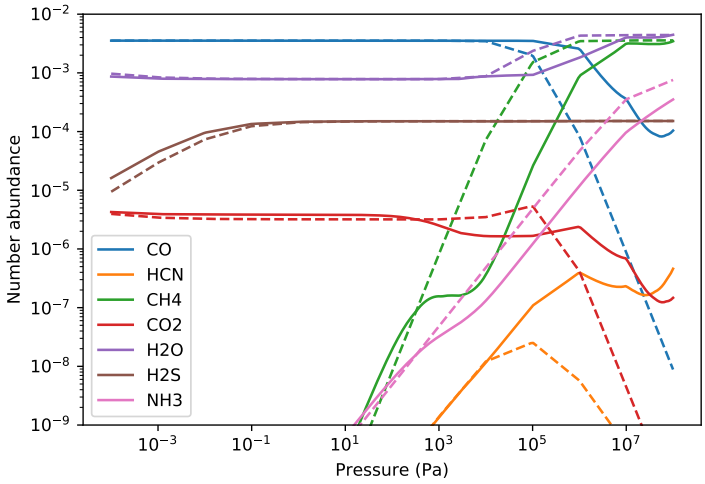


Figure 4.12: Number abundances of the most common molecules in the limb (dashed) and on the day side (solid), according to the best fit model. Hydrogen and helium, the dominant components of the atmosphere, are not shown.

more details). We find that the fractional scale height of the haze is a factor of a few larger than that of the gas. In effect, this means that haze particles are more abundant relative to gas in the upper atmosphere than in the lower atmosphere. This could possibly indicate that the haze is photochemical in nature, a possibility first suggested by Zahnle et al. (2009) and Pont et al. (2013). The fact that photochemical hazes can generate super-Rayleigh scattering slopes in the $T_{eq} = 1000 - 1500$ K regime was recently demonstrated by Ohno & Kawashima (2020), who showed that

strong eddy diffusion gives rise to a ρ_{haze}/ρ_{gas} ratio that increases with height, which in turn leads to a steep spectral slope.

Comparison with previous retrievals

Many authors have attempted to use retrievals to constrain the atmospheric properties of HD 189733b, starting with Madhusudhan & Seager (2009). Here, we review the most recent retrievals, including Lee et al. (2014) and Pinhas et al. (2019) in transmission and Lee et al. (2012) in emission, and compare their results to ours. Although Benneke (2015) separately performed a retrieval on HD 189733b's WFC3 transit spectrum, the limited wavelength range of these data prevented them from obtaining meaningful constraints on the atmospheric metallicity.

Lee et al. (2014) performed a retrieval on the 0.3 – 10 μm transmission spectrum data reported in Pont et al. (2013). Because this study was performed prior to the publication of the WFC3 data, it used NICMOS spectroscopy to constrain the shape of HD 189733b's near-infrared transmission spectrum. As part of this retrieval they explored several different potential aerosol species, including $MgSiO_3$, Mg_2SiO_4, astronomical silicate (a mixture of siliciate grains commonly seen in interstellar space), $MgSiO_3$, NaS, and tholins, all of which they argued might plausibly form in the upper regions of HD 189733b's atmosphere. They found that a vertically uniform aerosol layer stretching from 0.1 mbar to 10 bar filled with monodispersed particles smaller than 0.1 μm provides a good fit to the spectrum for all aerosol compositions. This matches well with our updated fits, which indicate that the data are consistent with a large range of imaginary refractive indices and favor particles with a mean radius smaller than 14 nm, distributed with a scale height much larger than that of the gas. Lee et al. (2014) also reported a constraint of $0.02-20\times10^{-4}$ on the abundance of water after accounting for the uncertainties introduced by different aerosol assumptions. Our best fit model is on the upper end of this range with a photospheric water abundance of 10^{-3}, but it is consistent with the water abundances from this study for both Mg_2SiO_4 and tholin aerosols.

Pinhas et al. (2019) use AURA to perform a retrieval on the transit spectrum from STIS, ACS, WFC3, and IRAC 3.6/4.5 (0.35–4.5 μm). They obtain a water abundance of $\log(X_{H_2O}) = -5.04^{+0.46}_{-0.30}$, which is 1.8% the equilibrium value at solar elemental abundances, and claim a strong detection of water depletion. We computed the water abundance at 10 mbar in our models, taking this to be representative of the photospheric pressure, and obtained $\log(X_{H_2O}) = -2.5 \pm 0.3$. Our result is

substantially discrepant with Pinhas et al. (2019): it is marginally super-solar (see their Figure 2) and comparable to the measured atmospheric C/H ratio for Jupiter (e.g., Lodders 2003, their Figure 6).

The cause of this discrepancy is not clear. In terms of data, Pinhas et al. (2019) include only the 3.6 and 4.5 μm *Spitzer* transit depths, while we also include transit depths at 5.8 μm, 8.0 μm, and 24 μm. The most significant difference in the methodology is that Pinhas et al. (2019) allows much more freedom than our retrieval. We use equilibrium abundances, while they fit for the abundances of 6 individual atoms and molecules. We adopt an isothermal limb, while they adopt the Madhusudhan (2019) parameterization of the limb T/P profile and fit for all 6 parameters. In total, they have 19 free parameters, while our transit-only retrieval has only 9. The high number of free parameters in Pinhas et al. (2019) may allow them to find a better fit to the data, one with sub-solar water abundance. On the other hand, the flexibility also puts them in more danger of over-fitting and of finding physically unrealistic compositions or T/P profiles.

We next consider previously published fits to HD 189733b's dayside emission spectrum. Lee et al. (2012) used optimal estimation to perform a retrieval on all published eclipse observations, with wavelengths ranging from 1.45 – 24 μm. They did not include the WFC3 eclipse observations (Crouzet et al., 2014), which were published after that study, and instead included NICMOS observations spanning a similar wavelength range (Swain et al., 2008). Using these data, they found a mixing ratio of $0.9 - 50 \times 10^{-4}$ for water, $3 - 150 \times 10^{-4}$ for carbon dioxide, and $< 0.4 \times 10^{-4}$ for methane, implying a C/O ratio of 0.45–1. The error ranges they derived for CO were so broad that they could not provide meaningful estimates of its abundance. Comparing to Figure 4.12, we see that our water abundance of 10^{-3} is fully consistent with these results, as is our low methane abundance of $\sim 10^{-7}$ at $P \sim 0.1$ bar. However, our model has several times less CO_2. Lee et al. (2012) observed that previous studies preferred much smaller CO_2 abundances, including Line et al. (2010) (10^{-7} to 10^{-5}), Swain et al. (2009) (10^{-7} to 10^{-6}), and Madhusudhan & Seager (2009) ($7 - 700 \times 10^{-7}$). They concluded that it is their *HST*/NICMOS data that caused the fits to prefer a high CO_2 abundance. As discussed above, some studies have questioned the reliability of the NICMOS results (Gibson et al., 2012; Deming & Seager, 2017), which also sometimes appear to contradict subsequent WFC3 observations (Deming et al., 2013). A lower CO_2 abundance would also be more physically plausible, as equilibruim chemistry predicts that

it should be relatively rare at the low atmospheric metallicities preferred by our model, and disequilibrium models including both photochemistry and quenching do not appreciably increase the predicted CO_2 abundance (e.g., Moses et al. 2013; Steinrueck et al. 2019.

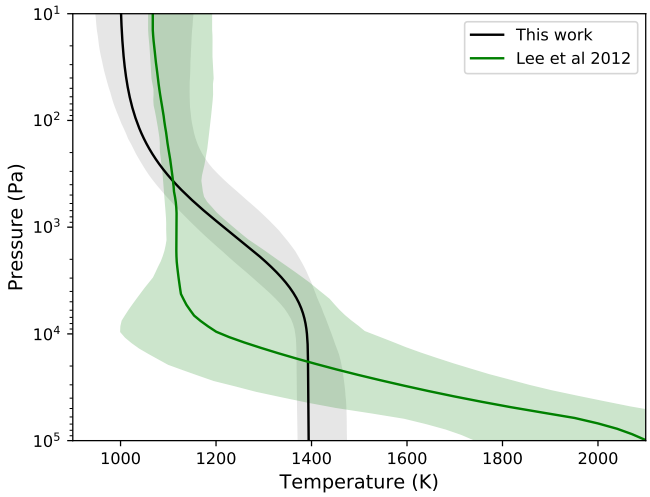


Figure 4.13: Comparison of our retrieved T/P profile (black) and that of Lee et al. (2012) (green).

In addition to abundances, Lee et al. (2012) also retrieved a T/P profile (their Figure 1), which we compare to our T/P profile in Figure 4.13. Although the two profiles are very discrepant at higher pressures (>∼1 bar), it is important to note that constraints on the T/P profile at pressures higher than 1 bar are imposed by the finite range of shapes allowed by the T/P profile parameterization, not by observational data. This is because emission spectroscopy cannot probe those depths, as shown by the contribution function (Figure 4.10). At low pressures the two profiles are consistent, despite the different shapes.

Importance of individual molecules

`PLATON` calculates the abundances and line opacities of 28 molecules. To ascertain which molecules are important, we re-calculate the best fit transit and eclipse spectra with the opacities of individual molecules zeroed out to see how χ^2 changes. We find that the transit spectrum is dominated by opacity from H_2O, CO_2, H_2S, and CH_4. Removing all other molecular opacities only increases χ^2 by 0.7. Starting from a reference point consisting of these four molecules, we remove each molecule in turn and calculate the resulting $\Delta\chi^2$, obtaining 1.4 for H_2S, 2.9 for CO_2, 3.6 for CH_4,

and 92.9 for H_2O. Zeroing the aerosol opacity yields $\Delta\chi^2 = 3780$. We conclude that water is by far the dominant molecule shaping the transit spectrum, with aerosol scattering as the most important opacity source overall. This is no surprise, as both the strong scattering slope at short wavelengths and the water feature in the WFC3 bandpass are obvious by eye (Figure 4.7).

The eclipse spectrum worsens by only $\Delta\chi^2 = 0.6$ when the opacity is zeroed for all molecules except H_2O, CO_2, H_2S, CH_4, and CO. Starting from a reference point consisting of these five molecules, we remove each molecule in turn and calculate the resulting $\Delta\chi^2$, obtaining 0.2 for CH_4, 8.5 for H_2S, 56.6 for CO, 81.4 for CO_2, and 551 for H_2O. Thus, we conclude that the emission spectrum contains information on more molecules than the transit spectrum, with H_2S, CO, CO_2, and H_2O all acting as important opacity sources, although water is still dominant.

The tests above reveal the contribution of different molecules to the best-fit transit and eclipse spectra. They do not reveal the significance with which individual molecules are detected in the fits, because many of the features induced by a molecule—a little more absorption here, a little less absorption there—can be mimicked by changes in the free parameters. To quantify the detection significance, we ran a series of retrievals on the transit and eclipse data where we zeroed out the opacity of one molecule at a time and calculated the resulting Bayesian evidence z. The log of the Bayes ratio (indicating the relative preference for the full model versus one without that molecule) is then given by the difference in $\ln(z)$ when compared to the retrieval where all molecular opacities were included. In transit, $\Delta \ln(z)$ was -0.9 for H_2S, 0.5 for CH_4, -0.8 for CO_2, and -9.9 for H_2O, with a margin of error on $\ln(z)$ equal to $\sim$0.2 for all retrievals. We conclude that the transit spectra only provide strong evidence for H_2O, with a Bayes factor of 20,000; all other molecules have a Bayes factor less than 3. In eclipse, $\Delta \ln(z)$ was 0.3 for H_2S, -0.7 for CO, -1.2 for H_2O, and -1.6 for CO_2, with a similar margin of error. We conclude that the eclipse spectrum does not strongly favor the existence of any one molecule. These results are consistent with intuition: there is a visually obvious water feature in the WFC3 transit spectrum, but no molecular features can be seen in the eclipse spectrum (which is predominately composed of broadband photometric points) at any wavelength.

Aerosol properties

Many authors have proposed a Rayleigh scattering haze to explain the optical transmission spectrum of HD 189733b (e.g. Lecavelier Des Etangs et al. 2008; Gibson

et al. 2012; Pont et al. 2013), and this is born out by our retrievals. Our posteriors indicate that a clear atmosphere with a zero number density of haze particles is ruled out to much greater than 3σ significance, and the posterior distribution of particle sizes puts the particles firmly in the Rayleigh regime ($r \lesssim \lambda/10$) for optical wavelengths. In fact, our posterior on the mean particle sizes pushes up against 1 nm, the lower end of the prior—indicating that arbitrarily small particles are allowed by the data. This means that no constraint on the haze composition is possible, as the scattering slope is always $\frac{dR_p}{d\log\lambda} = -4H$ regardless of composition. The lack of a lower limit on the mean particle size also implies that there is no upper limit on the particle density, as there is a perfect degeneracy between the two variables in the Rayleigh regime.

Our findings are consistent with the conclusions of previous studies (i.e. Gibson et al. 2012; Pont et al. 2013), which required the inclusion of a Rayleigh scattering haze in order to reproduce HD 189733b's infrared transmission spectrum.

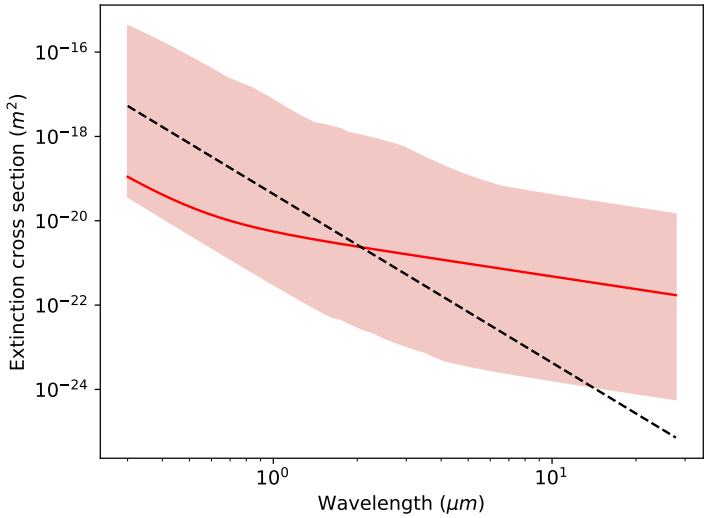


Figure 4.14: Extinction cross section of the haze particles in our best fit model (solid red), along with the 1σ range of extinction cross sections from our retrieval. The dashed solid line falls off as λ^{-4}, and is plotted for reference.

In Figure 4.14, we show that the aerosol extinction cross section falls as λ^{-4} at very short wavelengths, but shifts to λ^{-1} at longer wavelengths. This slow dropoff of extinction cross section at high wavelengths is due to aerosol absorption. To understand this, we can write down simple analytic expressions for scattering and absorption cross sections as a function of wavelength in the Rayleigh regime. These are (Mishchenko et al., 2002):

$$\sigma_{sca} = \frac{2^7 \pi^5 r^6}{3\lambda^4} \left| \frac{m^2 - 1}{m^2 + 2} \right|^2 \tag{4.15}$$

$$\sigma_{abs} = \frac{8\pi^2 r^3}{\lambda} Im\left(\frac{m^2 - 1}{m^2 + 2} \right). \tag{4.16}$$

$$\tag{4.17}$$

Writing the complex refractive index m as $n + \mathrm{i}k$, and assuming $k << n$:

$$\begin{aligned}
\sigma_{sca} &= \frac{2^7 \pi^5 r^6}{3\lambda^4} \frac{4k^2 n^2 + (n^2 - k^2 - 1)^2}{4k^2 n^2 + (n^2 - k^2 + 2)^2} \\
&\approx \frac{2^7 \pi^5 r^6}{3\lambda^4} \frac{(n^2 - 1)^2}{(n^2 + 2)^2} \\
\sigma_{abs} &= \frac{8\pi^2 r^3}{\lambda} \frac{6nk}{(n^2 - k^2 + 2)^2 + 4k^2 n^2} \\
&\approx \frac{8\pi^2 r^3}{\lambda} \frac{6nk}{(n^2 + 2)^2}.
\end{aligned}$$

One can see from these equations that aerosol scattering falls off with wavelength much faster than aerosol absorption, causing extinction to be dominated by scattering at short wavelengths and absorption at long wavelengths. Since the absorption cross section is proportional to the imaginary component of the refractive index, even a small imaginary component increases long wavelength extinction by orders of magnitude compared to a real refractive index.

The importance of k, the imaginary component of the refractive index, poses a challenge for our model. Many different cloud species have been proposed for this planet. Lee et al. (2016) used a GCM simulation to model condensate clouds and found that the clouds are dominated by silicate materials such as $MgSiO_3$ at mid-high latitudes, but TiO_2 and SiO_2 dominate in equatorial regions. Lavvas & Koskinen (2017) considered photochemical hazes and found that soot-composition aerosols provided a good match to HD 189733b's transmission spectrum. This study reported $k \sim 0.5$ at 500 nm for soot, while Kitzmann & Heng (2018) reported $3.7{\times}10^{-5}$ for glassy $MgSiO_3$ at the same wavelength, $5.1{\times}10^{-4}$ for TiO_2, and $1.7{\times}10^{-5}$ for SiO_2. In addition to these species-dependent variations in k, the k for each species also varies drastically (and uniquely) with wavelength. For example, k rises from $1{\times}10^{-4}$ to nearly 1 over the wavelength range $2 - 9$ μm for $MgSiO_3$, and SiO_2 exhibits a

similar behavior before dropping 2 orders of magnitude within 2 μm. In light of these uncertainties, we chose to fit for a wavelength-independent k value rather than fixing it to the theoretical prediction for a given cloud species.

We argued earlier that the presence of aerosol has a negligible effect on HD 189733b's dayside emission spectrum. We check the validity of this assumption using our best-fit model. We find that when we include the best-fit aerosol model from the transmission spectrum in our calculation of HD 189733b's dayside emission spectrum the resulting eclipse depth values change less than ~0.1 ppb, with a corresponding change in χ^2 of only 2×10^{-5}. This is unsurprising, as the photospheric pressure is lower for transmission spectrum than it is for eclipse spectrum by a factor of $\sqrt{2\pi R/H} \sim 50$ (Fortney, 2005); this is apparent when we compare the transmission spectrum contribution function in Figure 4.11 to the emission spectrum contribution function in Figure 4.10. At higher pressures, the mixing ratio of aerosol is lower because our retrievals prefer an aerosol scale height that is greater than the gas scale height. This causes aerosol absorption to be an important source of opacity at low pressures only. We note that the dayside is also expected to be hotter than the terminator, making it less likely that the condensate clouds detected at the terminator would persist in this region. Even if the dayside is in reality partly cloudy, the dayside emission would be dominated by clear regions because they have deeper and hotter photospheres, which emit more radiation.

Validity of equilibrium chemistry

`PLATON` assumes equilibrium chemistry. Since chemical reaction timescales decline rapidly with temperature, the colder a planet, the more disequilibrium chemistry matters. HD 189733b lies in a regime where disequilibrium chemistry may be important. Multiple studies have explored disequilibrium chemistry on this planet (Line et al., 2010; Venot et al., 2012; Moses et al., 2013; Agúndez et al., 2014; Blumenthal et al., 2018; Steinrueck et al., 2019). Venot et al. (2012) considered 1D models with both UV photochemistry and vertical mixing, and found negligible differences in the resulting transmission and emission spectra as compared to equilibrium models. Their Figure 10 shows that the disequilibrium-induced brightness temperature discrepancy is at most several Kelvin (less than 1%), while the transit radius discrepancy is at most ~30 ppm. Blumenthal et al. (2018) also model the emission spectrum under equilibrium and non-equilibrium conditions, finding no detectable difference even with *JWST* (their Figure 3). Other studies report changes in abundance as a result of disequilibrium chemistry, but do not compare the result-

ing spectra to the equilibrium model predictions. For example, Moses et al. (2013) find that the water abundance only becomes discrepant for P < 1 microbar, while the HCN abundance is enhanced by orders of magnitude for P <0.5 bar. In contrast to these 1D models, which predict relatively small changes in HD 189733b's observed transmission and emission spectrum, Steinrueck et al. (2019) calculated emission spectra for HD 189733b using general circulation models where they fixed the ratio of CH_4 to CO across the planet to mimic the effect of transport-induced quenching. They find that disequilibrium chemistry due to horizontal transport changes the emission spectrum by up to 20% for a heavily CO-dominated atmosphere (CH_4/CO = 0.001): there is a systematic offset of 10% in addition to wavelength-dependent discrepancies of order several percent.

To test whether disequilibrium chemistry may be important for our data, we performed retrievals on the transit and eclipse spectra with the vertical quench pressure as a free parameter. We quenched all molecules in our first trial, but only CH4 and CO in our second trial (in accordance with Morley et al. 2017), with no change in the following conclusions. Quenching did not result in a better fit ($\Delta\chi^2 \sim 0$) for either our transmission or emission spectra. In both cases, the posterior distribution of the quench pressure pushes up against the lower bound of the prior—0.1 mbar for the transit spectrum and 1 mbar for the eclipse spectrum. Additionally, the posterior distributions of the other parameters did not appreciably change in these retrievals. For example, the metallicity posterior shifted by 0.1σ for the transit retrieval and 0.5σ for the eclipse retrieval; the C/O ratio posterior shifted by 0.06σ for both transit and eclipse retrievals. We conclude that the effects of disequilibrium chemistry are below our detection threshold, in good agreement with the predictions of the 1D models from Venot et al. (2012) and Blumenthal et al. (2018).

The effect of starspots

HD 189733 is an active K dwarf with spots that cover a few percent of its surface. When the planet crosses a prominent spot it creates a readily identifiable deviation in the transit light curve shape; this makes it straightforward to identify and mask such events in high signal-to-noise *HST* observations (e.g. Sing et al., 2011; Pont et al., 2013). However, the presence of unocculted spots during a transit also biases the retrieved transit depth in a way that can mimic the effect of scattering as discussed in McCullough et al. (2014). If the fractional spot coverage varies from one transit epoch to another, the relative effect on the transit depth will vary as well. An additional complication is faculae, bright regions ~100 K hotter than

their surroundings. For a K1.5 star like HD 189733, faculae can cover 17-40% of the surface, partially cancelling the effect due to spots (Rackham et al., 2019). The large faculae fraction is a mixed blessing. On one hand, it becomes hard to correct for the effects of faculae because we cannot assume faculae are unocculted. On the other hand, the faculae covering fraction is so large that the planet is likely to cross at least one faculae and one non-faculae region during every transit, thus mitigating the spectral bias from unocculted faculae. We therefore neglect the effects of faculae and only consider spots.

Pont et al. (2013) attempted to correct for the effects of time-varying spot coverage by using photometric data from the Automated Patrol Telescopes (Henry, 1999) to estimate the apparent brightness at the epoch of each transit observation. Although these data made it possible to estimate the magnitude of the variations in spot coverage, they do not provide an estimate of the overall spot coverage fraction because the star is not necessarily spotless at maximum brightness. The authors resolve this problem by assuming an average spot coverage fraction of 1%, based on three pieces of evidence: the frequency of spot-crossing events during *HST* transits (Sing et al., 2011), stochastic starspot simulations by Aigrain et al. (2012), and the lack of features in the transmission spectrum (i.e. Mg H line, stellar sodium line) caused by abundant starspots. They argue that these lines of evidence make it unlikely that the spot coverage fraction is much above 2%.

McCullough et al. (2014) use the same data to argue for a higher starspot fraction of 4%. First, they emphasize that the starspot fraction derived from spot crossing events would be an underestimate if most starspots are in the polar regions, where the planet does not transit. Second, they use Equation 14 of Aigrain et al. (2012) to derive a lower bound on the starspot fraction:

$$\delta >\approx \frac{\Psi_{max} - \Psi_{min} + \sigma}{\Psi_{max} + \sigma}.$$

Since the measured difference between the maximum flux (Ψ_{max}) and minimum flux (Ψ_{min}) is around 4% and the rotational modulation (σ) is of order one percent, $\delta >\approx 0.04$. However, this derivation really computes a lower bound on the maximum starspot fraction over a period of time, whereas the 1-2% figure quoted by Pont et al. (2013) is the average starspot fraction. McCullough et al. (2014) go on to argue that if the starspot fraction were 4.3%, the majority of the increased apparent transit

depth in the UV compared to IR could be explained by unocculted starspots, and there would be no need to invoke scattering from aerosol.

To shed light on this issue, we ran another retrieval with `PLATON` where the spot coverage fraction is allowed to vary as a free parameter. We fix the spot temperature to 4250 K, the temperature derived by Sing et al. (2011) from observations of spot occultations. Since we used the transit depths from Pont et al. (2013), which already corrected for the effects of starspots using APT photometry and an assumed 1% baseline starspot fraction, our fitted spot coverage fraction is in reality an excess spot fraction above this 1% baseline. We find an excess spot coverage fraction of $1.8^{+0.7}_{-1}$%, which in turn implies a total average starspot fraction of $2.8^{+0.7}_{-1}$%. This figure is intermediate between the 1-2% argued for by Pont et al. (2013) and the 4% argued by McCullough et al. (2014). However, the inclusion of spot coverage fraction as a free parameter does little to modify the posteriors for the other model parameters. The median values of all parameters are consistent with the values from the fiducial retrieval (Table 4.6) to better than 1σ. We therefore conclude that our derived haze properties, including the mean particle size, particle number density, and fractional scale height, are insensitive to our assumptions about the spot coverage fraction.

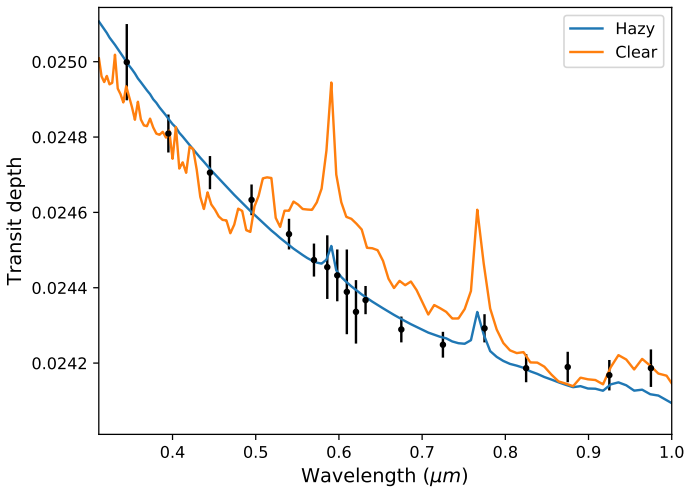


Figure 4.15: `PLATON` model of a hazy atmosphere on HD 189733b, compared to a clear atmosphere with a starspot-induced slope corresponding to a spot coverage fraction of 6%. Note the strong atomic and molecular features in the clear atmosphere, especially the far wings of the sodium and potassium absorption lines.

If unocculted starspots and haze can both introduce a slope in HD 189733b's optical transmission spectrum, why is there not a degeneracy between the two? To answer this question, we plotted a transmission spectrum with no aerosol but with a high spot coverage fraction of 6%, as shown in Figure 4.15. We found that the transmission spectrum has strong atomic and molecular absorption features even at short wavelengths ($\lambda < 1$ μm), the most prominent of which are the far wings of the Na doublet at 589 nm and of the K doublet at 770 nm. These features are not seen in the observational data, which are nearly featureless at optical wavelengths.

A consistency check for composition

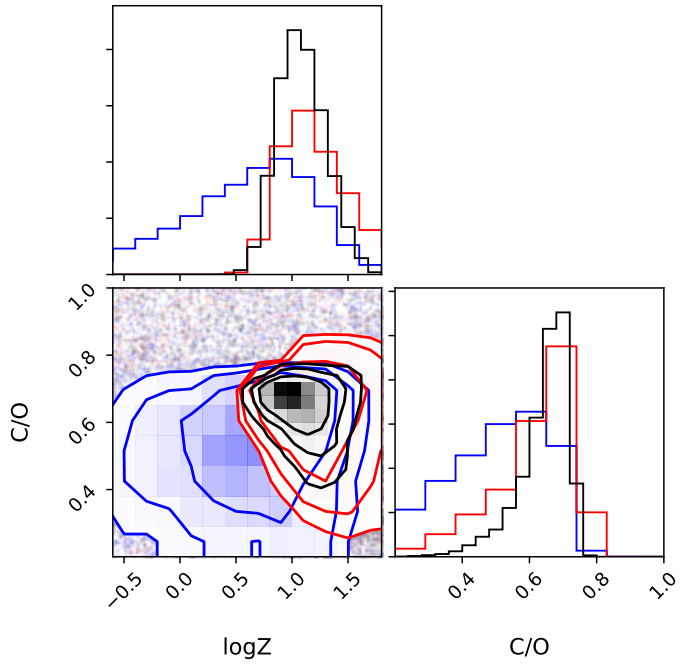


Figure 4.16: Posterior distributions for the transit-only retrieval (blue), eclipse-only retrieval (red), and combined retrieval (black).

Although we fit HD 189733b's transmission and emission spectra jointly, in reality the models used to fit these two spectra are largely independent of each other. The transmission spectrum determines the isothermal limb temperature and aerosol properties, while the emission spectrum sets the dayside T/P profile. The planetary radius is technically constrained by both transmission and emission spectra, but most of the statistical power comes from the transmission spectrum. The only two parameters that are comparably constrained by both spectra are the atmospheric

metallicity and C/O ratio, which determine the chemistry for both the limb and the day side.

To illustrate the relative contributions of transmission and emission spectra to HD 189733b's inferred atmospheric composition, we ran a transit-only and an eclipse-only retrieval. In the transit-only retrieval, the planetary radius, metallicity, C/O ratio, isothermal limb temperature, aerosol properties, and WFC3 offset were free parameters. In the eclipse-only retrieval, the metallicity, C/O ratio, dayside T/P profile parameters, and WFC3 offset were free parameters. Figure 4.16 shows the resulting posterior distributions for metallicity and C/O from these two retrievals, with the distributions from the combined retrieval overplotted. The three retrievals give fully consistent constraints on both parameters. The emission spectrum puts a slightly tighter constraint on metallicity, but both spectra place comparable constraints on the C/O ratio.

High resolution studies

HD 189733b has been a favorite target for high resolution spectroscopy. VLT/CRIRES detected water on its dayside (Birkby et al., 2013), as well as both water and CO on its terminator (Brogi et al., 2016). TNG/GIANO and CARMENES also detected water in the transit spectrum (Brogi et al., 2018; Alonso-Floriano et al., 2019). Carbon monoxide has been detected on its dayside by both Keck/NIRSPEC (Rodler et al., 2013) and VLT/CRIRES (de Kok et al., 2013). The dayside detections with CRIRES were confirmed by Cabot et al. (2019), who additionally report a high confidence (5.0σ) detection of HCN, with a HCN mixing ratio of 10^{-6} yielding peak detection significance.

We test the robustness of the water and HCN detections with `PLATON`. We replicate the methodology in Cabot et al. (2019) to reduce the CRIRES L band (3.18–3.27 μm) data. We then generate a high-resolution `PLATON` model of HD 189733b assuming the best fit parameters to the low resolution data, with only the line opacity of the molecule in question included. The model eclipse spectrum is cross correlated with the data to look for a signal with the expected radial velocity drift of the planet. We detect H_2O and HCN at a significance of 4.3σ and 4.8σ, respectively, in line with the results of Cabot et al. (2019). These results are shown in Figure 4.17.

While both molecular detections seem robust, they may not be. Standard high resolution analysis methods involve optimizing many parameters to maximize the detection significance. For Cabot et al. (2019), these are the number of SYSREM

(principal component analysis to remove telluric features) iterations for each of the 4 detectors, the planetary orbital velocity, the systemic velocity offset, the percentage of wavelengths to mask due to low atmospheric transmission, and the percentage of wavelengths to mask due to high variability. They report from injection-recovery tests that this optimization procedure yields false positives as high as 4.0σ 30% of the time.

To reduce the bias introduced by the optimization procedure, we fix the orbital velocity to 152.5 km/s (Brogi et al., 2016) and restrict the systemic velocity offset to ± 1 km/s. We then estimate the bias introduced by the optimization process by a bootstrap-inspired procedure. We randomly select, with replacement, 48 spectra from the original list of 48 spectra to form a new list of 48 spectra in random order. We then apply the same analysis used for the original set of spectra to the new set. This random selection and ordering, combined with the radial acceleration of the planet (amounting to 2 pixels/spectrum), means that the lines in the template will rarely match up with the planetary absorption lines. We expect that the detection significance of all molecules in this scenario should be 0σ, and the magnitude of the recovered signals therefore allows us to estimate the bias introduced by the optimization steps.

We run this bootstrap procedure 1000 times, gathering the "detection significance" from each run. We find that our optimization procedure, despite exploring fewer free parameters than Cabot et al. (2019), returns an average bias of 1.1σ for water and 2.9σ for HCN. Returning to the original analysis, our bias-corrected detection significance is then 3.2σ for water, and 1.9σ for HCN. We conclude that the water detection is statistically significant, but the HCN detection is not.

In order to reduce the magnitude of this bias, we modified our analysis to minimize the size of the parameter space we optimize over. Instead of optimizing the two masking parameters, we fix them to reasonable values: we mask wavelengths where the atmospheric transmission is under 30%, and where the standard deviation of the wavelength across all spectra is in the top 10% of all standard deviations for that detector. Instead of optimizing over the four-dimensional space consisting of the number of SYSREM iterations for each detector, we optimize the number of iterations for each detector in turn before summing their weighted CCFs. In addition, we include data from all detectors, instead of excluding the second detector due to the strong telluric absorption at those wavelengths.

These changes turn a 6 dimensional optimization problem into four 1-dimensional optimization problems, each of which can have one of 10 discrete values, drastically decreasing the potential for bias. We also apply a fourth order Butter high-pass filter to both the template and the data with a cutoff frequency of 0.01 pixel^{-1}, further reducing the potential for low-frequency systematics to create false signals. These simplifications have the additional benefit of making our optimization code much faster. In this new analysis, we find that water is detected at 3.8σ with an optimization bias of 0.7σ from bootstrapping, resulting in a corrected significance of 3.1σ. HCN is detected at 2.8σ with an optimization bias of 1.3σ, with a corrected significance of 1.5σ. These results are in good agreement with our previous conclusions: the water detection is secure, but the HCN detection is statistically insignificant.

Why does our HCN measurement consistently exhibit higher levels of bias than water? Although a detailed investigation of this question is outside the scope of this paper, we offer some speculations based on the differing statistical properties of the water and HCN line lists. Water has a large number of weak lines irregularly spaced across the band, while HCN has a smaller number of very strong lines with a regular periodic spacing. Thus, for HCN, the cross correlation function is dominated by strong lines spanning a smaller number of pixels, making overfitting more likely. The cross correlation function for water depends on contributions from many pixels, making it more difficult to overfit.

The periodic nature of the HCN lines means that the same problem occurs in Fourier space as in wavelength space. Cross correlation is mathematically equivalent to taking the Fourier transform of the template, multiplying it by the conjugate of the Fourier transform of the data, and inverse Fourier transforming the product. For templates with more periodic features, the Fourier transform of the template is dominated by a few high peaks. These peaks may by chance coincide with unsubtracted periodic systematics, resulting in a spurious signal. The Fourier transform for the water template is more evenly distributed, and is therefore less prone to this problem.

There are also physical reasons to doubt the HCN detection. The detections of H_2O and CO are fully consistent with our best fit model to the low resolution data, which predicts that these molecules should be the most abundant active gases in the atmosphere at pressures lower than 1 bar. These same models predict that the spectral signature of HCN should be significantly weaker than the signals from the more abundant H_2O and CH_4. CH_4 has a slightly higher absorption cross section between $3.18 - 3.27$ μm, the wavelength range CRIRES covers. While H_2O

has a lower cross section over these wavelengths, it is many orders of magnitude more abundant. Therefore, we would not expect HCN to be observable unless its abundance exceeds that of methane and is at least ∼10% that of water; this would be many orders of magnitudes higher than predictions from equilibrium chemistry models for this planet.

One way around this difficulty is to invoke disequilibrium chemistry, as HCN abundances are enhanced both by transport-induced quenching and by photochemistry. Moses et al. (2011) simulated HD 189733b and found that its HCN abundance is enhanced by disequilibrium processes. Encouragingly, they found HCN abundances close to 10^{-5}, similar to the 10^{-6} inferred observationally by Cabot et al. (2019). However, they find that CH_4 abundances are enhanced by the same processes, leaving the CH_4 abundance greater than or similar to the HCN abundance at typical photospheric pressures of 100 mbar. They also find that the water abundance is not changed except at very high altitudes.

The dominance of CH_4 in the models raises the question of whether the high-resolution CRIRES data shows any evidence of CH_4. We searched for methane using the same CRIRES data by utilizing the methane line list by Rey et al. (2017), which is complete at these high temperatures and has accurate line positions suitable for high resolution studies. We find only a 1.2σ signal, which is not significant. This result is shown in Figure 4.17.

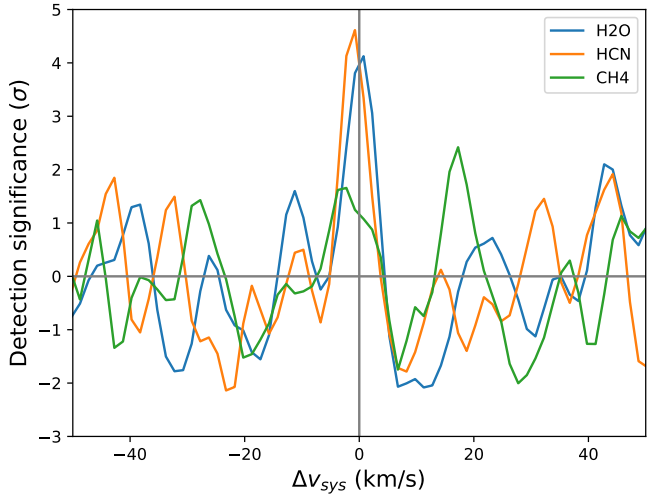


Figure 4.17: Detection significance of H_2O, HCN, and CH_4 in the high resolution emission spectrum of HD 189733b. H_2O and HCN are seemingly clearly detected, while CH_4 is not. The model used to cross correlate with data was generated by `PLATON` using the parameters of the best fit model from the low resolution retrieval. Despite the seeming robustness of the HCN detection, bootstrap analysis reveals that its actual significance is 1.5–1.9σ.

4.7 Conclusion

A new and improved `PLATON` is available for download.[3] It now comes with an eclipse depth calculator, updated opacities, joint transit-eclipse retrieval capability, and correlated-k capability. In addition, we provide high resolution opacity data, making line-by-line calculations possible for the first time. We demonstrate `PLATON`'s new capabilities by using it to simultaneously analyze the best available *HST* and *Spitzer* transit and eclipse depths for the archetypal hot Jupiter HD 189733b. To our knowledge, this is the first published retrieval on this comprehensive data set, as well as the first published joint retrieval that includes both transmission and emission spectroscopy.

Our resulting inferences for the properties of HD 189733b's atmosphere are qualitatively similar to—but more constraining than—those of previous authors. We find that the data favors a haze with a mean particle radius less than 14 nm. Our fiducial T/P profile indicates that the planet is consistent with a zero-albedo object with perfect heat redistribution. We find that the atmosphere is of moderately super-solar ($7 - 21\times$ solar) metallicity and constrain the C/O ratio to lie between $0.47 - 0.69$,

[3]Latest version: `https://github.com/ideasrule/platon`; version corresponding to this paper: `https://doi.org/10.5281/zenodo.3923090`.

consistent with the solar value, but possibly lower than the stellar value. This planet has one of the tightest metallicity constraints ever measured, with only WASP-127b and WASP-39b being comparable (see Figure 22 of Spake et al. 2019). In our best-fit model, CO and H_2O are the most abundant absorbing species at photospheric pressures, consistent with the detection of H_2O in *HST*/WFC3 spectroscopy, and the detection of both molecules in high resolution spectroscopy.

We explore the effects of stellar activity using a retrieval in which the starspot coverage fraction is allowed to vary as a free parameter, and find a best-fit starspot coverage fraction of $1.8^{+0.7}_{-1}$%. Even when this coverage is allowed to vary, our fit still requires the presence of a haze with much the same properties as the fiducial retrieval in order to create a featureless optical transit spectrum.

HD 189733b has exceptional observational data unmatched by any other exoplanet in quality, quantity, or wavelength range. Much of this data was collected by *Spitzer* or by now-defunct instruments on *HST*, and can never be replicated. Our retrieval demonstrates what kinds of properties can be inferred, and to what precision, for the most observationally favorable hot Jupiters in the pre-*JWST* era. We have come close to testing the prediction by Espinoza et al. (2017) that enhanced atmospheric metallicity is inevitably associated with sub-stellar C/O ratio—in fact, the main obstacle was the uncertain stellar C/O ratio. In addition, although the data for HD 189733b are already far more accurate than our models—the 3.6, 4.5, and 8.0 eclipse depths have errors of 1-2%, whereas we have demonstrated that small changes in `PLATON`'s (and TauREx's) algorithm change the eclipse depth by several percent—we cannot fully take advantage of this accuracy in a retrieval. This is because a retrieval has enough free parameters to fit the 3.6, 4.5, and 8.0 eclipse depths to arbitrary accuracy, and the lack of any molecular features in emission contributes to wide posteriors on atmospheric parameters despite the exceptional data. *JWST* will be able to accurately measure multiple emission bands from multiple molecules across a large wavelength range, spurring spur the development of more sophisticated and more accurate models than the current state of the art.

4.8 Acknowledgements

M.Z. acknowledges Plato (Greek:ΠΛΑΤΩΝ) for his clear and thought-provoking dialogues, which ought to be exemplars of good writing for academics everywhere. We also thank Michael R. Line for helpful advice. Support for this work was provided by *HST* GO programs 13431, 13665, and 14260.

Software: `numpy (van der Walt et al., 2011), scipy (Virtanen et al., 2020), matplotlib (Hunter, 2007), emcee (Foreman-Mackey et al., 2013), dynesty (Speagle, 2020), corner (Foreman-Mackey, 2016), nose, Travis-CI`

4.9 Appendix

The following HD 189733b parameters in Table 4.7 were fixed during the retrieval:

Table 4.7: Fixed planetary parameters in HD 189733b retrieval

Parameter	Value
T_s	5052 K
R_s	0.751 $R_\odot$
M_p	1.129 M_J
a	0.03142 AU

All parameters are taken from Stassun et al. 2017 except the last, which is from Southworth 2010

Tables 4.8 and 4.10 list the transit and eclipse depths we used in the retrieval, along with the papers they were taken from.

Table 4.8: Adopted transit depths. Sources: (1) Pont et al. 2013; (2) McCullough et al. 2014

λ_{min}(μm)	λ_{max}(μm)	Depth (ppm)	Error (ppm)	Source
0.32	0.37	24999	101	1
0.37	0.42	24809	50	1
0.42	0.47	24706	44	1
0.47	0.52	24633	41	1
0.52	0.56	24542	41	1
0.56	0.58	24473	44	1
0.58	0.592	24455	84	1
0.592	0.604	24433	69	1
0.604	0.615	24389	112	1
0.615	0.626	24336	84	1
0.626	0.638	24367	37	1
0.65	0.7	24289	34	1
0.7	0.75	24249	34	1
0.75	0.8	24292	37	1
0.8	0.85	24186	37	1
0.85	0.9	24190	40	1
0.9	0.95	24168	40	1
0.95	1.0	24186	50	1
1.0	1.17	24062	68	1
1.1184	1.1374	23962	73	2
1.1372	1.1562	24047	67	2
1.156	1.175	24078	105	2
1.1748	1.1938	24035	87	2
1.1936	1.2126	23961	80	2
1.2123	1.2313	23955	70	2
1.2311	1.2501	23884	56	2
1.2499	1.2689	24000	62	2
1.2687	1.2877	23863	61	2
1.2875	1.3065	23987	69	2
1.3062	1.3252	23961	60	2
1.325	1.344	23982	66	2
1.3438	1.3628	24134	55	2
1.3626	1.3816	24149	61	2
1.3814	1.4004	24091	63	2
1.4001	1.4191	24215	77	2
1.4189	1.4379	24199	64	2
1.4377	1.4567	24108	71	2
1.4565	1.4755	24018	67	2
1.4752	1.4942	24188	75	2
1.494	1.513	23941	62	2
1.5128	1.5318	24097	61	2
1.5316	1.5506	24002	62	2
1.5504	1.5694	24010	72	2

Table 4.9: Adopted transit depths, cont.

λ_{min}(μm)	λ_{max}(μm)	Depth (ppm)	Error (ppm)	Source
1.5691	1.5881	24100	87	2
1.5879	1.6069	23963	75	2
1.6067	1.6257	23916	98	2
1.6255	1.6445	24062	84	2
3.2	3.9	24047	84	1
4.0	5.0	24155	109	1
5.0	6.4	23951	207	1
6.4	9.3	24056	105	1
23.5	24.5	23898	291	1

Table 4.10: Adopted eclipse depths. Sources: (1) Crouzet et al. 2014; (2) Kilpatrick et al. 2020; (3) Charbonneau et al. 2008; (4) Agol et al. 2010

λ_{min}(μm)	λ_{max}(μm)	Depth (ppm)	Error (ppm)	Source
1.1184	1.1374	0	47	1
1.1372	1.1562	78	50	1
1.156	1.175	124	45	1
1.1748	1.1938	93	44	1
1.1936	1.2126	89	43	1
1.2123	1.2313	51	50	1
1.2311	1.2501	64	42	1
1.2499	1.2689	99	42	1
1.2687	1.2877	80	42	1
1.2874	1.3064	149	41	1
1.3062	1.3252	127	41	1
1.325	1.344	108	41	1
1.3438	1.3628	75	41	1
1.3626	1.3816	69	41	1
1.3813	1.4003	56	42	1
1.4001	1.4191	104	42	1
1.4189	1.4379	94	42	1
1.4377	1.4567	67	42	1
1.4565	1.4755	32	43	1
1.4753	1.4943	105	43	1
1.494	1.513	228	43	1
1.5128	1.5318	211	43	1
1.5316	1.5506	143	44	1
1.5504	1.5694	99	45	1
1.5691	1.5881	135	45	1
1.5879	1.6069	80	46	1
1.6067	1.6257	32	46	1
1.6255	1.6445	110	74	1
3.2	4.0	1481	34	2
4.0	5.0	1827	22	2
5.1	6.3	3100	340	3
6.6	9.0	3440	36	4
13.5	18.5	5190	220	3
20.8	26.1	5980	380	3

Table 4.11: ExoMol line lists used in PLATON v5

Molecule	List name	N_{lines}	T_{max}(K)	Reference
C_2H_4	MaYTY	49,673,223,799*	700	Mant et al. (2018)
CO	Li2015	125,496*	9000	Li et al. (2015)
H_2CO	AYTY	12,648,694,479*	1500*	Al-Refaie et al. (2015)
H_2S	AYT2	115,623,180*	2000*	Azzam et al. (2016)
H_2O	POKAZATEL	5,550,587,708*	∞^b*	Polyansky et al. (2018)
HCl	Yueqi	2588	?	Li et al. (2013)
HCN	Harris	34,418,408*	?	Barber et al. (2013)
MgH	MoLLIST	14,179*	?	GharibNezhad et al. (2013)
NH_3	CoYuTe	1,135,240,003*	1500*	Coles et al. (2019)
NO	NOname	2,280,366*	?	Wong et al. (2017)
OH	MoLLIST	54,276	?	Brooke et al. (2016)
PH_3	SAlTY	16,931,647,841*	1500*	Sousa-Silva et al. (2014)
SH	SNaSH	81,348*	5000	Yurchenko et al. (2018)
SiH	SiGHTLY	1,724,841*	5000	Yurchenko et al. (2017)
SiO	EBJT	254,675*	9000	Barton et al. (2013)
SO_2	ExoAmes	1,402,257,689*	2000	Underwood et al. (2016)
TiO	ToTo	295,086,011[a]	5000	McKemmish et al. (2019)
VO	VOMYT	277,131,624	5000	McKemmish et al. (2016)

[(a)] All 5 isotopologues combined. Each isotopologue has 58-60 million lines.

[(b)] This line list is complete

*These numbers disagree with those in the ExoMol database's def files

References

Agol, E., Cowan, N. B., Knutson, H. A., et al. 2010, ApJ, 721, 1861

Agúndez, M., Parmentier, V., Venot, O., Hersant, F., & Selsis, F. 2014, A&A, 564, A73

Agúndez, M., Venot, O., Iro, N., et al. 2012, A&A, 548, A73

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147

Al-Refaie, A. F., Changeat, Q., Waldmann, I. P., & Tinetti, G. 2019, arXiv e-prints, arXiv:1912.07759

Al-Refaie, A. F., Yachmenev, A., Tennyson, J., & Yurchenko, S. N. 2015, MNRAS, 448, 1704

Ali-Dib, M. 2017, MNRAS, 467, 2845

Allard, F., Homeier, D., & Freytag, B. 2012, Philosophical Transactions of the Royal Society of London Series A, 370, 2765

Allard, N. F., Spiegelman, F., Leininger, T., & Molliere, P. 2019, A&A, 628, A120

Allard, N. F., Spiegelman, F., & Kielkopf, J. F. 2016, A&A, 589, A21

Alonso-Floriano, F. J., Sánchez-López, A., Snellen, I. A. G., et al. 2019, A&A, 621, A74

Amundsen, D. S., Tremblin, P., Manners, J., Baraffe, I., & Mayne, N. J. 2017, A&A, 598, A97

Azzam, A. A. A., Tennyson, J., Yurchenko, S. N., & Naumenko, O. V. 2016, MNRAS, 460, 4063

Barber, R. J., Strange, J. K., Hill, C., et al. 2013, MNRAS, 437, 1828

Barstow, J. K., Aigrain, S., Irwin, P. G. J., et al. 2014, ApJ, 786, 154

Barton, E. J., Hill, C., Czurylo, M., et al. 2017, J. Quant. Spec. Radiat. Transf., 203, 490

Barton, E. J., Yurchenko, S. N., & Tennyson, J. 2013, MNRAS, 434, 1469

Beaulieu, J. P., Carey, S., Ribas, I., & Tinetti, G. 2008, ApJ, 677, 1343

Benneke, B. 2015, arXiv e-prints, arXiv:1504.07655

Birkby, J. L. 2018, arXiv e-prints, arXiv:1806.04617

Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, MNRAS, 436, L35

Blecic, J. 2016, ArXiv e-prints, 1604.02692

Blecic, J., Dobbs-Dixon, I., & Greene, T. 2017, ApJ, 848, 127

Blecic, J., Harrington, J., & Bowman, M. O. 2016, ApJS, 225, 4

Blumenthal, S. D., Mandell, A. M., Hébrard, E., et al. 2018, ApJ, 853, 138

Booth, R. A., & Ilee, J. D. 2019, MNRAS, 487, 3998

Bouanich, J.-P., Blanquet, G., Walrand, J., & Lepère, M. 2003, Journal of Molecular Spectroscopy, 218, 22

Bouchy, F., Udry, S., Mayor, M., et al. 2005, A&A, 444, L15

Brogi, M., de Kok, R. J., Albrecht, S., et al. 2016, ApJ, 817, 106

Brogi, M., Giacobbe, P., Guilluy, G., et al. 2018, A&A, 615, A16

Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2013, ApJ, 767, 27

Brooke, J. S. A., Bernath, P. F., Western, C. M., et al. 2016, J. Quant. Spec. Radiat. Transf., 168, 142

Cabot, S. H. C., Madhusudhan, N., Hawker, G. A., & Gandhi, S. 2019, MNRAS, 482, 4422

Chachan, Y., Knutson, H. A., Gao, P., et al. 2019, AJ, 158, 244

Charbonneau, D., Knutson, H. A., Barman, T., et al. 2008, ApJ, 686, 1341

Coles, P. A., Yurchenko, S. N., & Tennyson, J. 2019, MNRAS, 490, 4638

Cridland, A. J., van Dishoeck, E. F., Alessi, M., & Pudritz, R. E. 2019, A&A, 632, A63

Crouzet, N., McCullough, P. R., Deming, D., & Madhusudhan, N. 2014, ApJ, 795, 166

de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, A&A, 554, A82

Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, The Astrophysical Journal, 644, 560

Deming, D., Wilkins, A., McCullough, P., et al. 2013, ApJ, 774, 95

Deming, L. D., & Seager, S. 2017, Journal of Geophysical Research (Planets), 122, 53

Désert, J. M., Sing, D., Vidal-Madjar, A., et al. 2011, A&A, 526, A12

Espinoza, N., Fortney, J. J., Miguel, Y., Thorngren, D., & Murray-Clay, R. 2017, The Astrophysical Journal, 838, L9

Evans, T. M., Pont, F., Sing, D. K., et al. 2013, ApJ, 772, L16

Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fortney, J. J. 2005, MNRAS, 364, 649

Fortney, J. J., Lupu, R. E., Morley, C. V., Freedman, R. S., & Hood, C. 2019, The Astrophysical Journal, 880, L16

Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, ApJS, 214, 25

Freedman, R. S., Marley, M. S., & Lodders, K. 2008, ApJS, 174, 504

GharibNezhad, E., Shayesteh, A., & Bernath, P. F. 2013, MNRAS, 432, 2043

Gibson, N. P., Pont, F., & Aigrain, S. 2011, MNRAS, 411, 2199

Gibson, N. P., Aigrain, S., Pont, F., et al. 2012, MNRAS, 422, 753

Goody, R., West, R., Chen, L., & Crisp, D. 1989, J. Quant. Spec. Radiat. Transf., 42, 539

Gordon, I. E., Rothman, L. S., Hill, C., et al. 2017, J. Quant. Spec. Radiat. Transf., 203, 3

Grillmair, C. J., Burrows, A., Charbonneau, D., et al. 2008, Nature, 456, 767

Grimm, S. L., & Heng, K. 2015, ApJ, 808, 182

Guillot, T. 2010, A&A, 520, A27

Guo, X., Crossfield, I. J. M., Dragomir, D., et al. 2020, AJ, 159, 239

Hayes, J. J. C., Kerins, E., Awiphan, S., et al. 2020, MNRAS

Haynes, K., Mandell, A. M., Madhusudhan, N., Deming, D., & Knutson, H. 2015, ApJ, 806, 146

Hedges, C., & Madhusudhan, N. 2016, MNRAS, 458, 1427

Helling, C., Lee, G., Dobbs-Dixon, I., et al. 2016, MNRAS, 460, 855

Henry, G. W. 1999, PASP, 111, 845

Huitson, C. M., Sing, D. K., Vidal-Madjar, A., et al. 2012, MNRAS, 422, 2477

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

John, T. L. 1988, A&A, 193, 189

Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, Icarus, 328, 160

Kempton, E. M.-R., Lupu, R., Owusu-Asare, A., Slough, P., & Cale, B. 2017, PASP, 129, 044402

Kilpatrick, B. M., Kataria, T., Lewis, N. K., et al. 2020, AJ, 159, 51

Kipping, D. M., & Tinetti, G. 2010, MNRAS, 407, 2589

Kirk, J., López-Morales, M., Wheatley, P. J., et al. 2019, AJ, 158, 144

Kitzmann, D., & Heng, K. 2018, MNRAS, 475, 94

Kitzmann, D., Heng, K., Oreshenko, M., et al. 2020, ApJ, 890, 174

Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007, ApJ, 655, 564

Knutson, H. A., Charbonneau, D., Cowan, N. B., et al. 2009, ApJ, 690, 822

Knutson, H. A., Lewis, N., Fortney, J. J., et al. 2012, ApJ, 754, 22

Komacek, T. D., Jansen, M. F., Wolf, E. T., & Abbot, D. S. 2019, ApJ, 883, 46

Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, ApJ, 793, L27

Lacis, A. A., & Oinas, V. 1991, J. Geophys. Res., 96, 9027

Lavvas, P., & Koskinen, T. 2017, ApJ, 847, 32

Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., & Sing, D. 2008, A&A, 481, L83

Lee, G., Dobbs-Dixon, I., Helling, C., Bognar, K., & Woitke, P. 2016, A&A, 594, A48

Lee, G., Helling, C., Dobbs-Dixon, I., & Juncher, D. 2015, A&A, 580, A12

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

Lee, J.-M., Irwin, P. G. J., Fletcher, L. N., Heng, K., & Barstow, J. K. 2014, The Astrophysical Journal, 789, 14

Li, G., Gordon, I. E., Rothman, L. S., et al. 2015, ApJS, 216, 15

Li, G., Gordon, I. E., Roy, R. J. L., et al. 2013, J. Quant. Spec. Radiat. Transf., 121, 78

Line, M. R., Liang, M. C., & Yung, Y. L. 2010, ApJ, 717, 496

Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, ApJ, 775, 137

Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, AJ, 152, 203

Lines, S., Mayne, N. J., Boutle, I. A., et al. 2018, A&A, 615, A97

Lodders, K. 2003, ApJ, 591, 1220

Lupu, R. E., Zahnle, K., Marley, M. S., et al. 2014, ApJ, 784, 27

Madhusudhan, N. 2018, Atmospheric Retrieval of Exoplanets (Cham: Springer International Publishing), 2153–2182

Madhusudhan, N. 2019, ARA&A, 57, 617

Madhusudhan, N., Knutson, H., Fortney, J. J., & Barman, T. 2014, Protostars and Planets VI, 739

Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24

Mant, B. P., Yachmenev, A., Tennyson, J., & Yurchenko, S. N. 2018, MNRAS, 478, 3220

Marley, M. S., & Robinson, T. D. 2015, ARA&A, 53, 279

McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, ApJ, 791, 55

McKemmish, L. K., Masseron, T., Hoeijmakers, H. J., et al. 2019, MNRAS, 488, 2836

McKemmish, L. K., Yurchenko, S. N., & Tennyson, J. 2016, MNRAS, 463, 771

Mishchenko, M. I., Travis, L. D., & Lacis, A. A. 2002, Scattering, absorption, and emission of light by small particles

Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67

Morello, G., Waldmann, I. P., Tinetti, G., et al. 2014, ApJ, 786, 22

Morley, C. V., Knutson, H., Line, M., et al. 2017, AJ, 153, 86

Moses, J. I., Madhusudhan, N., Visscher, C., & Freedman, R. S. 2013, ApJ, 763, 25

Moses, J. I., Visscher, C., Fortney, J. J., et al. 2011, ApJ, 737, 15

Ngo, N. H., Tran, H., & Gamache, R. R. 2012, J. Chem. Phys., 136, 154310

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJ, 743, L16

Ohno, K., & Kawashima, Y. 2020, ApJ, 895, L47

Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, MNRAS, 482, 1485

Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, MNRAS, 480, 2597

Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, MNRAS, 385, 109

Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, MNRAS, 432, 2917

Rackham, B. V., Apai, D., & Giampapa, M. S. 2019, AJ, 157, 96

Reuter, D. C., & Sirota, J. M. 1993, J. Quant. Spec. Radiat. Transf., 50, 477

Rey, M., Nikitin, A. V., & Tyuterev, V. G. 2017, ApJ, 847, 105

Richard, C., Gordon, I. E., Rothman, L. S., et al. 2012, J. Quant. Spec. Radiat. Transf., 113, 1276

Rodler, F., Kürster, M., & Barnes, J. R. 2013, in European Physical Journal Web of Conferences, Vol. 47, European Physical Journal Web of Conferences, 11003

Sharp, C. M., & Burrows, A. 2007, ApJS, 168, 140

Showman, A. P., & Polvani, L. M. 2011, ApJ, 738, 71

Shulyak, D., Rengel, M., Reiners, A., Seemann, U., & Yan, F. 2019, A&A, 629, A109

Sing, D. K., Désert, J. M., Lecavelier Des Etangs, A., et al. 2009, A&A, 505, 891

Sing, D. K., Pont, F., Aigrain, S., et al. 2011, MNRAS, 416, 1443

Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049

Sousa-Silva, C., Al-Refaie, A. F., Tennyson, J., & Yurchenko, S. N. 2014, MNRAS, 446, 2337

Southworth, J. 2010, MNRAS, 408, 1689

Spake, J. J., Sing, D. K., Wakeford, H. R., et al. 2019, arXiv e-prints, arXiv:1911.08859

Speagle, J. S. 2020, MNRAS, 493, 3132

Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, AJ, 153, 136

Steinrueck, M. E., Parmentier, V., Showman, A. P., Lothringer, J. D., & Lupu, R. E. 2019, ApJ, 880, 14

Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, Nature, 464, 1161

Swain, M. R., Vasisht, G., & Tinetti, G. 2008, Nature, 452, 329

Swain, M. R., Vasisht, G., Tinetti, G., et al. 2009, ApJ, 690, L114

Tennyson, J., & Yurchenko, S. 2018, Atoms, 6, 26

Teske, J. K., Cunha, K., Smith, V. V., Schuler, S. C., & Griffith, C. A. 2014, ApJ, 788, 39

Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, Nature, 448, 169

Todorov, K. O., Deming, D., Burrows, A., & Grillmair, C. J. 2014, ApJ, 796, 100

Underwood, D. S., Tennyson, J., Yurchenko, S. N., et al. 2016, MNRAS, 459, 3890

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Venot, O., Hébrard, E., Agúndez, M., et al. 2012, A&A, 546, A43

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Waldmann, I. P., Rocchetto, M., Tinetti, G., et al. 2015, ApJ, 813, 13

Woitke, P., Helling, C., Hunter, G. H., et al. 2018, A&A, 614, A1

Wong, A., Yurchenko, S. N., Bernath, P., et al. 2017, MNRAS, 470, 882

Yurchenko, S. N., Al-Refaie, A. F., & Tennyson, J. 2018, A&A, 614, A131

Yurchenko, S. N., Bond, W., Gorman, M. N., et al. 2018, MNRAS, 478, 270

Yurchenko, S. N., Sinden, F., Lodi, L., et al. 2017, MNRAS, 473, 5324

Yurchenko, S. N., & Tennyson, J. 2014, MNRAS, 440, 1649

Zahnle, K., Marley, M. S., & Fortney, J. J. 2009, arXiv e-prints, arXiv:0911.0728

Zhang, H., Nakajima, T., Shi, G., Suzuki, T., & Imasu, R. 2003, Journal of Geophysical Research: Atmospheres, 108

Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2019, PASP, 131, 034501

Chapter 5

NO ESCAPING HELIUM FROM 55 CNC E

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, Antonija Oklopcic, and Renyu Hu. No Escaping Helium from 55 Cnc e. *AJ*, 161(4):181, April 2021. doi: 10.3847/1538-3881/abe382. M.Z. conceived the project, wrote the observing proposal, obtained the data, carried out the data analysis and the 1D simulations, and wrote the paper.

5.1 Abstract

We search for escaping helium from the hot super Earth 55 Cnc e by taking high-resolution spectra of the 1083 nm line during two transits using Keck/NIRSPEC. We detect no helium absorption down to a 90% upper limit of 250 ppm in excess absorption or 0.27 mÅ in equivalent width. This corresponds to a mass loss rate of less than $\sim 10^9$ g/s assuming a Parker wind model with a plausible exosphere temperature of 5000-6000 K, although the precise constraint is heavily dependent on model assumptions. We consider both hydrogen- and helium-dominated atmospheric compositions, and find similar bounds on the mass loss rate in both scenarios. Our hydrodynamical models indicate that if a lightweight atmosphere exists on 55 Cnc e, our observations would have easily detected it. Together with the non-detection of Lyman α absorption by Ehrenreich et al. (2012), our helium non-detection indicates that 55 Cnc e either never accreted a primordial atmosphere in the first place, or lost its primordial atmosphere shortly after the dissipation of the gas disk.

5.2 Introduction

The observed radius distribution of sub-Neptune-sized planets is bimodal, with peaks at $< 1.5\ R_\oplus$ and 2-3 $R_\oplus$ (Fulton et al., 2017). This bimodality can be explained if the observed population of sub-Neptune-sized planets formed with several $M_\oplus$ rocky cores and hydrogen-rich atmospheres, which were then stripped away from the most highly irradiated planets (i.e. Lopez & Fortney 2013; Ginzburg et al. 2018; see Owen 2019 for a literature review). The high inferred core densities in these models argue strongly for formation inside the ice line, and the semi-major axis distribution of this population of planets is well-matched by in situ formation models (Lee & Chiang, 2017). While the prevailing evidence at the moment appears

to favor in situ (or at least nearby) formation, the arguments proposed to date are by no means definitive as they rely on indirect model-based inferences. If the mass loss models used to simulated the observed radius distribution are incomplete or rely on incorrect assumptions, our conclusions about these planets may be incorrect as well. It is therefore important to have observational data to nail down theoretical models.

In this study, we focus on one of the most observationally favorable transiting super-Earths currently known: 55 Cnc e. 55 Cnc is a binary system with a K0 main sequence star and a M dwarf companion, separated by 1000 AU. The primary is a bright (V=5.95) star whose activity and rotation rate indicate it is very old, probably around 10 Gyr. It has at least 5 planets of various sizes and orbital distances, including e, a superheated $R = 1.88 \pm 0.03 R_\oplus$ (Bourrier et al., 2018) super-Earth with a period of 0.74 days and the only planet known to transit. The mean density of e as derived from radial velocities and transit depths is 6.7 g/cm^3, suggesting a rocky interior with an atmosphere contributing up to a few percent of the planet radius (Bourrier et al., 2018), although it is also consistent with a small iron core with a silicate mantle and no water or gas layer. At first glance 55 Cnc e seems unlikely to host a primordial hydrogen and helium rich atmosphere as it has a relatively high equilibrium temperature and correspondingly high predicted escape rate (Valencia et al., 2010). However, there are large theoretical uncertainties in mass loss models, driven in part by the uncertain X-ray and extreme ultraviolet spectrum of the star (Owen, 2019). It has also been suggested that this planet might have a lava ocean on its dayside that could outgas enough trapped hydrogen to form a thick secondary atmosphere (Chachan & Stevenson, 2018).

Current observational constraints on 55 Cnc e's atmosphere are also conflicting. Some observations, such as the non-detection of hydrogen Lyman α absorption by Ehrenreich et al. (2012) or the variations in infrared emission reported by Demory et al. (2016a), seem to indicate that it is unlikely to host a substantial hydrogen-rich atmosphere. Other observations suggest the opposite, including the planet's small measured day-night temperature gradient (Demory et al., 2016b), the tentative evidence for sodium and calcium absorption (Ridden-Harper et al., 2016), and the detection of a strong absorption feature in the planet's 1.1–1.7 μm transmission spectrum from *Hubble Space Telescope* (*HST*) WFC3, which has been attributed to HCN (Tsiaras et al., 2016). Regardless of its source, the large amplitude of the WFC3 absorption feature can only be matched by a hydrogen- or helium-dominated

atmosphere, because an atmosphere dominated by heavier elements would have a much smaller scale height and correspondingly weaker absorption features during transit. If 55 Cnc e does have a significant low mean molecular weight atmosphere, its ability to retain this atmosphere in the face of ongoing mass loss would place strict constraints on the magnitude of relevant mass loss processes, with correspondingly wide-reaching implications for our understanding of the overall population of short-period super-Earths and sub-Neptunes.

In this paper, we use the helium 1083 nm metastable triplet (Oklopčić & Hirata, 2018) to search for evidence of helium outflow from 55 Cnc e. Unlike previous Lyman α observations, this helium triplet is easily accessible using ground-based telescopes, and has been used to detect the extended atmospheres of multiple exoplanets. Most detections have been around Jupiter radius planets such as WASP-107b (i.e. Spake et al. 2018), with the 4 $R_\oplus$ GJ 3470b being the smallest planet with a detection to date (Ninan et al., 2020; Palle, E. et al., 2020). For 55 Cnc e, our observations of the 1083 nm helium line supplements the Ehrenreich et al. (2012) Lyman alpha observations in shaping our understanding of the exosphere. The Ehrenreich et al. (2012) observations took place at a stellar activity minimum, when the star's X-ray flux was 2–3 times lower than during our observations (see Section 5.7); this decreased X-ray flux might have suppressed the mass loss rate during these observations. In addition, some studies have proposed that mass loss over many Gyr can preferentially remove hydrogen from the atmosphere of a small planet while leaving helium behind (Hu et al., 2015; Malsky & Rogers, 2020). A helium dominated atmosphere would be consistent with the molecular weight inferred from the *HST* transit spectrum (Tsiaras et al., 2016). Such an atmosphere might be undetectable in Lyman α, but easy to detect using metastable helium. Finally, because Lyman α is such a strong line, the observed signal is dependent on the behavior of the diffuse exosphere far from the planet. The atmospheric absorption signal in the helium 1083 nm triplet is typically much weaker than the absorption signal in the Lyman-α line, but IR measurements are significantly more precise than UV measurements due to the much higher photon flux and (unlike Lyman α) we are able to observe the cores of the lines. This means that helium observations are sensitive to gas at smaller radii and with lower outflow velocities than Lyman α observations, making it easier to compare to mass loss models. It is for these three reasons—the differing stellar XUV irradiation, the possibility of a helium-dominated atmosphere, and the complementary sensitivities—that observations of the helium line are meaningful even for small planets with existing Lyman alpha non-detections.

We describe our observations in Section 2, data reduction pipeline in Section 3, analysis in Section 4, outflow models in Section 5, and scientific implications in Section 6.

5.3 Observations

We observed two transits of 55 Cnc e in Y band using the upgraded NIRSPEC instrument on Keck (Martin et al., 2018): one on December 4, 2019 at 13:56 UTC (barycentric), and one on December 18, 2019 at 13:48 UTC (barycentric). These observations used the 0.288 x 12 arcsec slit, giving NIRSPEC a resolution of 37,500, with a FWHM sampling of 3 pixels. All observations were performed with 60 second exposure times in an ABBA nod pattern to facilitate background subtraction. Details of the observations are given in Table 5.1.

Table 5.1: Keck/NIRSPEC observations

Parameter	12/04/19 (UTC)	12/18/19 (UTC)
Obs. duration	3.2 h	3.7 h
Number of obs	107	136
Obs before transit	23	40
Obs during transit	51	56
Obs after transit	33	40
Airmass before transit	1.1	1.04
Airmass during transit	1.02	1.01
Airmass after transit	1.03	1.11
Obs efficiency	56%	61%
Avg. SNR*	660	640

Average SNR is calculated per spectral pixel per exposure, based on the estimated error from optimal extraction.

Observation conditions were stable with good seeing during the first night. The seeing was worse and more variable during the second night, decreasing the per-pixel count rate on the detector and making it possible to use a more efficient observing strategy of 20 second subexposures with 3 co-adds instead of 15 second subexposures with 4 co-adds. The marginally increased observing efficiency could not fully compensate for the higher seeing, leading the second night to have a marginally worse SNR per spectrum.

For unknown reasons and at unpredictable times, the telescope would fail to nod. We encountered this issue on each of our four half-nights with NIRSPEC—two in April 2019, and two in December 2019. Sometimes it would not move at all; at other times, it would jump out of the slit, either by nodding too far, or by adding

a perpendicular component to the nod. When this happened, we would wait for the exposure to finish (NIRSPEC does not allow us to stop during the middle of an exposure), re-center the star, and re-start the nod sequence. We then discard the exposure, together with the previous exposure if it is necessary for background subtraction. On 12/04, nodding failed once during the transit and we lost 4 exposures (because we did not notice the failure immediately), corresponding to 7 minutes of observation time. On 12/18, nodding failed twice during the transit and once after the transit, resulting in a loss of 6 exposures. Other than these nodding failures, our observing sequence is continuous on both nights.

5.4 Data reduction

We calibrated the raw images and extracted 1D spectra for each order using a custom Python pipeline designed for the upgraded NIRSPEC. We describe each step of this process below.

Crosstalk removal

In its current configuration, NIRSPEC is divided into 32 readout channels, with 64 rows per channel. The rows of adjacent channels are read simultaneously, but in reverse order: row 0 of channel 0 should be read at the same time as row 63 of channel 1. The simultaneous readout causes crosstalk signals between each pair of channels. We follow the approach described in George et al. (2020) and initially assume that the crosstalk signal $X_{k,j,i}$ in NIRSPEC is proportional to the derivative of the signal in the source channel:

$$X_{k,j,i} = a_{k,j}(S_{k,i} - S_{k,i-1}), \tag{5.1}$$

where channel k is the source of the crosstalk, channel j is the destination, i is the row number, and a is a 32×32 matrix of scaling factors whose elements are of the order 400 ppm. The crosstalk introduced is therefore on the order a few hundred electrons for signals on the order of 1,000,000 electrons, but varies widely for different k, j pairs.

Unfortunately, there are no calibration data for NIRSPEC that can help to characterize the crosstalk, as we require a high flux derivative along the spatial axis. We therefore use archival NIRSPEC observations of 55 Cnc in L band from April 2 and 16, 2019, to measure the crosstalk matrix a. On each night, we took ~50

minute-long exposures, each with very high flux (~300,000 electrons per spectral pixel). To measure the crosstalk matrix a, we masked out pixels with a count greater than 70 ADU (200 electrons) and used a linear least squares fit to estimate a and the uncertainty on a. The uncertainty is high where the source channel is un-illuminated and/or the target channel has few un-illuminated pixels, and low where the source channel is illuminated but the target channel is not. We perform a weighted average over the estimates of a from each exposure on each night to obtain our best estimate for the matrix.

Relying on science data to measure the crosstalk matrix is less than ideal because channels that have no trace do not produce a measurable crosstalk signal in other channels; the trace dominates the crosstalk from other channels in channels where the trace is nearly horizontal and close to the middle. Nevertheless, a visual inspection indicates that this imperfect approach to crosstalk subtraction is sufficient to remove the crosstalk signal from our data (Figure 5.1).

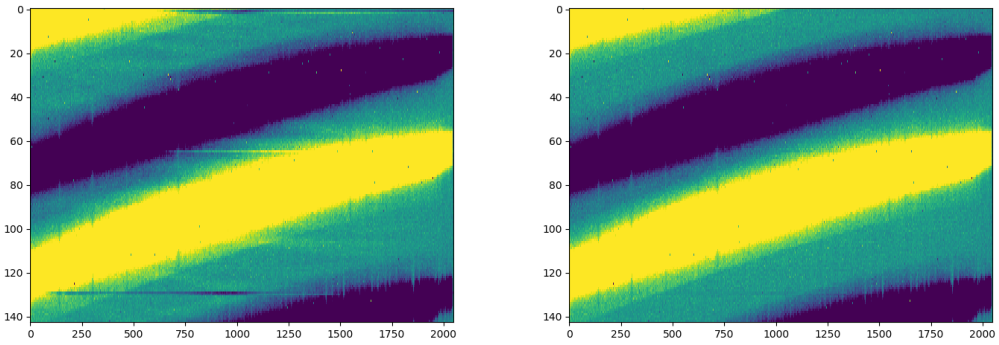


Figure 5.1: A portion of a raw A-B frame, showing the order containing the helium line. Top: no calibration corrections. Bottom: crosstalk subtracted. Notice that the subtle horizontal ripple pattern is gone, as are the two sharp horizontal lines.

We find that for NIRSPEC, Equation 5.1 is less accurate when j and i are of different parity. By carefully examining the crosstalk patterns, we found that the crosstalk signal is proportional not to $S_{k,i} - S_{k,i-1}$, but to $S_{k,i+2} - S_{k,i+1}$. In physical terms, this means that the 2nd rows of odd channels (not the 0th rows) are read at the same time as the 63rd rows of even channels. Similarly, the 1st row corresponds to the 62nd, the 2nd to the 61st, and so on, until the 63rd row of odd channels are read simultaneously with the 2nd rows of even channels. Therefore, rows 0 and 1 of every channel are being read out when their sister channels of opposite parity are not being read out.

We believe that this mismatch in the readout pattern causes another detector artifact: the anomalous rows, seen in Figure 5.1 as horizontal lines. Anomalous rows occur when the row number modulo 128 is equal to 0, 1, or 64. The first two correspond to rows 0 and 1 of even channels; the last corresponds to row 0 of odd channels. We speculate that the anomalous rows could be caused by crosstalk: when row 0 or 1 is being read out, the corresponding readout line in opposite parity channels could be carrying signals of much higher amplitude than image data, causing much higher crosstalk than normal.

In the absence of photons, the anomalous rows are mostly, but not entirely, consistent across the entire image. Thus, all 32 rows divisible by 128 are similar; all rows whose remainder is 1 when divided by 128 are similar; and all rows whose remainder is 64 when divided by 128 are similar. We create a template of these 3 different categories of anomalous rows by identifying regions of the detector which received few photons, either because they are in between orders, or because they are within an order but far from the trace. At each column, the template is equal to the median of the rows fitting this criterion. The template is subtracted from all relevant rows. This method works very well, but not perfectly—a faint hint of under-subtraction can be seen in Figure 5.1 (bottom) for row 128.

Image calibration

On the first (second) night, we took 80 (63) flat fields, each with an exposure time of 4.4 s per co-add and 20 co-adds. In total, we collected 750 million (590 million) electrons per pixel, ensuring that the photon error in the flat fields is far below the photon noise in the observational data. To calibrate the flats, we took 19 (3) darks, each with an exposure time of 4.4 s per co-add and 20 co-adds. We create a master dark by median stacking the individual dark frames. In each individual dark, pixels that deviate from the image-wide mean by more than 5σ are marked as bad pixels. Pixels marked as bad in more than half of the individual darks are marked as bad in the master dark.

We create a master flat by taking the median of the individual flats. Prior to combining, we subtract the crosstalk from each flat frame using the algorithm described in the previous subsection, subtract the master dark, and divide by the median flux. We identify the order containing the helium line and mask out everything else. as there is no need to extract spectra from other orders. The relevant order is fitted with a polynomial that is 5th order with in x and 3rd order in y, then divided by the

polynomial, in order to bring the values of all good pixels close to 1. We then create a mask of bad pixels, including both the bad pixels identified in the master dark, and the pixels with a flat value lower than 0.5 or higher than 1.5. The final master flat, together with the bad pixel mask, are saved in a FITS file.

Finally, we correct the raw science frames using the master flat. For each $A_1B_1B_2A_2$ nod, we create four difference images: $A_1 - B_1$, $B_1 - A_1$, $B_2 - A_2$, and $A_2 - B_2$. Each difference image is crosstalk subtracted, multiplied by the gain of g=2.85 e/ADU, and divided by the master flat F. We next construct a variance image for each differenced frame, which indicates the uncertainty in the measured flux at each individual pixel location. This variance image includes the photon noise from the star, the sky background, and the detector read noise. The background b, read noise N_R, and total variance V are computed as follows:

$$b = g\frac{A + B - |A - B|}{F} \tag{5.2}$$

$$N_R = \sqrt{2}\sqrt{\frac{N_{coadds}}{N_{reads}}}N_{R,0} \tag{5.3}$$

$$V = g\frac{|A| + |B|}{F^2} + N_R^2, \tag{5.4}$$

where $N_{R,0} = 44e^-$, $N_{reads} = 4$ for all of our observations, $g = 2.85$ e^-/ADU is the gain, and $N_{coadds} = 4$ on our first night and $N_{coadds} = 3$ on most of our second night. The iterative bad pixel algorithm in REDSPEC 3.0[1] (Kim et al., 2015), which we ported to Python, is used to identify and repair bad pixels. This algorithm identifies bad pixels with a variant of local sigma clipping. In total, it typically identifies and repairs 1000 hot pixels and 1000 cold pixels in a given image. Since the identification is limited to the ~100 pixel tall order containing the helium line, 2000 bad pixels represents 1% of all pixels. The bad pixels identified by the algorithm are combined with the bad pixels identified in the master dark and master flat to create a master bad pixel mask. The difference image, variance image, background image, and bad pixel mask are all saved in a FITS file.

Optimal extraction

After calibrating the images, we extract the 1D spectrum from each 2D spectral trace. We first determine the position of the trace in each image. For every column

[1] https://www2.keck.hawaii.edu/inst/nirspec/redspec.html

(corresponding to one wavelength), we fit a Gaussian to the pixel values to estimate the trace position We then fit a 5th order polynomial to the trace locations as a function of column number. The residuals in this final fit are typically smaller than 0.01 pixels. Accuracy is not paramount because we only use the trace to identify regions of the image very far from the trace, in order to mask them out.

After determining the position of the trace, we perform optimal extraction using a variant of the method described in Horne (1986). The original Horne (1986) algorithm assumes that the wavelength axis is aligned with the columns, and the spatial axis is aligned with the rows. Unfortunately, on NIRSPEC neither the wavelength nor spatial axes are aligned with either the rows or the columns. The axes are also not perpendicular to each other, and neither axis is straight, as can be seen in Figure 5.1. The typical way around this problem is to rectify the order spatially and spectrally by interpolation onto a rectilinear grid. This works well enough for low SNR data, but for our exceptionally high SNR data it introduces aliasing artifacts at the 0.1–1% level.

We instead use a rolling window approach where we assume that each column corresponds to one wavelength—a reasonable assumption, given that most of the flux is concentrated within 5 pixels of the center of the trace. Regions more than 15 pixels from the center of the trace are masked out to avoid interference from neighboring orders or cosmic rays. For each column, we take a window 81 pixels wide and 72 pixels tall, and fit the profile for every row. We chose 81 pixels as the width because we want the window to be small enough for the trace to deviate vertically by less than one pixel, but big enough to fit the profile accurately. We chose a 72 pixel window in the y (cross-dispersion) direction because the trace is 50 pixels lower on the left end of the detector than on the right, and we needed a window large enough to encompass not just the center of the trace, but also the wings across the full width of the image.

In Horne (1986), the spatial profile is estimated by dividing each column by its sum. This runs into problems in regions of high telluric absorption, where the spectral flux approaches zero. We therefore fit the product of the spectrum and a Chebyshev polynomial model of the profile directly to the observed data. Mathematically, if i represents the order of the Chebyshev polynomial T_i and j is the column number, we look for the x that minimizes $Ax = b$ where A and b are:

$$A_{i,j} = S_j T_i(\frac{j - N/2}{N})/\sigma_j \tag{5.5}$$

$$b_j = I_j/\sigma_j, \tag{5.6}$$

$N = 72$ is the size of the window and $\frac{j-N/2}{N}$ normalizes the column numbers to range from -1 to +1. After obtaining the least squares solution to x, the profile can be computed as:

$$P_j = \sum_{i=0}^{5} x_i T_i(\frac{j - N/2}{N}), \tag{5.7}$$

where 5 is the the maximum order of Chebyshev polynomials we fit.

We note that this is mathematically equivalent to fitting monomials. If we had replaced $T_i(y)$ with y^i in all the equations above, we would arrive at an identical P_j. The advantage of Chebshev polynomials comes from the numerical stability and robustness of the linear algebra solver. y^i becomes extremely small when i is large for all but the boundary values -1 and 1, causing numerical problems, whereas the Chebshev polynomials are well behaved for every order. The Chebshev polynomials are also mutually orthogonal; the monomials are not. In practice, it appears that neither of these matter for our low order profile fit, but since Chebysheb fitting is theoretically superior, we adopt it as our preferred form of polynomial fitting throughout our code.

We make several additional modifications to the Horne (1986) algorithm. The background cannot be estimated from the difference image alone, so we instead use the estimate from Equation 5.2. Instead of identifying bad pixels solely from the image using sigma clipping, we start with the bad pixel mask from the calibration stage (Subsection 5.4) and identify additional bad pixels using the image. Finally, we use an extremely high threshold $\sigma_{clip} = 12$ to reject bad pixels during the optimal extraction iterations. We require this high threshold because of a phenomenon that we observe with our high signal-to-noise data: the four-leaf clover pattern in the residuals image, shown in Figure 5.2. This pattern occurs wherever the spectrum is changing rapidly—namely, in the vicinity of deep stellar or telluric lines. The pattern traces the cross derivative $\frac{dF}{dxdy}$, and is more pronounced where the trace is highly tilted with respect to the horizontal. This clover pattern indicates that optimal extraction does not give optimal results in the vicinity of deep lines. Regions without

deep lines—including the vicinity of the 1.08333 μm helium line—are minimally affected.

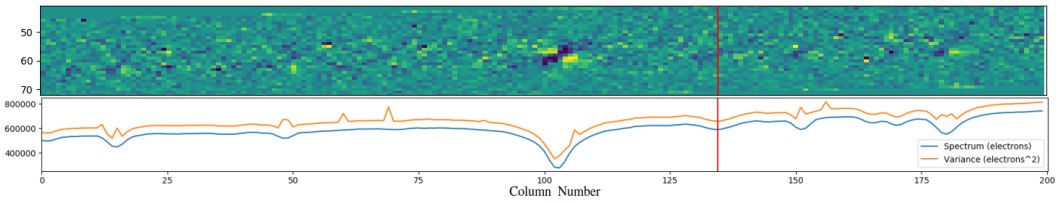


Figure 5.2: Optimal extraction results. Top: the residuals of optimal extraction, scaled to the standard deviation of each pixel and clipped from -5 to 5σ. If optimal extraction worked perfectly, this image would consist of Gaussian-distributed noise with a mean of 0 and a standard deviation of 1. The four-leaf clover pattern in the middle of the residuals image corresponds to a deep stellar line at 1.0830 um. The position of the helium 1.08333 um line is marked in red. Although a stellar helium line is present it is much shallower than the 1.0830 um line, and there does not appear to be any strong residual structure in this region. Bottom: optimally extracted spectrum with accompanying variance. The spikes in variance are due to masks on bad pixels.

We initially thought that the clover pattern was due to the misalignment between the image axes and the wavelength/spatial axes, or due to the non-orthogonality between the wavelength and spatial axes. We tried many different optimal estimation schemes that do not make the rectilinear assumptions of Horne (1986), including Marsh (1989) and an unpublished algorithm[2], in an attempt to get rid of the clovers. However, we eventually realized that the clover pattern is caused by the inseparability of the 2D PSF. The optimal extraction algorithm assumes that $PSF_{\lambda,x} = S(\lambda)P(x)$ for some spectrum $S(\lambda)$ and profile $P(x)$. A 2D Gaussian with its axes aligned with the wavelength and spatial axes is separable, but a tilted 2D Gaussian is not. Bolton & Schlegel (2010) simulate spectra under the assumption of PSF non-separability, apply optimal extraction to the simulated spectra, and obtain clover patterns strikingly similar to ours (see their Figure 1, right). The non-separability of the PSF means that only algorithms like spectro-perfectionism (Bolton & Schlegel, 2010) that adopt a 2D PSF can avoid the clovers. Spectro-perfectionism is not necessary for our work here because the helium line is not in a region where the spectrum varies rapidly.

[2]`https://code.obs.carnegiescience.edu/Algorithms/ghlb/at_download/file`

Wavelength solution

After extracting the 1D spectra, we determine the wavelength solution for each spectrum. We first create a template containing stellar and telluric lines at known wavelengths, with the stellar lines shifted to account for Earth's velocity relative to the star on that night. We adopted a $T_{eff} = 5200K$, $log(g) = 4.5$, and [M/H]=0.0 PHOENIX model spectrum (Husser et al., 2013) for the star. We used a telluric transmission spectrum from the Gemini website[3], which assumes a precipitable water vapor of 1.6 mm and an airmass of 1.5. This spectrum was calculated using ATRAN (Lord, 1992). We multiply the stellar spectrum by the telluric transmission and downsample to instrumental resolution (R=37,500) to get our final template.

We parameterize the wavelength solution as a third-order polynomial function of the normalized column number, $j' = \frac{j-N/2}{N}$:

$$\lambda(j') = \sum_{i=0}^{3} c_i T_i(j'), \tag{5.8}$$

where T_i are the Chebyshev polynomials. Similarly, we parameterize the continuum as a fifth-order polynomial of j', and multiply it by the template. We then use `scipy`'s differential evolution minimizer to minimize χ^2, which we calculate as the difference between the observations and the continuum-adjusted template interpolated using the proposed wavelength solution coefficients. A visual inspection of the resulting fit indicates that the line positions match to better than a pixel.

5.5 Analysis

After extracting the 1D spectra, we place the data from each night on a uniform wavelength grid and remove signals not related to the planet. This includes instrumental effects like detector fringing, telluric absorption lines, and absorption lines from the star itself. We then shift each spectrum into the planetary rest frame and quantify the amount of excess absorption in the helium triplet during transit. We describe each step of this process in detail below.

Better ephemeris

Due to the high radial acceleration of the planet during transit, an accurate ephemeris is necessary to shift spectra into the planet frame. We calculate an updated ephemeris by combining the epoch derived by Demory et al. (2016a) using 4 *Spitzer* 4.5 μm

[3]https://www.gemini.edu/observing/telescopes-and-sites/sites

transits, the epoch derived by Sulis et al. (2019) using 143 MOST transits, and 35 individual transit timings from TESS. We obtain the TESS transit timings using a procedure similar to that described in Dai et al. (2019). In short, we downloaded the *TESS* photometry from the Mikulski Archive for Space Telescopes (MAST). Using the archival ephemeris, we isolated data within a wide window of three times the archival transit duration. We fitted each individual transit with the `BATMAN` (Kreidberg, 2015) package and a quadratic function of time to account for local stellar variability. We adopted quadratic limb darkening and imposed Gaussian priors with widths of 0.3 centered around the theoretical values from `EXOFAST`[4] (Eastman et al., 2013). We then fitted all *TESS* transits globally after removing the stellar variability component. With this global model as a template, we revisit each individual transits allowing only the mid-transit time and local quadratic function to vary. This process is iterated a few times until convergence. Finally, the individual *TESS* transit epochs are fitted together with the archival transit times. The updated ephemeris is $P = 0.73654604 \pm 1.6 \times 10^{-7}$d and $T_0 = 2458723.38328 \pm 0.00014$ ($\mathrm{BJD_{TDB}}$) with negligible covariance C=-10^{-14}d. The covariance C is defined such that the prediction error at epoch E from this ephemeris is:

$$\sigma_T^2 = \sigma_{T_0}^2 + E^2\sigma_P^2 + 2CE. \tag{5.9}$$

We chose the initial epoch T_0 to make C as close to 0 as possible, allowing the last term to be ignored. The prediction error evaluates to 13 seconds at the time of our December 2019 observations.

Making a spectral grid

The first step is to linearly interpolate all spectra on a given night onto a common wavelength grid. We choose a wavelength grid spanning 1.0826 to 1.0840 μm, which encompasses the locations of the three lines in the helium triplet. We select a resolution of 110,000 for our grid, approximately matching the native pixel resolution. The process of interpolation introduces covariances between adjacent wavelength bins, artificially smoothing out the interpolated spectrum. We keep track of the covariance matrix during our analysis, and eventually use it to calculate likelihoods.

[4] `astroutils.astronomy.ohio-state.edu/exofast/limbdark.shtml`.

We do an initial check for helium absorption during the transit by dividing all spectra from both nights into two categories, in-transit and out-of-transit. We take the mean of the in-transit spectra to create a master in-transit spectrum, and the mean of the out-of-transit spectra to create a master out-of-transit spectrum. We then calculate the excess absorption in the stellar frame as $F_{in}/F_{out} - 1$. The two master spectra and the excess absorption are plotted in Figure 5.3. There does not appear to be any detectable increase in absorption during the transit at the position of the helium line.

Residuals image

Since we are interested in fractional changes in the spectrum, we take the natural log of the spectral grid and subtract the mean of every row and column, producing what we call the residuals image. Every pixel in the residuals image approximately represents the fractional flux change at that epoch and wavelength from the mean spectrum. Taking the natural log also has the advantage of linearizing the effect of changing airmass. The observed flux is roughly $F(\lambda) = F_{vac}(\lambda)e^{-\alpha z}$ where z is the airmass, so $\ln F(\lambda) = \ln F_{vac}(\lambda) - \alpha z$.

We show the residuals image through various steps of the pipeline in Figure 5.4 (top). The most striking feature in the median subtracted image are the vertical bars caused by fringing, which are spaced 20 pixels apart. A fast Fourier transform (FFT) of each spectrum reveals a prominent peak at a frequency of 0.052 pixel^{-1}. To remove the fringing, we use `scipy` to apply a second order Infinite Impulse Response (IIR) notch digital filter with a frequency of 0.052 pixel^{-1} and a quality factor of 15. We apply the filter twice, which is sufficient to suppress the peak in the FFT without bringing the spectral power substantially below that of neighboring frequencies. The subsequent panel in Figure 5.4 shows the residuals image after fringing correction, demonstrating that the notch filter has effectively removed the fringing.

The most prominent features in the fringing-corrected residuals image (Figure 5.4) are the strongly variable columns. These are deep telluric and stellar lines that vary for a variety of reasons, including the suboptimal performance of optimal extraction in deep lines (discussed in Figure 5.4), the inaccuracy of interpolation across deep lines, and the time-variable telluric absorption. The line immediately to the right of center in the residuals image, for example, is a water absorption line at 1.08351 μm. Its dependence on airmass is clear: as the night progresses, 55 Cnc first rises,

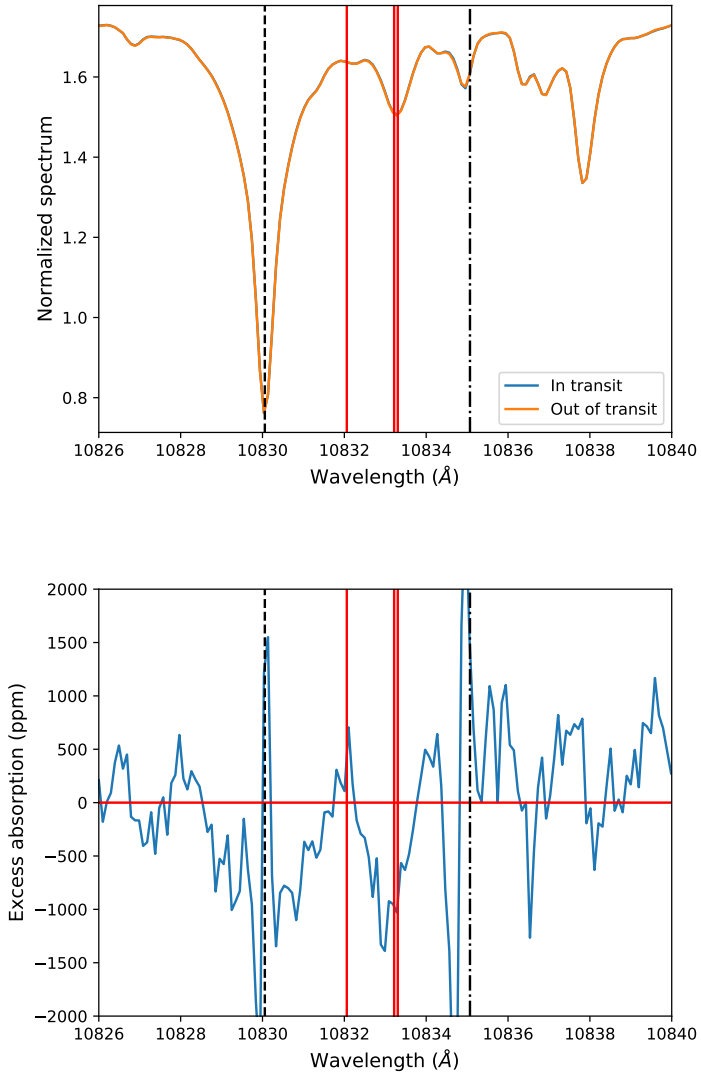


Figure 5.3: Excess absorption. Top: in-transit vs. out-of-transit spectra in the stellar rest frame, calculated using data from both nights. Bottom: excess absorption ($F_{in}/Fout - 1$), plotted in the stellar frame; fringing has been removed with a bandstop filter, but additional corrections for telluric and instrumental effects described in Subsection 5.5 have not been applied. The red vertical lines mark the locations of the three helium lines. The dashed vertical black lines represent strong stellar lines, while the dash-dotted vertical line is a telluric water line. The most prominent features in the excess absorption plot are the continuum variation, and the sudden spikes/dips at the position of the strong lines. The latter is due to the poor performance of optimal extraction in the vicinity of strong lines. Helium absorption would manifest as a spike in the vicinity of the red vertical lines, which is not seen.

then sets. The line is dark at the beginning, brightens until the airmass reaches its minimum, and dims again. The other strong line is a Si I line at 1.0830057 μm.

We correct for variability in telluric absorption using SYSREM (Mazeh et al., 2007). SYSREM is a generalization of Principal Component Analysis that takes into account the errors on the data. Like PCA, it identifies eigenvectors and eigenvalues which, when linearly combined, best explain the residuals image. When we apply SYSREM to our residuals image, the first component has eigenvalues that closely track the airmass, showing that the algorithm is successfully identifying telluric variability.

After subtracting the first principal component identified by SYSREM, we compute the standard deviation of each column and mask the most variable columns. This removes the prominent telluric line and all 4 of the most prominent stellar lines. One iteration of SYSREM is not sufficient to remove the continuum variation seen in Figure 5.4. We therefore remove it by fitting a third order polynomial with respect to column number for each row and subtracting the polynomial. The residuals image, averaged across both nights, is shown in the bottom-most panel of Figure 5.4. In Figure 5.5, we additionally show a plot of excess absorption in the stellar frame. There is no sign of helium absorption in either plot.

Shifting into the planetary rest frame

55 Cnc has an extremely high orbital speed of 230 km/s, causing it to accelerate by 130 km/s over the duration of the transit. At every epoch, we use the barycentric Julian date to compute the radial velocity of the planet relative to the star, shift the residual spectrum accordingly, and resample onto the common wavelength grid using linear interpolation. The covariances introduced by linear interpolation are properly computed and propagated. We then take the mean of all the shifted in-transit residual spectra to arrive at the final excess absorption spectrum. The average spectrum across both nights is taken to be the fiducial excess absorption spectrum.

The excess absorption from both nights, as well as the averaged excess absorption, is shown in Figure 5.6. The combined spectrum has a standard deviation of 146 ppm, smaller than the planet's white light transit depth of 350 ppm, although some amount of correlated noise is present. No sign of helium absorption can be seen.

The bottom panel of Figure 5.6 shows what the excess absorption spectrum would look like with a 1300 ppm helium absorption signature. The absorption profile was taken from the 2.5D hydrodynamic model with the weakest absorption (solid green curve in Figure 5.15). This figure shows that the reduction process subtracts out at most 20% of the signal, and that even the 2.5D hydrodynamic model with the weakest absorption is ruled out by our data.

Constraints on absorption

To quantify the constraint on excess absorption that our data provide, we used the nested sampling code `dynesty` (Speagle, 2020). We modelled the data as a scaled and vertically shifted version of the 2.5D hydrodynamic model with the weakest absorption (solid green curve in Figure 5.15, corresponding to a H-dominated atmosphere with 10^{-10} envelope fraction and 1 year dispersal timescale). The observed excess absorption spectrum (Fig. 5.6) is truncated to 10,831–10,835 Å to avoid the need to model any unsubtracted low-frequency variations in the data. Instead, the low-frequency variations are accounted for by vertically shifting the model to match the data.

We compute a covariance matrix for the excess absorption spectrum using the errors from the optimal extraction algorithm propagated forward through the pipeline. This matrix accounts for the covariances caused by the two linear interpolations—first onto a common wavelength grid, and second into the planetary rest frame. To account for unmodelled systematics, we multiply the covariance matrix by a free parameter e^2, where $e > 1$. The log likelihood is given by:

$$\ln L = -\frac{1}{2} r^T K^{-1} r - \frac{1}{2}|K| - \frac{N}{2}\ln 2\pi, \tag{5.10}$$

where K is the covariance matrix and r is the residuals.

Figure 5.7 shows the results of our nested sampling run. The free parameters are the vertical offset of the model, the maximum excess absorption (A), and the error multiple e whose square multiplies the covariance matrix. Our data is consistent with zero excess absorption, and the best-fit error inflation parameter e indicates that our errors are likely underestimated by approximately 25%. The preferred vertical offset is consistent with zero, as expected from visual inspection. Our fit places a 90% upper limit on A of 250 ppm, corresponding to an equivalent width of 0.27 mÅ. To see whether A is significantly non-zero, we performed a nested sampling run with the amplitude removed as a free parameter. The resulting log Bayesian evidence log(Z) was 1.3 higher in this run, indicating that the data prefers a model with no absorption.

The constraint on A can also be expressed in scale heights, where the scale height $H_{eq} = \frac{k_B T_{eq}}{\mu g}$ is computed assuming an equilibrium temperature of 2000 K and a hydrogen-dominated atmosphere. This metric was first proposed by Nortmann et al.

(2018). We obtain an upper limit on the increase in apparent planet radius at 1083 nm: $\delta_{Rp}/H_{eq} < 11$ with $H_{eq} = 350$ km, which we put into context in Figure 5.8 by plotting the EUV flux and δ_{Rp}/H_{eq} of all planets with helium detections or non-detections. The EUV flux experienced by 55 Cnc e is obtained from Salz et al. (2016), who calculate $F_{EUV} = 7.4$ W/m^2 by applying the scaling relation of Linsky et al. (2014) to the Lyman alpha luminosity. Following Kasper et al. (2020), we conservatively adopt 3x error bars on the EUV flux on either end, as different EUV reconstruction techniques give results that are discrepant by an order of magnitude in some cases (i.e. Salz et al. 2015b; Drake et al. 2020). Figure 5.8 shows that, relative to other planets with helium measurements, 55 Cnc e is a highly EUV-irradiated planet with a tight upper limit that is far below the δ_{Rp}/H_{eq} of successful detections for other planets.

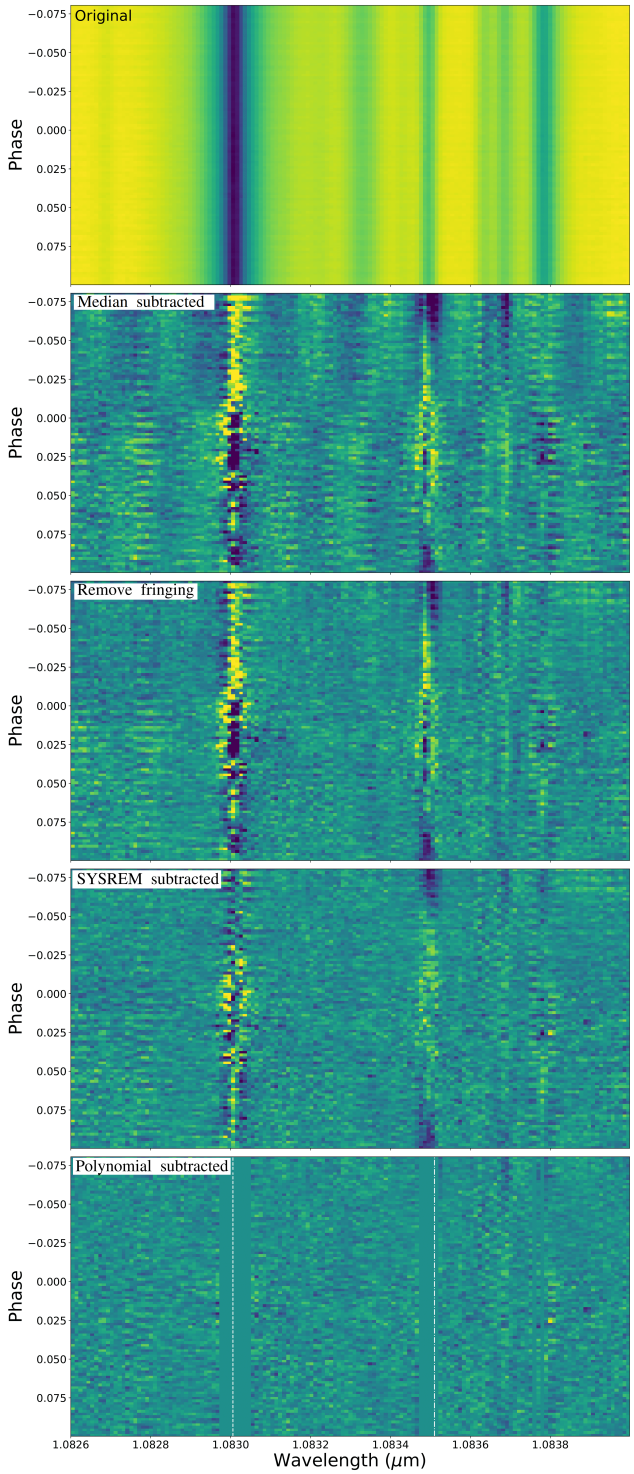


Figure 5.4: Pipeline steps: the original spectra; after removing fringing; after removing one SYSREM component; after masking variable lines and subtracting off continuum variations; after subtracting off continuum variations by fitting a polynomial to each spectrum. In the last panel, the vertical white lines mark a stellar Si line at 1.0830057 μm (left) and a strong water line at 1.08351 μm (right).

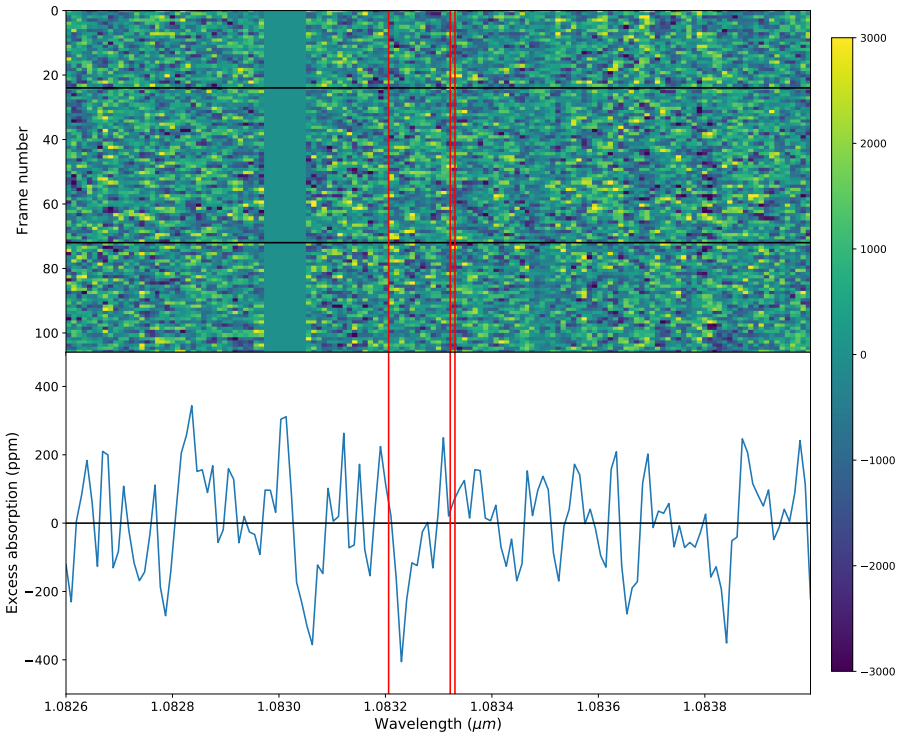


Figure 5.5: Excess absorption. Top: Combined residuals image for both nights, showing excess absorption (ppm) in the stellar rest frame after all pipeline steps except shifting into the planetary rest frame. The top and bottom horizontal black lines mark the beginning and end of transit, respectively. The vertical red lines mark the helium triplet positions. No helium absorption is evident. Bottom: excess absorption during transit ($1 - F_{in}/F_{out}$) computed from the final residuals image for both nights.

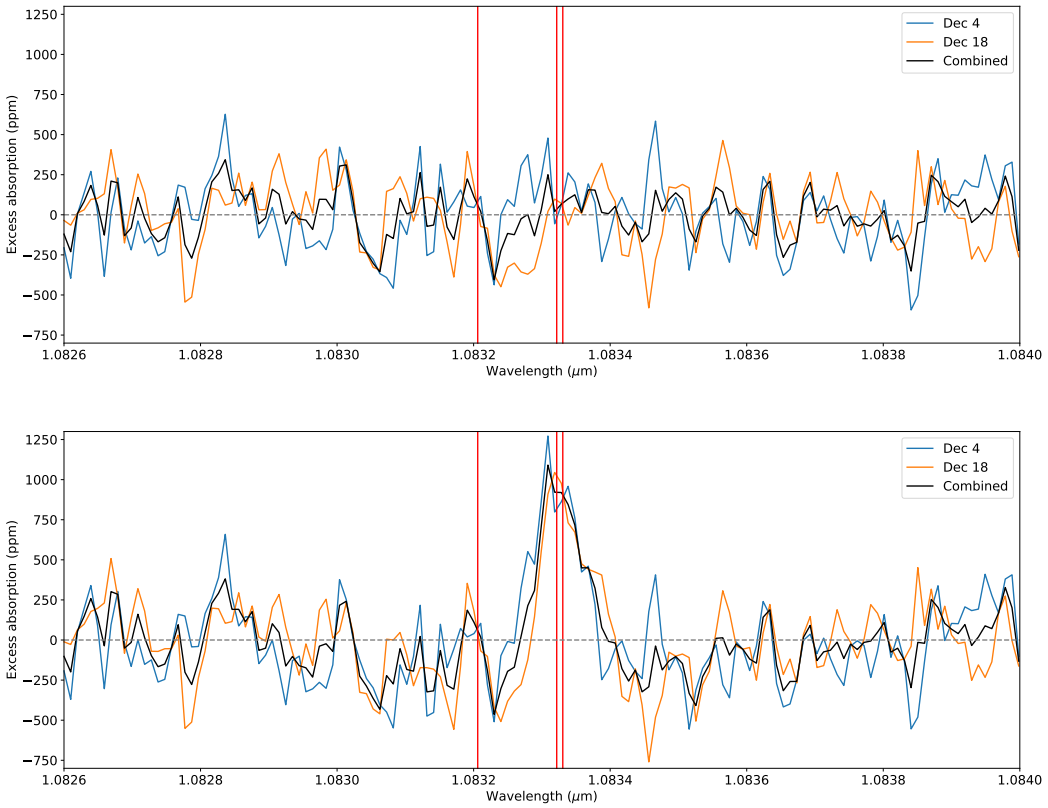


Figure 5.6: Excess absorption and injection test. Top: excess absorption during the transit in the planet frame as a function of wavelength, averaged over both nights. Helium absorption from the planet would result in a positive signal at the position of the red vertical lines. Due to the nature of ground-based high resolution spectra, the white light transit cannot be detected. Bottom: same as above, except an absorption signal was injected into the spectra before any reduction was undertaken. The injected signal has a maximum excess absorption of 1300 ppm, corresponding to the 2.5D hydrodynamic model with the weakest absorption (dashed blue curve in Figure 5.15).

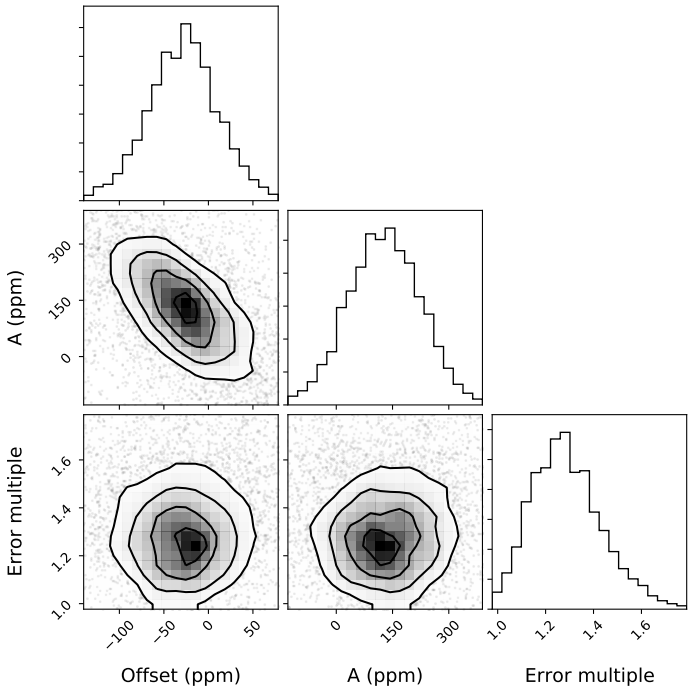


Figure 5.7: 2D posteriors from our fit to the data, constraining the maximum excess absorption (A). A is 1.3σ from 0.

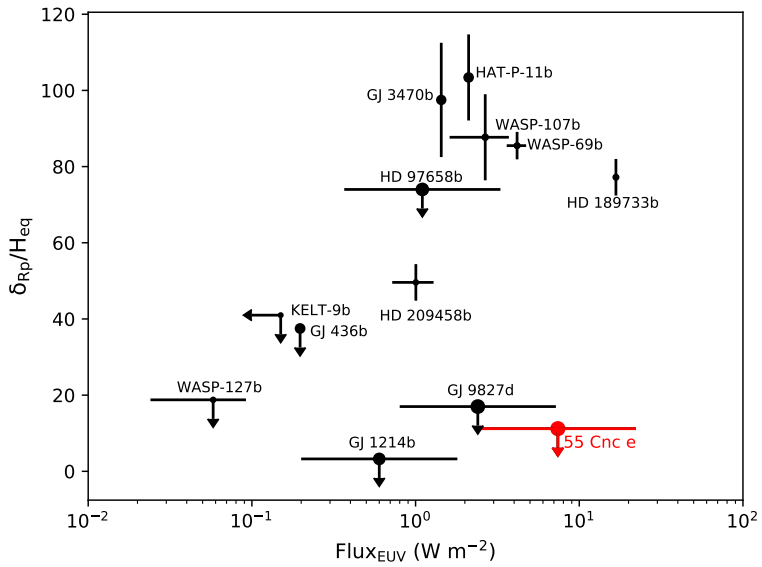


Figure 5.8: 55 Cnc e in context, with the size of each point inversely proportional to the planetary radius. The y-axis shows δ_{Rp}/H_{eq}, the ratio of the increase in apparent radius at 1083 nm to the scale height at equilibrium temperature. The only super Earth with a helium measurement or non-detection, 55 Cnc e is highly irradiated and has a very tight observational constraint compared to its peers. Data for other planets was taken from Kasper et al. 2020. We did not include the helium non-detections of AU Mic b (Hirano et al., 2020), WASP-52b (Vissapragada et al., 2020), K2-100b (Gaidos et al., 2020) or K2-25b (Gaidos et al., 2020), which are less sensitive (the upper limits on δ_{Rp}/H_{eq} are all greater than 100).

5.6 Modeling

We next turn to helium outflow models to explore what our non-detection of helium absorption might mean for 55 Cne e's atmospheric composition and corresponding mass loss rate. We start by performing order-of-magnitude estimates of the mass loss rate from analytical formulae. Next, we use three independent models to interpret the observational result. First, an isothermal Parker wind model (Oklopčić & Hirata, 2018) puts constraints on the 2D parameter space of temperature and mass loss rate, but does not constrain either parameter independently. Our second model, The PLUTO-CLOUDY Interface (TPCI), can model the outflow in a 1D fashion given the stellar XUV spectrum and predict the mass loss rate, temperature profile, and absorption spectrum. Our third and most sophisticated model is a 2.5D model (Wang & Dai, 2018) which combines ray-tracing radiative transfer, real-time non-equilibrium thermochemistry, and hydrodynamics to model the outflow, assuming it is symmetric about the star-planet axis. This model can also predict the mass loss rate, temperature profile, and absorption spectrum, given an assumed atmospheric composition and envelope fraction.

We compare the TPCI and 2.5D predictions directly to observations. Since the Parker wind model requires both the mass loss rate and the exosphere temperature as input parameters, we take a typical exosphere temperature from the TPCI model as a reasonable estimate, and use it within the framework of the Parker wind model to constrain the mass loss rate.

For convenience, we compile the take-away results of the models we considered in Table 5.2. In many of these models, we consider a range of possible parameters and atmospheric compositions. The range of values we give in the table reflect the range of parameters and compositions.

Table 5.2: Summary of Model Predictions

Model	Dimensions	$\dot{M}$ (g/s)	Peak Absorption (ppm)
Energy-limited	0	$>2\times10^{9}$	N/A
Semi-empirical	0	8×10^{9}–1×10^{11}	N/A
Parker wind	1	N/A	900-1800*
TPCI	1	0.75–1.1×10^{10}	800–1000
2.5D	2.5	0.70–1.5×10^{10}	1400–2400

*Assumes the mass loss rate and exosphere temperature predicted by TPCI

Analytic estimates of escape rate

Before running complex hydrodynamic simulations, it is useful to perform rough analytical estimates of the escape rate to illustrate the typical magnitudes involved and their dependence on stellar and planetary quantities. The simplest way to estimate the escape rate is to assume that it is limited by the X-ray and extreme UV stellar radiation hitting the planet. The energy per unit mass required to escape the planet's gravity well is $dE/dm = -GM_p/R_{XUV}$, while the rate at which XUV radiation impinges upon the planet is $dE/dt = \frac{L_{XUV}}{4\pi a^2}\pi R_p^2$. Assuming a fraction η of the energy goes toward driving mass loss, we obtain:

$$\frac{dm}{dt} = \frac{\eta}{4}\frac{R_p^3 L_{XUV}}{GM_p a^2}. \tag{5.11}$$

R_p is the radius of the XUV photosphere, but we take it to be the optical transit radius in order to obtain a conservatively low mass loss rate. We obtain $L_{XUV} = 10^{27.70}$ erg/s from Salz et al. (2016), who in turn obtain it by applying the scaling law in Linsky et al. (2014) to the star's Lyman alpha luminosity. We assume a low efficiency $\eta = 0.15$, in line with Monte Carlo heating models (Shematovich et al., 2014; Ionov & Shematovich, 2015), and obtain $2{\times}10^9$ g/s. Because the value of η and L_{XUV} are both highly uncertain, this should be regarded as an order-of-magnitude estimate only.

A slightly more sophisticated approach to estimating the mass loss rate is to adopt the semi-empirical expressions of Wang & Dai (2018), who derived their expression from 2.5D numerical simulations of mass loss. Their equation is:

$$\begin{aligned}\dot{M} = 4.5 \times 10^{-10} M_\oplus/yr \cdot (\frac{R_{EUV}}{5R_\oplus})^2 \\ \max(F', F'^{0.6})\min(1, M'^{-0.5})\end{aligned} \tag{5.12}$$

$$M' = \frac{M_c}{10M_\oplus} \tag{5.13}$$

$$F' = \frac{F_{EUV}}{5 \times 10^{14} erg/s/\text{Å}}. \tag{5.14}$$

For 55 Cnc, the X-ray luminosity is only 10% of the total XUV luminosity, so the terms EUV and XUV are more or less interchangeable. We use three methods to

determine R_{EUV}, the apparent radius of the planet in the EUV. First, we use the radius implied by the optical transit depth: $1.9R_{\oplus}$. This radius leads to a mass loss rate of $8{\times}10^9$ g/s. Second, we use the Hill radius, $R_{Hill} = a(\frac{M_p}{3M_s})^{1/3} = 7.5R_{\oplus}$. This leads to a mass loss rate of $1{\times}10^{11}$ g/s. These two mass loss rates are the lower and upper limits of what is reasonable under the semi-empirical framework of Wang & Dai (2018).

The third method for estimating R_{EUV} is to follow Wang & Dai (2018) in assuming that the EUV photosphere is at $\rho = 10^{-13}$ g cm^{-3}. To find the radius that corresponds to this density, we assume the atmosphere is isothermal, and need to find one point with known r and ρ. Wang & Dai (2018) (following Owen & Jackson 2012) pick the radiative-convective boundary (RCB), and use a parameterized opacity to calculate the position of the RCB. Unfortunately, this calculation involves many unknown quantities, such as the Kelvin-Helmhotz timescale and the envelope fraction. It may also be particularly ill-suited to thin atmospheres, where the radiative-convective boundary may be very close to the surface and at a much higher temperature than T_{eq}. We therefore assume that P=100 mbar corresponds to the radius inferred from the optical transit depth, 1.9 $R_{\oplus}$. 100 mbar is the approximate photosphere of transit observations in the optical for solar metallicity planets with thick atmospheres, making this guess more accurate than trying to estimate the RCB radius and pressure.

Assuming r(P=100 mbar) = 1.9 $R_{\oplus}$, we obtain $\rho_{phot} = \frac{P\mu}{k_B T_{eq}} = 1.4 \times 10^{-6}$ g/cm^3 and $\beta \equiv \frac{GM_c\mu}{R_p k_B T_{eq}} = 37$, and can calculate R_{EUV}:

$$R_{EUV} = \frac{R_{phot}}{1 + \beta^{-1} \ln(\rho_{EUV}/\rho_{phot})} \tag{5.15}$$

$$= 3.4R_{\oplus}. \tag{5.16}$$

This implies $\dot{M} = 3{\times}10^{10}$ g/s.

1D Parker wind model

Now that we have a rough sense for the expected mass loss rate for a hydrogen- and helium-rich atmosphere, we can ask whether mass loss rates in this range are definitively ruled out by our helium non-detection. We use the Parker wind model of Oklopčić & Hirata (2018) to translate our helium non-detection into a joint constrain on the mass loss rate and the exosphere temperature. This study treats the outflow as an isothermal, radially symmetric Parker wind from a hydrogen-dominated atmosphere. The model calculates the population levels of singlet and triplet states by

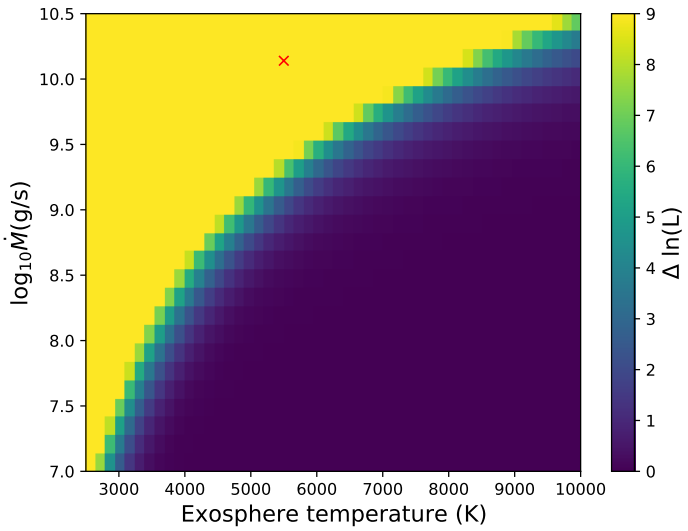


Figure 5.9: The confidence with which each combination of mass loss rate and exosphere temperature is ruled out, according to the 1D Parker wind model of Oklopčić & Hirata (2018). The red x indicates the mass loss rate and approximate exosphere temperature predicted by the non-isothermal Salz et al. (2016) model.

balancing recombination, photoionization, collisional (de)-excitation, and radiative decay as a function of altitude. Using the helium number density and the population level of the triplet state as a function of altitude, it then calculates the transit depth as a function of wavelength.

We simulate absorption spectra for a 2D grid of mass loss rates ranging from 10^7 to 10^{11} g/s, and exospheric temperatures ranging from 2500 to 9000 K, in order to better quantify the limits our data place on the planet's present-day mass loss rate. For each combination of mass loss rate and exosphere temperature, we compute an excess absorption spectrum and calculate the corresponding log likelihood of the observational data given the model. We then compute Δln(L): the difference between the ln(L) of model versus that of a zero absorption model.

We show the resulting $\Delta ln(L)(\dot{M}, T_0)$ in Figure 5.9. In these models, higher temperature exospheres have faster outflow velocities and correspondingly weaker helium absorption for a given mass loss rate. At the planetary surface, for example, a 5000 K exosphere with a mass loss rate of 1.4×10^{10} g/s requires a 0.16 km/s wind; a 10,000 K exosphere with the same mass loss rate requires a 3 km/s wind. A faster wind implies lower density at the same mass loss rate, decreasing helium absorption. This is both because there are fewer helium atoms per cubic volume, and because the fraction of helium atoms in the triplet state is lower. The latter, in

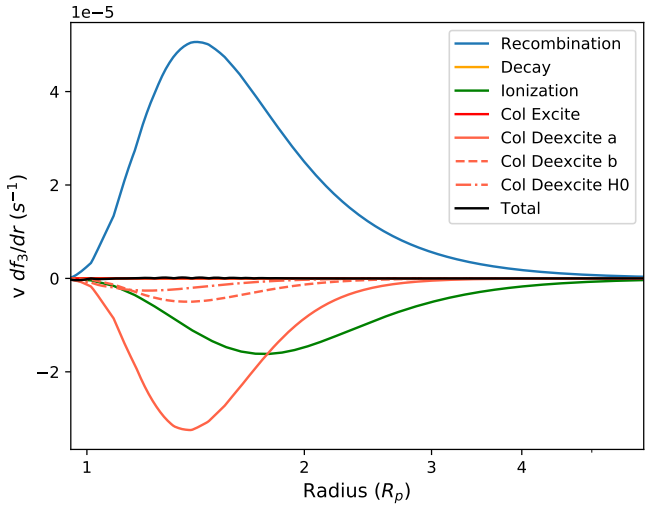


Figure 5.10: Production and destruction rates of triplet helium due to various processes. From a Parker wind model of 55 Cnc e with a mass loss rate of 1.38×10^{10} g/s and exosphere temperature of 5000 K.

turn, is because the triplet state is primarily populated by recombination (see Figure 5.10), and a lower density means a lower recombination rate. For these reasons, the mass loss constraint is less stringent at higher temperatures.

The Parker wind model does not provide a way to estimate the exosphere temperature, which makes it impossible to constrain the mass loss rate without the help of another model. The TPCI model predicts a peak exosphere temperature of 5000-6000 K (Salz et al. 2016; see next subsection for details). At $T_0 = 5000$ K, the mass loss rate is constrained to $\dot{M} < 10^{9.1}$ g/s ($\Delta ln(L) < 7$), or $\dot{M} < 10^{8.8}$ g/s ($\Delta ln(L) < 3$). At $T_0 = 6000$ K, the mass loss rate is constrained to $\dot{M} < 10^{9.5}$ g/s ($\Delta ln(L) < 7$), or to $\dot{M} < 10^{9.2}$ g/s ($\Delta ln(L) < 3$), where $\Delta ln(L)$ thresholds of 7 and 3 correspond to likelihood ratios of 1100 and 20, respectively. In summary, the Parker wind model gives an upper limit on the mass loss rate of $\sim 10^9$ g/s for exospheric temperatures predicted by TPCI.

Lastly, we use this model to gain intuition on the physical processes that determine the triplet helium fraction, and therefore the helium absorption strength. Figure 5.10 plots, as a function of radius, the production and destruction rates of triplet helium due to the processes considered by Oklopčić & Hirata (2018): recombination, radiative decay, ionization, collisional excitation/deexcitation with electrons, and collisional deexcitation with neutral hydrogen atoms. Recombination is the most important production mechanism, while destruction is due to a combination of

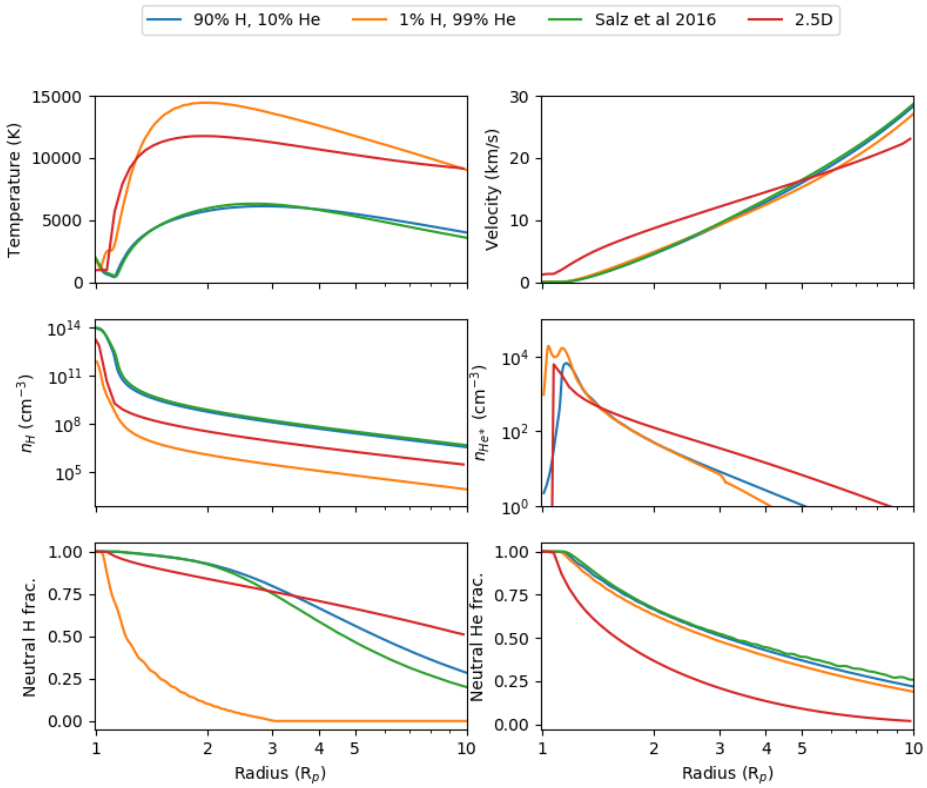


Figure 5.11: Profiles of various physical quantities for the hydrogen- and helium-dominated TPCI models in addition to the 2.5D model. The low-amplitude oscillations in the helium-dominated profiles are numerical artifacts resulting from the non-zero advection length. In green are the results from Salz et al. 2016, who compute all quantities plotted here except for the helium triplet density.

collisional deexcitation with electrons and ionization: the former is dominant close to the planet while the latter is dominant far from the planet, as one would expect. Collisional excitation and radiative decay are negligible. The production and destruction rates very nearly cancel out, indicating that the triplet helium fraction is mostly in equilibrium and that advection is not significant.

1D PLUTO-CLOUDY hydrodynamic model

In this section we use a 1D, spherically symmetric radiative-hydrodynamical simulation to predict the exospheric temperature structure, mass loss rate, and metastable helium absorption signal for 55 Cnc e under several different scenarios. Salz et al. (2016) previously used this model to simulate hydrogen-rich exospheres for several planets, including 55 Cnc e, but did not calculate the predicted absorption signal from metastable helium. The Salz et al. (2016) model predicts an exospheric tem-

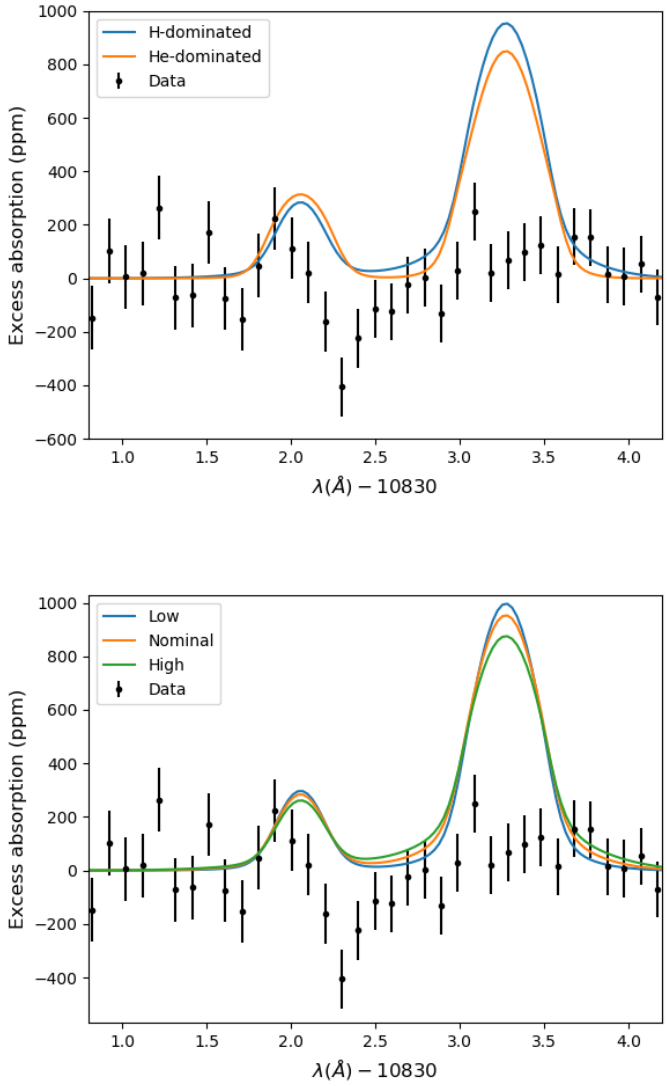


Figure 5.12: TPCI model results. Top: TPCI hydrogen-rich (90% H, 10% He) and helium-rich (1% H, 99% He) models are both ruled out by our observational data. Bottom: increasing or decreasing the stellar XUV flux by a factor of 2 does not significantly change the absorption signal in the TPCI hydrogen-rich model.

perature of 3000-6000 K and an outflow velocity at large distances of 10-15 km/s. It predicts a mass loss rate of 1.4×10^{10} g/s, corresponding to 0.9% of the planetary mass per Gyr. As the paper shows, the predicted Lyman-α signal is consistent with the non-detection by Ehrenreich et al. (2012). This mass loss rate would imply that 55 Cnc e started life as a sub Neptune with a hydrogen/helium envelope of ~10% by mass.

We use TPCI to reconstruct the Salz et al. 2016 model and calculate the corresponding helium absorption signal. This also allows us to explore other compositions, including a helium-dominated atmosphere.

This model couples PLUTO, a hydrodynamics code that can work in 1, 2, or 3 dimensions (Mignone et al., 2007), and CLOUDY (Ferland et al., 2013), a 1D plasma simulation and spectral synthesis code. PLUTO and CLOUDY are linked in The PLUTO-CLOUDY Interface (TPCI; Salz et al., 2015a). In TPCI, CLOUDY calculates the equilibrium chemistry and ionization state of the medium given a radiation field, computes the heating and cooling rates of the new state, and feeds this information to PLUTO. PLUTO heats or cools the medium appropriately, evolves the medium hydrodynamically, and provides the new state to CLOUDY, after which the cycle restarts. Both PLUTO and CLOUDY are sophisticated, publicly available, and general-purpose codes that have been applied to a variety of astrophysical problems, ranging from exoplanet mass loss to the magnetic fields of neutron stars (PLUTO; e.g. Sur et al. 2020) to high-redshift gamma ray bursts (CLOUDY; e.g. Shaw & Ferland 2020). Salz et al. (2016) adopt a 1D spherically symmetric model, and do not include the planetary magnetic field. CLOUDY includes the 30 lightest elements, from hydrogen to zinc, and accounts for many physical processes, including radiative and collisional ionization/recombination, inner shell ionization, and charge exchange. However, Salz et al. (2016) assume a purely atomic hydrogen and helium atmosphere.

As input to the PLUTO-CLOUDY model, the authors use the X-ray luminosity measured by XMM-Newton in April 2009, namely $4.6{\times}10^{26}$ erg/s between 5–100 Å (Sanz-Forcada et al., 2011). This is 2.4 times lower than the flux measured by Chandra, which was obtained simultaneously with the *HST* Lyman alpha observations in March/April 2012 (Ehrenreich et al., 2012). This is likely because March/April 2012 was right at the minimum of the 10.5 ± 0.5 year stellar activity cycle (Bourrier et al., 2018), while April 2009 was 3 years from minimum and 2.3 years from maximum. Because our December 2019 observations were almost exactly one stellar cycle after April 2009, we expect the star's X-ray properties to be similar to those observed by Sanz-Forcada et al. (2011) and adopted by Salz et al. (2018), and more favorable for mass loss measurements than the conditions encountered by Ehrenreich et al. (2012).

Because we do not know exactly what settings or what version of TPCI they used we are unable to replicate their results exactly, but we match their temperature profile to better than 500 K and their mass loss rate to within 25% accuracy (see Figure 5.11).

This is far smaller than the factor of several uncertainty they report as the inherent model error (their Table 2), resulting from factors such as the neglected 3D structure (4x uncertainty), uncertain irradiation strength (3x), and neglect of magnetic fields (2x). Following Salz et al. (2016), we neglect molecules and elements other than hydrogen and helium in our simulation.

We find that the advection length—roughly speaking, the resolution of CLOUDY's advection calculations— is a crucial parameter for these models. Smaller values lead to more accurate results, but take longer to converge. Large values lead to spurious spatial oscillations in the temperature and ionization state. We adopt an advection length of 0.15 planetary radii, which we find is small enough that the oscillations in temperature are of order 0.2% for the helium-dominated run and 0.06% for the hydrogen-dominated run. We run the simulations for 150 days of model time (roughly 1000 sound crossing times) while neglecting advection. We then turn on advection, which slows down the run by a factor of ~80, and let the simulation run for another 150 days of model time. We monitor the evolution of the temperature, density, velocity, and mass loss rate (calculated as $4\pi r^2 \rho(r) v(r)$) profiles to verify that the simulation has in fact converged, with temperatures fluctuating by less than 50 K and the mass loss rate by less than 0.1%.

After reproducing the hydrogen-dominated model from Salz et al. (2016), we run a second TPCI simulation for a helium-dominated atmosphere with 99% He and 1% H by number to explore a scenario where slow mass loss over many Gyr fractionated the atmosphere.

Since CLOUDY computes the level populations of every species, we configure it to report the number density of helium atoms in the triplet state at every radial coordinate. We use this number density profile, in addition to the temperature and velocity profiles, to compute the excess absorption spectrum.

Figure 5.11 compares the profiles for temperature, velocity, hydrogen density, helium triplet density, and neutral fraction for the hydrogen- and helium-dominated TPCI models. Using these profiles, we computed the mass loss rate as $\dot{M} = 4\pi r^2 \rho v / 4$. The division by 4 follows Salz et al. (2016) and is meant to account for the 3D nature of the outflow, as we simulate only the sub-stellar point. We obtain a mass loss rate of 1.1×10^{10} g/s for the hydrogen-dominated scenario and 7.5×10^{9} g/s for the helium-dominated scenario.

In Figure 5.12, we compare the predicted excess absorption from our TPCI models to the observed excess absorption spectrum. Taking into account the interpolation-induced covariance between the data points, we find that a zero absorption model is preferred over the TPCI hydrogen-dominated model by $\Delta\ln(L) = 126$, and preferred over the helium-dominated model by $\Delta\ln(L) = 100$. Although these estimates do not account for the error due to systematics or variable tellurics, it can be seen visually that the predicted absorption for both models is ruled out.

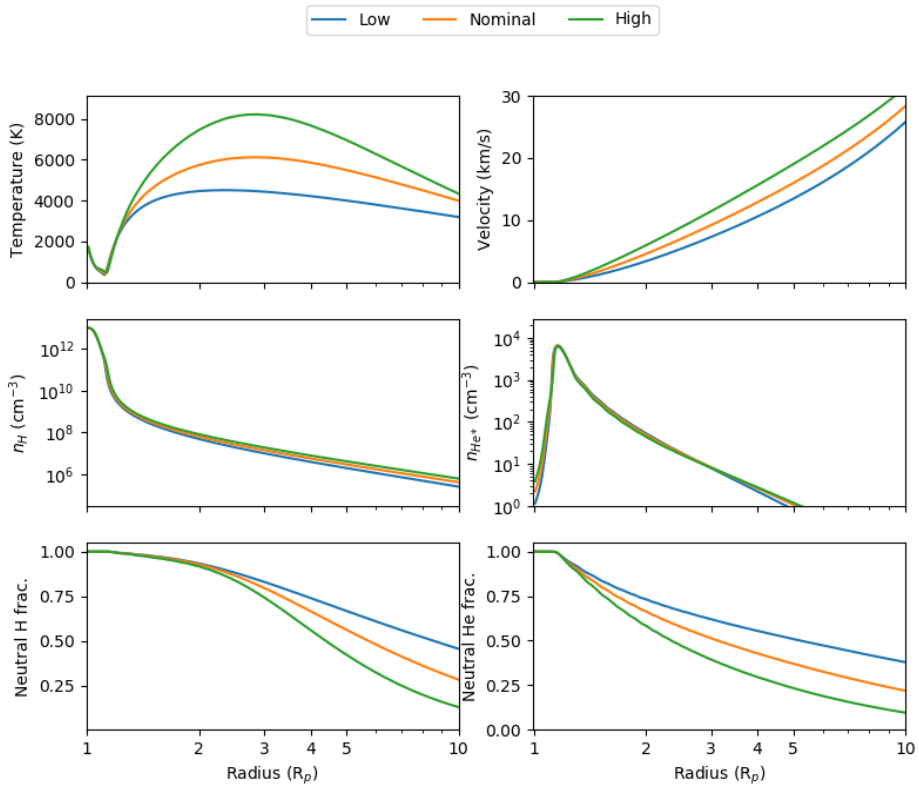


Figure 5.13: Profiles of various physical quantities for hydrogen-dominated TPCI models of different stellar XUV flux: nominal, 2x nominal (high), and half nominal (low).

Critical to all photoevaporation models is the stellar EUV luminosity, which cannot be directly observed and must be reconstructed from X-ray and/or Lyman-α flux. However, different EUV reconstruction techniques give results that are discrepant by an order of magnitude in the worst-case scenario (i.e. Salz et al. 2015b; Drake et al. 2020), although inactive main sequence stars like 55 Cnc e can be modelled more accurately. To determine the effect that uncertainties in the EUV luminosity have on the helium absorption signal, we ran TPCI simulations for the hydrogen-dominated scenario at twice and half the nominal stellar flux. The profiles of various physical

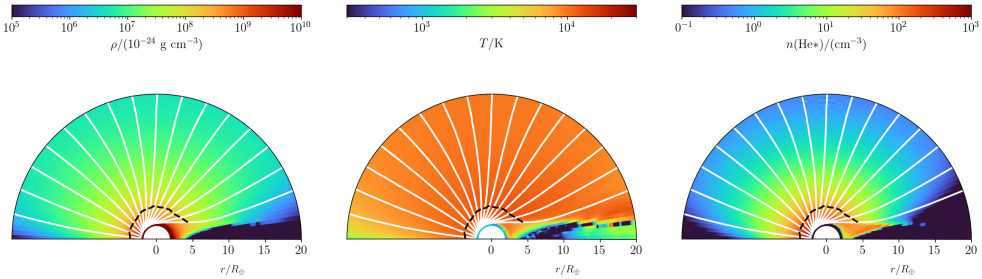


Figure 5.14: Density, temperature, and triplet helium density from the fiducial 2.5D model. The star is toward the left, and the simulation volume is represented by revolving the semicircles about their straight edge. The white lines are streamlines, while the dashed black lines represent the sonic surface.

quantities with respect to radius is shown in Figure 5.13 for the low-flux, nominal, and high-flux scenarios, while the helium absorption signals are compared in Figure 5.12.

The high flux scenario has a mass loss rate 3 times higher than the low flux scenario, implying a slightly sub-linear scaling of mass loss rate with respect to flux. Doubling the stellar flux substantially increases the exosphere temperature and outflow velocity while substantially decreasing the neutral fraction of both hydrogen and helium. The hydrogen number density is moderately higher at higher flux, but the triplet helium number density is remarkably insensitive to flux. This insensitivity means that the excess absorption is barely affected by uncertainties in the stellar flux—it is only 20% lower in the high-flux scenario than in the low-flux scenario.

2.5D hydrodynamic simulations

Our final and most sophisticated model utilizes the approach outlined in Wang & Dai (2018), which combines ray-tracing radiative transfer, real-time non-equilibrium thermochemistry, and hydrodynamics based on the higher-order Godunov method code `Athena++` (Stone et al., 2020). Photoevaporation is inherently not a spherically symmetric phenomenon, since it is driven by stellar XUV flux and the star is only in one direction. Compared to 1D LTE models, this axisymmetric (2.5D) model better captures the anisotropy of the outflow pattern, while the non-LTE thermochemistry self-consistently predicts the mass loss rate and the line profiles. Like Oklopčić & Hirata (2018) and Wang & Dai (2020), metastable helium is added as a chemical species and key reactions that form and destroy this species are included in the thermochemical network. The model incorporates a total of 26

species and 135 reactions, including various relevant heating and cooling processes (e.g. photoionization, photodissociation of atomic hydrogen and Lyman-α cooling).

Our simulations are done in a spherical coordinate system centered on the planet, with the polar axis pointing from the center of the planet to the host star. The simulation domain are in the radial and polar directions [r,θ], while symmetry is assumed in the ϕ direction. Photons for the ray-tracing calculation are divided into five energy bins: (1) $h\nu$ = 2 eV for infrared, optical and near ultraviolet (NUV) photons, (2) $h\nu$ = 7 eV for "soft" far ultraviolet (FUV) photons, (3) $h\nu$ = 12 eV for the Lyman-Werner band FUV photons that can photodissociate molecular hydrogen but cannot ionize them, (4) $h\nu$ = 20 eV for "soft" extreme ultraviolet (soft EUV) photons that can ionize hydrogen but *not* helium, (5) $h\nu$ = 40 eV for hard EUV (and soft X-ray; denoted by "hard EUV" hereafter for simplicity) photons that ionize hydrogen *and* helium. Photon fluxes in each energy bin are determined according to the planet's orbital separation and the corresponding luminosities for a typical K star, according to the reviews in Oklopčić (2019). The EUV flux for this model star is 10,926 erg s^{-1} Å^{-1}, 30% higher than what we adopted in the 1D TPCI simulations—a negligible difference given the inherent uncertainty in EUV flux. In addition to the opacities caused by photochemical reactions, we also added an effective opacity term to all bands, particularly in the optical band $h\nu$ = 2 eV where our opacity calculation did not include the Thomson cross-section $\sigma/\mathrm{H} \simeq 6.7 \times 10^{-25}$ cm^2.

We note that the models we present here do not account for the effect of the stellar wind on the predicted mass loss rate. In previously published 1D models (i.e. Murray-Clay et al. 2009), the stellar wind acts as a simple scalar suppression force. In Wang & Dai (2020), they discuss the effect of the stellar wind on planetary outflows using the same hydrodynamic model presented here. They find that in higher dimensions, the stellar wind only suppresses the day side mass loss; the outflow is simply redirected towards the night side of the planet, forming a comet-like tail. In these models, adding a solar-like stellar wind (we note that 55 Cnc e is an old, sun-like star) only changes the 3D mass loss rate by a few percent, corresponding to a change in the equivalent width of He absorption of less than 10% (see Table 2 in Wang & Dai, 2020). In short, stellar wind may alter the line profile of He absorption, but it is unable to quench the outflow in three dimensions and is therefore unlikely to be the reason behind the non-detection of He absorption for 55 Cnc e.

Item	Value
Simulation domain	
Radial range	$1.89 \leq (r/R_\oplus) \leq 20$
Latitudinal range	$0 \leq \theta \leq \pi$
Resolution ($N_{\log r} \times N_\theta$)	144×96
Atmospheric properties*	
$\rho(R_\mathrm{p})$	10^{-7} g cm^{-3}
$T(R_\mathrm{p}) \simeq T_{eq}$	1990 K
Radiation flux [photon cm $^{-2}$ s^{-1}]	
2 eV (IR/optical)	1.1×10^{21}
7 eV (Soft FUV)	4.3×10^{16}
12 eV (Lyman-Werner FUV)	1.5×10^{13}
20 eV (Soft EUV)	1.2×10^{13}
40 eV (Hard EUV)	1.6×10^{14}
Initial abundances $[n_\mathrm{X}/n_{atom}]$ †	
H_2	0.455
He	0.091
H_2O	1.8×10^{-4}
CO	1.4×10^{-4}
S	2.8×10^{-5}
Si	1.7×10^{-6}

$*$: $R_\mathrm{p} \simeq 1.89\ R_\oplus$ is the size of the rocky planet core.
†: Dust grains are not included, since $T_{eq} \simeq 1990$ K is well above the dust sublimation temperature.

Table 5.3: Setups of the fiducial 2.5D numerical models of the evaporating atmosphere

A typical planetary atmosphere consists of a convective interior and a quasi-isothermal exterior (e.g. Rafikov, 2006), but our numerical tests found that even a thin convective isentropic layer will cause the whole atmosphere to become over-dense and unbounded for such a close-in low-mass planet (the boil-off regime, see discussions in Owen & Wu 2016 and Wang & Dai 2018). We therefore set up the model atmospheres with a quasi-isothermal layer directly above the rocky core. In contrast, the TPCI model did not set a solid surface or impose a truncation of the gas reservoir; instead, it sets pressure and density boundary conditions at the inner radius. We summarize the key quantities that define the 2.5D fiducial model in Table 5.3. In addition to the fiducial model, which has a density of $\rho(R_\mathrm{p}) = 10^{-7}$ g cm^{-3} at the rocky surface, we also considered models with $\rho(R_\mathrm{p}) = 10^{-5}$ g cm^{-3} and

10^{-9} g cm^{-3}, respectively. As with the TPCI models, we also considered a scenarios with a He-dominated atmosphere (1% hydrogen and 99% helium by atom number) spanning the same three surface pressures.

Figure 5.14 shows the 2D profiles of density, temperature, and triplet helium density, along with streamlines and the sonic surface. As can be seen, the anti-stellar side of the planet has drastically different physical conditions from the star-facing side. However, due to its compactness, much of colder, less dense material on the anti-stellar side would not block any more light during transit than the planet itself. Except for this region, the rest of the simulation domain is largely spherically symmetric, especially in the number density of triplet helium, which directly determines the helium absorption signature.

Figure 5.11 shows the radial profiles of temperature, velocity, hydrogen number density, triplet helium number density, neutral H fraction, and neutral He fraction for the fiducial model (hydrogen dominated, $\rho(R_p) = 10^{-7}$ g cm^{-3}), showing a fairly typical supersonic photoevaporative outflow that carries metastable helium atoms. The assumed EUV flux of 55 Cnc e produces a relative abundance of metastable helium in 55 Cnc e's atmosphere of $\sim 10^{-7}$, nearly identical to the abundance predicted by TPCI. Despite this low abundance, the transmission spectra in Figure 5.15 still have clearly recognizable excess absorption with amplitudes of a few thousand ppm. The absorption is still greater than 1000 ppm even for the thinnest atmosphere ($\rho(R_\mathrm{p}) = 10^{-9}$ g cm^{-3}). The mass-loss timescales of these models, summarized in Table 5.4, range between years and thousands of years. The mass loss rates are within 50% of those predicted by TPCI for both hydrogen and helium dominated atmospheres. These mass-loss rates are time-averaged after the simulation reaches a quasi-steady state after dozens of dynamical timescales elapsed since the start of the simulations. The non-detection of helium absorption is therefore in good agreement with the short dispersal timescales for these atmospheres, implying that 55 Cnc e would not have been able to retain a primordial atmosphere for very long after the dispersal of the gas disk.

Model No.	$\rho(R_\mathrm{p})$ (g cm^{-3})	M_{env}/M_p $\times 10^{-9}$	$P(R_p)$ (mbar)	$\dot{M}$ 10^{10} g s^{-1}	τ^a_{disp} (yr)	$\langle W_\lambda \rangle^b$ (10^{-3})	$\Delta \ln(L)^c$
H/He							
1	10^{-5}	98	500	1.5	5000	2.4	748
2*	10^{-7}	0.98	5	0.84	86	1.4	243
3	10^{-9}	0.0098	0.05	0.85	0.9	1.4	252
He-Dom.							
4	10^{-5}	98	400	1.1	4000	3.3	712
5	10^{-7}	0.80	4	0.94	65	2.0	538
6	10^{-9}	0.008	0.04	0.70	0.9	1.6	342

a: Dispersal timescale, defined by the current atmospheric mass divided by the steady-state mass loss rate.
b: Dimensional equivalent width of excess absorption, time-averaged from the end of ingress through the start of egress.
c: Difference in log likelihood between this model and a model with zero absorption. A bigger number means a worse fit.
$*$: Fiducial model. (§5.6).

Table 5.4: 2.5D simulations: Parameters and results

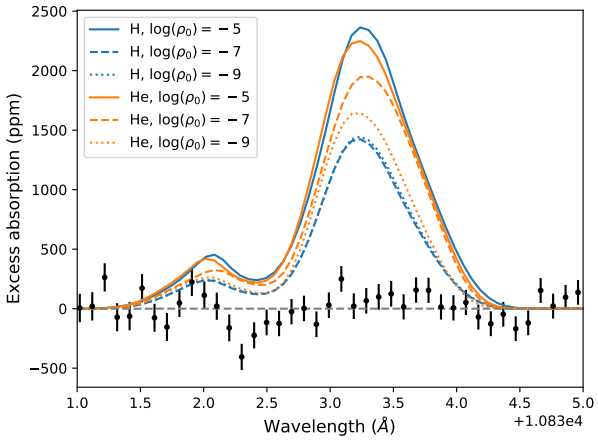


Figure 5.15: Metastable helium transmission spectra of model atmospheres for the simulations described in §5.6 and Tables 5.3, 5.4, presenting the time-averaged (from the end of ingress through the start of egress) excess absorption. The horizontal dashed line indicates zero excess absorption for reference. Different reference densities (measured at the planet surface R_p) are marked by different line styles, with $log(\rho_0) = -5, -7, -9$ corresponding to roughly 500, 5, and 0.05 millibar, respectively. Orange lines represent helium-dominated models, while blue lines represent hydrogen-dominated models.

5.7 Discussion

Now that we have explored the limits that can be placed on 55 Cnc e's atmosphere from our helium observations alone, we next consider how these results relate to other published observations of 55 Cnc e's atmosphere. We first focus on Ehrenreich et al. (2012), which presents a Lyman α transit of 55 Cnc e from *HST*/STIS. Lyman α and the helium 1083 nm line are both good probes of the outflowing atmosphere, but the former is sensitive only to the high speed portions of the outflow while the latter can only probe the low-speed portions of the outflow where metastable helium exists. Ehrenreich et al. (2012) measures the transit depth between 1215.36 Å and 1215.67 Å, obtaining a value of 0.3 ± 2.4%. From this, they constrain the mass loss rate to be below $3{\times}10^8$ g/s (3σ). They interpret their data using the Bourrier & Lecavelier des Etangs (2013) model, which is a 3D particle-based simulation that takes into account radiation pressure, photoionization, the stellar wind, and the effects of self-shielding for both stellar photons and protons. The inclusion of outward forces from radiation pressure and stellar wind gives the outflow a trailing cometary tail: a consequence of the Coriolis acceleration $-2\mathbf{\Omega} \times \mathbf{v}$. This model has 5 tunable parameters: the mass loss rate, the stellar EUV flux, and three stellar wind parameters: bulk velocity, temperature, and density. (By contrast, our 2.5D model is a hydrodynamics code which predicts the mass loss rate given the EUV flux, but does not take the stellar wind into account and cannot model a tail.) As with all Lyman alpha observations of even the nearest stars, the line core is completely absorbed by the interstellar medium, and only the far wings are visible. This means that the detectability of an absorption signal depends as much on the highly uncertain kinematic structure of the outflow as on the quantity of outflowing gas.

Even though we quote a slightly higher upper limit on the mass loss rate ($\sim 10^9$ g/s from the isothermal Parker wind model) than Ehrenreich et al. (2012) ($3{\times}10^8$ g/s), our models are in fact consistent with their Lyman alpha non-detection even at mass loss rates that are definitively ruled out by our helium non-detection. Figure 12 of Salz et al. (2016) compares the Lyman alpha data to their TPCI simulation of 55 Cnc e under the assumption of a hydrogen-dominated atmosphere, finding that the two are consistent. As we have seen, however, the helium absorption signal predicted by TPCI for the hydrogen-dominated atmosphere case is highly inconsistent with our non-detection. In addition, a 1D model is not ideal for modelling Lyman-α because the strength of the line absorption, combined with interstellar absorption that eliminates the line core, make the kinematic structure of the outflow crucially important in predicting the observable signal. For this reason, we used the same

code we used to run our 2.5D models to run a helium-dominated 3D hydrodynamic model (resembling model 5 in Table 5.4, corresponding to an intermediate mass helium atmosphere). We found that just 1% hydrogen in the outflow is enough to give a 12% excess absorption depth in the Lyman α line core, but the line core is not observable due to interstellar absorption.

To determine what is observable, we start with the stellar intrinsic Lyman alpha profile, corrected for interstellar absorption, that Ehrenreich et al. (2012) provide in Figure 6. We model the observed out-of-transit spectrum by convolving the absorbed profile with the line spread function of STIS, as provided by STScI[5]. We model the observed in-transit spectrum by multiplying the post-ISM-absorption profile by the predicted Lyman alpha absorption profile from the exosphere and convolving the product with the line spread profile. The resulting excess absorption spectrum for our helium-dominated 2.5D model is shown in Figure 5.16. We did not show the hydrogen-dominated case as the high-opacity region extends well beyond the domain of our simulation. It is clear that the predicted Lyman alpha absorption is fully consistent with the observations. As a caveat, we note that absorption is often seen in the high-velocity wings of Lyman-α observations, which could be due to charge exchange with the stellar wind (Holmström et al., 2008; Tremblin & Chiang, 2013). Since our model does not include stellar wind, it may severely underestimate the absorption at observable wavelengths. Were the stellar wind to be included, it is possible that the Lyman alpha non-detection would become a stronger constraint on mass loss than the helium non-detection.

As noted in the previous section, the observations by Ehrenreich et al. (2012) coincided with the stellar activity minimum. As far as we know, nobody observed 55 Cnc e in either Lyman alpha or He I 1083 nm during the stellar maximum in mid 2017. Our observations took place in between said maximum and the next minimum in late 2022. The star's X-ray luminosity is 2.4 times higher in this phase than it was in 2012, making it conceivable that any potential outflow from the planet would have increased in strength relative to the epoch of the Lyman α observations during the stellar minimum. Both the TPCI models and the 2.5D simulations assume the higher X-ray flux relevant for the epoch of our helium observations, and as a result they may over-predict the magnitude of the Lyman-α absorption at the epoch of the *HST* observations. Since the predicted Lyman-α absorption from these higher

[5] https://www.stsci.edu/hst/instrumentation/stis/performance/spectral-resolution

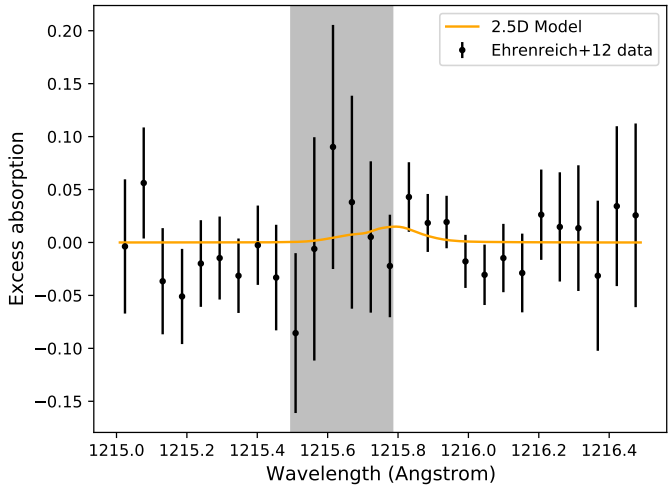


Figure 5.16: Lyman alpha observations of Ehrenreich et al. (2012), compared to predictions of our helium-dominated 2.5D model (model 5 in Table 5.4). Within the gray shaded region, the intrinsic stellar flux is fully absorbed by the ISM, but the observed flux is non-zero because of the broad instrumental line spread profile.

X-ray flux models is already undetectable, there is no need to recalculate the models with a lower X-ray flux level.

Prior to our observations, it was also conceivable that 55 Cnc e might have had a helium-dominated atmosphere. However, our observations show no evidence of helium absorption, and our 2.5D simulations show that the mass loss rates are much too fast for the planet to keep either a hydrogen- or helium-dominated atmosphere. Fast mass loss also makes it difficult to create a helium-dominated atmosphere in the first place, because escape rates much faster than the diffusion-limited rate do not significantly fractionate the elements (Hu et al., 2015).

Our observations appear to contradict those of Tsiaras et al. (2016), who analyzed two transits of 55 Cnc e observed with *HST*/WFC3, and found an upward-sloping spectral feature consistent with HCN absorption in a lightweight atmosphere. They perform a free retrieval on the transmission spectrum with TauREx and find a mean molecular weight μ of 2–6 (their Figure 9). The atmospheric scale height is inversely proportional to μ; thus, the amplitude of features in a transmission spectrum is also inversely proportional to μ. $\mu = 2 - 6$ is consistent with a hydrogen dominated atmosphere ($\mu = 2.3$) or a helium dominated atmosphere ($\mu \approx 4$), but excludes atmospheres dominated by heavier molecules (i.e. water with $\mu = 18$, N_2 with $\mu = 14$, O_2 with $\mu = 16$). If the atmosphere is not hydrogen or helium dominated,

absorption features from any molecule, including HCN, would be undetectable. If μ = 16, for example, the scale height would be 47 km (assuming equilibrium temperature) and the change in transit depth corresponding to one scale height would be 3 ppm. This lies far below the detection threshold of the observations reported in Tsiaras et al. (2016). We conclude that our non-detection of escaping helium, together with the non-detection of escaping hydrogen by Ehrenreich et al. (2012) and the strong irradiation of the planet over its long life, make it unlikely that the features detected by Tsiaras et al. (2016) in the *HST*/WFC3 spectrum are planetary in nature.

Two more pieces of evidence complicate the picture: the significant phase offset in the *Spitzer* 4.5 μm phase curve (Demory et al., 2016b), and the year-to-year variability in the secondary eclipse depth in this same band (Demory et al., 2016a). To explain the eclipse variability, one could invoke a magma world with no atmosphere, or an extremely tenuous mineral atmosphere in equilibrium with the molten surface (Ito et al., 2015), in which case ejecta from volcanic eruptions could periodically shroud the surface. The mass and radius of 55 Cnc e are consistent with a world without an atmosphere (Bourrier et al., 2018). However, the phase offset suggests a thick, high molecular weight atmosphere. This point was discussed in Angelo & Hu (2017), who suggested that an N_2-dominated atmosphere would be consistent with the *Spitzer* data. The possible composition of a nitrogen-dominated atmosphere and the observability of spectral features are explored in Miguel (2019) and Zilinskas et al. (2020). However, such a thick atmosphere would be unlikely to change substantially on year-long timescales. Our non-detection of helium is agnostic to all potential high mean molecular weight atmospheres, as helium is not expected to be a significant component of secondary (e.g., non-primordial) atmospheres. While photodissociation at the top of a water-rich atmosphere might create a detectable Lyman-α signal, high resolution spectral observations rule out water-rich atmospheres (volume mixing ratio > 0.1%) with a mean molecular weight of <= 15 g/mol at a 3σ confidence level (Jindal et al., 2020).

5.8 Conclusion

We observed two transits of 55 Cnc e using Keck/NIRSPEC to look for metastable helium absorption in the 1083 nm line. We found no absorption greater than 250 ppm (90% upper limit), and used three independent models to interpret this result. First, an isothermal Parker wind model (Oklopčić & Hirata, 2018) puts constraints on the temperature and mass loss rate (Figure 5.9), with the mass loss rate

constrained to less than $\sim 10^9$ g/s for exosphere temperatures of 5000-6000 K. This exosphere temperature is obtained from our second model, The PLUTO-CLOUDY Interface (TPCI), which can model the outflow in a 1D fashion given the stellar XUV spectrum. TPCI predicts a mass loss rate of 1.1×10^{10} g/s for a hydrogen-dominated atmosphere and 7.5×10^9 g/s for a helium-dominated atmosphere, both of which result in absorption several times stronger than what is (not) observed (Figure 5.12). Our third and most sophisticated model is a 2.5D model (Wang & Dai, 2018) which combines ray-tracing radiative transfer, real-time non-equilibrium thermochemistry, and hydrodynamics to model the outflow, assuming it is symmetric about the star-planet axis. Even for extremely thin atmospheres with dispersal timescales of millenia or less, the model still predicts high mass loss rates of $\sim 10^{10}$ g/s for both hydrogen and helium dominated atmospheres (see Table 5.4), which result in 1500–2500 ppm excess absorption—many times higher than what is observed (Figure 5.15). Although the significant model uncertainties must be kept in mind, our observations provide strong evidence against the existence of a low mean molecular weight primordial atmosphere on 55 Cnc e.

If 55 Cnc e instead possesses a high mean molecular weight secondary atmosphere, detection via transit spectroscopy will be extremely challenging. Ultimately, we believe that emission spectroscopy with next-generation telescopes—*JWST*, TMT, ELT, and GMT—represents the best path forward. This planet's high dayside temperature (~2700 K) makes it a particularly favorable target for emission spectroscopy, and the magnitude of spectral features seen in emission is independent of the mean molecular weight of the atmosphere. While the brightness of 55 Cnc makes it feasible to search for relatively small signals, systematics that are negligible for low SNR targets become important when the photon noise is small. In this paper, we saw that crosstalk, fringing, and the inseparability of the 2D PSF each required special handling; observations with both *JWST* and next-generation ground-based telescopes will likely encounter similar technical challenges. Despite these challenges, atmosphere modeling studies suggest that it is indeed possible to detect high mean molecular weight atmospheres for 55 Cnc e using next-generation telescopes (e.g., Zilinskas et al. 2020). To date, atmospheric absorption features have only been detected for planets with relatively massive, hydrogen-rich atmospheres. Detecting a high mean molecular weight atmosphere around a high density planet like 55 Cnc e would provide invaluable insights into the nature and origin of the broader population of short-period super-Earths.

Software: `numpy (van der Walt et al., 2011), scipy (Virtanen et al., 2020), matplotlib (Hunter, 2007), dynesty (Speagle, 2020), corner (Foreman-Mackey, 2016)`

5.9 Acknowledgements

M.Z. would like to acknowledge Joe Ninan, who was indispensable in deciphering the mysteries of NIRSPEC systematics. He also acknowledges Jacob Bean for sending most of the data plotted in Figure 5.8.

References

Angelo, I., & Hu, R. 2017, AJ, 154, 232

Bolton, A. S., & Schlegel, D. J. 2010, PASP, 122, 248

Bourrier, V., & Lecavelier des Etangs, A. 2013, A&A, 557, A124

Bourrier, V., Dumusque, X., Dorn, C., et al. 2018, A&A, 619, A1

Chachan, Y., & Stevenson, D. J. 2018, ApJ, 854, 21

Dai, F., Masuda, K., Winn, J. N., & Zeng, L. 2019, ApJ, 883, 79

Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2016a, MNRAS, 455, 2018

Demory, B.-O., Gillon, M., de Wit, J., et al. 2016b, Nature, 532, 207

Drake, J. J., Kashyap, V. L., Wargelin, B. J., & Wolk, S. J. 2020, ApJ, 893, 137

Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83

Ehrenreich, D., Bourrier, V., Bonfils, X., et al. 2012, A&A, 547, A18

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109

Gaidos, E., Hirano, T., Wilson, D. J., et al. 2020, MNRAS, 498, L119

Gaidos, E., Hirano, T., Mann, A. W., et al. 2020, MNRAS, 495, 650

George, E. M., Tulloch, S. M., Ives, D. J., & ESO Detector Group. 2020, Journal of Astronomical Telescopes, Instruments, and Systems, 6, 011003

Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759

Hirano, T., Krishnamurthy, V., Gaidos, E., et al. 2020, ApJ, 899, L13

Holmström, M., Ekenbäck, A., Selsis, F., et al. 2008, Nature, 451, 970

Horne, K. 1986, PASP, 98, 609

Hu, R., Seager, S., & Yung, Y. L. 2015, ApJ, 807, 8

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6

Ionov, D. E., & Shematovich, V. I. 2015, Solar System Research, 49, 339

Ito, Y., Ikoma, M., Kawahara, H., et al. 2015, ApJ, 801, 144

Jindal, A., de Mooij, E. J. W., Jayawardhana, R., et al. 2020, AJ, 160, 101

Kasper, D., Bean, J. L., Oklopčić, A., et al. 2020, AJ, 160, 258

Kim, S., Prato, L., & McLean, I. 2015, REDSPEC: NIRSPEC data reduction

Kreidberg, L. 2015, PASP, 127, 1161

Lee, E. J., & Chiang, E. 2017, ApJ, 842, 40

Linsky, J. L., Fontenla, J., & France, K. 2014, ApJ, 780, 61

Lopez, E. D., & Fortney, J. J. 2013, ApJ, 776, 2

Lord, S. D. 1992, A new software tool for computing Earth's atmospheric transmission of near- and far-infrared radiation, NASA Technical Memorandum 103957

Malsky, I., & Rogers, L. A. 2020, ApJ, 896, 48

Marsh, T. R. 1989, PASP, 101, 1032

Martin, E. C., Fitzgerald, M. P., McLean, I. S., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Proc. SPIE, 107020A

Mazeh, T., Tamuz, O., & Zucker, S. 2007, Astronomical Society of the Pacific Conference Series, Vol. 366, The Sys-Rem Detrending Algorithm: Implementation and Testing, ed. C. Afonso, D. Weldrake, & T. Henning, 119

Mignone, A., Bodo, G., Massaglia, S., et al. 2007, ApJS, 170, 228

Miguel, Y. 2019, MNRAS, 482, 2893

Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, ApJ, 693, 23

Ninan, J. P., Stefansson, G., Mahadevan, S., et al. 2020, ApJ, 894, 97

Nortmann, L., Pallé, E., Salz, M., et al. 2018, Science, 362, 1388

Oklopčić, A. 2019, arXiv e-prints, arXiv:1903.02576

Oklopčić, A., & Hirata, C. M. 2018, ApJ, 855, L11

Owen, J. E. 2019, Annual Review of Earth and Planetary Sciences, 47, 67

Owen, J. E., & Jackson, A. P. 2012, MNRAS, 425, 2931

Owen, J. E., & Wu, Y. 2016, ApJ, 817, 107

Palle, E., Nortmann, L., Casasayas-Barris, N., et al. 2020, A&A, 638, A61

Rafikov, R. R. 2006, ApJ, 648, 666

Ridden-Harper, A. R., Snellen, I. A. G., Keller, C. U., et al. 2016, A&A, 593, A129

Salz, M., Banerjee, R., Mignone, A., et al. 2015a, A&A, 576, A21

Salz, M., Czesla, S., Schneider, P. C., & Schmitt, J. H. M. M. 2016, A&A, 586, A75

Salz, M., Schneider, P. C., Czesla, S., & Schmitt, J. H. M. M. 2015b, A&A, 576, A42

Salz, M., Czesla, S., Schneider, P. C., et al. 2018, A&A, 620, A97

Sanz-Forcada, J., Micela, G., Ribas, I., et al. 2011, A&A, 532, A6

Shaw, G., & Ferland, G. J. 2020, MNRAS, 493, 5153

Shematovich, V. I., Ionov, D. E., & Lammer, H. 2014, A&A, 571, A94

Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, Nature, 557, 68

Speagle, J. S. 2020, MNRAS, 493, 3132

Stone, J. M., Tomida, K., White, C. J., & Felker, K. G. 2020, ApJS, 249, 4

Sulis, S., Dragomir, D., Lendl, M., et al. 2019, A&A, 631, A129

Sur, A., Haskell, B., & Kuhn, E. 2020, MNRAS, 495, 1360

Tremblin, P., & Chiang, E. 2013, MNRAS, 428, 2565

Tsiaras, A., Rocchetto, M., Waldmann, I. P., et al. 2016, ApJ, 820, 99

Valencia, D., Ikoma, M., Guillot, T., & Nettelmann, N. 2010, A&A, 516, A20

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Vissapragada, S., Knutson, H. A., Jovanovic, N., et al. 2020, AJ, 159, 278

Wang, L., & Dai, F. 2018, ApJ, 860, 175

Wang, L., & Dai, F. 2020, arXiv e-prints, arXiv:2101.00045

Zilinskas, M., Miguel, Y., Mollière, P., & Tsai, S.-M. 2020, MNRAS, 494, 1490

Chapter 6

MASS LOSS FROM TWO MINI NEPTUNES

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, Leonardo A. dos Santos, Luca Fossati, Gregory W. Henry, David Ehrenreich, Yann Alibert, Sergio Hoyer, Thomas G. Wilson, and Andrea Bonfanti. Detection of Ongoing Mass Loss from HD 63433c, a Young Mini-Neptune. *AJ*, 163(2):68, February 2022. doi: 10.3847/1538-3881/ac3f3b. M.Z. conceived the project, wrote the observing proposals, obtained the data, carried out the fiducial data analysis, and wrote the paper.

6.1 Abstract

We detect Lyman α absorption from the escaping atmosphere of HD 63433c, a $R = 2.67R_\oplus$, $P = 20.5$ d mini Neptune orbiting a young (440 Myr) solar analogue in the Ursa Major Moving Group. Using *HST*/STIS, we measure a transit depth of $11.1 \pm 1.5\%$ in the blue wing and $8 \pm 3\%$ in the red. This signal is unlikely to be due to stellar variability, but should be confirmed by an upcoming second transit observation with *HST*. We do not detect Lyman α absorption from the inner planet, a smaller $R = 2.15R_\oplus$ mini Neptune on a 7.1 d orbit. We use Keck/NIRSPEC to place an upper limit of 0.5% on helium absorption for both planets. We measure the host star's X-ray spectrum and MUV flux with XMM-Newton, and model the outflow from both planets using a 3D hydrodynamic code. This model provides a reasonable match to the light curve in the blue wing of the Lyman α line and the helium non-detection for planet c, although it does not explain the tentative red wing absorption or reproduce the excess absorption spectrum in detail. Its predictions of strong Lyman α and helium absorption from b are ruled out by the observations. This model predicts a much shorter mass loss timescale for planet b, suggesting that b and c are fundamentally different: while the latter still retains its hydrogen/helium envelope, the former has likely lost its primordial atmosphere.

6.2 Introduction

Mass loss shapes exoplanet demographics and atmospheric properties. The observed radius distribution of close-in sub-Neptune-sized planets is bimodal, with peaks at < 1.5 $R_\oplus$ and 2-3 $R_\oplus$ (Fulton et al., 2017; Fulton & Petigura, 2018). This bimodality can be explained by scenarios in which the observed population of sub-Neptune-

sized planets formed with a few $M_{\texttt{Earth}}$ rocky cores and hydrogen-rich atmospheres, which were then stripped away from the most highly irradiated planets (e.g., Lopez & Fortney 2013; Owen & Wu 2017; Lehmer & Catling 2017; Mills & Mazeh 2017). However, it is possible that the gap is caused by mass loss driven by the forming protoplanet's own cooling luminosity (Ginzburg et al., 2018; Gupta & Schlichting, 2019). It has also been proposed that the radius valley is primordial (Lee & Connors, 2021).

An escaping atmosphere can be detected in absorption when the planet transits in front of its host star. Because of the abundance of hydrogen and the strength of the Lyα line, Lyα exospheres can absorb a very large fraction of starlight during transit. The Neptune-sized GJ 436b, for example, has a transit depth of 56% in the Lyα blue wing and 47% in the red wing (Lavie et al., 2017). The second most abundant element of escaping primordial atmospheres is helium. In 2018, Spake et al. (2018) detected helium absorption from a transiting exoplanet for the first time. The He I 1083 nm line is observable from the ground and has a much higher photon flux compared to Lyα. It also has its own challenges: the signal size is much smaller, and only active early M to late G stars have the right high-energy spectrum to maintain a suitably high triplet ground state population (Oklopčić & Hirata, 2018). These two probes provide complementary information. Because ISM absorption wipes out the Lyα core, Lyα probes high-velocity hydrogen in the tenuous outer reaches of the escaping atmosphere. The metastable helium line is not absorbed by the ISM and probes low-velocity helium in the denser part of the exosphere, closer to the planet surface (∼3 planetary radii vs. ∼12 planetary radii).

The young mini Neptune regime is the most critical for understanding the processes behind the radius gap, yet mass loss has never been securely detected from planets of this size in either Lyα or helium. The smallest planet with a secure detection in either wavelength is the Neptune-mass GJ 3470b, with a radius of 3.9 $R_\oplus$ and a mass of 13 $M_\oplus$(Bourrier et al., 2018). However, it is not young, with a rotation period of 20 days and an age of 1–4 Gyr (Biddle et al., 2014). The other planets with definite Lyα detecions—GJ 436b (Lavie et al., 2017), HD 189733b(Bourrier et al., 2013), HD 209458b (Vidal-Madjar et al., 2008)—are even bigger and older. Of the planets detected in helium, none are mini Neptunes or smaller, and none are indisputably younger than 1 Gyr. WASP-107b (Močnik et al., 2017) has a gyrochronological age of 0.6 ± 0.2 Gyr but an evolutionary track age estimate of 8.3 ± 4.3 Gyr, illustrating the difficulty of measuring ages for isolated stars.

The scarcity of successful measurements is due to the many conditions necessary for Lyα absorption or helium absorption to be detectable. For both wavelengths, we need a transiting exoplanet with a hydrogen-rich atmosphere on a tight orbit around a star of at least moderate activity. Interstellar Lyα absorption saturates for even the closest stars, making it hard to observe planetary absorption beyond 50 pc and almost impossible beyond 100 pc. Triplet helium absorption requires a high population of triplet helium, which in turn requires a star that has a high extreme UV to mid UV ratio, which is optimally achieved for K type stars, but not impossible around active G stars. Very few currently known transiting planets fit these criteria, and few of those are young mini Neptunes.

The G5 star HD 63433 (TOI 1726) is a young (414 ± 23 Myr) and nearby (22 pc) solar analogue (M=1 $M_{\odot}$), a member of the Ursa Major moving group (Mann et al., 2020). In keeping with its young age and 6.4 d rotation period, it has an exceptionally high X-ray luminosity. The Second ROSAT All-sky Survey measured its 0.1–2.4 KeV X-ray flux to be $F_{\rm X} = 1.4 - 1.8 \times 10^{-12}$ erg s^{-1} cm^{-2}. The XMM-Newton slew survey (Saxton et al., 2008) measured its 0.2–12 KeV flux as 4.4 ± 1.4 and 1.6 ± 0.8 erg s^{-1} cm^{-2} on two separate visits. Taking $1.5{\times}10^{-12}$ to be representative, the corresponding stellar X-ray luminosity is 10^{29} erg s^{-1}, 40 times higher than the average solar X-ray luminosity (Judge et al., 2003). In addition, the star's negative radial velocity (−16 km/s), together with the positive radial velocity of the local interstellar cloud in that direction (22 km/s), gives an unusually clear view of the blue Lyα wing and a glimpse of the core. Luckily, the core and blue wing are where we expect the most planetary absorption: planetary outflows have a typical speed of ∼2 times the sound speed (or ∼20 km/s), and the stellar wind pushes the outflow toward the observer.

Inside this intense X-ray environment reside two mini Neptunes, both discovered by TESS (Mann et al., 2020): a 2.15 $R_{\oplus}$ planet on a 7.1 day orbit, and a 2.67 $R_{\oplus}$ planet on a 20.5 day orbit. Although they do not have measured masses, previous radial velocity and transit timing studies have found that even mature planets in this size range typically have low densities consistent with the presence of volatile-rich envelopes (e.g., Rogers 2015; Wolfgang & Lopez 2015; Hadden & Lithwick 2017). If these planets do have hydrogen/helium envelopes, their young age and the high X-ray luminosity of their host star point to the likelihood of ongoing mass loss. HD 63443b, in particular, is at an orbital period where there are more super Earths (1-1.5 $R_{\oplus}$) than mini Neptunes (2-3 $R_{\oplus}$), as shown in Figure 6 of Fulton & Petigura

(2018). Although it may currently have a gaseous envelope, this envelope is likely to be stripped away, moving it into the super Earth population.

HD 63433 is a uniquely favorable target for mass loss studies. Its young age, high activity, negative radial velocity, and close-in mini Neptunes provide ideal conditions for probing mass loss in the most critical regime. The existence of two mini Neptunes in the same system allows us to test hydrodynamical models by comparing their predictions for the two planets to observations: a comparative approach that has hitherto been impossible. No closer transiting mini Neptune host younger than 1 Gyr is known, let alone one with the other desirable properties to boot.

To study this system, we marshalled a variety of space and ground telescopes to characterize the star's high-energy spectrum and look for absorption from the escaping upper atmospheres in the Lyα line and the helium line. We describe our observations and data reduction in Section 6.3, our analysis in Section 6.4, our modelling of the star in Section 6.5, and our modelling of the planetary exospheres in Section 6.6. After comparing models to observations and discussing the broader context of our work in Section 6.7, we conclude in Section 6.8.

6.3 Observations and data reduction

We characterize the extended atmospheres and corresponding present-day mass loss rates for both planets by measuring the wavelength-dependent transit depth when the planet passes in front of its host star. We observe transits of both planets in the Lyα line with the Space Telescope Imaging Spectrograph on *Hubble Space Telescope* (*HST*/STIS) (Woodgate et al., 1998), and in the 1083 nm helium triplet with the updated NIRSPEC on Keck (Martin et al., 2018). We then compare the measured absorption during transit to predictions from mass loss models for each planet. In order to create these models, we must have a good knowledge of the high energy spectrum of the host star, which drives the outflows in our models. We use XMM-Newton to characterize the X-ray spectrum of the star, and estimate its extreme UV spectrum using scaling relations based on the reconstructed stellar Lyα emission flux. We also use archival data from ROSAT, which observed the star in 1990 as part of the ROSAT All-Sky Survey, and optical monitoring data from the T3 0.40m Automatic Photoelectric Telescope (APT) at Fairborn Observatory, to characterize the star's long-term variability and activity cycle.

***HST*/STIS**

With *HST*/STIS, we obtained two 9-orbit transit observations of the Lyα line with the MAMA detector (program 16319, PI: Michael Zhang). Because the South Atlantic Anomaly (SAA) prevents more than 5-6 consecutive orbits of observation, we observe as many orbits as we can in the vicinity of the transit, take no data for the 7-9 South Atlantic Anomaly-affected orbits, and observe the remaining 3-4 orbits after the gap. All orbits except the first contain 2523 s of science exposure time in TIME-TAG mode using the G140M grism with a central wavelength of 1222 Å and a slit width of 52 x 0.1″. The first orbit in the pre-gap and post-gap segments contain only 1515 s of science exposure time because target acquisition and acquisition peak-up occur during these orbits. In all orbits, wavelength calibration occurs after the science exposure, during occultation.

On Oct 29/30, 2020 UTC, *HST* observed a transit of planet c, with 5 orbits near transit and 4 orbits after the gap. On Jan 28/29, 2021 UTC, it observed a transit of planet b with the same configuration. On Mar 19, 2021, it attempted to observe six consecutive orbits bracketing a second transit of b, but STIS entered safe mode before the observations were to begin and all data were lost. On the following day, it successfully observed a 3-orbit baseline.

For our analysis of these data, we rely on `stistools` 1.3.0, a Python package provided by Space Telescope Science Institute that contains several relevant functions for data reduction. We start with the tag files, lists of photons that encode the time of arrival and position on the detector. First, we use `inttag` to turn the photon lists into raw images by accumulating the photons into 315 second subexposures (303 s for the first orbit). The first orbit of every visit contains 5 subexposures, while subsequent visits contain 8. The photon wavelengths are Doppler corrected prior to accumulation to account for *HST*'s orbit around the Earth.

Second, we use `calstis` to perform standard data reduction tasks, including subtracting the dark image, flat fielding, rejecting cosmic rays, and wavelength calibration. Wavelength calibration is performed using the wavelength calibration files taken during occultation, which contain lamp lines but no astrophysical signal. These files allow `wavecal`, a component of `calstis`, to assign a wavelength to every pixel.

The last step is spectral extraction. `x1d` attempts to locate the spectrum in the spatial direction but often fails because it excludes the region around the Lyα line to remove geocoronal emission—a process which also removes almost all stellar flux. Instead,

we locate the spectrum ourselves by summing the columns in each row, subtracting a smoothed version of the sums to remove skyglow variations, and fitting a Gaussian to the 30 pixels closest to the peak. After receiving the spectrum location, `x1d` sums up the pixel values in extraction windows 1 pixel wide and 11 pixels high, centered on the computed trace location. To compute the background, it uses two 5 pixel high windows, 40 pixels from the trace on either side. Unlike the spectral extraction window, the background extraction windows are tilted to account for the tilt of the iso-wavelength contours. Finally, `x1d` subtracts the background from the gross flux to get the net flux.

After `x1d` extracts the spectrum for every subexposure, we interpolate the spectrum onto a common wavelength grid for all subexposures. The grid has a linear spacing of 0.053 Å, matching the pixel scale of the detector.

Keck/NIRSPEC

Our NIRSPEC/Keck data (program C261) were taken on Dec 30, 2020 (transit of planet c) and Jan 7, 2021 (transit of planet b). All observations were in *Y* band in the high resolution mode. On Dec 30, the transit was in progress when our observations started. We collected 3.4 h of in-transit observations and 5 hours of post-transit baseline. On Jan 7, the transit started toward the end of our observations. We collected 5.4 h of pre-transit and 2.6 h of in-transit observations. On both nights, we used the 12 x 0.432″ slit, giving the spectrograph a resolution of 25,000. The sky was clear, and the seeing (1–1.5″) was poor but typical of this time of year. In addition, the telescope suffered from wind shake on Dec 30, further broadening the line profile in the spatial direction. We achieved a typical SNR of 400 per spectral pixel in 60-second exposures, all taken in the ABBA nod pattern to eliminate background. Because we only used one coadd per exposure, we achieved a high observation efficiency of 77%.

We calibrated the raw images and extracted 1D spectra for each order using a custom Python pipeline designed for the upgraded NIRSPEC. The pipeline is described in detail in Zhang et al. (2021), but we summarize it here. First, we subtract crosstalk from each raw frame. Then, we create a master flat, identifying bad pixels in the process. We use this master flat to compute a calibrated A-B difference image for each A/B pair. After identifying the spectral trace containing the 1083 nm lines, we use optimal spectral extraction to obtain 1D spectra along with their errors. We create a template from model telluric lines and a model stellar spectrum, shifted

in wavelength to account for the star's average Earth-relative radial velocity during that night. We then use this template to derive the wavelength solution for each individual spectrum.

After extracting the 1D spectra, we place the data from each night on a uniform wavelength grid and remove signals not related to the planet. We do this using SYSREM, which can be thought of as Principal Component Analysis with error bars (Mazeh et al., 2007). After removing the first principal component, we shift to the planetary frame, divide the data up into in-transit and out-of-transit portions and compare the portions to search for planetary absorption. Removing more principal components worsens the self-subtraction problem, already severe with a single component (see Section 6.4).

XMM-Newton

On March 26 2021, XMM-Newton observed the star for 6 ks (XMM prop. ID 088287, PI: Michael Zhang). XMM-Newton has 3 European Photon Imaging Camera (EPIC) detectors with different technologies (2 MOS, 1 pn), 2 Reflection Grating Spectrometers, and an Optical Monitor, all of which observe the target simultaneously. We configured the EPIC cameras to observe with the medium filter and small window, giving us 97% observing efficiency on the two MOS CCDs and 71% efficiency on the one pn CCD. These observations measure the star's X-ray spectrum, which plays an important role in driving photoevaporative mass loss. We configured the Optical Monitor to observe the star in the UVM2 filter ($\lambda = 231 \pm 48$ nm) for 2.7 ks and the UVW2 filter ($\lambda = 212 \pm 50$ nm) for 2.9 ks. These observations measure the star's mid ultraviolet flux, which can photoionize metastable helium but not create it, and therefore tend to decrease helium absorption in the metastable 1083 nm line. Although these observations are not simultaneous with the Lyα and helium mass loss observations, they are within 6 months, while the P_{cyc}/P_{rot} vs. P_{rot} relationship derived by Suárez Mascareño et al. (2016) implies a stellar cycle period of 5.5 years (with $\sim 40\%$ uncertainty).

To analyze XMM-Newton data, we download the raw Observation Data File (ODF) and use the Science Analysis System (SAS)[1] provided by the XMM-Newton team to reduce it. We run `xmmextractor`, thereby going from ODF to spectra with default settings and no human intervention. For the Optical Monitor, SAS produces the light curve of the star in the UVW2 and UVM2 filters, the two mid ultraviolet filters

[1] `https://www.cosmos.esa.int/web/xmm-newton/sas`

we selected. For each of the two Reflection Grating Spectrometers (RGSs), SAS produces two spectra (first and second order) and other data products, which we do not use because RGS has only a tenth of the throughput of the EPIC detectors in addition to substantial background; no stellar signal is visible in the data. For each of the three EPIC detectors, SAS generates the light curve, the background-subtracted spectrum, the Redistribution Matrix File (RMF), and the Ancillary Response File (ARF). The ARF gives the effective area of the detector as a function of photon energy, while the RMF gives the probability of a photon being detected in each channel as a function of photon energy. In optical astronomy terminology, the ARF gives the throughput multiplied by aperture area, while the RMF gives the wavelength-dependent line spread profile.

With XMM, as with ROSAT, the RMF is nearly diagonal for high energies, but is far from diagonal at low energies, where most of the stellar flux resides. In addition, the ARF is highly energy dependent for the two EPIC MOS detectors, although not for the pn-detector. These factors mean it is impossible to simply plot the measurements and see what the X-ray spectrum looks like. Rather, it is necessary to have a forward model and fit the parameters to find the best match to the data, taking into account the RMF and ARF. To do this fitting, we use the interactive tool `xspec` 12.11.1 (Arnaud, 1996).

ROSAT

To analyze the ROSAT All-Sky Survey (RASS) data for HD 63433, we download the data for the relevant sector[2] and reduce it using HEAsoft 6.28[3] by following the guide "ROSAT data analysis using `xselect` and FTOOls."[4] We define the source region to be a circle centered on the source with a radius of 200 arcseconds. Following the advice of Belloni et al. (1994), we define two circular background regions on either side of the source along the scan direction, both 800 arcseconds away and with a radius of 200 arcseconds. Using `xselect`, we extract the source spectrum and the background spectrum from the events list. We download the RMF for the PSPC-C detector[5] and use `pcarf` (part of the ROSAT subpackage of FTOOLS) to generate the ARF. The image file has negative EXPOSURE, DETC (deadtime correction),

[2] `https://heasarc.gsfc.nasa.gov/FTP/rosat/data/pspc/processed_data/900000/rs931219n00/`

[3] `https://heasarc.gsfc.nasa.gov/docs/software/lheasoft/`

[4] `https://heasarc.gsfc.nasa.gov/docs/rosat/ros_xselect_guide/xselect_ftools.html`

[5] `https://heasarc.gsfc.nasa.gov/docs/rosat/pspc_matrices.html`

and ONTIME header values to indicate that there is no unique value: the image is pieced together from scanning observations, and the effective exposure time is different at each pixel. We use the exposure map to determine the correct exposure time (437 s), and set the correct EXPOSURE, DETC, and ONTIME on the source and background files. Finally, we use `xspec` to load the source, background, RMF, and ARF, and analyze the data in the same way as the XMM observations.

6.4 Analysis of transit data

New ephemerides

As part of a CHEOPS Guaranteed Time Observation (GTO) program CH_PR100031, one transit of each planet was observed on UT Dec 9/10, 2020 (c) and UT Nov 25/26, 2020 (b). We combine these data with sector 20 TESS observations taken from December 24, 2019, to January 20, 2020 to refine the ephemeris so that we can predict the transit midpoint to ~1 minute accuracy at the time of our *HST* and Keck observations. The new ephemeris is significantly more precise than the old TESS-only estimate, which had an accuracy of ~30 min during the epochs of our *HST* and Keck observations. For c, we obtain $T_0 = 2458844.05824 \pm 0.00048$ BJD and $P = 20.543888^{+0.000046}_{-0.000045}$ d. For b, we obtain $T_0 = 2458916.45142^{+0.00030}_{-0.00032}$ BJD and $P = 7.107789 \pm 0.000010$ d.

Lyα absorption during transit

We first examine the UV data to search for signs of Lyα absorption during the transits of planets b and c. In Figure 6.1, we show the spectral sequence from each *HST* visit. For planet c, a clear decrease in the blue wing flux can be seen during the planetary transit. For planet b, the blue wing does not appear markedly different during transit. The red wing is more than four times dimmer than the blue wing, making it harder to see any planetary absorption in these 2D plots.

Previous studies have established that *HST* Lyα observations exhibit modulations in flux within each orbit, which has been attributed to telescope breathing (e.g. Kimble et al. 1998; Ehrenreich et al. 2015). According to Kimble et al. (1998), thermal variations over the course of each spacecraft orbit move the secondary mirror, which changes the focus, which leads to 10-20% variations in slit loss for the smallest slits. However, 10-20% variations in flux are observed for Ly α data taken in the 0.05″ slit (e.g. Ehrenreich et al. 2015), the 0.1″ slit (e.g. most observations in Lavie et al. 2017), and the 0.2″ slit (e.g. Garcia Muñoz et al. 2020). This insensitivity to slit

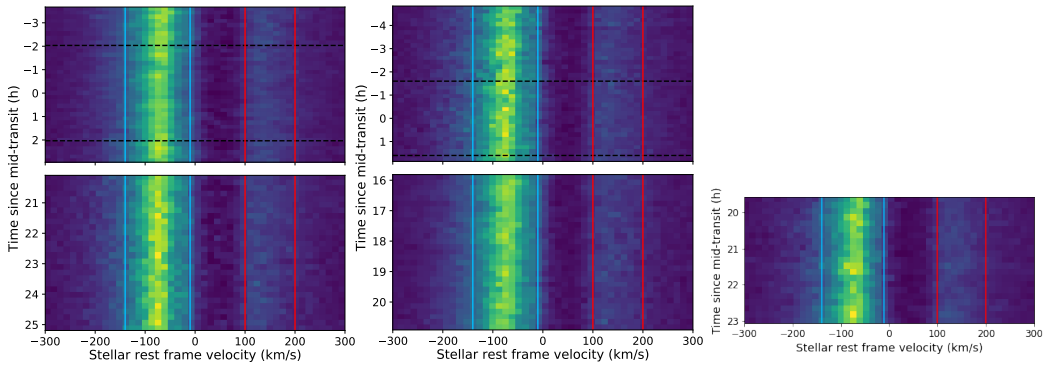


Figure 6.1: *HST*/STIS spectra for planet c (left) and planet b (center and right for the earlier and later observations). Flux increases from dark blue to yellow. The white horizontal gap represents the 7–9 orbit hiatus due to the South Atlantic Anomaly (SAA) crossing. Gaps due to Earth occultation are not shown to avoid visual distraction; the y axis is therefore not strictly accurate because it assumes time increases uniformly from spectrum to spectrum. The horizontal dashed black lines indicate the beginning and end of the white light transit. The region between the blue (red) vertical lines is what we define as the blue (red) wing. They correspond to [−140,−10] km/s and [100,200] km/s.

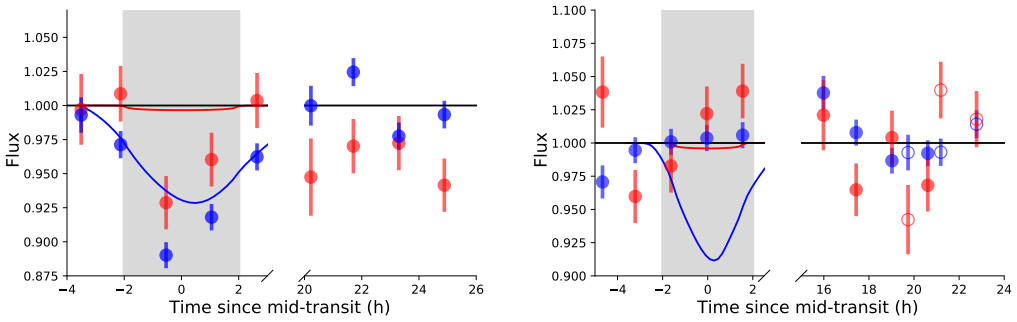


Figure 6.2: Lyα light curves for planet c (left) and b (right), in the red and blue wings, compared to the predictions from our 3D hydrodynamic model. For b, the solid circles represent the first visit, and the open circles represent the failed second visit. The error bar for each orbit is its photon noise. The grey region represents the white light transit.

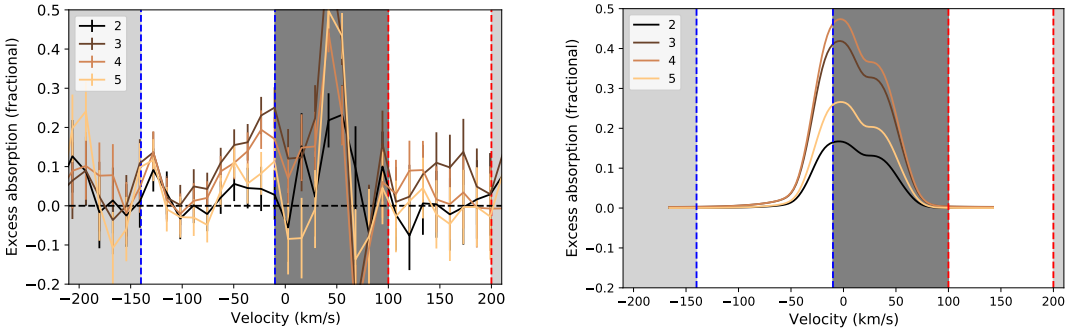


Figure 6.3: Per-orbit excess absorption spectrum for planet c, in the observations (top) and the fiducial model (bottom). The first orbit is used to compute the out-of-transit spectrum. The white regions are the blue and red wings, the light grey regions are the far wings, and the dark grey region has low flux because of interstellar absorption. The second and third orbits, in addition to the second half of the second orbit, are within the white light transit.

size indicates that intra-orbit flux variations are probably not the primary cause of the breathing effect.

We attempt to correct for the instrumental flux variations by decorrelating against a variety of different variables, including time since beginning of orbit, the centroid position of the blue wing, the latitude and longitude of the telescope, and the focus of the telescope as estimated from temperature sensors[6]. However, we find that none of these instrumental noise models can reduce the scatter in our light curves. Although there is clearly a correlation between flux and orbital phase in our observations, the shape of this trend varies from orbit to orbit. The measured flux increases steeply with time for the first orbit in a visit, but the increase becomes less pronounced in future orbits, until it becomes flat in the fourth or fifth orbits. As a result, we cannot remove this effect by detrending with a simple function of orbital phase as other papers do (e.g. Bourrier et al. 2013).

Since we are unable to effectively remove these intra-orbit flux variations, we instead bin our light curves into a single point for each spacecraft orbit. Figure 6.2 shows the resulting light curves for the integrated red and blue wings for all three transit observations. We calculate a photon noise of 1.4% for the blue wing during the first orbit and 1.0% for subsequent orbits; for the red wing, it is 2.8% for the first orbit and 2.0% for subsequent orbits. We conservatively adopt the first-orbit error for

[6] `https://www.stsci.edu/hst/instrumentation/focus-and-pointing/focus/hst-focus-model`

all orbits, since instrumental systematics and stellar variability undoubtedly inflate the noise beyond the photon limit. Using these error bars, we find that the excess absorption during the transit of planet c is 11.1 ± 1.5% in the blue wing and 8 ± 3% in the red wing. For planet b, we place a 2σ upper limit on the in-transit absorption of 3% in the blue wing and 4% in the red wing. Incidentally, the standard deviation of the blue fluxes for all orbits other than the 5 bracketing the transit of c is 1.8%; the standard deviation of the red fluxes for these same orbits is 3.1%. This indicates our inflated error bars of 1.4% and 2.8% are not far from the mark.

We illustrate the wavelength dependence of the absorption from planet c by plotting the excess absorption spectrum, $1 - F/F_{\rm out}$, for each orbit in Figure 6.3. Initially, we tried using the post-SAA segment of the observations for the out-of-transit spectrum. However, we noticed that this introduced significant correlated noise into the excess absorption spectrum, which we attributed to changes in the intrinsic stellar spectrum in the 18 hours between the two segments. This variability can also be seen in the red wing light curve in Figure 6.2, which shows that the red wing is notably dimmer in the post-SAA segment than in the out-of-transit orbits of the pre-SAA segment. Unfortunately, this variability makes the post-SAA orbits much less useful as a baseline than we had originally hoped. We considered using the average of the first and fifth orbits for the out of transit spectrum, but the blue wing light curve shows that the fifth orbit might contain planetary absorption. We therefore opted to use the first orbit for the out of transit spectrum, but note that this first orbit often has an anomalous flux level when compared to later orbits, although it is still a better baseline than the average of the post-SAA spectra. The resulting plot provides a useful illustration of the progression of the excess absorption spectrum from orbit to orbit, but we should not place too much weight on the absolute value of each spectrum.

Examining Figure 6.3, we see that the excess absorption is highest in the region of the blue wing near −10 km/s (i.e., closest to the line center). The absorption in this region increases steadily from each orbit to the next until orbit 3, after which it decreases in orbit 4, and decreases again in orbit 5, but does not decrease to 0. The excess absorption in orbit 3 is 26 ± 6%, more than double the wing-integrated excess absorption of 11%. Reassuringly, the excess absorption spectrum decreases blueward of −10 km/s until it is indistinguishable from 0 at −100 km/s. It is concerning that the excess absorption in the red wing is highest around 150 km/s, not at the low-velocity edge of 100 km/s where we might expect to see it. However,

the magnitude of the measured absorption in the red line is lower (8% vs. 11%) and the noise is much higher, making the data statistically consistent with a flat or slightly declining absorption spectrum in the red wing.

Stellar variability in Lyα

Since HD 63433 is a young star, it is important to determine whether stellar variability could explain the absorption signal. Only the visit containing the transit of planet c shows significant blue wing variability, which we are ascribing to planetary absorption. Among the other 4 visits, which contain a total of 16 orbits, the blue wing is remarkably stable, with a standard deviation of 1.8%. The only outlier is the first orbit of the second visit of the successful b observation. However, the first orbit in a visit is expected to be more variable than the others. The science exposure is shorter, giving rise to 30% higher photon noise. It also starts at a later *HST* orbital phase, violating our logic for looking at orbit-aggregated data points—namely that repeatable systematics that depend on orbital phase will be averaged out. Finally, it takes *HST* approximately one orbit to thermally relax after pointing to a new target, and the first orbit of STIS exoplanet observations at optical wavelengths is routinely discarded because of the higher systematics (e.g. Huitson et al. 2013; von Essen et al. 2020).

We can place this level of stellar variability in context by comparing to observations in the literature. Llama & Shkolnik (2016) used disk-resolved Lyα images of the Sun to estimate that for stars with solar activity levels we would expect to see the measured transit depth vary by 0.8% due to activity-induced measurement error, and are unlikely to see more than 1.5% variability. However, HD 63433 is more active than the Sun. Following Kulow et al. (2014), we examine observations of the CII line by Loyd & France (2014). The CII line has a formation temperature similar to that of Lyα, making it a good tracer of variability in this line. We compare to observations of Pi UMa, a G1.5 star with a fast rotation period (P=4.89 d) which, like HD 63433, is a member of the Ursa Major Moving Group. Loyd & France (2014) found that the mean-normalized excess noise on 60 s timescales for this star was 3.2%. For the 28 Myr G1.5 star EK Dra, the excess noise was less than 1%; for the 13 Myr G1.5 star HII1314, it was less than 5.6%. If we assume that the stellar variability has a comparable magnitude on several hour timescales (i.e., the duration of a transit) we might expect HD 63433 to vary by a few percent. This

would suggest that it is unlikely that stellar variability caused the 11% decrease in brightness in the blue wing during the transit of planet c.

Lyα observations of exoplanet hosts are somewhat less encouraging. Bourrier et al. (2017b) saw a 20% dip in Lyα during the transits of sub-Earths Kepler-444e and f, as well as a 40% dip when no known planet was transiting. Although Kepler-444 is an old (11 Gyr) K star, the authors couldn't exclude the possibility that the observed variability was due to stellar activity. Bourrier et al. (2017a) observed HD 97658, an old K star, with STIS over three visits (15 orbits in total) and found that, during the second visit, the Lyα flux declined by 20% over a period of several hours. This decline did not coincide with the white light transit and did not have a clear transit-like shape. It is unclear whether these variations are due to stellar variability or instrumental artifacts.

After considering the totality of the evidence, we conclude that the blue wing absorption from planet c is very likely to be planetary. It occurs at the expected time and becomes stronger as one approaches the core of the line, in accordance with physical expectations. In every other visit, the blue wing flux is remarkably stable. Our *HST* program will observe a second transit of planet c to see if the signal re-appears, which would provide a definitive confirmation of its planetary origin. Unfortunately, due to an alignment between c's orbital period and the visibility period imposed by the South Atlantic Anomaly, the next observing window is unlikely to occur before 2023.

The red wing absorption detection is more tentative. The red wing is much fainter than the blue wing, and has a correspondingly high level of photon noise. The out-of-transit variability in this wing appears to be higher, and the post-SAA visit for planet c is almost as low as the lowest in-transit data point. The excess absorption spectrum also appears to rise toward higher velocities, in contravention of theoretical expectations, although the rise is not statistically significant. On the other hand, the timing and shape of the transit light curve is strikingly similar to that of the blue wing. We consider it more likely than not that the red wing absorption is real, but without a second transit observation, the detection remains tentative at best.

A search for absorption in other UV lines

We can use these same spectra to search for planetary absorption in the Si III line at 1206.5Å and the two N V lines at 1238.8Å and 1242.8Å. For c, we see a marginal transit-like signal of 6 ± 3% in Si III but no transit-like signal in N V (1.5 ± 4%). The

Si III line is known to be highly variable (dos Santos et al., 2019), and we measure a relative flux of 0.89 ± 0.02 for this line in the last orbit, 11% lower than in the preceding orbits. For planet b, the Si III line is even more variable, and we see no transit-like feature. We calculate an excess absorption of $1.5 \pm 2.5\%$ in Si III and $-1.5 \pm 4\%$ in N V.

Independent analyses of the Lyα data

The fiducial analysis reported above was performed by the first author (Michael Zhang, MZ). Two separate analyses were performed by co-authors Luca Fossati (LF) and Leonardo dos Santos (LDS) using independent pipelines. There was no communication between the co-authors during these independent analyses other than to agree on a common velocity range in which to look for absorption: $[-140,-10]$ km/s in the blue wing and [100,200] km/s in the red wing. The results of these independent analyses are plotted in the Appendix (Section 6.11). All three analyses show a clear blue wing absorption signal from c, and no red or blue wing absorption from b. The alternative analyses show no red wing absorption from c. This is likely due to their 3x higher scatter (see Appendix), but the non-detection of red wing absorption in these alternative analyses nevertheless underscores the tentative nature of the detection in the fiducial analysis.

Helium absorption during transit

We next examine the Keck data to search for signs of helium absorption during the transits of planets b and c. After the processing steps described in Subsection 6.3, we are left with a residuals image of size N_{epochs} by $N_{wavelengths}$. Each pixel in the residuals image approximately represents the fractional flux change at that epoch and wavelength from the mean spectrum. We shift these spectra to the planetary frame, combine all out-of-transit spectra into a master out-of-transit spectrum, combine all in-transit spectra to a master in-transit spectrum, and subtract the master in-transit spectrum from the master out-of-transit spectrum. Figure 6.5 shows the resulting excess absorption spectrum for each planet.

We see clear evidence of stellar activity in Figure 6.4, and in the line-integrated line fluxes (Figure 6.6). The stellar 10833 Å helium lines, which trace chromospheric activity, are variable on both nights. During the transit of planet c on the first night, the lines experience a bump in brightness starting around 3 hours after mid-transit and then fall to even lower values after the bump. During the second night, the

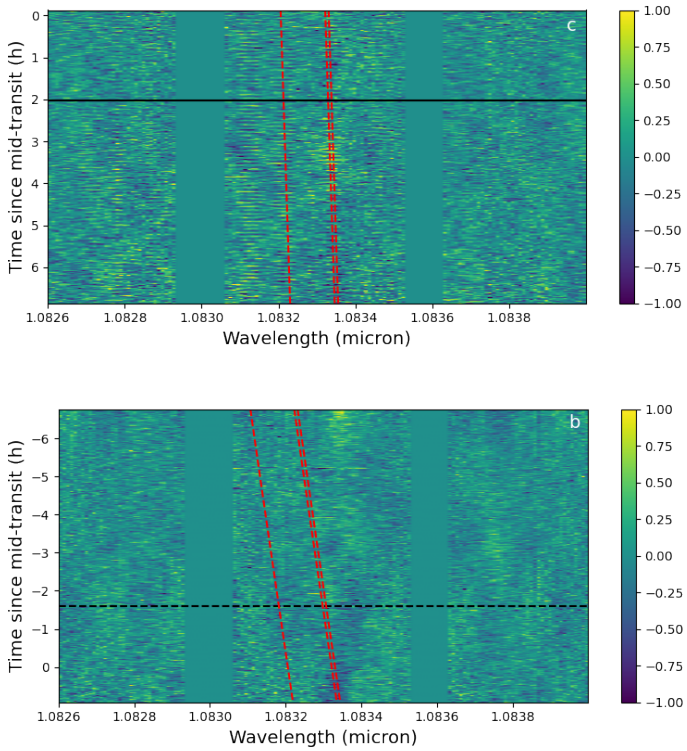


Figure 6.4: Residuals images after SYSREM, for planet c (top) and b (bottom). Colors indicate relative brightness change in percent. The dashed red lines indicate the wavelengths of the helium triplet in the planetary frame. The dashed horizontal black line indicates the beginning of transit, while the solid horizontal black line indicates the end of transit. Note the darkening during the transit of b, which we ascribe to stellar variability (see text).

helium lines exhibit a more complicated behavior: they start high, then decline for 2 hours, rise again, decline again, rise a third time, and finally decline to their lowest level over the night. The peak-to-peak amplitude of this variability is around 1% on the first night and 2% on the second night.

We conclude that the 0.2% excess absorption signal in the helium lines for planet c (Figure 6.5) is most likely due to stellar variability, not planetary absorption. It is apparent in Figure 6.4 that the excess absorption signal is caused by the darkening of the stellar spectrum around the 10833Å helium lines in the final hour of the observations. A close examination of the darkened portion of the spectrum shows that the darkening started, not at the beginning of the transit, but an hour afterwards. In addition, the darkened portion of the spectrum does not follow the radial velocity of the planet. On the contrary, it moves toward shorter wavelengths as time progresses. This is the opposite of what we would expect a planetary absorption

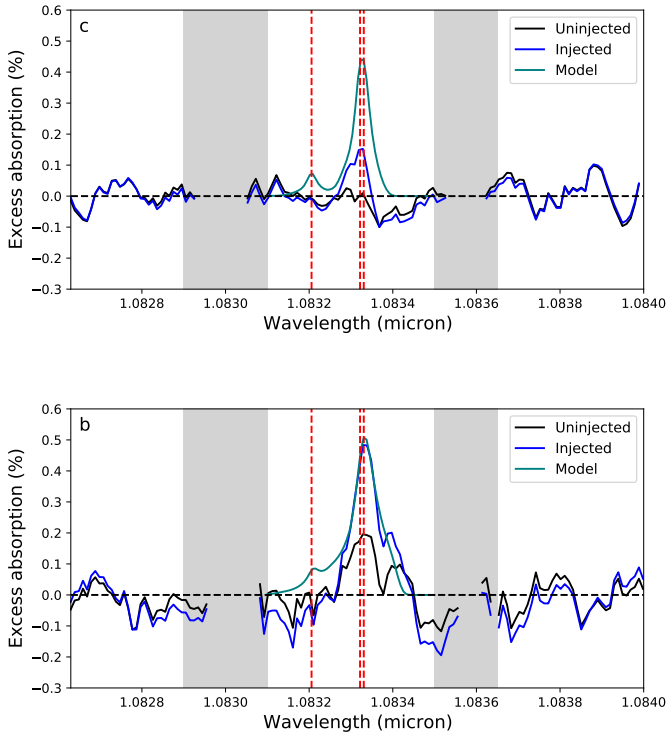


Figure 6.5: In black is the excess absorption spectrum in the planetary frame for c (top) and b (bottom). The vertical red lines indicate the positions of the helium triplet. The absorption from planet b is probably stellar variability (see text, and Figure 6.4). The injected models (teal) are from the 3D hydrodynamic simulations described in Section 6.6, and the observed excess absorption spectrum after injection is shown in blue. The grey regions are locations of strong stellar and telluric lines.

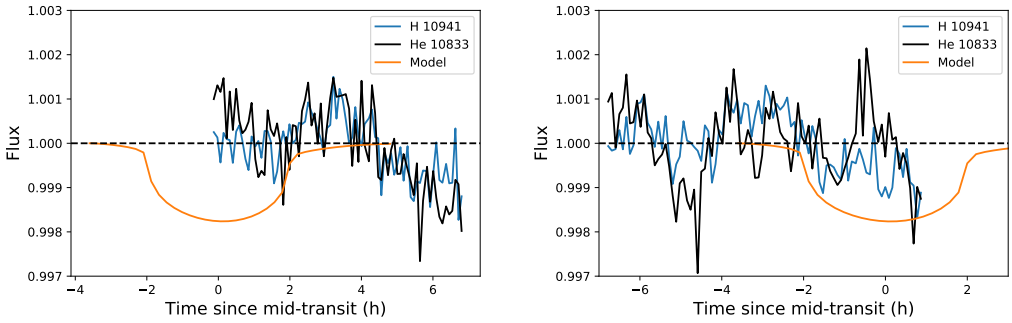


Figure 6.6: Light curves of the HeI 10833 Å and HI 10941 Å lines, integrated in a 3.2 Å bandpass, for planet c (top) and b (bottom). To normalize out non-stellar variations, the line fluxes are divided by the continuum at 10825 Å (HeI) or 10937 Å (HI), also integrated in a 3.2 Å bandpass. In orange we plot the predictions of the 3D hydrodynamic model.

signal to do. Finally, 0.2% is smaller than the observed amplitude of the variability in the stellar helium lines on both nights. Although we cannot rule out a planetary or hybrid planetary-and-stellar explanation for the observed absorption signal, the data are entirely consistent with a purely stellar explanation.

We next explore whether or not we might be able to model out some of this stellar variability using other chromospheric lines in our spectrum. The most promising candidates are two lines in the hydrogen Paschen series: γ (n=6 $\rightarrow$ 3, 10941 Å) and δ (n=7 $\rightarrow$ 3, 10052 Å). We use the data from the second night, which had more complex stellar behavior than the first, as our test case. We find that the γ line has a time-varying behavior similar to that of the helium lines (Figure 6.6), but with a lower overall amplitude. The δ line is also variable, but this variability does not appear to be correlated with the variability in the helium lines. Ultimately, neither line displayed a strong enough correlation with the helium lines to enable an effective correction for stellar activity. However, the similar behaviors of the Paschen γ and helium lines provide additional support for our conclusion that the variability in Figure 6.4 is likely stellar and not planetary in origin.

We next consider what limits we can place on the magnitude of helium absorption during the transits of planets b and c. Planet c barely accelerates during its transit, making it hard to disentangle planetary signals from the stellar and telluric variability that SYSREM is meant to subtract. As a result, we expect significant self-subtraction from our analysis pipeline. Planet b accelerates more and should experience less self-subtraction. Figure 6.5 illustrates this phenomenon: after injecting an artificial helium absorption signal, our measured excess absorption spectrum is only 35% the size of the injected signal (i.e., 65% self-subtraction). For planet b, the injection-recovery test indicates a self-subtraction of 40%.

Due to stellar variability, we cannot assume statistical independence between epochs and use the standard statistical methods to compute an upper limit on the helium excess absorption. We can, however, arrive at a reasonable guess by examining the observed stellar variability during each of the two nights and its corresponding effect on the excess absorption spectrum. The helium lines never deviate by more than 1% from the median on either night, and even if they did, it is unlikely that the stellar variability would line up with the planetary transit. For planet b, the acceleration of the planet is significant enough to place the planetary absorption lines outside of the stellar lines at the beginning of the transit, further decreasing the impact of stellar variability. A 1% planetary absorption would be reduced to

0.5% due to self-subtraction, but this would still be readily detectable in the excess absorption spectrum in Figure 6.5. We test this by injecting a planetary signal with an amplitude of 1% into the data, running it through the pipeline, and examining the intermediate outputs. Even in the pre-SYSREM stage (before any self-subtraction happens), the planetary signal is clearly visible above the amplitude of the stellar variability. The planetary signal remains obvious when we reduce the peak excess absorption of the injected signal to 0.5%, but its final amplitude in that case would be comparable to the amplitude of the stellar variability during the transit of planet c. We therefore conclude that the peak excess absorption must be less than 1% for both planets with high confidence, and less than 0.5% with medium confidence.

6.5 Understanding the star

In order to model mass loss, we need to know the star's intrinsic spectrum at high energies. Heating from the X-ray and extreme UV (EUV) flux drives the outflow in our models, while the stellar Lyα line profile, in combination with the instrumental line spread profile, is necessary to predict the observed absorption using the models. For the star's X-ray spectrum, we use the XMM-Newton data described in §6.3, and combine these data with older observations at X-ray and optical wavelengths to characterize the star's long-term variability and activity cycle. For the star's UV spectrum, we use a combination of scaling relations and, when appropriate, observations of the Sun's UV spectrum.

X-ray spectrum and stellar variability

XMM-Newton spectrum

We analyze the XMM-Newton EPIC observations using the `xspec` package. To get the underlying X-ray spectrum, we fit a model equal to the sum of two APEC emission models. These models (Smith et al., 2001) assume an optically thin, collisionally ionized plasma with the temperature, metallicity, redshift, and normalization as free parameters. We fix the redshift to 0 (EPIC's velocity resolution is >6000 km/s), require the two model components to have the same metallicity, and let the two temperatures, two normalization factors, and the global metallicity vary freely. We fit the data by minimizing the W statistic, the analogue of χ^2 for a distribution corresponding to the difference of two Poisson distributions (namely source and background).

Figure 6.7 shows the EPIC data, the best fit to the data obtained by `xspec`, and the intrinsic spectrum implied by the best fit parameters. We ran a Markov Chain Monte

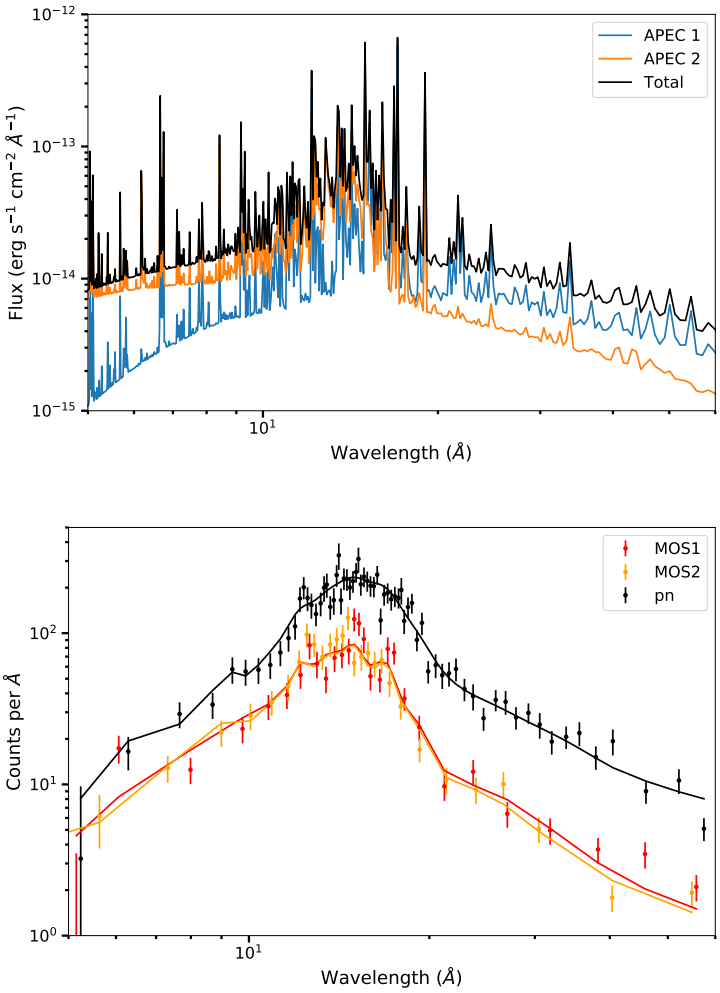


Figure 6.7: X-ray observations. Top: Best-fit intrinsic stellar X-ray spectrum. The total spectrum as observed from Earth (black) is the sum of a low-temperature component (APEC 1) and a high-temperature component (APEC 2). Bottom: the XMM EPIC spectra measured by the three detectors (MOS1, MOS2, pn) and the folded models, which take into account the instrumental RMF and ARF (the X-ray equivalents of LSF and throughput). For clarity, the data are binned so that each bin, except those on the edges, contains at least a 3σ detection.

Table 6.1: Model parameters for XMM-Newton data

Parameter	Value
Metallicity	0.44 ± 0.07
kT_1 (keV)	0.38 ± 0.02
EM_1 (cm^{-3})	$3.8 \pm 0.5 \times 10^{51}$
kT_2 (keV)	0.83 ± 0.04
EM_2 (cm^{-3})	$2.1^{+0.3}_{-0.4} \times 10^{51}$
Flux* (erg/s/cm^2)	$1.25^{+0.03}_{-0.02} \times 10^{-12}$

*Derived, not a fit parameter. For the range 5-100 Å (0.124-2.48 keV).

Carlo (MCMC) fit using the Metropolis-Hastings algorithm and a chain length of 10,000 to estimate the range of parameters consistent with the data. We plotted the chain to ensure convergence, which occurred within the first 1000 samples. Table 6.1 shows the resulting 1D MCMC posteriors. The metallicity is reported with respect to the solar abundances of Asplund et al. (2009). We find a dominant component with a temperature of 0.38 keV (4.4 MK), with a slightly sub-dominant component at 0.83 keV (9.6 MK). The emission measures are on the high end compared to other moderately active G8-K5 dwarfs (Wood & Linsky, 2010), but of the same order of magnitude. The best-fit metallicity is subsolar, and because the photospheric metallicity of the star is roughly solar ([M/H] = -0.09 ± 0.08; Mann et al. 2020), it is substellar as well. This is reminiscent of the findings of Poppenhaeger et al. (2013), who observed the moderately active K dwarf HD 189733 A with Chandra. In that study they fit the O, Ne, and Fe abundances separately, obtaining values of 0.31 ± 0.02, 0.25 ± 0.13, and 0.64 ± 0.05 (relative to solar), respectively. We carried out an analogous fit where we allowed the O, Ne, and Fe abundances to vary freely and obtained 0.67 ± 0.1, $0.22^{+0.40}_{-0.05}$, and $0.97^{+0.05}_{-0.18}$, respectively. These results are due to an effect called the first ionization potential (FIP) bias: for many inactive and moderately active stars, elements with high FIP (e.g. C, O, N, Ne) are depleted in the corona compared to low FIP elements (e.g. Mg, Si, Fe). FIP bias was seen by Wood & Linsky (2010) in 5 out of 7 moderately active G8-K5 dwarfs. In these stars, the coronal abundances of C, O, N, and Ne were all lower than the photospheric abundances. FIP bias is also seen in the solar corona, although there, low FIP elements are enhanced by a factor of ~ 3 and high FIP elements generally have photospheric abundances (Feldman & Widing, 2002). It has been suggested that the FIP bias may arise from wave ponderomotive forces on the upper chromosphere (Laming, 2004, 2017). For our purposes here, it is sufficient to reconstruct the

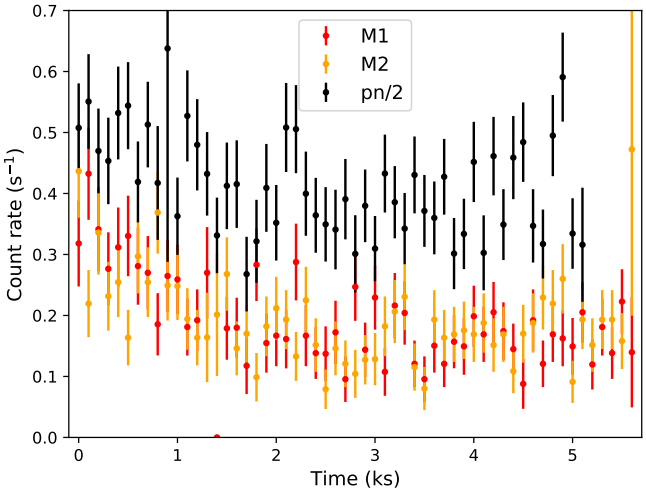


Figure 6.8: Background-subtracted X-ray light curves of HD 63433 recorded by the three EPIC cameras (MOS1, MOS2, pn). The pn light curve is divided by 2 for clarity.

intrinsic X-ray spectrum of the host star and we therefore leave further analysis of the FIP bias to interested stellar astronomers.

Our fit to the EPIC data constrains the time-averaged 5–100 Å (0.124-2.48 keV) flux to $1.25^{+0.03}_{-0.02} \times 10^{-12}$ erg/s/cm^2. Restricting the range to the observable range of both XMM and ROSAT, we find a 0.2–2.4 keV flux of $1.12 \pm 0.02 \times 10^{-12}$ erg/s/cm^2. These error bars are deceptively small, as the X-ray spectra of active stars can vary significantly in time. Figure 6.8 shows the X-ray light curve captured by the three EPIC cameras. HD 63433 is brighter in the first 2 ks of observation than in the remaining 4 ks, with the pn-flux declining by 25% and the two MOS fluxes declining by 50%. The difference in observed amplitude is likely due to the different characteristics of the two types of detectors. The MOS detectors' sensitivities drop off more sharply toward low energies (<1 keV), where most of the star's X-ray flux resides, than the pn-detector. The observed variability in the X-ray light curve means that, absent simultaneous observations, we cannot know the X-ray flux during the epoch of our hydrogen or helium observations to better than ~25%.

Long-term X-ray variability from ROSAT data

We evaluate the magnitude of the stellar X-ray variability over longer timescales using archival ROSAT data from 1990. These data have a much lower SNR than the XMM-Newton data, as a result of the lower effective area of the detector and

the shorter exposure time. Whereas XMM's EPIC cameras captured 3600 X-ray photons, ROSAT's PSPC-C captured only 86. As with XMM, we fit the data with two summed APEC emission models. In order to prevent the fit from wandering off to unphysical parts of parameter space, we fix the metallicity to the value derived from XMM and constrain kT_1 to lie between 0 and 0.5 keV and kT_2 to lie between 0.5 and 1.0 keV.

We find that the shape of the unfolded ROSAT spectrum is consistent with the shape of the unfolded XMM spectrum. We derive a 0.124–2.48 keV flux of 1.5–1.9 $\times 10^{-12}$ erg/s/cm^2 and a 0.2–2.4 keV flux of 1.25–1.71 $\times 10^{-12}$ erg/s/cm^2. This is in line with the flux reported by the Second ROSAT all-sky survey (2RXS) source catalog (Boller et al., 2016) for both a power-law fit (1.39 $\times 10^{-12}$ erg/s/cm^2) and a blackbody fit (1.77 $\times 10^{-12}$ erg/s/cm^2). We conclude that the star's X-ray flux appears to have been 33% higher during the ROSAT observation than during the XMM observation, but the two measurements are consistent at the 2σ level.

Long-term optical variability from the APT data

We use ground-based optical photometric monitoring data to evaluate where HD 63433 was in its activity cycle during the epochs of our *HST* and Keck observations. HD 63433 has been monitored since 1998 in the Johnson V and B photometric pass bands with the Tennessee State University T3 0.40 m Automatic Photoelectric Telescope (APT) at Fairborn Observatory in southern Arizona. Our observations cover 23 observing seasons from 1998-99 to 2020-21, although there are relatively few observations in the last three seasons due to APT scheduling changes and instrument problems. While the typical season runs from early October to late April, the 2019-2020 data only span March 15, 2020 to April 19, 2020, and the 2020-2021 data only span October 12, 2020 to November 25, 2020.

Our measurements of HD 63433 were made differentially with respect to the comparison star HD 64465 (HIP 38677; F5). A second star, HD 63432 (HIP 38231; A2), was used to check the stability of the relative photometry using HD 64465. Details of the robotic telescopes and photometers, observing procedures, and data reduction can be found in Henry (1999) and Fekel & Henry (2005). Figure 6.9 shows the Variable minus Comparison and Check minus Comparison APT light curves. The lower two panels show small variability in the Chk-Cmp light curves. Their seasonal means vary over a range of 0.006 and 0.008 mag in the V and B, respectively. However, the Var-Cmp seasonal means in the upper two panels vary

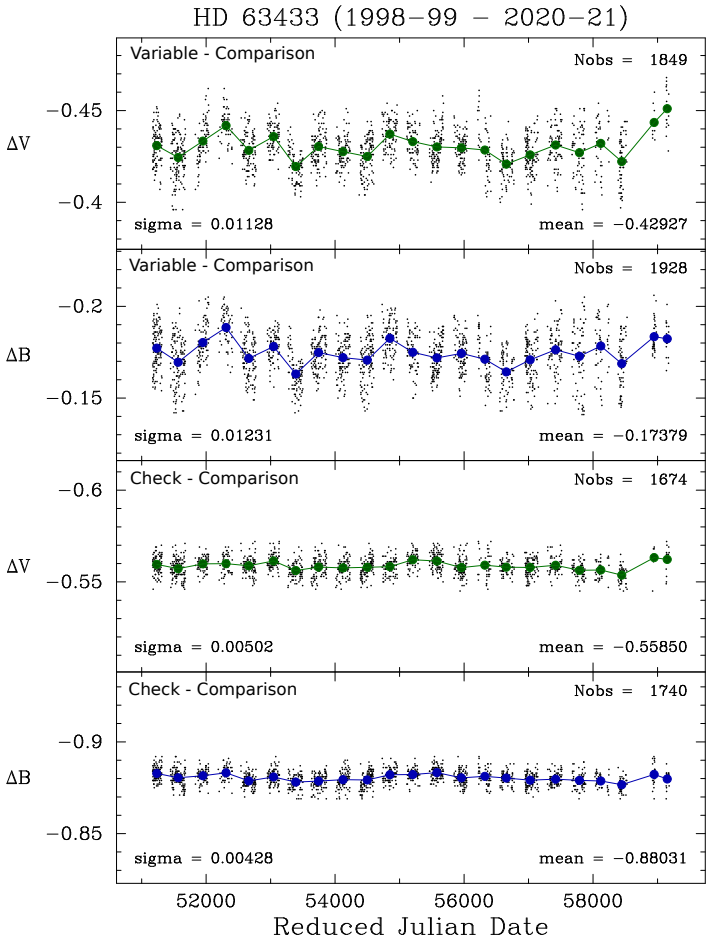


Figure 6.9: Light curve for HD 63433 ('Variable') in the Johnson V (green; top panel) and B (blue; upper middle) pass bands normalized using a comparison star. Nightly observations are plotted as small grey points, while the annual means are shown as larger colored circles. An equivalent light curve for a less active star, HD 63432 ('Check'), is shown in the lower middle and bottom panels to demonstrate the photometric stability of these data.

over a much larger range of 0.025 and 0.022 mag, demonstrating that most of the variability seen in the Var-Cmp light curves is intrinsic to HD 63433.

Keeping the data limitations in mind, the APT light curve does appear to show that HD 63433 was anomalously bright in 2020 and 2021 in both filters, with a flux 1-2% above the 20 year average in V. In fact, it appears to be brighter than at any point since observations began. For active stars ($\log R'_{\rm HK} > -4.7$), the V and B brightness varies inversely with stellar activity: the more active the star, the more spots it has, and the dimmer it appears (Lockwood et al., 2007). Therefore, the APT data suggest that HD 63433 was unusually quiescent during the time of our

mass-loss observations, consistent with the marginally lower X-ray flux observed by XMM-Newton in March 2021 compared to ROSAT in 1990.

We computed a Lomb-Scargle Periodogram of the data and found a clear, narrow peak corresponding to the rotation period of the star. Both the V and B data indicate $P_{rot} = 6.413 \pm 0.002$ d, slightly lower than the 6.46 ± 0.01 d reported by Gaidos et al. (2000) using the APT data available then. The APT-inferred rotation period is far more precise than the one inferred from TESS and K2 data (Mann et al., 2020) because of the much longer baseline (1200 rotation periods vs. 3-4). The periodogram also shows a broad double peak around 800–1100 d, perhaps indicative of a short stellar cycle. However, while there is clear non-random behavior in the light curve on a timescale of years, no periodic stellar cycle is obvious by inspection. HD 63433's photometric variability is typical of stars younger than 2–3 Gyr, which have complex interannual variations that are often composed of multiple cycles, compared to the simple cycles of older stars (Oláh et al., 2016).

UV spectrum

Lyα profile

In order to translate our mass loss models into a prediction for the Lyα light curve during transit, we need a measurement of the star's intrinsic Lyα profile. The STIS observations do not directly tell us the intrinsic Lyα profile because the the line core is absorbed by the ISM, and the instrumental line-spread profile smears out the remaining flux. We reconstruct the intrinsic profile from our data using a hierarchical Bayesian model implemented in `stan` (Stan Development Team, 2018).

In principle there are an infinite number of intrinsic profiles that can fit the data, because the flux at the core of the line is unconstrained. Therefore, we need to utilize a prior on the line shape in order to reconstruct the height of the line core using the flux in the wings. In a previous survey of stellar Lyα emission, Wood et al. (2005b) used the profile of the observed Mg II h and k lines (2796 and 2804 Å) as their template for the Lyα line shape. Unfortunately, we have no such data for HD 63433. Instead, we start with the reconstructed Lyα profile of HD 165185 from Wood et al. (2005b), a star with the same spectral type as HD 63433 and a similar rotation period. We then allow `stan` to modify the profile as follows:

$$F(\lambda) = F_{init}(\lambda) + \Delta_1(\lambda) \tag{6.1}$$

$$\Delta_1(\lambda) = \text{cumsum}(\Delta_2(\lambda)) \tag{6.2}$$

$$\Delta_2(\lambda) \sim N(0, 5 \times 10^{-14}), \tag{6.3}$$

where F_{init} is the HD 165185 profile, cumsum is the cumulative sum, and N(0, 5×10^{-14}) is a normal distribution with a mean of 0 and standard deviation of $5{\times}10^{-14}$. The intrinsic spectrum of HD 63433 is that of HD 165185 plus differences, the differences are in turn the cumulative sum of second differences, and we impose a Gaussian prior on the second differences with a standard deviation of 5×10^{-14} erg s^{-1} cm^{-2} Å^{-1}. These equations allow `stan` to modify the HD 165185 profile to fit the HD 63433 data, but not arbitrarily: it enforces continuity in the modifications made to the profile, and penalizes large changes to avoid overfitting. This process is mathematically equivalent to L2 regularization.

We can obtain an independent constraint on the magnitude of the interstellar absorption by using the parameters derived by Dring et al. (1997) for two stars. β Gem and σ Gem are 1.3 and 2.2 degrees, respectively, from HD 63433. Despite having very different distances (10.3 pc and 37.5 pc), the two have indistinguishable N(HI) of $10^{18.26}$ and $10^{18.20}$ cm^{-2}. This is because the region within 10 pc has an abnormally high neutral hydrogen fraction compared to the rest of the 100 pc Local Bubble (Wood et al., 2005b). Dring et al. (1997) found that the sightline to these stars can be modelled by assuming two clouds: one at 21.7 km/s with a column density of $10^{18.027}$ cm^{-2} and a HI Doppler parameter of 12.35 km/s, and another at 32.5 km/s with a column density of $10^{17.801}$ cm^{-2} and a Doppler parameter of 11.0 km/s. (In practice, Dring et al. (1997) fit the two sightlines separately, but we averaged the results here because they are remarkably similar.) The more strongly absorbing cloud has a velocity consistent with theoretical expectations. The local interstellar cloud (LIC) is moving at 25.7 km/s in the direction of l=186°, b=-16°, according to high-resolution observations of local stars (Lallement et al., 1995), in good agreement with the flow of ISM particles through the solar system ([l=183°, b=-16°] at 26.3 km/s) (Witte, 2004). Projecting this velocity along the line of sight, we compute a radial velocity of 19.9 km/s. Combined with the -16 km/s radial velocity of the star, the 36 km/s difference is what strongly suppresses the red wing while keeping the blue wing unusually intact.

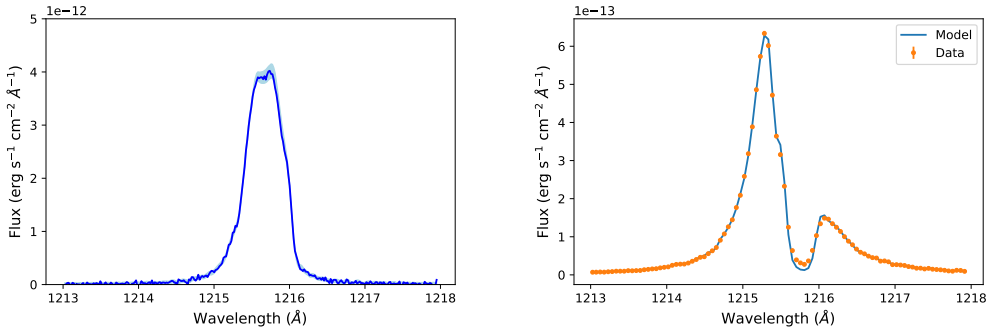


Figure 6.10: Stellar Lyα. Top: the reconstructed stellar Lyα profile (blue), with uncertainty indicated by the gray shading. Bottom: the model Lyα profile after interstellar absorption and convolution with the instrumental line spread profile (blue), compared to the data (orange). The data has error bars, but they are too small to see.

To summarize, our free parameters are the 282 second differences $\Delta_2(\lambda)$, while the interstellar absorption is fixed to that found by Dring et al. (1997). We run 10,000 iterations of 4 chains each, and check all five of the diagnostics provided by `stan` to ensure convergence: effective sample size, potential scale reduction factors, divergent transitions, percentage of transitions ending prematurely due to maximum tree depth, and E-BFMI (energy). We take the sample with the highest posterior probability to generate the fiducial Lyα profile. Figure 6.10 shows the resulting best-fit intrinsic Lyα profile. The model provides a close match to the data everywhere except at the center of ISM absorption, where the data are higher than the model. We speculate that this may be because the LSF provided by STScI does not have sufficiently strong wings. This was previously noted by Bourrier et al. (2017a), who fit their own LSF for the 52 x 0.05 arcsec slit.

We evaluate how sensitive the Lyα flux is to our choice of reconstruction method by repeating our analysis using a completely independent method similar to the one in Bourrier et al. (2017a). In this version, we model the stellar Lyα line as a Voigt profile, the instrumental line spread profile as a sum of two Gaussians, and the interstellar medium as a cloud with a single Gaussian velocity dispersion and velocity. We then use differential evolution to optimize the free parameters: the μ_*, σ_*, γ_*, and amplitude of the stellar line profile, the two standard deviations and one relative amplitude which characterize the LSF, and the velocity offset of the intervening cloud. The column density of the cloud is set to the sum of the column densities of the two clouds found by Dring et al. (1997), and the Doppler broadening

parameter is set to 12 km/s, very close to the values derived by Dring et al. (1997) for both clouds. With this method, we obtain a Lyα flux of 46 erg cm^{-2} s^{-1} at 1 AU, 18% lower than the fiducial value. Based on the fit error and Linsky et al. (2014), the Lyα flux we derive is probably not accurate to better than ~30%. The fit parameters are given in Table 6.2.

Table 6.2: Parameters derived from alternate fit to Lyα data

Parameter	Value
μ_*	1215.595Å
σ_*	0.198Å
γ_*	0.072Å
A_*	3.40×10^{-12} erg cm^{-2} s^{-1} Å^{-1}
$\sigma_{lsf,1}$	0.18Å
$\sigma_{lsf,2}$	0.10Å
$A_{lsf,2}/A_{lsf,1}$	0.26
v_{cloud}	16.8 km/s

Broader UV spectrum

We first estimate the shape of the stellar spectrum in the extreme ultraviolet: wavelengths shortward of the Lyα line but longward of 100Å. These photons ionize hydrogen, which deposits heat into the atmosphere and drives the outflow. There are currently no EUV telescopes, so we rely on the scaling relations obtained by Linsky et al. (2014) to estimate the EUV flux in 100 Å bins between 100–1170 Åbased on the star's measured Lyα flux. For the wavelength range 100–400 Å, these scaling relations were based on stellar observations with Extreme Ultraviolet Explorer (Bowyer et al., 1994) and Far Ultraviolet Spectroscopic Explorer (Sembach, 1999). For the wavelength range 400–1170 Å, they are based on the semi-empirical solar model of Fontenla et al. (2014).

Moving to longer wavelengths, the stellar flux between the Lyα line at 1216 Å and the triplet helium ionization limit of 2588 Å has special importance for helium observations. These photons are not energetic enough to ionize hydrogen or ground state helium, and therefore do not create the ions or electrons which, through recombination, create triplet state helium. However, they are energetic enough to destroy triplet ground state helium via ionization. This means that the level population of triplet helium, and the magnitude of the corresponding triplet helium absorption signal during transit, is sensitive to the value of the stellar flux in this wavelength range.

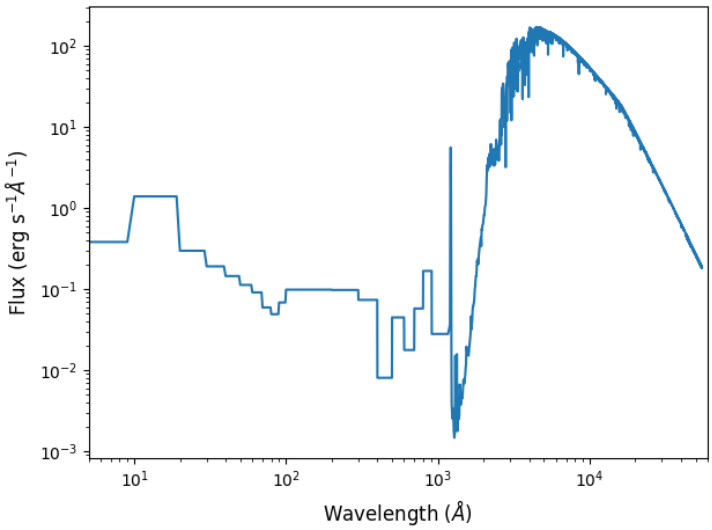


Figure 6.11: Fiducial stellar spectrum. We binned the X-ray data for better visibility.

HD 63433 is a solar analogue, so we adopt the solar MUV spectrum as measured by the Solar Radiation and Climate Experiment (SORCE) satellite.[7] We verify the applicability of this spectrum by comparing to data from XMM-Newton's Optical Monitor, which observed the star's MUV flux through two filters: UVM2 (λ = 231 ± 48 nm) and UVW2 (λ = 212 ± 50 nm). The OM measured $90.2 \pm 0.3 s^{-1}$ in UWM2 and $35.02 \pm 0.15 s^{-1}$ in UVW2. After correction for coincidence losses due to multiple photons hitting the detector in the same frame, we obtain $128.6 \pm 0.6 s^{-1}$ and $40.2 \pm 0.2 s^{-1}$, respectively. We used the solar spectrum and the filter transmission profiles to predict what OM would have seen if HD 63433 were an exact solar clone and obtained a count rate within 2% of the observed rate for UVM2 and within 7% for UVW2. We therefore conclude that the Sun's MUV spectrum accurately matches that of HD 63433, and adopt the Sun's spectrum between 1216 Å and 2588 Å.

Final reconstructed stellar spectrum

To recap, we obtain the X-ray spectrum by fitting a model to XMM EPIC data; the EUV spectrum using the scaling relations of Linsky et al. (2014); the Lyα spectrum by adopting ISM absorption parameters inferred from nearby stars and modifying a similar star's spectrum to fit the STIS data; the 1216–2588 Å spectrum by assuming it is identical to solar; and the NUV, optical, and IR spectrum from PHOENIX. The reconstructed spectrum is plotted in Figure 6.11.

[7] https://lasp.colorado.edu/home/sorce/data/

Table 6.3: Band-integrated fluxes

Band	Wavelengths (Å)	Flux at 1 AU (cgs)
X-ray	5 – 100	27 ± 14
EUV	100 – 912	91 ± 27
f Lyα	1214 – 1217	56 ± 17
MUV	1230 – 2588	2600 ± 150
Total	5 – 50,000	$1.02 \pm 0.04 \times 10^6$

In Table 6.3, we list the band-integrated fluxes of physical interest. In addition to the nominal values, we make an attempt to estimate the error bars. For the X-ray flux, we adopt 50% errors because of the significant variability we see in even our 6 ks XMM observation. The Lyα error is estimated based on Linsky et al. (2014), who state that the Lyα reconstruction process gives rise to 10–30% errors (we adopt 30%). The EUV error is also estimated based on Linsky et al. (2014), who find that the scaling relations we relied on to obtain our EUV spectrum are accurate to 30–60% (RMS) in 10 nm bins. We assume that the errors bin down when the total EUV flux is calculated, and adopt 30% as the final uncertainty. The MUV error is calculated from the 7% mismatch between the solar spectrum and the XMM OM photometry. The nominal value and the error for the bolometric flux are both calculated using the luminosity from Mann et al. (2020).

6.6 Mass loss modeling

Planet parameters

Parameter	Value
T_*	$5640K$
R_*	$0.912 \pm 0.07 R_\odot$
$M_{p,c}$	$7.3 M_\oplus$
$M_{p,b}$	$5.5 M_\oplus$
$R_{p,c}$	$2.67 R_\oplus$
$R_{p,b}$	$2.15 R_\oplus$
a_c	$0.1458 AU$
a_b	$0.0719 AU$
P_c	$20.5 d$
P_b	$7.1 d$

Table 6.4: Stellar and planetary properties. All values are from the discovery paper, Mann et al. (2020).

In order to set up our simulations of the outflow from each planet, we need to know the planets' radii, semimajor axes, masses and rotation periods. These are summarized

in Table 6.4. We obtain the first two from the discovery paper. Unfortunately, the youth and high activity of HD 63433 make it difficult to measure planet masses using the radial velocity technique. This is especially true for b, which has an orbital period close to the rotation period of the star. The planets are also not particularly close to any orbital resonances, making it unlikely that we could obtain dynamical mass constraints using transit timing variations. The assumed planet mass is a key ingredient for our models, as the predicted mass loss rate is exponentially sensitive to the assumed mass (e.g., Eq. 64 of Adams 2011: $\dot{m} \propto \exp(-GM_p/R_p c_s^2)$).

In the absence of any empirical mass constraints, we instead utilize a mass-radius relation derived from population-level studies of planets orbiting older (approximately greater than a Gyr) stars. In the discovery paper for this system, Mann et al. (2020) used the Chen & Kipping (2017) probabilistic forecasting relation to calculate an estimated mass of $5.5 \pm 2M_\oplus$ for b and $7.3 \pm 2M_\oplus$ for c. If we instead utilize the polynomial mass-radius relation from Wolfgang et al. (2016), we would predict a mass of $7.3 \pm 2M_\oplus$ for b and $9.7 \pm 2M_\oplus$ for c. If we use the mass-radius relation from Bashi et al. (2017), we would predict a mass of 5.6 $M_\oplus$ for b and 8.2 $M_\oplus$ for c. However, these mass-radius relations are all derived from observations of planets that are significantly older than HD 63433, whereas the planets in this system might still be inflated because they have lost a smaller fraction of their primordial atmospheres. We therefore adopt the lowest mass estimates, those of Chen & Kipping (2017), as the fiducial case for our models. Using the scaling relation from the previous paragraph, we estimate that an uncertainty of 2 $M_\oplus$ in the estimated planet masses translates to an uncertainty of $\sim$ 50% in mass loss rate for c and $\sim$ 80% for b, assuming a sound speed of 10 km/s.

We next consider whether or not HD 63433 b/c are likely to be tidally synchronized. The tidal synchronization timescale is (Guillot et al., 1996):

$$\tau_s = \frac{Q\omega_p M_p a^6}{G M_*^2 R_p^3}. \tag{6.4}$$

Adopting a tidal quality factor of Q=100 and an initial rotation rate of one Earth day, this evaluates to 0.02 Myr for b and 0.7 Myr for c (i.e., much less than the present-day age of the star). While Q is highly uncertain and the stronger gravity of a mini Neptune could reduce tidal dissipation by a factor of 2-3 (Efroimsky, 2012), it is difficult to get Q much above 1000 for a rocky planet (Clausen & Tilgner, 2015).

We therefore conclude that it is very likely that both planets have rotation periods equal to their orbital periods, and use this assumption in our models.

Important physical processes

Thermal mass loss is driven largely by stellar X-ray and EUV flux. These high energy photons are absorbed far above the optical photosphere, heating the thermosphere. The main cooling mechanism is emission of Lyα radiation by collisionally excited atoms (Murray-Clay et al., 2009), which scales with temperature as $\Lambda \propto \exp(-118,348K/T)$. The strong exponential dependence of the cooling rate keeps the temperature around a few thousand K, but below $\sim 10^4$ K. In the region of the outflow closest to the planet, the temperature rises with increasing radius. This is driven by the increased XUV heating and the decreased radiative cooling at larger separations. Eventually, the importance of these effects diminishes as the optical depth of the atmosphere above becomes effectively transparent to XUV, and adiabatic cooling causes the temperature to slowly drop.

Hydrodynamic outflows can be divided into three regimes, depending on the dominant mechanism for creating neutral hydrogen (Lampón et al., 2021). In the outflow, stellar flux blueward of the Lyman limit destroys neutral hydrogen by ionizing it, while recombination creates it. The neutral hydrogen population is also augmented by advection from lower in the atmosphere. There are three possible regimes (Lampón et al., 2021), depending on whether recombination or advection is the dominant neutral hydrogen creation mechanism. If recombination is dominant, the planet is in the recombination-limited regime, characterized by a narrow partially ionized zone and by low ($\sim$ 5%) efficiency in converting flux to kinetic energy (because the energy is radiated away by recombination). If advection is dominant, the planet is in the photon-limited regime, characterized by a very wide partially ionized zone and high flux-to-kinetic-energy conversion efficiency ($\sim$ 15%). If neither mechanism is clearly dominant, the planet is in the energy limited regime. As we will discuss below, our simulations show that both HD 63433 planets are in the photon limited regime (Figure 6.14). In this regime, the mass loss rate is high and neutral hydrogen is abundant in the outer regions, boosting the Lyα signal.

Helium absorption is from helium atoms in the metastable triplet ground state. Helium must stay in this triplet ground state, 19.8 eV above the singlet ground state, in order to absorb at 1083 nm. This level is populated, in most situations, by ionization followed by recombination, which ends with the recombining electron in

the triplet state $\sim 3/4$ of the time. Triplet helium can be destroyed by collisional de-excitation with electrons and neutral hydrogen in the lower atmosphere and by photoionization ($h\nu > 4.8$ eV) in the upper atmosphere. Other ways of producing and destroying triplet helium exist, such as collisional excitation and spontaneous radiative decay, but they are usually negligible. Close to the planet, no ionizing radiation penetrates, so nothing is ionized and there is no recombination to populate the triplet state. Going outwards, the electron density at first increases, causing the triplet state number density to increase to some maximum. Past this maximum, where the atmosphere is already largely ionized, $n(e)$ and $n(\mathrm{He}II)$ both decrease with r because of expansion. In this region, collisional de-excitation falls with r, but not enough to overcome the decreased production and the mostly constant photoionization of triplet helium, causing the triplet number density to fall. For a more detailed, quantitative overview of the physics of helium absorption in exoplanets, see Oklopčić & Hirata (2018).

Although we have discussed these processes as if they are radially symmetric, in reality outflows will have a non-spherical geometry that is shaped by the planet's immediate environment. The stellar wind and radiation pressure both work to push escaping gas away from the star. The Coriolis force then imparts a sideways force to the gas, creating a comet-like tail trailing the planet (Schneiter et al., 2007). If the planet has a magnetic field it can suppress the mass loss rate and confine the outflow (e.g. Adams 2011; Owen & Adams 2014; Khodachenko et al. 2015), but exoplanetary magnetic fields are poorly understood because they have never been observed. We simulate the outflow from both planets without including magnetic fields, and discuss the ways that magnetic fields might alter our predictions in §6.7.

3D hydrodynamic models

Our 3D models utilize the approach outlined in Wang & Dai (2018); Wang & Dai (2021), which combines ray-tracing radiative transfer, real-time non-equilibrium thermochemistry, and hydrodynamics based on the higher-order Godunov method code `Athena++` (Stone et al., 2020). We include a stellar wind with a mass-loss rate of $\sim 8\times$ solar, as calculated in §6.7, and a roughly solar velocity of 400 k$m\ s^{-1}$. Our models account for the hydrodynamic and thermochemical interactions between the stellar wind and the planetary outflow. These 3D models capture the anisotropy of the outflow pattern better than the 1D LTE models, while the non-LTE thermochemistry self-consistently predicts the mass loss rate and the line profiles. The model incorporates a total of 26 species and 135 reactions, including various relevant heating and

cooling processes (e.g. photoionization, photodissociation of molecular hydrogen, Lyman-α cooling, etc.). Metastable helium is a chemical species like the others, and key reactions that form and destroy this species are included in the thermochemical network. We refer the reader to Wang & Dai (2018) and Wang & Goodman (2017) for details of these reactions. The most important cooling processes in the models presented here are recombination, PdV work, and ro-vibrational cooling by H_2O/OH and CO, while the most important heating mechanism is photoionization of H and He (see Figure 3 in Wang & Dai 2018).

Starting from the stellar spectrum computed in Section 6.5, we group photons into seven energy bins for the ray-tracing calculation:

- $h\nu$ = 1.4 eV for infrared, optical, and near ultraviolet (NUV) photons, $h\nu$ = 7 eV for "soft" far ultraviolet (FUV) photons
- $h\nu$ = 12 eV for the Lyman-Werner band FUV photons, which can photodissociate molecular hydrogen but cannot ionize them (this is not to be confused with the Lyα line, which the band does not include)
- $h\nu$ = 16 eV for "soft" extreme ultraviolet (soft EUV) photons, which can ionize hydrogen but *not* helium
- $h\nu$ = 47 eV for hard EUV photons that ionize hydrogen *and* helium
- $h\nu$ = 300 eV for soft X-rays, which are abundant for an active star like HD 63433
- $h\nu$ = 3000 eV for hard X-rays

We note that this discretization of radiation artificially shrinks the vertical extent of the region where significant stellar XUV is deposited. If the radiation were not discretized, the photoionization cross section would drop with photon energy beyond the ionization energy, so that higher energies take over as the outflow becomes optically thick to lower energies. Unfortunately, this discretization is required in order to make our 3D model computationally tractable.

In addition to the opacities caused by photochemical reactions, we also include an effective opacity term in all bands, using dust/PAH as a proxy (see also Wang & Dai, 2018). This is particularly important in the optical band due to its heating

effects, as our opacity calculation did not include the Thomson cross-section $\sigma/\mathrm{H} \simeq 6.7 \times 10^{-25}$ cm^2.

A typical planetary atmosphere consists of a convective interior and a quasi-isothermal exterior (e.g. Rafikov, 2006). With the use of an interior model, we adjusted the envelope fraction to match the observed radius and assumed mass of the planets. We find an envelope mass fraction of $\sim 0.6\%$ for b and 2% for c, typical values for mini Neptunes. However, we note that without a precise mass measurement, the mass fraction of the H/He envelope cannot be precisely constrained. The precise envelope fraction is not important to modelling the outflow because it is the gravity and density at the optical transit radius that set the inner radial boundary condition, above which all photospheres of relevant bands of radiation are located. We summarize the key quantities that define the fiducial model in Table 6.5. Note that, to reduce the cost of simulation while keeping all necessary physical features, we only simulate the half-space above the orbital plane, and assume reflection symmetry over that plane.

Model results

Our models predict a mass loss rate of 0.11 $M_\oplus$/Gyr (2.1×10^{10} g/s) for c and 0.35 $M_\oplus$/Gyr (6.6×10^{10} g/s) for b. We can use our initial envelope fractions (2% for c and 0.6% for b) to calculate corresponding atmospheric mass loss timescales of 0.9 Gyr for planet c and 0.08 Gyr for planet b.

Figure 6.12 shows the orbital plane of our simulations. The outflow is initially somewhat spherically symmetric and still maintains this symmetry at $\sim 10R_\oplus$, the approximate photospheric radius for metastable helium absorption. Around 50 $R_\oplus$ it loses this symmetry as gas emanating from the day side is pushed toward the night side by the stellar wind and radiation pressure. As we move radially outward from the planet in a direction perpendicular to the planet-star axis, the temperature is initially equal to the planetary equilibrium temperature, rises to a peak of a few thousand Kelvin, declines slightly, and then jumps to a few million Kelvin as the outflow encounters the 1 MK stellar wind in a shock.

Figure 6.13 (top) shows the number densities of various species radially outward from the planet along the direction of orbital motion. The bottom pane shows the temperature, radial velocity, sound speed, and hydrogen ionization fraction. We find that the triplet helium density, which controls helium absorption, peaks around 200 cm^{-3} at 1.6 R_p (c) or 120 cm^{-3} at 2.1 R_p (b) before slowly declining farther

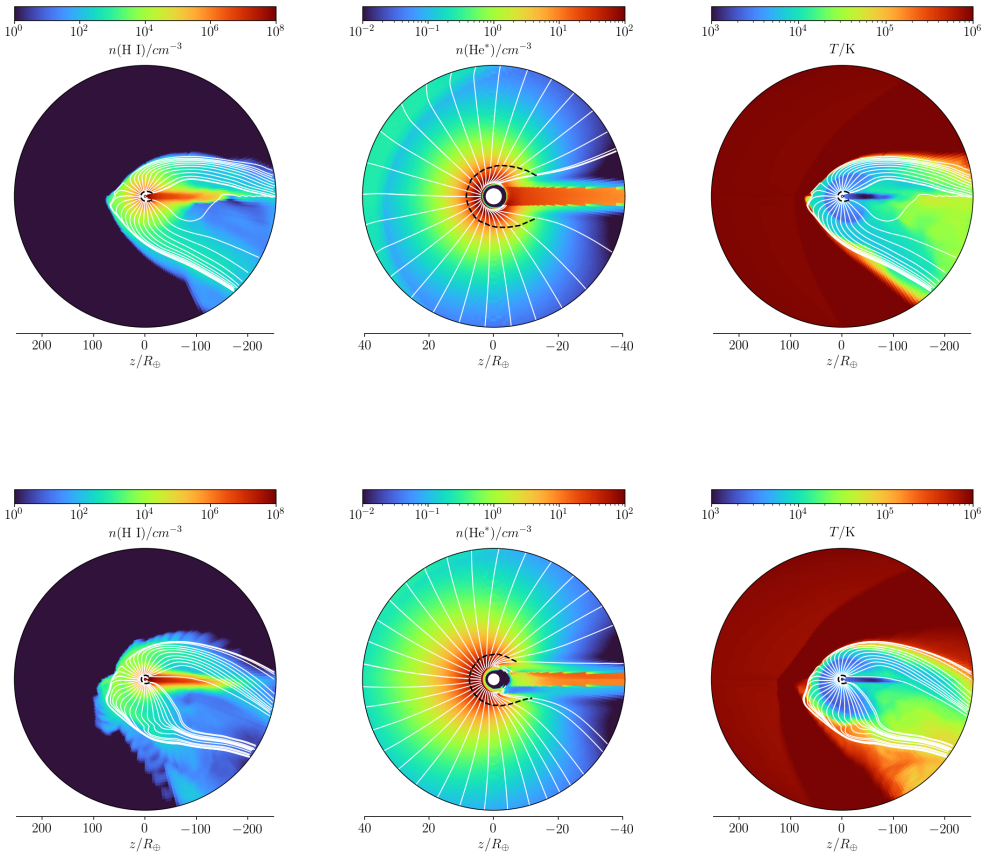


Figure 6.12: Neutral hydrogen number density, temperature, and triplet helium density from the time-averaged (over the last $\sim$ 10 kinematic timescales) fiducial 3D models, for c (top) and b (bottom). The star is toward the left, the velocity vector of orbital motion lies in the paper plane and points upwards. These plots show the profiles in the orbital plane. The white lines are the streamlines, while the dashed black lines represent the inner sonic surface. Note that the middle column showing metastable helium distribution zooms into the innermost 40 $R_\oplus$ region for clearer presentation.

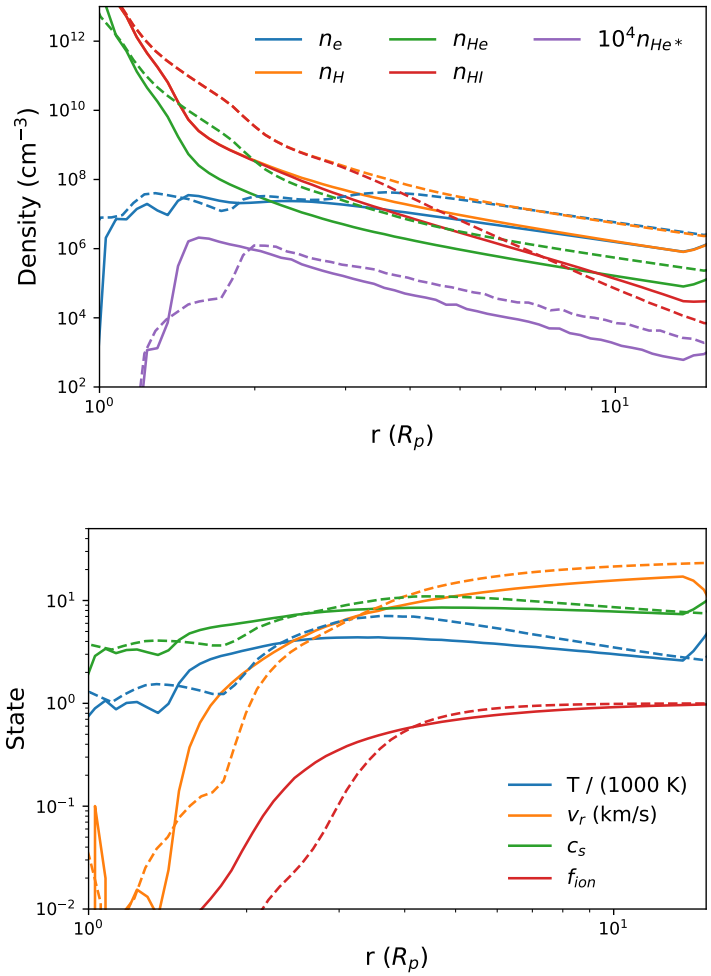


Figure 6.13: Radial profiles of key quantities along the radial lines pointing to the direction of orbital motion. Profiles are calculated for the time-averaged data (over the last 10 kinematic timescales of the simulations) for the fiducial models of planets c and b. Top: number densities of electrons (n_e), neutral helium (n_{He}), metatstable helium (n_{He^*}), neutral hydrogen (n_{H}), and ionized hydrogen (n_{HI}). Bottom: temperature T, radial velocity v_r, sound speed c_s, and hydrogen ionization fraction f_{ion}. Solid lines are for planet c; dashed lines for b.

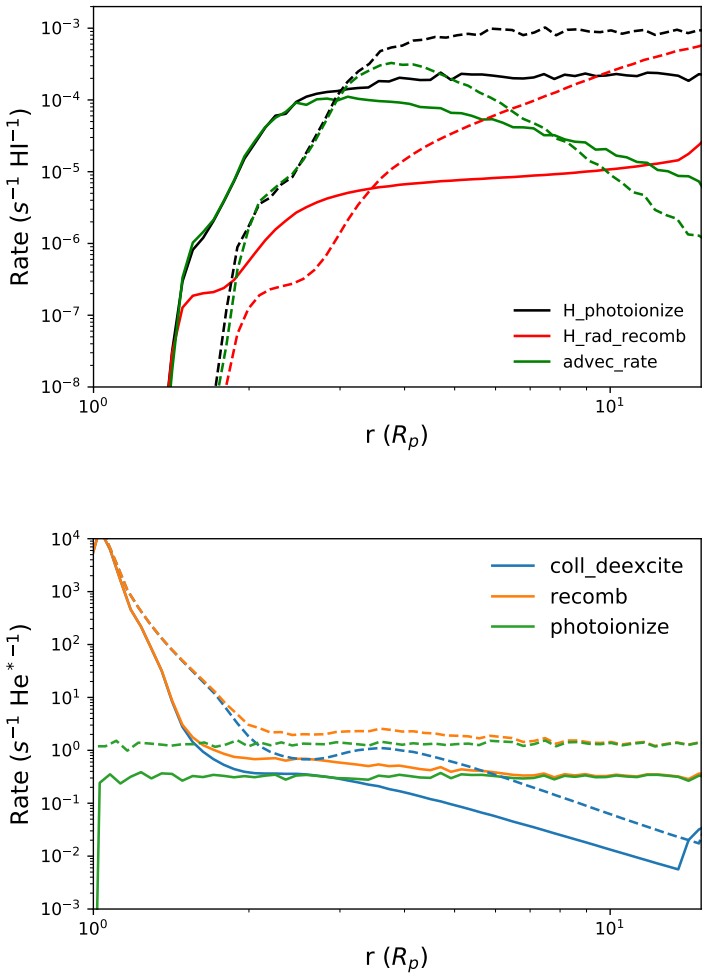


Figure 6.14: Similar to Figure 6.13, but for important reaction rates. Top: rates of important processes that create or destroy neutral hydrogen, namely photoionization, radiative recombination, and advection. Bottom: rates of processes that create or destroy metastable helium, namely collisional de-excitation, recombination, and photoionization [$He^* + h\nu(> 4.8\ eV) \rightarrow He + e^-$]. Rates for planet c are shown in solid lines, and those for planet b are in dashed lines. The rates are computed along the radial line pointing to the direction of the planets' orbital motion.

out. The neutral hydrogen density declines smoothly with increasing distance. The temperature rises to a maximum of 4000 K at 3 R_p (c) or 7000 K at 4 R_p (b) before slowly declining. The sound speed hovers around the 7-10 km/s typical of ionized hydrogen at several thousand Kelvin. At larger radii ($> 5R_p$), the outflow velocity asymptotes to $\sim$2 times the sound speed and the density falls roughly as r^{-2}, in accordance with the analytic Parker wind prediction. The neutral hydrogen fraction is nearly 1 at the surface, but declines to 50% around 3 R_p, falling to <3% at 15 R_p.

The rates we plot in Figure 6.14 show that triplet helium is created by recombination and destroyed predominantly by collisional de-excitation (smaller radii) or photoionization (larger radii), as discussed in §6.6. This is the same qualitative behavior seen for 55 Cnc e with a 1D PLUTO-CLOUDY model (Zhang et al., 2021), for some generic planets with a Parker wind model (Oklopčić, 2019), and for the gas giant WASP-107b with the same 3D model (Wang & Dai, 2021). Our models do include collisional excitation and radiative decay, but they are not plotted because they are negligible. Collisional excitation is negligible because the collisional excitation coefficient is 12 orders of magnitude smaller than the collisional deexcitation coefficient (Oklopčić & Hirata, 2018). Although singlet helium is typically $\sim 10^6$ times more abundant than triplet helium, a gap of 10^6 remains. Radiative decay is negligible because transitions between the triplet and singlet states are forbidden. This means that the Einstein A coefficient for the transition to the singlet ground state is very small (1.272×10^{-4} s^{-1}) and the decay timescale is very long ($\sim$ 2.2 hr). We can compare this decay timescale to the triplet helium production timescale, which is the triplet number density divided by the recombination rate. This timescale is on the order of 0.5 s. This short production timescale means that the triplet helium density is in local equilibrium, and is not significantly affected by advection except indirectly (advection carries neutral gas outward, reducing the electron number density).

Figure 6.14 shows the rates of various processes that create and destroy neutral hydrogen: photoionization, recombination, and advection. It can be seen that for both planets, advection dominates over recombination until the outermost regions of the outflow (11 R_p for c, 7 R_p for b). We conclude that both planets are in the energy limited regime for photoevaporative mass loss, where photoevaporation is relatively efficient because recombination does not radiate away all the incoming stellar high-energy radiation (Lampón et al., 2021). The rate for photoionization plateaus around 2.5 (3.5) R_p for planet c (b) because the outflow becomes optically thin to ionizing

radiation. This is the same general region where the flow becomes supersonic and where the hydrogen becomes predominantly ionized, making recombination and collisional de-excitation with electrons increasingly important. The triplet helium fraction, however, peaks well before this critical point (1.6 R_p for c and 2.1 R_p for b) before beginning a decline of $n_{He*} \propto r^{-3.5}$ for c and $n_{He*} \propto r^{-3}$ for b.

Our models predict a relatively symmetric transit light curve in both Lyα (Figure 6.2) and helium (Figure 6.6). This stands in contrast to Lyα observations of GJ 436b (Lavie et al., 2017) and helium observations of WASP-107b (Allart et al., 2019; Kirk et al., 2020), which show a much more delayed and extended egress. This happens because the stellar wind and radiation pressure both push the outflow away from the star, where the Coriolis acceleration $-2\mathbf{\Omega} \times \mathbf{v}$ slows its velocity relative to the planet's orbital motion. Naively, one would expect a stronger stellar wind to cause a longer tail and more asymmetric transit shape, but this is not the case in our models. We initially performed 3D simulations with a solar-strength wind and saw an asymmetrical helium transit, with peak absorption occurring 1.5 hours after the white light transit mid-point for planet c. When we switched to a more realistic 8× solar wind, we obtained the fiducial model presented here. In this version of the model the confining effect of the stellar wind overpowers the Coriolis force, accelerating the outflow and increasing the importance of inertial forces relative to Coriolis forces (parameterized by the Rossby number, Ro=vr/(2Ω)). This results in a more symmetric transit shape for the 8× solar wind case.

Parameter	Value
Mass loss	
$\dot{M}_{\rm b}$	0.35 $M_\oplus$ Gyr^{-1}
$\dot{M}_{\rm b,10\ M_\oplus}$	0.17 $M_\oplus$ Gyr^{-1}
$\dot{M}_{\rm c}$	0.11 $M_\oplus$ Gyr^{-1}
Simulation Domain	
Radial range (c)	$0.98 \leq (r/R_p) \leq 94$
Radial range (b)	$0.91 \leq (r/R_p) \leq 116$
Latitudinal range	$0 \leq \theta \leq \pi$
Azimuthal range	$0 \leq \phi \leq \pi$
Resolution ($N_{\log r} \times N_\theta \times N_\phi$)	$192 \times 128 \times 64$
Photon Flux at 1 AU* [cm $^{-2}$ s^{-1}]	
1.4 eV (IR/optical)	4.5×10^{17}
7 eV (Soft FUV)	1.8×10^{16}
12 eV (LW)	3.0×10^{11}
16 eV (Soft EUV)	1.2×10^{12}
47 eV (Hard EUV)	3.2×10^{11}
0.3 keV (Soft X-ray)	4.7×10^{10}
3 keV (Hard X-ray)	1.5×10^{9}
Initial Abundances [$n_{\rm X}/n_{\rm H}$]	
H_2	0.5
He	0.1
H_2O	1.8×10^{-4}
CO	1.4×10^{-4}
S	2.8×10^{-5}
Si	1.7×10^{-6}
Dust grains	1.0×10^{-9}
Dust/PAH Properties	
$\sigma_{\rm dust}$/H	8×10^{-22} cm^2
m_{dust}/m_{gas}	7×10^{-7}
Stellar Wind	
$\dot{M}$	$2 \times 10^{13} g/s$
$v_{\rm w}$	$400 km/s$
Temperature	$10^6 K$

*Divide by 0.1458^2 for c and 0.0719^2 for b

Table 6.5: Properties of the fiducial model

6.7 Discussion

Comparing model and data

Before comparing our model predictions to the observed magnitude of absorption during transit, it is useful to consider the implications of the atmospheric mass loss timescales in these models. Although these quantities are sensitive to the assumed mass, we can nonetheless draw some general conclusions. Planet c has a mass loss timescale of 0.9 Gyr in our models, while b has a mass loss timescale of 0.08 Gyr. The star has an estimated age of 0.4 Gyr, suggesting that b is unlikely to have retained a primordial atmosphere while it is at least plausible for c to have done so. At earlier times, the planets' puffier radii leads to lower gravity and a higher Roche radius, both leading to increased mass loss. This prediction is consistent with our non-detection of excess absorption in either Lyα or metastable helium absorption during transits of planet b. We explore the effect that the assumed mass of planet b has on its predicted atmospheric lifetime in more detail in §6.7.

If we set these arguments aside for the moment and assume that both planets host hydrogen-rich atmospheres, we can use our 3D models to predict the time-dependent absorption signal during transit. Figure 6.2 shows our model predictions for the observed Lyα absorption. We find that the model slightly underpredicts the amount of blue wing absorption for planet c. The model also predicts that the point of maximum absorption will occur slightly after mid-transit, whereas our data prefers a peak before mid-transit. The model predicts that there should be negligible absorption in the red wing bandpass; although we see weak evidence for red wing absorption in our data, the detection is not conclusive (see §6.4 for more details). For planet b, the model predicts an absorption depth in the blue wing of the Lyα line that is comparable in magnitude to that of planet c, with a shorter but symmetric transit shape. This absorption signal is conclusively ruled out by our data. As with c, the model predicts negligible absorption in the red wing; this is consistent with our non-detection.

For planet c, where we detect absorption in the blue wing of Lyα, we can compare the predicted wavelength-dependent shape of the absorption signal as a function of time to the observed line shape (Fig. 6.3). The magnitude of absorption in the observed spectrum rises from <5% to 25% as we move from -100 km/s to -10 km/s, while the model spectrum rises from <5% to 60% as we move from -50 km/s to -10 km/s. We conclude that the observed absorption signal is significantly more widely dispersed in velocity space than the modeled absorption. This phenomenon

of Lyα absorption being present far above the sound speed was first noted by Vidal-Madjar et al. (2003) for HD 209458b, and has also been observed in all four planets with definitive Lyα detections (HD 209458b, HD 189733b, GJ 436b, GJ 3470b). Holmström et al. (2008) proposed that the -140 km/s absorption signal seen for HD 209458b is caused by charge exchange between neutral hydrogen in the planetary outflow and solar wind protons, creating a population of neutral hydrogen atoms with a high velocity dispersion heading away from the star. We note that kinetic processes in the interaction region between the stellar wind and the planetary outflow might be difficult to model correctly due to the long mean free paths, which exceed 1 R_p when $n_H <\sim 10^4$ cm^{-3}. If future models deal with this process with more consistency, the observed signal might be properly explained by invoking stellar wind hydrogen ions neutralized and thermalized near the stellar wind-planetary outflow contact surface.

While the simulation explains the blue wing absorption of c somewhat well, it does not explain the red wing absorption at all. The simplest explanation for the discrepancy is that the observed red wing absorption is actually stellar variability. The absorption detection is marginal even from a pure SNR perspective, and there are other reasons to doubt the detection (Subsection 6.4), including the non-detection in the alternative analyses (see Appendix). If we assume the detection is real, the red wing absorption is difficult to explain purely hydrodynamically because it occurs at 100–200 km/s, far above the sound speed of ~10 km/s. Since some extreme red wing absorption can be obtained from energetic neutral atoms at the intersection of the planetary outflow and the stellar wind, one might think the problem can be solved by making the stellar wind faster (resulting in more kinetic energy being thermalized in the shock) and denser. However, this would also tend to confine the outflow and push more of the absorption to the blue wing, potentially worsening the discrepancy between model and data. Using an order-of-magnitude calculation, we also considered the possibility that the red wing absorption could be explained by the far wings of Lyman alpha. We considered a r=25 $R_\oplus$ uniform sphere with n(HI) = 10^6 cm^{-3} with T=7000 K and a mean atomic weight of 0.7 AMU, very roughly matching the simulated conditions in Figure 6.12. We found that the optical depth across the center of the sphere is 0.03 at 50 km/s, 0.006 at 100 km/s, and 0.003 at 150 km/s. Even an optical depth of 0.03 across the entire 25 $R_\oplus$ disk would increase the transit depth by only 0.2%. We conclude that if the absorption signal at 100-200 km/s is real, it is unlikely to be caused by gas moving at less than 50 km/s.

We next consider whether or not the predicted metastable helium absorption spectrum from these models is consistent with the upper limits from our NIRSPEC observations (see Fig. 6.5). For c, the model predicts 0.4% absorption, which would be reduced to 0.15% due to the self-subtraction when we remove the time-varying telluric signal. This absorption should be marginally detectable, but may plausibly be hidden by stellar activity. For b, the model predicts 0.5% absorption, which would be reduced to 0.3% by self-subtraction. This is slightly stronger than the 0.2% absorption in Figure 6.5, which we attribute to stellar variability. Our injection/recovery tests show that the predicted signal would have been detectable in both the residuals images and the excess absorption spectrum. In addition to transit-averaged excess absorption spectra, we also compare the line-integrated light curves to model predictions (Figure 6.6). Here, the model predicts a level of absorption considerably above the noise for both planets, but not considerably above the stellar variability we observe on both nights.

A higher mass cannot explain the non-detections for planet b

We next consider whether or not a higher assumed mass for planet b can explain the Lyα and helium non-detections. We run another 3D model with a core mass of 10 $M_{\oplus}$; this is 2.7 $M_{\oplus}$ higher than the predictions of all mass-radius relations we consider in §6.6, and high enough for runaway accretion during planet formation to occur. Aside from the increased core mass and correspondingly increased envelope fraction, all other aspects of the model were unchanged.

As expected, we find that increasing the assumed core mass does modestly reduce the magnitude of the predicted absorption signal during transit. The high mass model predicts a Lyα absorption depth in the blue wing of 7%, compared to 10% in our fiducial model. However, this signal is still too strong to accord with our observations. In the helium lines, the high mass model predicts a peak absorption of 0.35%, compared to 0.6% in the fiducial model. If we interpret the darkening of the helium line during the transit of planet b as planetary absorption instead of stellar variability, the observed signal would be comparable in magnitude to this prediction. If the helium signal is planetary, it should be detectable with follow-up observations from telescopes even smaller than Keck, as we were limited by stellar variability and not by photon noise.

Possible envelope compositions for planet b

As previously noted, planet b's short predicted mass-loss timescale and our non-detection of absorption from either hydrogen or helium suggests that this planet may have already lost its primordial hydrogen/helium envelope. Assuming an Earth-like composition, a bare core with the radius of b would have a mass of $M/M_\oplus = (R/R_\oplus)^4 = 21 M_\oplus$ (Lopez & Fortney, 2014). This core would be comparable in mass to the inferred cores of gas giant planets, and would have been highly susceptible to runaway accretion even at its present-day location in the inner disk (e.g., Lee, 2019). Rocky cores with radii between 2-3 $R_\oplus$ planets are quite rare among the sample of planets with measured densities, although not unheard of (Mocquet et al., 2014). This suggests that a small fraction of 10-40 $M_\oplus$ cores may avoid runaway gas accretion, presumably because they did not form until the disk was already dispersing (Lee, 2019).

If we are willing to consider high mean molecular weight envelopes, it is also possible to match planet b's observed radius with lower core masses. Mass-radius relations for water-rich planets suggest that they can have bulk densities similar to those of mini Neptunes with rocky cores and hydrogen envelopes with mass fractions of $1-2\%$ (e.g. Turbet et al. 2020; Aguichine et al. 2021; Nixon & Madhusudhan 2021). Outflows from planets with water-rich envelopes will contain hydrogen created by the photodissociation of water, and hence will still absorb in Lyα. Johnstone (2020) found that the energy efficiency of photoevaporation for water atmospheres ($\sim$10%) is similar to that of hydrogen/helium atmospheres. However, up to 7/8 of the outflowing mass would be oxygen atoms, which would reduce the hydrogen number density and the corresponding magnitude of the absorption signal in Lyα. Water worlds would not have significant amounts of helium in their atmospheres, consistent with our upper limit on helium absorption. More modelling is required to determine if our Lyα non-detection for b (or our detection for c) is consistent with a water world. We might also consider high mean molecular weight envelopes with other compositions, such as CO_2 or N_2, although these atmospheres would be more compact and would require correspondingly larger core masses. Kite & Barnett (2020) found that highly irradiated planets are likely to lose heavy gases alongside hydrogen and helium, but that a volcanically revived atmosphere is plausible for $T_{eq} <\sim 1000$ K planets around solar-mass stars. These compact atmospheres might cool efficiently, which would suppress the outflow and reduce the observational signature of Lyα, and there should be no helium because helium is insoluble in

lava. Such a metal-rich outflow might still be detectable in other FUV lines (Garcia Muñoz et al., 2021).

Other potentially important physical effects

Our models of hydrogen-rich outflows are ultimately unable to provide a satisfactory match to the observational data for either planet. While we can argue that planet b likely lost its primordial atmosphere, that is not the case for planet c. We therefore consider whether or not other physical effects, such as stellar or planetary magnetic fields, that we neglect in our models might have a significant effect on the magnitude or shape of the predicted outflows.

Magnetic fields

In order to evaluate the potential importance of magnetic fields, we first need to estimate the stellar magnetic field strength. Vidotto et al. (2014) measure the magnetic field strength of 73 stars ranging from 1 Myr to 10 Gyr and found that it is proportional to $t^{-0.655\pm0.045}$. We use this scaling relation to estimate the magnetic field of the 440 Myr HD 63433. We find that this star is predicted to have a magnetic field strength of approximately 5 times solar, or ∼13 G. This is consistent with Rosén et al. (2016), who used polarization data to measure the mean magnetic fields of five young Sun-like stars between 300 and 700 Myr and found $B_* \approx 20$G (with a range of 10-25 G), with no evidence for any age dependence. We adopt 20 G as the fiducial field strength and assume that, in the absence of a stellar wind, its strength falls off as r^{-3} (consistent with a dipole). It should be noted that the mean field of the Sun varies from 0.2 to 2 G across a solar cycle (Plachinda et al., 2011), and it is likely that HD 63433 also has a variable field strength.

If we assume the Elasser number $\Lambda = \frac{\sigma B^2}{\rho\Omega}$ is 1, where ρ is the density of the interior, Ω is the rotation rate, and σ is the conductivity, we can estimate the planetary magnetic fields. Assuming that the interiors of the HD 63433 planets have Earth-like densities and conductivities, the magnetic field should scale as the square root of the rotation rate. Rescaling Earth's magnetic field strength (0.25-0.65 G) to account for the slower rotation rates of the HD 63433 planets, we find that the planetary magnetic field is around 0.15 G for b and 0.08 G for c. We further assume that the planetary magnetic field is a dipole and falls off with distance as r^{-3}. We note, however, that the predicted magnetic field strengths of mini Neptunes are highly uncertain. For example, Christensen & Aubert (2006) numerically analyze dynamo

models and conclude that the Elasser number can range from 0.06 to 100 while the magnetic field is independent of rotation rate. Christensen et al. (2009) find using observations that B scales with $q_0^{2/3}$, where q_0 is the heat flux. If so, the magnetic field of the HD 63433 planets could be greater than that of Earth.

Having estimated the relevant magnetic field strengths, we explore their importance for the outflow. Following Owen & Adams (2014), we calculate the ratio of ram pressure to magnetic pressure:

$$\Lambda = \frac{2\dot{M}v}{B^2 r^2} \tag{6.5}$$

for both the stellar wind and the planetary outflow, where $\dot{M}$, v, and B are defined at radial distance r. For the stellar wind, we assume the speed is the same as it is for the Sun (400 km/s). The speed does drop as one approaches the Sun, but not dramatically; Venzmer & Bothmer (2018) predicted that the Parker Solar Probe would see a speed of 340 km/s at 0.16 AU and 290 km/s at 0.046 AU. What Parker actually measured near perihelion on January 17, 2021, when it was 0.10 AU from the Sun, was 250–320 km/s. To calculate the mass loss rate, we use the astrospheric observations of Wood et al. (2005a). This study found that the mass-loss rate scales with the X-ray flux as $\dot{M} \propto F_X^{1.34\pm0.18}$ until $F_X = 7 \times 10^5$ erg cm^{-2} s^{-1}, at which point it abruptly falls from ~80× solar to ~8× solar. This may be due to a large-scale change in magnetic topology, as extremely active stars tend to have a polar starspot while less active stars have more starspots at low latitudes. HD 63433 has an X-ray flux of $F_X = 1.4 \times 10^6$ (cgs), which is just past the transition point between these two regimes. We therefore adopt a value of 1.2×10^{13} g/s, corresponding to the lower value of 8 times solar mass loss rate. This value is similar to the 16× predicted by simulations in Cranmer (2017), but it is possible that the star's actual mass loss rate is closer to 80x solar.

Using this stellar mass loss rate, we can obtain an initial estimate for the ratio of ram to magnetic pressure if we assume that the stellar magnetic field is a dipole unaffected by the wind. We find that $\Lambda_* = 43$ for b and $\Lambda_* = 690$ for c. However, the very high Λ_* means that the stellar magnetic field is carried with the wind and should be nearly radial at the positions of the two planets. The stellar magnetic field at the location of the planet will therefore be much higher, and the corresponding Λ_* will be much lower, than this simple dipole model predicts.

We next calculate the corresponding Λ for the planets. Planet c loses mass at a rate of $3{\times}10^{10}$ g/s in our fiducial model. At 2 R_p, the outflow has a low ionization fraction and has not passed the sonic point. At this point, $v \sim 2$ km/s, giving $\Lambda_c = 10$. Planet b loses mass at a rate of $8{\times}10^{10}$ g/s in our fiducial model. At 2 R_p, the modeled outflow still has a low ionization fraction and has not passed the sonic point. The outflow has a velocity of $v \sim 0.8$ km/s, which translates to $\Lambda_b = 5$. Our choice of scaling law for the planetary magnetic field means that $\Lambda \propto r^4$ at constant v, but since v accelerates, Λ rises even more steeply with r. For comparison, a 0.4% transit depth corresponds to a photospheric radius of 2.9 R_p for b and 2.3 R_p for c, while a 10% transit depth corresponds to 15 and 12 R_p. This means the outflow becomes ram pressure dominated inside of the triplet helium photospheric radius, and well inside of the sonic radius. We conclude that the planetary magnetic field is probably insignificant in shaping the outflow for the fiducial magnetic field strengths. However, if the magnetic fields are several times larger than fiducial—which is easily possible, given the poor theoretical knowledge of planetary dynamos—these Λ values would decrease by a factor of 25-100, making the magnetic field highly significant in shaping the outflow. Most simulations show that magnetic fields should decrease the outflow rate and the signal strength (Owen & Adams, 2014; Khodachenko et al., 2015; Arakcheev et al., 2017), although Carolan et al. (2021) predicts an increase in the outflow rate.

We can use this same approach to estimate the radial distance from the planet where the stellar wind begins to overwhelm the planetary outflow. To compute this, we take the ratio of the ram pressures:

$$\Pi_{WW} = \frac{\dot{M}_*}{\dot{M}_p} \frac{v_*}{v_p} \frac{r^2}{a^2}. \tag{6.6}$$

At 4 R_p, we obtain $\Pi_{WW} = 1.3$ for c and 3.2 for b. These numbers are 3.0 and 1.6 at 15 R_p, indicating that the stellar wind plays a significant role in shaping the observed Lyα outflow. This is fully consistent with our simulations (Subsection 6.6).

Lastly, we compute the effect of the interplanetary magnetic field on the outflow. In the Parker model, the field is dragged along by the stellar wind, giving rise to a radial component that scales as r^{-2} and an azimuthal component that scales as r^{-1}, pointing opposite to the direction of stellar rotation (Parker, 1958). We use the measured radius dependence of the interplanetary field in the solar system

(Hanneson et al., 2020) to estimate the value of the stellar field at the positions of HD 63433b/c, assuming that the interplanetary field has the same proportionality with the stellar surface field in both systems. For b, we find a field strength of 0.02 G and an angle $\theta = \arctan(\omega_* a / v_*) = 16°$. For c, we find a field strength of 0.005 G and an angle of 31°. HD 63433 has a magnetic field 5 times stronger than the Sun's, and the predicted field strength at the location of planet c is approximately a factor of five larger than the magnetic field strength recorded by the Parker Solar Probe's FIELDS instrument during its June 2020 perihelion, when it was 0.13 AU from the Sun (similar to c's 0.14 AU semimajor axis). We therefore conclude that the distance scaling is in reasonable agreement with solar system observations. However, the angles predicted by the Parker model are somewhat less reliable; in this model the solar interplanetary field at Earth's distance should have a value of $\theta = -40°$, yet it is rarely within 10 degrees of that value. There are also frequent deviations below -60° and above 50° at both solar maximum and minimum (Tasnim & Cairns, 2016). If we use these nominal values to compute the ratio of the planetary ram pressure to the stellar magnetic pressure, we find values of 42 for c and 4.3 for b at 2 R_p. At 15 R_p, the ratios are 4.4 for c and 2.2 for b. We conclude that the interplanetary magnetic field likely has some influence on the very outermost portions of the outflow, but it is probably less significant than the ram pressure from the stellar wind.

It is important to remember that these calculations are order-of-magnitude estimates and nearly all of the quantities involved have significant uncertainties. The Sun's magnetic field is complex and variable; at Earth's distance, it has been shown to vary by a factor of a few over the course of the solar cycle. The B^2 dependence of magnetic pressure means that a factor of a few uncertainty in the magnetic field strength translates to an order of magnitude uncertainty in the magnetic pressure. Our estimated stellar wind mass-loss rate is based on a small number of indirect measurements in Wood et al. (2005a). It also falls close to a breakpoint in the scaling relation; it is possible that the mass-loss rate is closer to 80× solar than to 8× solar. The wind velocity of stars other than the Sun is unknown. There are no direct measurements of magnetic field strengths for extrasolar planets, and the field strengths of super-Earth cores may not scale simply as $\Omega^{1/2}$. We utilize model predictions for the mass loss rates of the two planets, which may vary by a factor of a few depending on our starting assumptions. Nevertheless, our fiducial values indicate that the stellar magnetic field is expected to be predominantly radial at the positions of both planets, with a significant tangential component pointing opposite the orbital direction. They also suggest that the stellar wind will have a significant

effect on the shape of the outflow, in good agreement with our 3D models. Lastly, they indicate that the interplanetary magnetic field may influence the the outflow at the relatively large separations probed by our Lyα observations, and that the planetary magnetic field likely has a negligible effect.

Radiation pressure

Our outflow models also neglect to consider the effect of radiation pressure. This effect has been invoked to explain detections of highly blueshifted Lyα absorption for multiple planets (e.g. Vidal-Madjar et al. 2003; Ehrenreich et al. 2012; Bourrier et al. 2014). However, we do not expect radiation pressure to be important in the HD 63433 system. Since the radiation pressure comes predominantly from Lyα, which is strongly absorbed by neutral hydrogen, we can calculate the radiation pressure using the Lyα flux inferred from observations. Assuming that all species are coupled, we find that $P_{\rm rad} = F_{Ly\alpha}/c = 9 \times 10^{-8}$ dynes/cm^2 at the semimajor axis of planet c. We compare this to the stellar wind ram pressure, $P_{\rm wind} = \dot{M}v/(8\pi a^2) = 4 \times 10^{-6}$ dynes/cm^2. The ratio between the two is $P_{\rm wind}/P_{\rm rad} = 45$. Since both pressures scale as a^{-2}, this ratio is the same for planet b. The stellar wind ram pressure is the lowest pressure that exists in our simulation. The total pressure is higher in the interaction region between the wind and the planetary outflow, and continues to increase as we move closer to the planet. As a result, radiation pressure is even less important in altering the hydrodynamics.

One might wonder if the ion-neutral collision cross section is high enough for the hydrodynamic assumption to be valid. For collisions between H and H+, $\langle \sigma v \rangle = 3 \times 10^{-9}$ cm^3 s^{-1} (Draine 2011, their Table 2.1). Assuming a v that corresponds to 10^4 K and a characteristic length scale of R_p, the hydrodynamic assumption corresponds to $\rho >\sim 10^{-18}$ g cm^{-3}. This density condition holds true for almost the entirely of the planetary outflow from both planets, although it begins to break down for c in the region outside the stream lines and the low-density interior of the extended tail. The proton-proton collision cross section is substantially larger, and imposes a less limiting condition of $\rho >\sim 10^{-20}$ g cm^{-3}. Our conclusions—that the hydrodynamic condition is mostly satisfied, and that radiation pressure is unlikely to significantly affect the photoevaporative winds—agree with those of Debrecht et al. (2020), who used a 3D hydrodynamic simulation to study the outflow from the hot Jupiter HD 209458b. However, we note that the hydrodynamic condition is not satisfied in the stellar wind in either our simulation or Debrecht et al.

(2020), and that more sophisticated modelling would be useful for the interaction region between the stellar wind and the planetary outflow.

6.8 Conclusion

In this study we use *HST* and Keck to look for escaping hydrogen and helium from two mini Neptunes orbiting a young solar analogue. We detect 11.5 ± 1.5 % Lyα absorption in the blue wing during a transit of planet c, but not during a transit of planet b. We do not detect excess helium absorption during transits of either planet, to a stellar-variability-limited upper limit of ∼0.5%. We use near-contemporaneous XMM-Newton data to characterize the stellar high-energy environment during our Lyα and helium observations, and combine these observations with ground-based optical monitoring data and archival ROSAT data to constrain the long-term stellar activity cycle. We use the measured stellar X-ray spectrum and an extrapolated extreme UV spectrum as inputs to 3D hydrodynamic models of hydrogen- and helium-rich planetary outflows, which we compare to the observational data.

For c, our hydrogen-rich models provide a reasonable match to the measured shape and depth of the transit light curve in the blue wing of the Lyα line. Similarly, their predictions for the magnitude of absorption in the metastable helium line are consistent with our upper limit for this planet. However, the observed blue wing absorption is less compact in velocity space than the simulated absorption. We speculate that this might be due to charge exchange with the stellar wind, which is difficult to model accurately with our hydrodynamic model, but which Tremblin & Chiang (2013) concludes can cause 10% Ly alpha absorption at 100 km/s. In addition, our models do not account for the effects of magnetic fields. While the effect of the planetary magnetic field on the outflow is likely negligible, the interplanetary field could play a subdominant role in confining and guiding the outflow beyond $\sim 10R_p$.

Our observations and models both suggest that planet b is fundamentally different from planet c. Our models predict that if planet b hosts a hydrogen-rich atmosphere it should also exhibit strong Lyα absorption during transit, but this is definitively ruled out by our observations. Our models also predict a detectable metastable helium absorption signal from planet b, which is inconsistent with our observational upper limit. The predicted mass loss timescale for planet c is longer than the age of the system, but the corresponding mass loss timescale for planet b is significantly shorter. This implies that c could have retained a primordial H/He atmosphere, while

b probably does not. If b is a rocky core, it would have to be unusually massive, but a water-rich composition with a high mean molecular weight atmosphere could explain both the radius and the Lyα non-detection. Fortunately, HD 63433 is a bright nearby star, and its planets are favorable targets for atmospheric characterization by *HST* and *JWST*. If planet b does host a hydrogen-dominated atmosphere, it may have detectable absorption from water and other molecules in its infrared transmission spectrum. If b hosts a high molecular weight atmosphere, detecting it in outflow may still be possible by looking at metal FUV lines (Garcia Muñoz et al., 2021).

Software: numpy (van der Walt et al., 2011), scipy (Virtanen et al., 2020), matplotlib (Hunter, 2007), stan (Stan Development Team, 2018), stistools, SAS, HEAsoft

6.9 Acknowledgements

This study was based on observations with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for this work was provided by NASA through grant number GO-16319 from STScI. This study also utilized data obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. L. dos Santos and D. Ehrenreich acknowledge that this project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (project Four Aces grant agreement No 724427), and it has been carried out in the frame of the National Centre for Competence in Research PlanetS supported by the Swiss National Science Foundation (SNSF). TGW acknowledges support from STFC consolidated grant number ST/R000824/1. SH acknowledges CNES funding through the grant 837319. GWH acknowledges long-term support of the APT program from NASA, NSF, Tennessee State University, and the State of Tennessee through its Centers of Excellence Program.

6.10 Appendix

In the published paper, we will make publicly available the reconstructed stellar spectrum, the reduced observations in both Lyα and helium, and the 1D profiles from our hydrodynamic model.

6.11 Alternate analyses of Ly alpha data

The fiducial analysis reported in the previous subsection was performed by the first author (Michael Zhang, MZ). Two other analyses were performed by two co-authors (Luca Fossati, LF and Leonardo dos Santos, LDS) using independent pipelines, and with no communication other than agreeing on a common velocity range in which to look for absorption: [−140,−10] km/s in the blue wing, [100,200] km/s in the red.

In LDS's analysis, he leveraged the time tag stream of events in the raw datasets to break down each *HST* orbit into ten subexposures to improve the temporal resolution of the time series. The subexposures are then reduced using `calstis` with its default settings. In order to correct for telescope breathing, he calculated the total flux between 1200-1248 Å (excluding the geocoronal contamination range) for each subexposure, and subsequently phase-folded the flux time-series to the orbital period of the space telescope ($P_{\rm HST}$ = 95.42 min). He tried to correct for the breathing effect by fitting this time series to a Fourier decomposition model for the systematic modulation with varying degrees and period equal to that of *HST*. This is the same approach used by Bourrier et al. (2017a) and references therein. For both visits of planet b, and the first visit of c, a first degree systematics model is favored over higher degree models and no systematics models by a ΔBIC > 10. A model with no systematics is favored for the second visit of c, since higher degrees do not significantly improve the BIC statistic. The uncertainties are calculated assuming that the observations are in the Poisson counting regime (see dos Santos et al. (2021)). LDS did not correct for potential modulation due to stellar activity, which is expected for young stars, in this analysis. This is very difficult to do without simultaneous observations at other wavelengths.

In LF's analysis, he first considered the calibrated 2-dimensional spectra providing information on the photon arrival time ("_tag.fits" files) and split each *HST* observation into 5 sub-exposures of equal exposure time. From each sub-exposure image, he extracted the stellar spectra using a slanted rectangular extraction box with an aperture of 40 pixels and the background employing an identical extraction box, but shifted upwards by 130 pixels. He then removed the relative background from the stellar spectra, phased each sub-exposure with *HST*'s orbit, and tried to correct for the breathing effect. To do this, he excluded the observations obtained during the first *HST* orbit of each visit, because the first *HST* orbit is notoriously affected by additional systematics. He modelled the breathing effect separately for

each *HST* visit as a polynomial of varying order, selecting the one that minimizes the Bayesian Information Criterion: BIC = $\chi^2 + k \log N$, where k is the number of free parameters and N is the number of data points. For the first and second visit covering the transit of planet b, he described the breathing effect with a second and third order polynomial, respectively, while for the first and second visit covering the transit of planet c, he described the breathing effect with a second and first order polynomial, respectively. For each *HST* visit, which was characterised by a distinct shape for the breathing effect, he applied the same correction to the Lyα fluxes for the blue and red wings.

All three authors found that breathing corrections do not significantly or consistently reduce the scatter of the orbit-averaged fluxes. MZ considered these corrections in his preliminary analyses but did not feel that they were justified. LF found that his corrections have virtually no effect on the blue wings, but reduce the scatter in the red wing of planet b and increase the scatter in the red wing of planet c. LDS finds that his corrections marginally increase the scatter for both wings of planet c, but marginally decrease the scatter for both wings of planet b. Given these results, we decided to exclude breathing corrections from the fiducial analysis presented in this study.

Figure 6.15 compares the light curves from the three analyses, all of which are shown with no breathing correction. The results are largely consistent, although we find that the fiducial analysis has a lower average scatter than the others. All three analyses show clear blue wing absorption during the transit of planet c. All three analyses find no evidence for absorption in either wing during the transit of b. The fiducial analysis shows tentative red wing absorption from c while the other two do not. This difference may be due to the significantly higher standard deviation in the light curves from the two alternative analyses (2.8% vs. 7.9% for LDS and 8.0% for LF). The conclusions above do not change when the fiducial analysis is compared to the breathing-corrected versions of LF's and LDS' analyses (Figure 6.16).

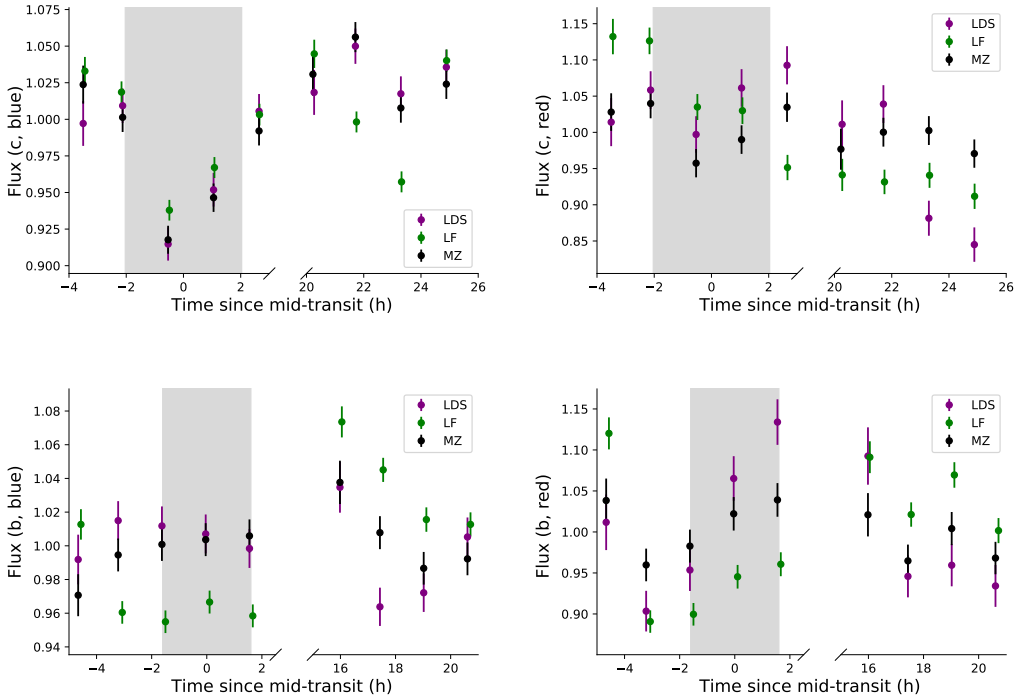


Figure 6.15: Comparison of independent analyses by MZ, LDS, and LF. The blue wing light curves are shown on the left, and the red wing light curves are shown on the right. Planet c is shown in the top panel, and planet b is shown in the bottom panel. The duration of the white-light transit is indicated by the grey shaded region. We exclude the data from the failed second observation of planet b in this comparison.

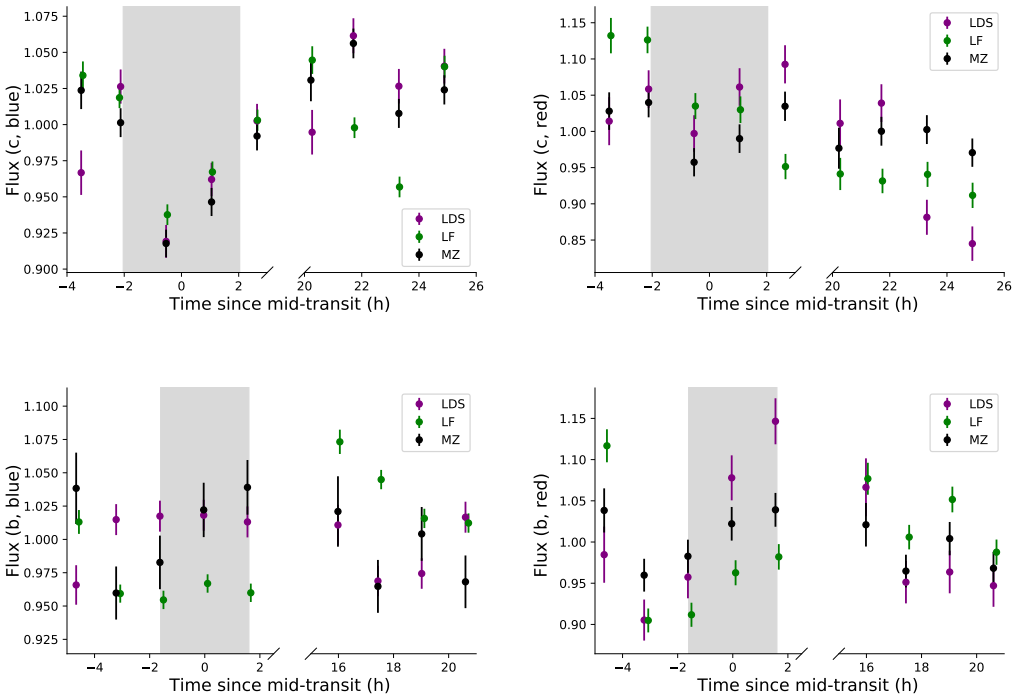


Figure 6.16: Same as Figure 6.15, but breathing corrections are included in the LF and LDS analyses. The fiducial analysis remains the same.

References

Adams, F. C. 2011, ApJ, 730, 27

Aguichine, A., Mousis, O., Deleuil, M., & Marcq, E. 2021, arXiv e-prints, arXiv:2105.01102

Allart, R., Bourrier, V., Lovis, C., et al. 2019, A&A, 623, A58

Arakcheev, A. S., Zhilkin, A. G., Kaigorodov, P. V., Bisikalo, D. V., & Kosovichev, A. G. 2017, Astronomy Reports, 61, 932

Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481

Bashi, D., Helled, R., Zucker, S., & Mordasini, C. 2017, A&A, 604, A83

Belloni, T., Hasinger, G., & Izzo, C. 1994, A&A, 283, 1037

Biddle, L. I., Pearson, K. A., Crossfield, I. J. M., et al. 2014, MNRAS, 443, 1810

Boller, T., Freyberg, M. J., Trümper, J., et al. 2016, A&A, 588, A103

Bourrier, V., Ehrenreich, D., King, G., et al. 2017a, A&A, 597, A26

Bourrier, V., Lecavelier des Etangs, A., & Vidal-Madjar, A. 2014, A&A, 565, A105

Bourrier, V., Lecavelier des Etangs, A., Dupuy, H., et al. 2013, A&A, 551, A63

Bourrier, V., Ehrenreich, D., Allart, R., et al. 2017b, A&A, 602, A106

Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., et al. 2018, A&A, 620, A147

Bowyer, S., Lieu, R., Lampton, M., et al. 1994, ApJS, 93, 569

Carolan, S., Vidotto, A. A., Hazra, G., Villarreal D'Angelo, C., & Kubyshkina, D. 2021, MNRAS, 508, 6001

Chen, J., & Kipping, D. 2017, ApJ, 834, 17

Christensen, U. R., & Aubert, J. 2006, Geophysical Journal International, 166, 97

Christensen, U. R., Holzwarth, V., & Reiners, A. 2009, Nature, 457, 167

Clausen, N., & Tilgner, A. 2015, A&A, 584, A60

Cranmer, S. R. 2017, ApJ, 840, 114

Debrecht, A., Carroll-Nellenback, J., Frank, A., et al. 2020, MNRAS, 493, 1292

dos Santos, L. A., Ehrenreich, D., Bourrier, V., et al. 2019, A&A, 629, A47

dos Santos, L. A., Bourrier, V., Ehrenreich, D., et al. 2021, A&A, 649, A40

Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton University Press)

Dring, A. R., Linsky, J., Murthy, J., et al. 1997, ApJ, 488, 760

Efroimsky, M. 2012, ApJ, 746, 150

Ehrenreich, D., Bourrier, V., Bonfils, X., et al. 2012, A&A, 547, A18

Ehrenreich, D., Bourrier, V., Wheatley, P. J., et al. 2015, Nature, 522, 459

Fekel, F. C., & Henry, G. W. 2005, AJ, 129, 1669

Feldman, U., & Widing, K. G. 2002, Physics of Plasmas, 9, 629

Fontenla, J. M., Landi, E., Snow, M., & Woods, T. 2014, Sol. Phys., 289, 515

Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109

Gaidos, E. J., Henry, G. W., & Henry, S. M. 2000, AJ, 120, 1006

Garcia Muñoz, A., Fossati, L., Youngblood, A., et al. 2021, ApJ, 907, L36

Garcia Muñoz, A., Youngblood, A., Fossati, L., et al. 2020, ApJ, 888, L21

Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759

Guillot, T., Burrows, A., Hubbard, W. B., Lunine, J. I., & Saumon, D. 1996, ApJ, 459, L35

Gupta, A., & Schlichting, H. E. 2019, MNRAS, 487, 24

Hadden, S., & Lithwick, Y. 2017, AJ, 154, 5

Hanneson, C., Johnson, C. L., Mittelholz, A., Al Asad, M. M., & Goldblatt, C. 2020, Journal of Geophysical Research: Space Physics, 125, e2019JA027139, e2019JA027139 10.1029/2019JA027139

Henry, G. W. 1999, PASP, 111, 845

Holmström, M., Ekenbäck, A., Selsis, F., et al. 2008, Nature, 451, 970

Huitson, C. M., Sing, D. K., Pont, F., et al. 2013, MNRAS, 434, 3252

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

Johnstone, C. P. 2020, ApJ, 890, 79

Judge, P. G., Solomon, S. C., & Ayres, T. R. 2003, ApJ, 593, 534

Khodachenko, M. L., Shaikhislamov, I. F., Lammer, H., & Prokopov, P. A. 2015, ApJ, 813, 50

Kimble, R. A., Woodgate, B. E., Bowers, C. W., et al. 1998, ApJ, 492, L83

Kirk, J., Alam, M. K., López-Morales, M., & Zeng, L. 2020, AJ, 159, 115

Kite, E. S., & Barnett, M. N. 2020, Proceedings of the National Academy of Sciences, 117, 18264

Kulow, J. R., France, K., Linsky, J., & Loyd, R. O. P. 2014, ApJ, 786, 132

Lallement, R., Ferlet, R., Lagrange, A. M., Lemoine, M., & Vidal-Madjar, A. 1995, A&A, 304, 461

Laming, J. M. 2004, ApJ, 614, 1063

—. 2017, ApJ, 844, 153

Lampón, M., López-Puertas, M., Czesla, S., et al. 2021, A&A, 648, L7

Lavie, B., Ehrenreich, D., Bourrier, V., et al. 2017, A&A, 605, L7

Lee, E. J. 2019, ApJ, 878, 36

Lee, E. J., & Connors, N. J. 2021, ApJ, 908, 32

Lehmer, O. R., & Catling, D. C. 2017, ApJ, 845, 130

Linsky, J. L., Fontenla, J., & France, K. 2014, ApJ, 780, 61

Llama, J., & Shkolnik, E. L. 2016, ApJ, 817, 81

Lockwood, G. W., Skiff, B. A., Henry, G. W., et al. 2007, ApJS, 171, 260

Lopez, E. D., & Fortney, J. J. 2013, ApJ, 776, 2

—. 2014, ApJ, 792, 1

Loyd, R. O. P., & France, K. 2014, ApJS, 211, 9

Mann, A. W., Johnson, M. C., Vanderburg, A., et al. 2020, AJ, 160, 179

Martin, E. C., Fitzgerald, M. P., McLean, I. S., et al. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII, ed. C. J. Evans, L. Simard, & H. Takami, 107020A

Mazeh, T., Tamuz, O., & Zucker, S. 2007, Astronomical Society of the Pacific Conference Series, Vol. 366, The Sys-Rem Detrending Algorithm: Implementation and Testing, ed. C. Afonso, D. Weldrake, & T. Henning, 119

Mills, S. M., & Mazeh, T. 2017, ApJ, 839, L8

Mocquet, A., Grasset, O., & Sotin, C. 2014, Philosophical Transactions of the Royal Society of London Series A, 372, 20130164

Močnik, T., Hellier, C., Anderson, D. R., Clark, B. J. M., & Southworth, J. 2017, MNRAS, 469, 1622

Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, ApJ, 693, 23

Nixon, M. C., & Madhusudhan, N. 2021, MNRAS, 505, 3414

Oklopčić, A. 2019, ApJ, 881, 133

Oklopčić, A., & Hirata, C. M. 2018, ApJ, 855, L11

Oláh, K., Kővári, Z., Petrovay, K., et al. 2016, A&A, 590, A133

Owen, J. E., & Adams, F. C. 2014, MNRAS, 444, 3761

Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29

Parker, E. N. 1958, ApJ, 128, 664

Plachinda, S., Pankov, N., & Baklanova, D. 2011, Astronomische Nachrichten, 332, 918

Poppenhaeger, K., Schmitt, J. H. M. M., & Wolk, S. J. 2013, ApJ, 773, 62

Rafikov, R. R. 2006, ApJ, 648, 666

Rogers, L. A. 2015, ApJ, 801, 41

Rosén, L., Kochukhov, O., Hackman, T., & Lehtinen, J. 2016, A&A, 593, A35

Saxton, R. D., Read, A. M., Esquej, P., et al. 2008, A&A, 480, 611

Schneiter, E. M., Velázquez, P. F., Esquivel, A., Raga, A. C., & Blanco-Cano, X. 2007, ApJ, 671, L57

Sembach, K. R. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 166, Stromlo Workshop on High-Velocity Clouds, ed. B. K. Gibson & M. E. Putman, 243

Smith, R. K., Brickhouse, N. S., Liedahl, D. A., & Raymond, J. C. 2001, ApJ, 556, L91

Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, Nature, 557, 68

Stan Development Team. 2018, Stan Modeling Language Users Guide and Reference Manual, Version 2.18.0

Stone, J. M., Tomida, K., White, C. J., & Felker, K. G. 2020, ApJS, 249, 4

Suárez Mascareño, A., Rebolo, R., & González Hernández, J. I. 2016, A&A, 595, A12

Tasnim, S., & Cairns, I. H. 2016, Journal of Geophysical Research: Space Physics, 121, 4966

Tremblin, P., & Chiang, E. 2013, MNRAS, 428, 2565

Turbet, M., Bolmont, E., Ehrenreich, D., et al. 2020, A&A, 638, A41

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Venzmer, M. S., & Bothmer, V. 2018, A&A, 611, A36

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J. M., et al. 2003, Nature, 422, 143

—. 2008, ApJ, 676, L57

Vidotto, A. A., Gregory, S. G., Jardine, M., et al. 2014, MNRAS, 441, 2361

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

von Essen, C., Mallonn, M., Hermansen, S., et al. 2020, A&A, 637, A76

Wang, L., & Dai, F. 2018, ApJ, 860, 175

Wang, L., & Dai, F. 2021, ApJ, 914, 99

Wang, L., & Goodman, J. 2017, ApJ, 847, 11

Witte, M. 2004, A&A, 426, 835

Wolfgang, A., & Lopez, E. 2015, ApJ, 806, 183

Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, ApJ, 825, 19

Wood, B. E., & Linsky, J. L. 2010, ApJ, 717, 1279

Wood, B. E., Müller, H. R., Zank, G. P., Linsky, J. L., & Redfield, S. 2005a, ApJ, 628, L143

Wood, B. E., Redfield, S., Linsky, J. L., Müller, H.-R., & Zank, G. P. 2005b, ApJS, 159, 118

Woodgate, B. E., Kimble, R. A., Bowers, C. W., et al. 1998, PASP, 110, 1183

Zhang, M., Knutson, H. A., Wang, L., et al. 2021, AJ, 161, 181

Chapter 7

ESCAPING HELIUM FROM TOI 560.01, A YOUNG MINI NEPTUNE

Michael Zhang, Heather A. Knutson, Lile Wang, Fei Dai, and Oscar Barragán. Escaping Helium from TOI 560.01, a Young Mini-Neptune. *AJ*, 163(2):67, February 2022. doi: 10.3847/1538-3881/ac3fa7. M.Z. conceived the project, wrote the observing proposals, obtained the data, carried out the data analysis and the 1D simulations, and wrote the paper.

7.1 Abstract

We report helium absorption from the escaping atmosphere of TOI 560.01 (HD 73583b), a $R = 2.8R_\oplus$, $P = 6.4$ d mini Neptune orbiting a young (~600 Myr) K dwarf. Using Keck/NIRSPEC, we detect a signal with an average depth of $0.68 \pm 0.08\%$ in the line core. The absorption signal repeats during a partial transit obtained a month later, but is marginally stronger and bluer, perhaps reflecting changes in the stellar wind environment. Ingress occurs on time, and egress occurs within 12 minutes of the white light egress, although absorption rises more gradually than it declines. This suggests that the outflow is slightly asymmetric and confined to regions close to the planet. The absorption signal also exhibits a slight 4 km/s redshift rather than the expected blueshift; this might be explained if the planet has a modest orbital eccentricity, although the radial velocity data disfavors such an explanation. We use XMM-Newton observations to reconstruct the high energy stellar spectrum and model the planet's outflow with 1D and 3D hydrodynamic simulations. We find that our models generally overpredict the measured magnitude of the absorption during transit, the size of the blueshift, or both. Increasing the metallicity to 100× solar suppresses the signal, but the dependence of the predicted signal strength on metallicity is non-monotonic. Decreasing the assumed stellar EUV flux by a factor of 3 likewise suppresses the signal substantially.

7.2 Introduction

There is a growing body of evidence suggesting that atmospheric mass loss dramatically shapes the population of close-in exoplanets detected by transit surveys. This population is dominated by planets with radii between $1 - 4$ $R_\oplus$ ('sub-Neptunes'), which have no solar system analogue. Smaller sub-Neptunes ($1 - 1.7$ Earth radii)

appear to have Earth-like bulk compositions and are commonly referred to as 'super Earths,' while larger sub-Neptunes ($2 - 3$ Earth radii) are called 'mini Neptunes' and have low bulk densities indicating the presence of volatile-rich envelopes that typically constitute a few percent of their total mass. The two populations are separated by a gap in the radius distribution where few planets reside (Fulton et al., 2017; Fulton & Petigura, 2018). Sub-Neptunes are challenging to characterize, but new telescopes, observing techniques, and numerical models are opening up this frontier to exploration.

It has been suggested that all sub-Neptunes were formed with hydrogen-rich envelopes, which were then stripped away from the most highly irradiated and least massive planets. For young, low-density planets on close-in orbits, photoevaporation can drive strong hydrodynamic outflows that rapidly remove hydrogen-rich gas (e.g., Owen & Wu 2017; Mills & Mazeh 2017). However, this mass loss may also be driven by the newly formed planet's own internal luminosity (Ginzburg et al., 2018; Gupta & Schlichting, 2019). An alternate explanation for the radius gap is that it has nothing to do with mass loss, but is instead because cores have a broad mass distribution, with the smaller cores having never accreted gas in the first place (Lee & Connors, 2021). It is also possible that some mini Neptunes have no hydrogen-rich envelopes at all, but instead formed with substantial water-rich envelopes (e.g. Mousis et al. 2020). This could dramatically change the mass loss rate, especially that of helium, which would have been already lost to space alongside the primordial hydrogen.

Empirical measurements of mass loss from young sub-Neptunes provide a critical test of these competing hypotheses. We expect that mass loss rates will be highest at relatively early times, when the star's high-energy flux is enhanced and the planet is still inflated (Owen, 2019). By searching for evidence of outflows during this critical early period, we can determine whether or not a young sub-Neptune planet possesses a hydrogen and helium-rich envelope. When an outflow is detected, it can be used to test and refine mass loss models, with the eventual goal of developing an accurate understanding of atmospheric evolution through the planet's life.

We can search for outflows by measuring the amount of absorption from hydrogen and/or helium during the transit, but the small size of mini Neptunes makes them challenging targets for this technique. In a recent study, we obtained the first measurement of Lyα absorption from a young mini Neptune, HD 63433c (Zhang et al., 2021b). This planet is one of two transiting mini Neptunes orbiting a 400

Myr solar analogue. Our non-detection of similar absorption from HD 63433b suggests that the inner planet may not possess a hydrogen and helium-rich envelope at all, in agreement with its shorter predicted atmospheric lifetime. HD 63433 is a particularly favorable target for Lyα observations, as it is an active nearby star that is moving towards us. This results in an unusually high flux in the near blue wing of the Lyα line, which is expected to contain most of the absorption signal from the planet's escaping atmosphere. To date, Lyα absorption has only been conclusively detected for four planets aside from HD 63433c (Vidal-Madjar et al., 2008; Bourrier et al., 2013; Lavie et al., 2017; Bourrier, V. et al., 2018), all of which are larger than Neptune.

The metastable helium triplet at 1083 nm can also be used to probe atmospheric mass loss (Oklopčić & Hirata, 2018; Spake et al., 2018), and has the advantage of being readily accessible to ground-based observatories. The strength of the absorption in this line is less than that in Lyα, at most several percent, but this precision is achievable with many infrared spectrographs. However, observations in this triplet are largely restricted to planets orbiting active K stars, which have the optimal UV spectral shape to produce a significant population of metastable helium (Oklopčić, 2019). To date there have been multiple detections of outflows from close-in gas giant planets using this line, including the hot Jupiter HD 189733b (Salz et al., 2018), the inflated Saturn WASP-107b (Spake et al., 2018; Allart et al., 2019), and the warm Neptune GJ 3470b (Palle et al., 2020). However, helium mass loss has never been securely detected for planets smaller than 4 $R_\oplus$ despite many attempts (i.e. Kasper et al. 2020; Gaidos et al. 2020; Gaidos et al. 2020; Zhang et al. 2021a,b). Most of these published non-detections are of planets around old and inactive stars (HD 97658b, 55 Cnc e, GJ 1214, GJ 9857d), and the few observations of young planets were not very sensitive (K2-100b, K2-25b). We also did not detect helium absorption from HD 63433c, despite the Lyα detection. We suspect that the unfavorable host star type (G5) and underestimated outflow confinement mechanisms combined to suppress the size of the expected absorption signal in this line, while the unexpectedly high stellar variability in the 1083 nm line decreased the sensitivity of our observations.

The Transiting Exoplanet Survey Satellite (TESS) has vastly increased the sample of small transiting planets that are amenable to atmospheric characterization, including the aforementioned HD 63433 system. TESS recently identified TOI 560.01 (HD 73583b), a 2.83 ± 0.10 $R_\oplus$, $10.1^{+3.2}_{-3.0} M_\oplus$ planet orbiting a K4V star with a 6.4

d period (Barragán et al., 2021). This star is young: the SuperWASP project's photometry reveals a robustly detected rotation period of 12 days, corresponding to a gyrochronological age of ~600 Myr. It is also close by, with a distance of 31.6 pc and a J band magnitude of 7.6. The age and spectral type of the star, the size of the planet, and the closeness of the system combine to make TOI 560.01 an exceptionally favorable mini Neptune for probing helium mass loss. In addition, the planet has an outer companion at P=18.9 d with a very similar radius, making comparative mass loss studies possible.

In this paper, we use Keck/NIRSPEC to measure the helium absorption signal from TOI 560.01 and XMM to measure the star's high energy spectrum. We then compare our helium measurement to predictions from 1D and 3D photoevaporative mass loss models. We describe the observations in §7.3, our analysis of the data in §7.4, our reconstruction of stellar properties in §7.5, and our modelling of the outflow in §7.6. We discuss in §7.7 and summarize our conclusions in §7.8.

7.3 Observations and data reduction

Keck/NIRSPEC

We used Keck/NIRSPEC to observe a full transit of TOI 560.01 on March 18, 2021 UTC and a partial transit on April 19, 2021 UTC (40% of T_{14}). All observations were obtained in Y band using the high resolution mode with the 12 x 0.432″ slit, resulting in a spectral resolution of 25,000. We used 60 second exposure times, and adopted an ABBA nod pattern to help subtract background. Because we only used one coadd per exposure, we achieved a high efficiency of 77% for these observations. On March 17, the sky was clear and the seeing was 0.6–0.7". We observed 0.7 h of pre-transit baseline, the 2.1 h transit, and 0.6 h of post-transit baseline. We achieved a typical SNR of 170 per spectral pixel in one 60 second exposure. On April 18, the seeing was better (0.5" at both beginning and end of night), but there were sporadic cirrus clouds which caused large fluctuations in the water column density. We observed 1.9 h of pre-transit baseline and 0.8 h of transit before the target sank too low to observe. On this second night, we achieved a higher average SNR of 200 per spectral pixel, but with more weather-induced variability.

We calibrated the raw images and extracted 1D spectra for each order using a custom Python pipeline designed for the upgraded NIRSPEC. The pipeline is described in detail in Zhang et al. (2021a), but we summarize it here, along with the target-specific differences in our reduction for TOI 560. First, the pipeline subtracts crosstalk from

each raw frame. Then, it calibrates the raw frames by computing a master flat, identifying bad pixels in the process. It uses the master flat to compute a calibrated A-B difference image for each A/B pair. The spectral trace containing the 1083 nm line (spanning 10,803–11,008 Å) is identified, and we perform optimal spectral extraction to obtain spectra along with their errors. A template is computed from a model stellar spectrum and model tellurics, with the stellar spectrum shifted in wavelength to account for the star's average Earth-relative radial velocity during that night. We use the template to derive the wavelength solution for each individual spectrum.

After extracting the 1D spectra, we remove tellurics. In Zhang et al. (2021a), we did this by running SYSREM, which can be thought of as Principal Component Analysis with error bars (Mazeh et al., 2007). However, we have subsequently found that while SYSREM performs excellently in removing tellurics, it also removes half the planetary signal unless the planet, like 55 Cnc e, undergoes extreme radial acceleration during the observation (Zhang et al., 2021b). Furthermore, due to an unfortunate coincidence, a strong telluric water line falls right on top of the helium line during both nights of our observations. We therefore opted for a more conservative and physically motivated telluric removal method. We used `molecfit` (Smette et al., 2015), which models tellurics using a meteorological model for Earth's temperature-pressure profile at the time and place of observation, but allows the user to fit for the rapidly varying water column density. We fit for the water column density using narrow wavelength ranges containing strong telluric lines while avoiding any stellar lines. We also fit for the continuum with a line. The line spread profile is fixed to a Gaussian with a FWHM of 3.5 pixels, a value we settled upon after several fits to different spectra (with the FWHM as a free parameter) converged upon similar values. 3.5 pixels is close to the theoretical value of 3 pixels for the slit we used. After fitting the water lines, `molecfit` produces a telluric corrected spectrum across the full wavelength range of the order containing the helium line, which we adopt for the rest of the analysis.

Having obtained wavelength calibrated and telluric corrected spectra, we interpolate all spectra onto a common wavelength grid with a uniformly logarithmic spacing of $\lambda/110,000$ and a range of 10,810–10,850 Å. We remove fringing by applying a notch filter twice, using exactly the same parameters described in Zhang et al. (2021a). We divide each spectrum by the continuum, take the logarithm of the entire spectral grid, and subtract the mean of every row and column from that row and column.

This results in a residuals grid: a $N_{obs} \times N_{wav}$ grid of numbers representing the relative deviation of a pixel from the mean for that row and column. For every column (wavelength), we subtract the mean of the out-of-transit part of the residuals image for that column; we then invert the residuals image. The residuals image now gives excess absorption relative to the out-of-transit baseline. However, there are continuum variations that cause structure in this image. We mask out the helium line, mask out the strong lines because optimal extraction performs poorly with them Zhang et al. (2021a), fit a 3rd order polynomial to each row (epoch) with respect to wavelength, and subtract off the polynomial.

XMM-Newton

On April 21, XMM-Newton measured the star's X-ray and MUV spectrum, which are crucial for modeling photoevaporative mass loss and predicting the metastable helium population. We observed the system for a total of 13 ks as part of XMM prop. ID 088287 (PI: Michael Zhang). We configured the EPIC cameras to observe with the medium filter and small window, giving us 97% observing efficiency on the two MOS CCDs and 71% efficiency on the one pn CCD. We configured the Optical Monitor to observe the star in the UVM2 filter (231 ± 48 nm) for 7.2 ks and the UVW2 filter (212 ± 50 nm) for 4.4 ks. Although these observations are not simultaneous with the helium mass loss observations, they were taken a month after the first helium observation, a short time compared to typical stellar activity cycles.

To analyze XMM-Newton data, we download the raw Observation Data File (ODF) and use the Science Analysis System (SAS) provided by the XMM-Newton team to reduce it. We run `xmmextractor`, to obtain spectra from the ODF with default settings. For the Optical Monitor, SAS reports the count rate in the UVW2 and UVM2 filters, the two mid ultraviolet filters we selected. For each of the three EPIC detectors, SAS generates the light curve and the background-subtracted spectrum.

For the pn detector, we find that the automatic reduction by `xmmextractor` leads to significantly negative fluxes at 0.23–0.28 keV, which is unphysical. Therefore, we use SAS to manually reduce the pn data by defining the source region as a circle 17.5″in radius, and the background region as an annulus centered on the source with an inner radius of 20″and an outer radius of 35″. Defined in this way, the spectrum no longer has significantly negative fluxes around 0.25 keV.

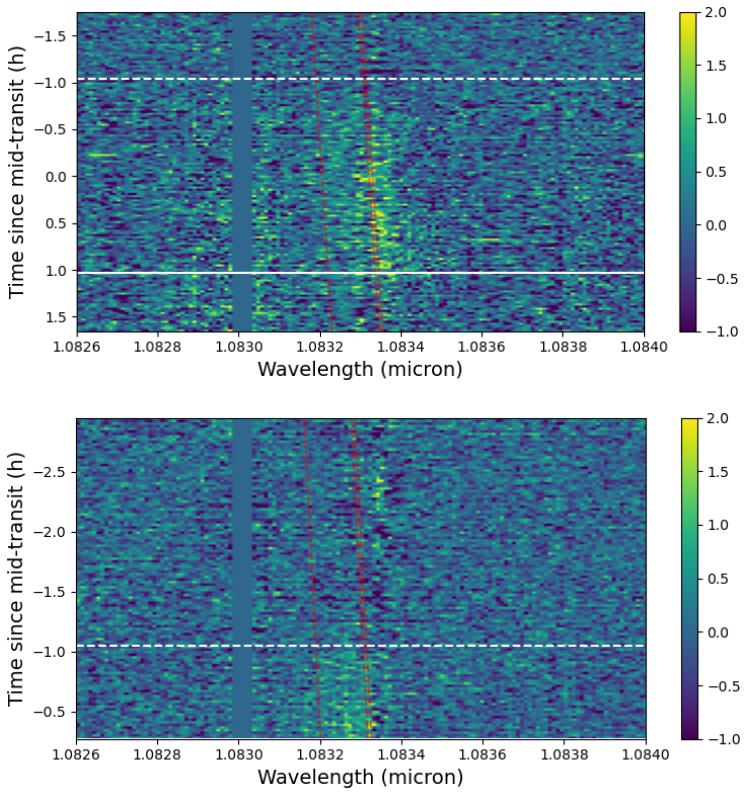


Figure 7.1: Percent excess absorption from TOI 560.01 as a function of time and wavelength, for the first (left panel) and second (right panel) nights of observation. The dashed white line indicates the beginning of the white light transit, while the solid white line indicates the end. The red lines show the wavelengths of planetary helium absorption. At 10,830 Å is a deep stellar Si I line, which, like other strong lines, we mask as part of our analysis because optimal extraction deals poorly with very strong lines (Zhang et al., 2021a).

7.4 Analysis of helium transit observations

Figure 7.1 shows the excess absorption in the 1083 nm metstable helium triplet as a function of time and wavelength, for both nights of observation with Keck/NIRSPEC. To compute the excess absorption, we adopt a baseline that includes all pre-ingress spectra and (for the first night) all spectra taken more than 12 minutes after egress. Ingress and egress are computed purely based on the white light ephemeris. Figure 7.2 shows the band-integrated light curve for each night in the helium line. We detect strong absorption starting from the white light ingress, reaching a maximum of 1.7% half an hour after the midpoint of the white light transit, and declining quickly after white light egress. The excess helium absorption begins on time on both nights, but ends ∼10 min late on the first night, the only night we captured the egress. The band-integrated excess absorption is ∼0.7%, and the equivalent

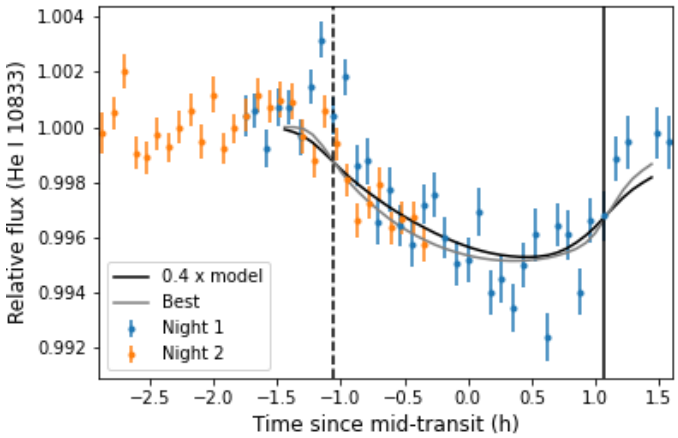


Figure 7.2: Light curve of the helium line (top) in a 1.5Å bandpass centered on 10833.27Å. The dashed black line marks the beginning of the white light ingress, while the solid black line marks the end of white light egress. We overplot the predicted light curves from the fiducial and best fit Microthena models as solid black and grey lines, respectively. We rescale the amplitude of the fiducial model light curve by a factor of 0.4 to match the amplitude of the observed signal.

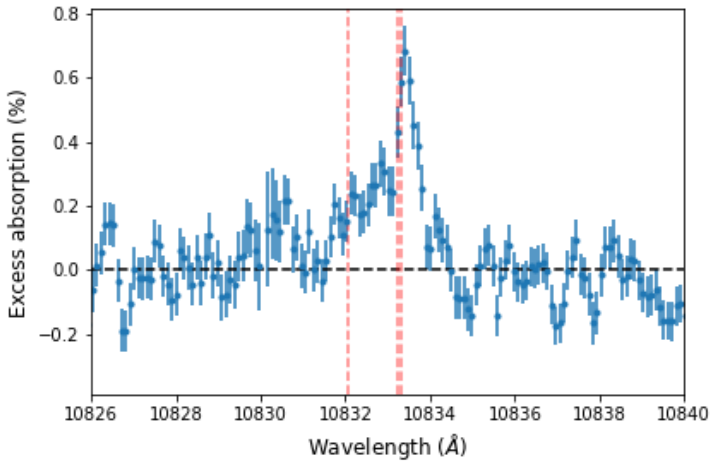


Figure 7.3: Excess absorption spectrum from the full transit observation on night 1, plotted in the planetary frame.

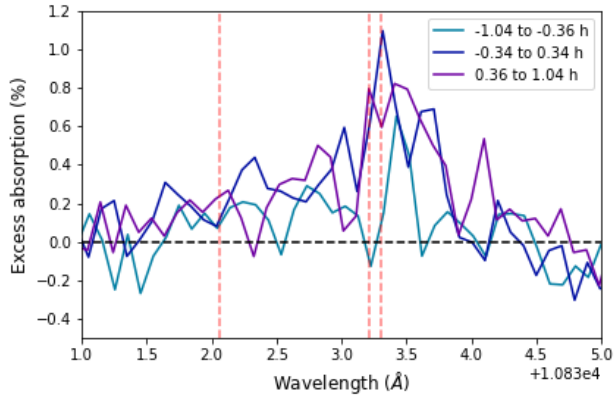


Figure 7.4: Excess absorption spectrum for each third of the transit on night 1. Wavelengths are in the planetary frame.

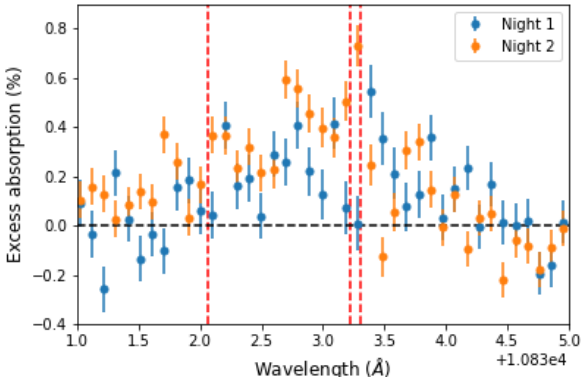


Figure 7.5: Comparison of the excess absorption spectrum on both nights, averaged over only the in-transit phases observed on night 2 (first 40% of transit).

width of the planetary absorption is 7 ± 0.4 mÅ. The absorption spans at least 1.5Å, corresponding to a velocity spread of 40 km/s—far higher than the escape velocity from the planetary surface of 15 km/s, or the escape velocity from the approximate helium absorption radius ($\sim 4R_p$) of 7.5 km/s. This shows that the atmosphere is escaping, because velocity dispersion is the only significant broadening mechanism: natural broadening is of order 0.01 km/s, the equatorial rotational velocity of the planet is 0.2 km/s, and pressure broadening is $\sim 10^{-11}$ km/s at picobar pressures because it is ~10 km/s at atmospheric pressure (Niermann et al., 2010). For unknown reasons, the helium flux rises by 0.2% right before ingress on the first night. Since we do not see the same brightening on the second night, and since there is no plausible way the planetary outflow can cause the star to brighten before the transit, we attribute this brightening to stellar activity. The helium line is well known to

be a tracer of chromospheric activity, which is more intense in younger stars. In previous observations we monitored the young G5 star HD 63433 for two nights, and saw $>\sim 0.2\%$ variability on both nights (Zhang et al., 2021b).

In Figure 7.3, we examine the wavelength-dependence of the transit-averaged excess absorption signal from the first night. We find that it is redshifted by ~ 0.14 Å relative to the radial velocity of the planet, corresponding to a velocity of ~4 km/s. The 1D absorption spectrum peaks at $0.68 \pm 0.08\%$ and declines quickly in the red wing but slowly in the blue wing, consistent with gas being pushed towards the observer by radiation pressure or stellar wind. This extended tail of blue-shifted absorption is similar to that observed for WASP-107b (Allart et al., 2019), a planet with a far more extended egress in the integrated helium light curve. In Figure 7.4, we divide the transit into thirds and show the excess absorption spectrum for each third. Consistent with the other plots, there is slightly redshifted absorption in each part, with the final third having much stronger absorption than the first third.

In Figure 7.5, we compare the excess absorption spectrum for the two nights averaged over the in-transit phases observed on the second night. We find that the absorption observed on the second night appears slightly stronger and more blueshifted than it did during the equivalent time window on the first night. The two nights are otherwise consistent. To quantify the significance of the differences between the two nights, we used nested sampling as implemented by `dynesty` (Speagle, 2020) to fit a Gaussian to the excess absorption spectra plotted in Figure 7.5, with 3 free parameters: amplitude, standard deviation, and mean. We find a mean position of 10833.06 ± 0.15 Å on night 1 and $10,832.70 \pm 0.06$ Å on night 2, a difference of 0.36 ± 0.16 Å (10 ± 4 km/s in velocity space). The amplitude was $0.22 \pm 0.04\%$ for night 1 and $0.40 \pm 0.04\%$ on night 2, for a difference of $0.18 \pm 0.06\%$. These statistical tests confirm what visual inspection shows: on night 2, the helium absorption was marginally stronger and marginally bluer. Further observations are necessary to determine whether this variability is due to underestimated error bars, a consequence of poorly understood stellar variability, or a change in the properties of TOI 560.01's outflow.

On the second night of observation, the stellar helium line was slightly narrower (by 8%) than on the first night, but of indistinguishable depth (to within 1.5%) and position (to within 0.5 km/s). Minor differences in the stellar line do not affect our results for the planetary excess absorption, which is computed by comparing in-transit and out-of-transit spectra on the same night.

Possible causes of the observed redshift

The redshifted absorption peak seen during the first night (Figure 7.1 and 7.3) is unusual. Aside from HAT-P-32b (Czesla et al., 2021), none of the gas giant planets with spectrally resolved helium observations have exhibited a similar redshift (e.g., Salz et al., 2018; Allart et al., 2019; Palle et al., 2020), and material flowing away from the star should appear blueshifted during transit. If the observed redshift is due to the geometry of TOI 560.01's atmospheric outflow, it would suggest that TOI 560.01 has unusual outflow properties, perhaps implying an unexpected weak stellar wind. However, we must first ascertain whether the apparent redshift is real.

The apparent redshift is unlikely to be due to wavelength calibration uncertainties. NIRSPEC does experience wavelength drift over the course of a night (Kasper et al., 2020), but our wavelength calibration is performed independently on every spectrum, so drift does not affect us. In addition, we computed the wavelength solutions with an alternate method that uses the median observed spectrum instead of the theoretical spectrum as a template. Comparing the results, we conclude that the wavelength solution for each spectrum is accurate to at least 1 km/s.

The apparent redshift is also not due to ephemeris uncertainties. The ephemeris predicts the transit mid-point on the first night with a 1σ uncertainty of 1 minute, during which time the planet accelerates by 0.07 km/s. This is much smaller than the observed redshift of ∼4 km/s. Similarly, uncertainties on the stellar mass and semimajor axis can only change the acceleration of the planet by $\sim 7\%$. Since the planet accelerates from −4.5 km/s to +4.5 km/s during the transit, a 7% change is insufficient to account for the observed redshift.

The final source of uncertainty is the eccentricity. To obtain the radial velocity of the planet, we assumed a perfectly circular orbit. However, Mills et al. (2019) used the transit durations of 1000 Kepler planets, combined with accurate stellar radii from the California-Kepler Survey and Gaia, to statistically infer typical eccentricities. They obtained a mean eccentricity of 0.05 for systems with multiple transiting planets. The star-planet distance changes at a maximum rate of $K_p e$ where K_p is the orbital speed; since K_p = 102 km/s for TOI 560.01, an eccentricity of 0.05 can cause an apparent helium signal redshift of up to 5.1 km/s. Barragán et al. (2021) constrained the eccentricity to $0.10^{+0.08}_{-0.07}$ using a joint fit to the radial velocity and light curve, which is consistent with both 0 and 0.05. However, their radial velocity fit results in a planetary radial velocity of $-7.4^{+9.3}_{-7.0}$ km/s at mid-transit, which is in the wrong direction, albeit consistent with a 4 km/s redshift to within 2σ. We

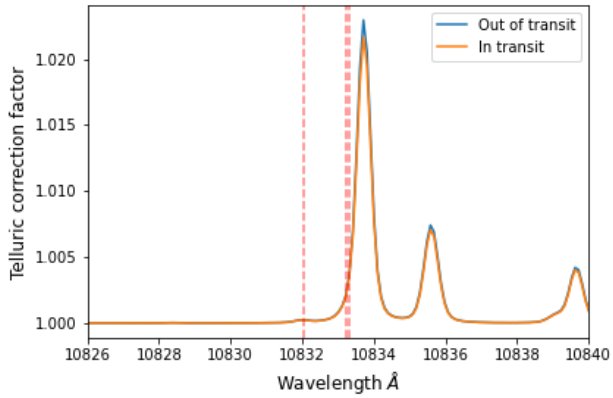


Figure 7.6: Median telluric correction factors inferred by molecfit, for the in-transit and out-of-transit spectra. The wavelengths of the helium triplet are indicated with vertical red lines.

are therefore unable to differentiate between a redshift caused by a non-zero orbital eccentricity and one caused by the geometry of the outflow.

Telluric absorption

As we mention in Subsection 7.3, a telluric water absorption line overlaps with the red end of the stellar helium line, and we use molecfit to correct for this telluric absorption. One might wonder whether imperfect telluric correction impacts our analysis, and in particular, whether it might be the cause of the apparent redshift in the absorption signal.

Figure 7.6 shows the median correction that molecfit applied to the in-transit and out-of-transit spectra. Although the total correction was, at maximum, 2%—which is comparable to the highest excess absorption from the planet—the correction is nearly identical for the in-transit and out-of-transit spectra, with a difference of only 0.12%. However, the out-of-transit spectra span a larger range of airmass than the in-transit spectra, so comparing only the median correction factors does not tell the whole story.

To get a better idea of the effects of telluric correction, we looked at two other telluric lines redward of the helium line: one with a slightly lower absorption depth (1.5%), and one with a significantly higher absorption depth (4%). Figure 7.7 shows the excess absorption spectrum (both 2D and 1D) in the region of wavelength space around these lines. For both lines, there are no statistically significant features in either plot. In the 1D spectrum, there does appear to be small (0.1%) dips redward

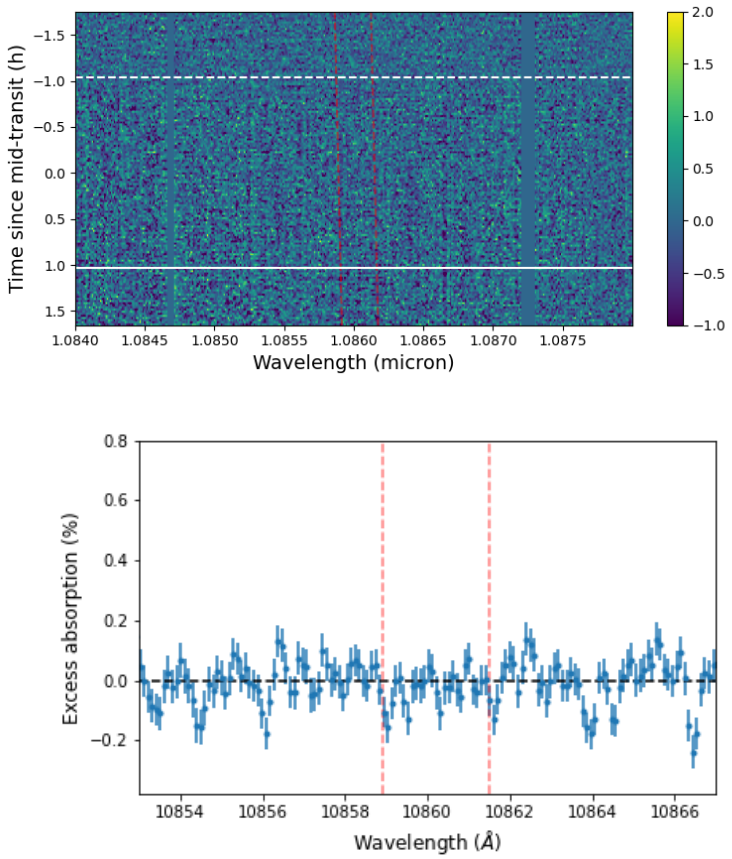


Figure 7.7: To explore the effect of telluric correction on the helium signal, which overlaps with a telluric line, we choose two other telluric lines: one slightly weaker, and one significantly stronger. We show the 2D excess absorption spectrum in the stellar frame (top) and the 1D excess absorption spectrum in the planetary frame (bottom), with the wavelengths of the telluric lines marked in red; the bluer line is the weaker. These plots are analogous to Figure 7.1 and Figure 7.3.

of both telluric lines. These dips are probably coincidental, but if not, they would only strengthen our conclusion that the helium absorption is redshifted.

RM and CLV

The Rossiter-Mclaughlin (RM) effect and center-to-limb variations (CLV) plague high-resolution transit spectroscopy of giant exoplanets. They introduce pseudo-signals which can be a substantial portion of the helium absorption signal for planets like HD 189733b (Salz et al., 2018).

The RM effect occurs when the planet or its escaping atmosphere blocks out a portion of the rotating stellar disk. Because the portion it blocks out has a non-zero rotational velocity, the star appears to experience a radial velocity change during

the course of the transit. Combining the radius and rotational period of TOI 560, we calculate a rotational speed of 2.8 km/s. If escaping atmosphere blocks 0.7% of the limb, an apparent radial velocity shift of (0.7%)(2.8) = 0.02 km/s is created. To estimate the effect this has on the excess absorption, we multiply by the maximum derivative of the relative flux in the vicinity of the helium line, $\frac{dln(F)}{dv}$ = 0.02 km^{-1} s. We obtain 0.04%, well below the average excess absorption of 0.7%, and below our noise.

The CLV effect occurs because line centers experience different limb darkening from the continuum. Standard stellar spectra libraries such as PHOENIX (Husser et al., 2013) and MARCS (Van Eck et al., 2017) do not model the chromosphere, and therefore do not include the helium line. Without knowing how the helium line depth varies across the stellar surface, it is impossible to precisely model the CLV, but we can do an approximate calculation. de Jager et al. (1966) reported that for the Sun, the center ($\theta = 0°$) has a 10833Å line depth of 6% while the edge ($\theta = 70°$) has a line depth of 10%. TOI 560 has a much deeper 10833Å line, at 30%. If the center of the star has a line depth of 20% and the edge has a depth of 40%, when the escaping atmosphere blocks 1.7% of the center of the stellar disk, the stellar pseudo-signal would increase the apparent helium signal by (10%)(1.7%) = 0.17%. This is not negligible, but it is a small fraction of the total signal. It is not possible under any circumstances for the entire helium signal to be due to CLV because the white light transit depth is only 0.16%, far smaller than the measured helium signal.

7.5 The star

Age

We estimate the star's age with gyrochronology. The star has a rotation period of 12.2 ± 0.3 d and a mass of 0.71 ± 0.02$M_\odot$ (Barragán et al., 2021). It has a B-V color of 1.112 ± 0.001, from the Hipparcos input catalog (Turon et al., 1993). Using Schlaufman (2010), which relates age to stellar mass and rotation period, we obtain an age of 635 ± 40 Myr. Using Mamajek & Hillenbrand (2008), which relates age to B-V color and rotation period, we obtain 540 ± 80 Myr. The error bars for the first estimate are purely statistical, with no model uncertainty, while the error bars for the second estimate include the statistical uncertainties of the model parameters.

X-ray observations

We analyzed the XMM data using the approach described in Zhang et al. (2021b), which we summarize here. We used `xspec` to fit a model consisting of two com-

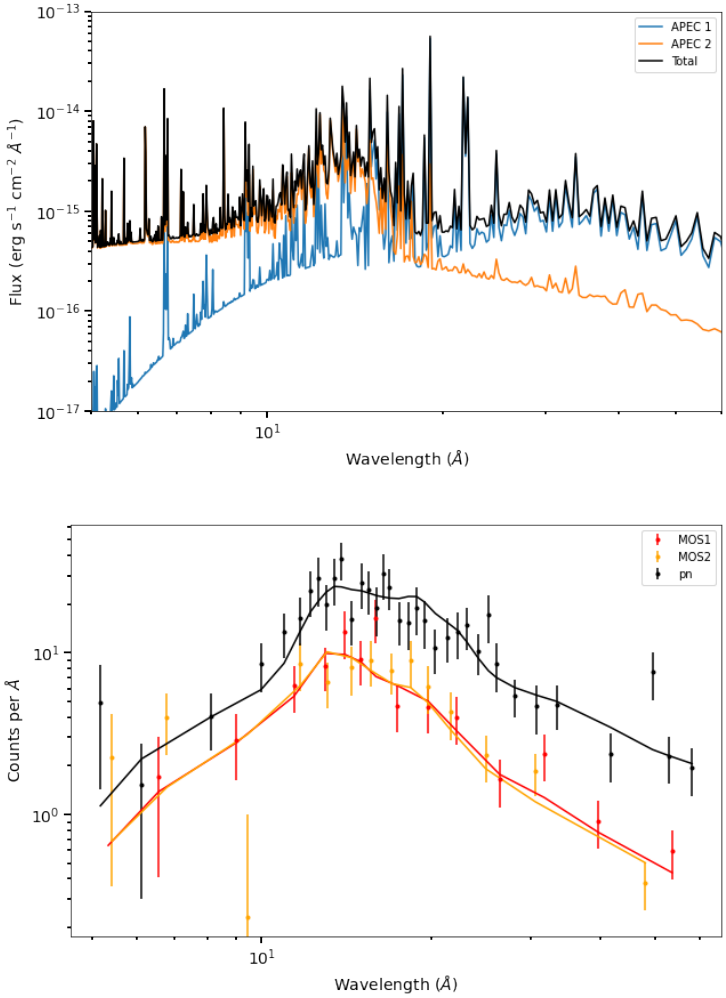


Figure 7.8: X-ray observations. Top: Best fit model spectrum from our XMM EPIC observations of TOI 560. The total spectrum as observed from Earth (black) is the sum of a low-temperature component (APEC 1) and a high-temperature component (APEC 2). Bottom: the data and the folded models, which take into account the instrumental throughput and line spread function.

ponents of optically thin, collisional plasma in equilibrium. We also tried one and three components, but found that two components minimizes the Bayesian Information Criterion. We show the data and fitted model in Figure 7.8, and list the corresponding model parameters in Table 7.1. The model is a good fit to the data for all three EPIC detectors, and the fit parameters are consistent with expectations for a moderately active star. As discussed in Wood & Linsky (2010) and Zhang et al. (2021b), coronal metallicities are often lower than the equivalent photospheric metallicities.

Table 7.1: Model parameters for XMM-Newton data

Parameter	Value
Metallicity	0.60 ± 0.26
kT_1 (keV)	0.25 ± 0.03
EM_1 (cm^{-3})	$7.3 \pm 2.5 \times 10^{50}$
kT_2 (keV)	0.94 ± 0.13
EM_2 (cm^{-3})	$2.2 \pm 0.7 \times 10^{50}$
Flux* (erg/s/cm^2)	$1.05 \pm 0.06 \times 10^{-13}$

*Derived, not a fit parameter. For the range 5–100 Å (0.124–2.48 keV).

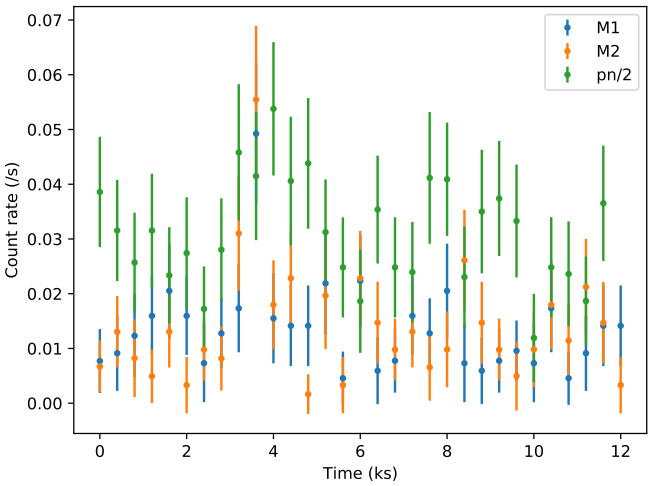


Figure 7.9: Background-subtracted X-ray light curves of TOI 560, recorded by the three EPIC cameras.

From the model fits, we derive a 5–100 Å flux of $1.05 \pm 0.06 \times 10^{-13}$ erg/s/cm^2. However, the star's intrinsic variability is likely to be much higher than the reported precision for this flux measurement. Figure 7.9 shows the light curve measured by the EPIC detectors during the 12 ks observation. As is typical for X-ray observations of active stars, the light curve is variable at the 20% level on kilosecond time scales.

Lyα and EUV

The star's extreme UV flux (roughly 100–912 Å) is the predominant driver of photoevaporative mass loss. Unfortunately, no EUV telescopes currently exist, and strong interstellar absorption makes it difficult to accurately measure EUV flux from even the closest stars. To obtain the EUV spectrum, we use the scaling relations of Linsky et al. (2014), which give the EUV flux in 100 Å bins with respect to the Lyα flux. Unfortunately, there are also no measurements of the star's Lyα flux. To

compute the flux, we use Linsky et al. (2013), which provides a scaling relation with respect to the rotation period and another with respect to the X-ray flux, with mean dispersions of 32% and 22%, respectively. For TOI 560, both scaling relations give the same Lyα flux: 18 erg s^{-1} cm^{-2} at 1 AU.

MUV

The 1230–2588 Å stellar flux (which we will call "MUV") ionizes triplet state helium, but does not contribute to producing it because it cannot ionize hydrogen or helium. It therefore plays a crucial role in determining the size of the metastable helium population. To quantify the MUV luminosity, we observed the star in two MUV filters using XMM-Newton's Optical Monitor. The Optical Monitor recorded a count rate of 0.654 ± 0.011 s^{-1} in the UVM2 filter and 0.565 ± 0.015 s^{-1} in the UVW2 filter. Correcting for the minor coincidence losses, we obtain count rates of 0.679 ± 0.011 s^{-1} and 0.585 ± 0.015 s^{-1}, respectively.

As a starting point for the UV spectrum, we used a PHOENIX model (Husser et al., 2013) for a star with T_{eff} = 4600 K, log(g)=4.5, and [M/H]=0. We then rescaled this model to reflect the radius of the star, its distance, and the effective areas of the filters as a function of wavelength, and used it to predict the count rate in each filter. We obtain $0.41 s^{-1}$ for UVM2 and $0.45 s^{-1}$ for UVW2, substantially below the measured values. To rectify the mismatch, we multiplied the PHOENIX UV flux (1230–3500 Å) by a factor of 1.67, bringing the predicted count rates (0.677 and 0.586 s^{-1}) into exact alignment with our measurements.

Reconstructed stellar spectrum

Table 7.2: Band-integrated fluxes

Band	Wavelengths (Å)	Flux at 1 AU (cgs)
X-ray	5 – 100	4.4 ± 2.2
EUV	100 – 912	14 ± 4
Lyα	1214 – 1217	18 ± 4
MUV	1230 – 2588	11 ± 2
Total	5 – 50,000	$2.47 \pm 0.13 \times 10^5$

Figure 7.10 shows the reconstructed stellar spectrum, while Table 7.2 lists the corresponding band-integrated fluxes. The error bars on the band-integrated fluxes, which are very approximate, are computed using the methodology of Zhang et al. (2021b). We find that the ratio of MUV to XUV (X-ray + EUV) flux for this star is 0.6, which is relatively low. For the young solar analogue HD 63433, this ratio

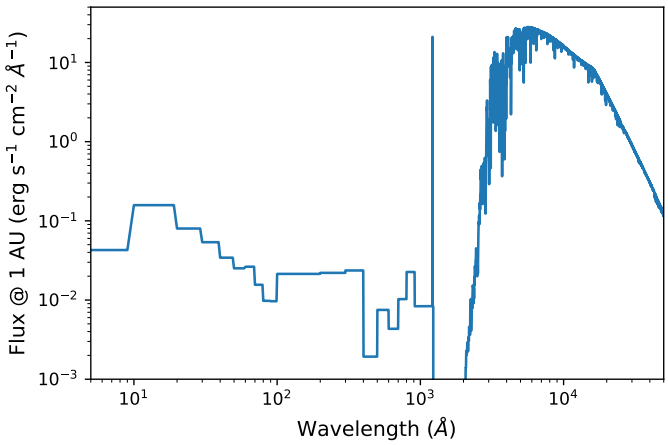


Figure 7.10: Fiducial stellar spectrum used in our mass loss modeling. We bin the X-ray data for better visibility in this figure. The errors associated with the spectrum are listed in Table 7.2, and were computed using the methodology of Zhang et al. (2021b).

is 63 (Zhang et al., 2021b). As expected, K-type stars like TOI 560 have lower MUV-to-XUV ratios than G stars (Oklopčić, 2019). Low MUV-to-XUV ratios are more favorable for helium observations because the MUV flux destroys triplet helium while the XUV flux, although capable of destroying triplet helium, also ionizes hydrogen and helium. This creates the electrons and ionized helium which then recombine to produce triplet state helium.

Stellar wind

The stellar wind shapes the planetary outflow by confining it and pushing it away from the star, creating a Coriolis force that can lead to the formation of a comet-like tail. However, the density and speed of the solar wind are highly variable with time, and there are few observational constraints on winds from other stars. To infer the stellar wind conditions, we rely on Wood et al. (2005), which used astrospheric absorption to characterize the mass loss rates of a handful of nearby stars. This absorption occurs where the stellar wind collides with the interstellar medium and charge exchange creates a population of hot neutral hydrogen, which then absorbs the stellar Lyα emission line. In order to translate a measured Lyα profile into a mass loss rate, one must accurately model the intrinsic stellar Lyα emission, the population of energetic neutral atoms, and the neutral hydrogen of the interstellar medium along the line of sight, none of which are trivial problems. Despite these challenges, Wood et al. (2005) were able to use their measured mass loss rates to

derive a scaling relation with the star's X-ray flux. We use this relation to obtain an estimated mass loss rate of 11 times solar for TOI 560. We further assume that the wind speed is comparable to that of the Sun (400 km/s), and use these values for our fiducial 3D models described in §7.6.

The mass loss rate we derived is consistent with the new data presented by Wood et al. (2021) (their Figure 10). However, the new data also shows that among the 7 GK dwarfs with an X-ray flux similar to TOI 560, the range in inferred mass loss rate per unit surface area is over 200x. This means it is unfortunately impossible to predict the stellar mass loss rate to even an order of magnitude with any confidence.

7.6 Modeling

We model the outflow using two different hydrodynamic codes: The PLUTO-CLOUDY Interface (TPCI), a 1D code developed by Salz et al. (2015a); and Microthena, a 3D code developed by Wang & Dai (2018). Both codes have been extensively used to study photoevaporation (e.g. Salz et al. 2015b, 2016; Kasper et al. 2020; Zhang et al. 2021a for TPCI; Wang & Dai 2020, 2021; Zhang et al. 2021b for Microthena). As we discuss below, we view the 3D simulations as a more accurate predictor of the observed helium signal during transit, because photoevaporation is fundamentally a 3D problem. Different areas of the planet receive different stellar irradiation levels, and the stellar wind cannot be ignored for young and active stars. Nevertheless, we find it useful to run both models for two reasons. First, 1D models much simpler than TPCI are widely used to interpret helium observations, most commonly variants of the Parker wind model proposed by Oklopčić & Hirata (2018) (e.g. Kasper et al. 2020). By running both 1D and 3D models for TOI 560.01 and comparing their predictions, we can evaluate the magnitude of the error introduced by assumptions of radial symmetry for this planet. Second, there are aspects of the problem that TPCI handles better than our 3D models, which must use a simplified physics framework in order to remain computationally tractable. In particular, Microthena divides radiation into just seven energy bins and only computes atomic line cooling from a small number of species, while TPCI propagates the entire continuum and computes line cooling from all neutral and ionized species. We therefore take advantage of TPCI's sophisticated radiative transfer to explore the qualitative effects of increasing metallicity (up to 100× solar) on the absorption signature. This is a topic that has been relatively unexplored in the literature to date, but TOI 560.01's small size and high bulk density mean that it could easily host a metal-rich atmosphere.

1D TPCI models

TPCI is a combination of two sophisticated and widely used codes: the hydrodynamic solver PLUTO (Mignone et al., 2007), and the plasma simulation and spectral synthesis code CLOUDY (Ferland et al., 2013). CLOUDY computes the equation of state and heating and cooling rates and provides these to PLUTO; PLUTO evolves the fluid using the Euler equations and provides the new conditions to CLOUDY, restarting the cycle.

We set up the simulations following the methodology of Zhang et al. (2021a), which in turn is based on the methodology of Salz et al. (2015b). Briefly, we set the lower boundary condition at a radius of 2.9 $R_\oplus$, a particle number density of 10^{14} cm^{-3}, and a pressure computed from the number density and a pressure of $P = nk_BT_{eq} = 10.2$ dynes/cm^2 (assuming $T_{eq} = 740$ K). The planetary mass is fixed at 9 $M_\oplus$, derived from the mass-radius relation of Chen & Kipping (2017) and consistent with the RV-measured mass of $10.06^{+3.24}_{-2.98} M_\oplus$. The stellar spectrum is derived in Section 7.5. In Zhang et al. (2021a), we did not include any metals or molecules; here, we continue to neglect molecules, but include metals with a solar abundance greater than 10^{-5}: C, N, O, Ne, Mg, Si, S, Fe. For our four simulations, the abundances of these metals are scaled from solar by a factor of 1, 10, 30, and 100, and the mean molecular weight is adjusted accordingly. At Z=1, hydrogen makes up 92% of the atoms, helium 8%, and the metals less than 0.1%. By weight, the ratios are 74%, 25%, and 1.3%. At Z=100, even though the metals comprise only 9% of the atoms, they comprise 57% of the mass; hydrogen and helium are only 32% and 11%, respectively. At this metallicity, hydrogen and helium become minor species by mass.

We evolve each simulation for at least 100 time units without advection, where 1 unit is approximately the sound crossing time of $R_p/(10$ km/s). We then evolve the simulation for at least 10 time units with advection turned on (100 units for Z=1, 10, and 30), at which point we perform a visual inspection to ensure convergence has been reached. We use the conditions at the last timestep to compute the helium absorption depth. Since the outflow is 3D and we only simulate the substellar point, we must rescale the model predictions to approximate the 3D geometry. Stone & Proga (2009) showed that the mass loss rate is higher by a factor of 4 in the spherically symmetric simulation, and we therefore adopt this as our rescaling factor, as was done by Salz et al. (2016). We implement the rescaling by dividing both the velocities and the densities by two, because this gives a helium line width

comparable to the observed width. Given that photoevaporation is inherently a 3D process, there is no "right" way to obtain a transit spectrum from a 1D simulation.

Figure 7.11 shows how the the resulting excess absorption spectrum varies as a function of metallicity, while Table 7.3 list the mass loss rate, peak absorption, and equivalent width of the absorption. Despite our rescaling to account for 3D effects, we find that our 1D models systematically overpredict the measured helium absorption signal, which peaks at 1.7%. The model with a metallicity of 100× solar has the weakest predicted helium absorption signal. However, Figure 7.11 shows that the predicted helium absorption signal does not decrease monotonically with increasing metallicity. In fact, the helium signals for the 10× and 30× solar metallicity models are stronger than for the solar metallicity model. The mass loss rate follows a similar pattern: we obtain 0.32, 0.42, 0.37, and 0.13 $M_\oplus$/Gyr for the 1×, 10×, 30×, and 100× metallicities, respectively. The velocity at the edge of our simulation (15 R_p) decreases with increasing metallicity, from 29 km/s to 20 km/s, while the location of the sonic point increases slightly with metallicity, with a minimum value of 2.5 R_p at Z=1 and a maximum value of 3.1 R_p at Z=100. The density is fairly similar: 50% higher for Z=10 and 30 than for Z=1, 40% lower for Z=100 than for Z=1.

While the density in the far regions is similar between the 1× and 100× solar metallicity models, n_e is 2.5× lower in the 100× solar model, while $n_{\mathrm{He}II}$ is 6× lower. Since triplet helium is primarily produced by recombination at a rate of $n_e n_{\mathrm{He}II}\alpha$ and destroyed by photoionization, and the recombination coefficient α is a factor of two lower in the higher metallicity case due to the higher temperature (Pequignot et al., 1991), one would expect the triplet helium density in the outer regions to be 30× lower—in approximate agreement with the TPCI models. We ascribe the low n_e at high metallicity to the lower mass density, and to the lower number of easily ionizable electrons per unit mass. For example, if we define easily ionizable as having an ionization energy smaller than 54 eV (the second ionization energy of helium), hydrogen and helium have 1 easily ionizable electron per amu, carbon has 0.25, neon has 0.1, and iron has 0.07. The lower $n_{\mathrm{He}II}$ at high metallicity is due to a combination of helium being less abundant (11% by mass vs. 25%) and the second-ionized fraction being higher (77% vs 56%), leaving fewer helium atoms in the singly ionized state. The latter, in turn, is due to a combination of n_e being 2.5× lower, and the temperature-dependent helium recombination coefficient being a factor of two lower. In short, helium absorption drops dramatically at high

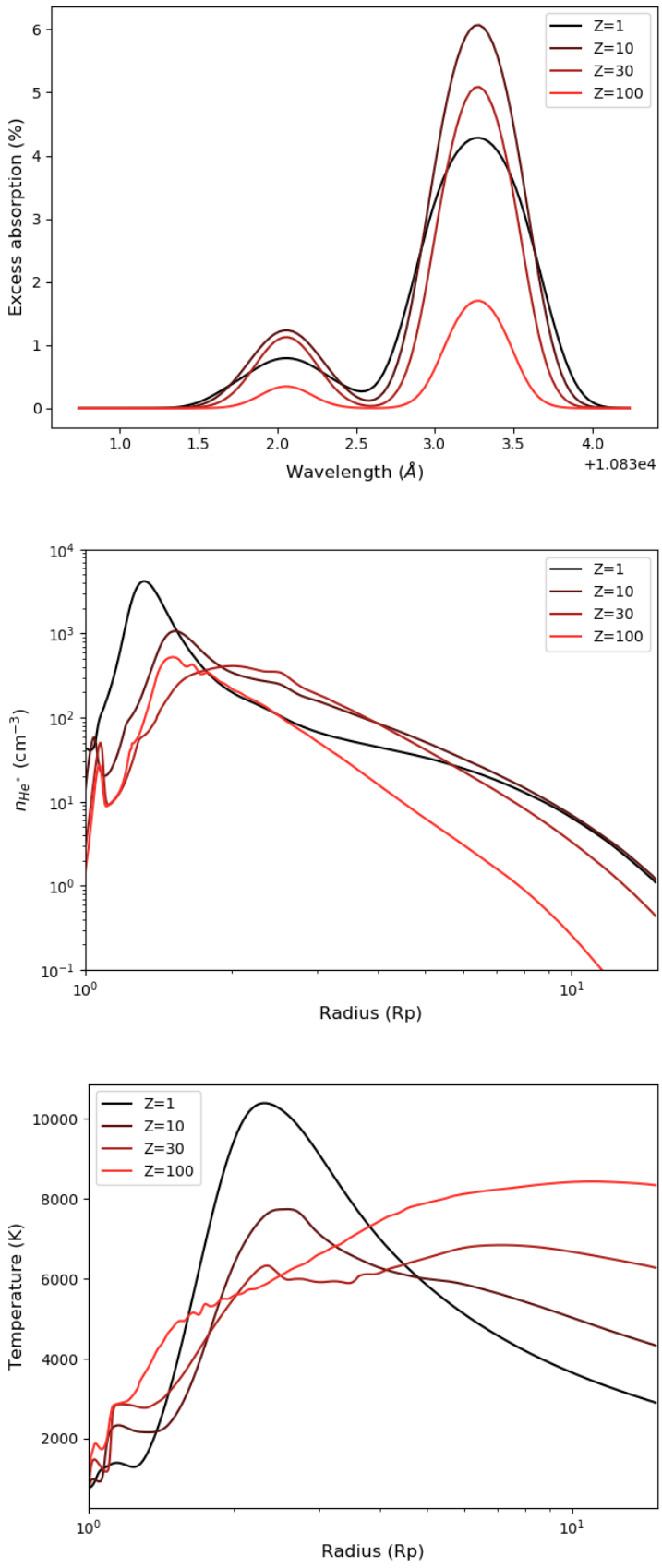


Figure 7.11: Excess absorption spectrum (top), triplet helium number density (middle), and temperatures (bottom) predicted by TPCI as a function of metallicity.

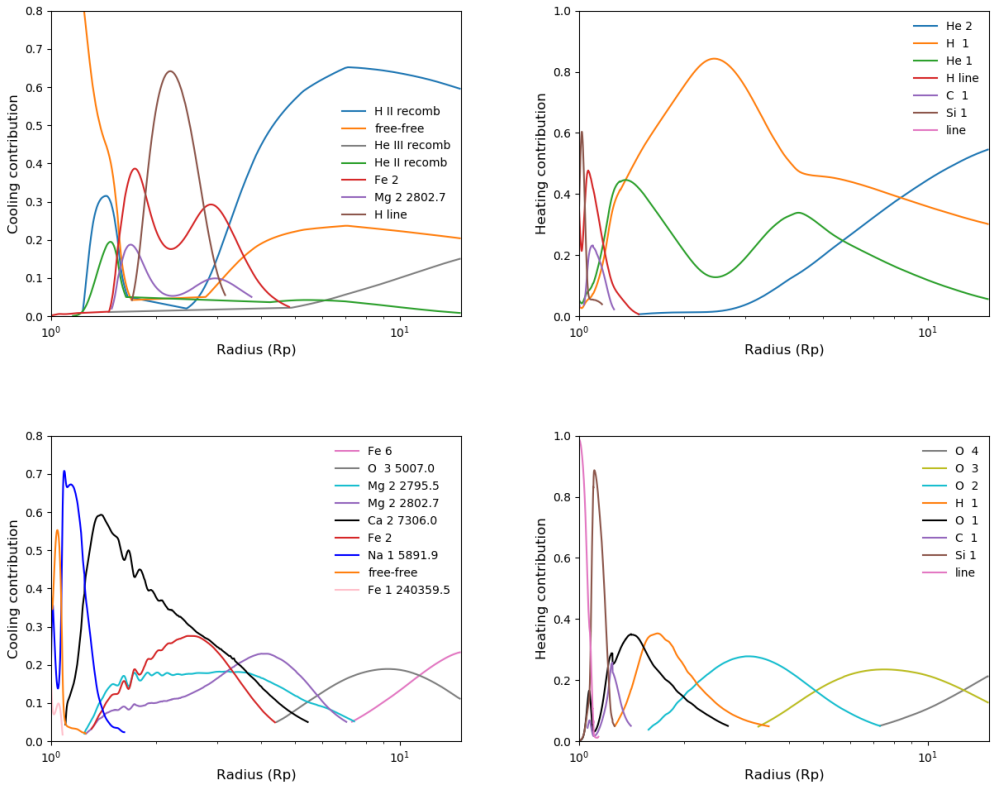


Figure 7.12: Cooling and heating. Top row: radiative cooling (left panel) and radiative heating (right panel) contributions for our 1D solar metallicity TPCI model. Bottom row: cooling (left panel) and heating (right panel) contributions for our 100× solar metallicity TPCI model. The labels are as provided by TPCI. For cooling, a label like O 3 5007.0 means the OIII line at 5007Å. For heating, a label like O 3 means photoionization of OIII, while "line" refers to line heating. While hydrogen and helium are dominant in both heating and cooling at Z=1, a variety of metals in a variety of ionization states dominate heating and cooling at Z=100.

metallicities largely because metals begin dominating hydrogen and helium by mass, but also due to changes in heating and cooling rates.

The temperature profile is governed by the equation of energy conservation (Murray-Clay et al., 2009):

$$\rho v \frac{\partial}{\partial r}\left[\frac{kT}{(\gamma-1)\mu}\right] = \frac{k_B T v}{\mu}\frac{\partial \rho}{\partial r} + \Gamma + \Lambda, \tag{7.1}$$

where the left-hand term represents advection and the three right-hand terms represent adiabatic cooling, radiative heating, and radiative cooling. We plot all four

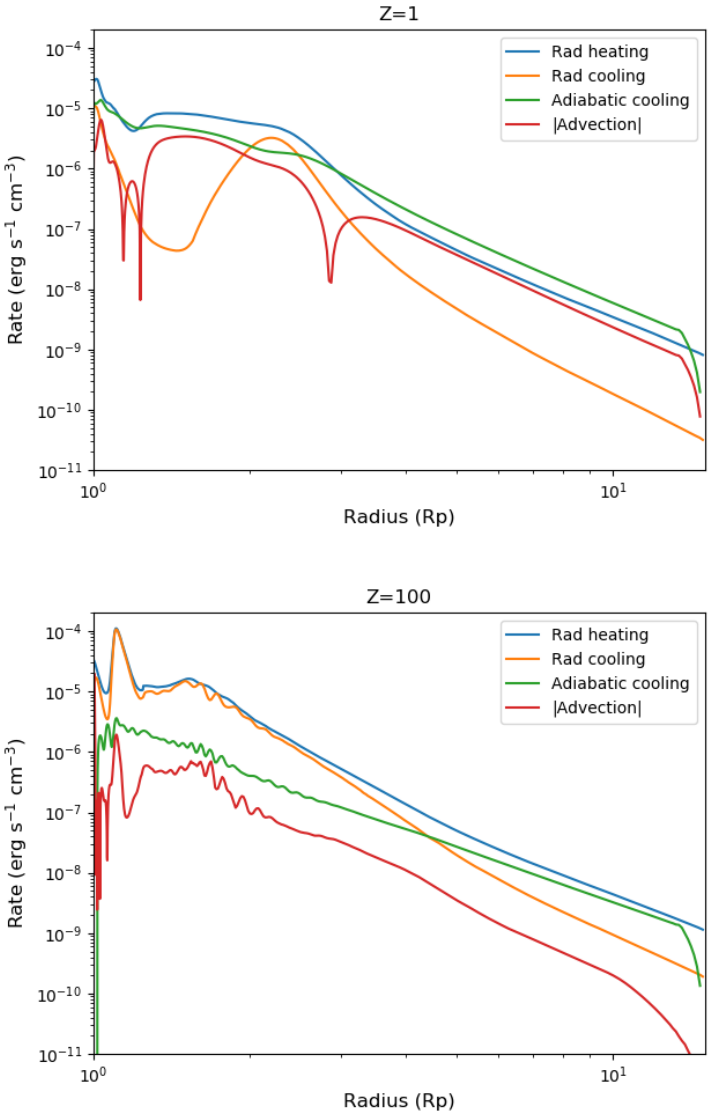


Figure 7.13: Comparison of the four terms of the energy conservation equation (Equation 7.1). The advection term switches sign, so we plot the absolute value. The oscillations in the Z=100 plot are numerical artifacts. The radiative heating and cooling terms are further dissected in Figure 7.12.

terms in Figure 7.13 for the Z=1 and Z=100 simulations. At Z=1, adiabatic cooling is more important than radiative cooling at nearly all radii; at Z=100, radiative cooling is more important than adiabatic below 4 R_p.

We plot the radiative heating and cooling rates in Figure 7.12, which reveals a complex interplay between different processes. In the solar metallicity model, cooling is dominated by free-free emission below 1.6 R_p, hydrogen line emission between 2 and 2.9 R_p, and recombination after 3.5 R_p. Metal cooling still matters in this model—Fe II compromises 30% of cooling from 1.6–2.0 R_p and 2.9–3.4 R_p, for example—but hydrogen is clearly the most important coolant. As the metallicity increases, metals become more and more prominent coolants. The cooling in the 100× solar metallicity model is dominated by a diverse array of metal lines, including Na I, Ca II, Mg II, O III, and Fe VI, among many others. In the solar metallicity model, the heating at radii beyond 1.3 R_p is dominated by a combination of photoionization of HI, HeI, and (beyond 9 R_p) HeII; below 1.3 R_p, it is dominated by a combination of line heating and photoionization of metals like Si I. The same pattern holds for the 10× solar metallicity model. At a metallicity of 30× solar, HI photoionization is only the most significant heating contribution from 1.9 to 4.5 R_p. At a metallicity of 100× solar, it is only the most significant contributor from 1.7 to 2.4 R_p, with a diverse array of metals in various ionization states (like O II and Si I) dominating at larger radii.

These changes in heating and cooling cause changes in the temperature profile shown in Figure 7.11. In the solar metallicity model, the temperature profile peaks at $2.49R_p$ and 10,000 K, declining to 2800 K at $15R_p$. As the metallicity increases, the peak moves to lower temperatures and larger radii, while the temperature in the outer regions increases. By the time the metallicity reaches 100× solar, the temperature peak has become more of an asymptote, as the temperature rises to 8000 K at 6 R_p and then stays nearly constant beyond that point. Consistent with this behavior, the total non-adiabatic cooling is significantly lower than the total heating in the solar metallicity model, but the two become comparable in the outer regions ($> 6R_p$) in the highest metallicity model.

3D Microthena hydrodynamic models

Our 3D models utilize the approach outlined in Wang & Dai (2018), Wang & Dai (2020), and Zhang et al. (2021b), which combines ray-tracing radiative transfer, real-time non-equilibrium thermochemistry, and hydrodynamics based on the higher-

order Godunov method code Athena++ (Stone et al., 2020). As in Zhang et al. (2021b), we perform radiative transfer by dividing photons into seven energy bins for the ray-tracing calculation: 1.4, 7, 12, 16, 47, 300, and 3000 eV, meant to represent IR/optical/NUV, FUV, Lyman-Werner photons (which photodissociate molecular hydrogen), soft EUV, hard EUV, soft X-rays, and hard X-rays, respectively.

We first run the model for the fiducial parameters: a planet mass of 9 $M_\oplus$ (consistent with the RV mass and mass-radius relation) and radius 2.9 $R_\oplus$, the stellar spectrum derived in Section 7.5, a solar metallicity atmosphere, and a stellar wind with a solar velocity (400 km/s) and an 11× solar mass loss rate (see Section 7.5). We then selectively vary every parameter except the planet mass in an attempt to find a best fit model. Microthena takes about one day to run a model on a GPU cluster, so running more than a handful of models is infeasible.

The top panel of Figure 7.14 shows the predicted absorption from the fiducial model with respect to time and wavelength, while the bottom panel compares the predicted average in-transit excess absorption spectrum to the observations. This fiducial model significantly over-predicts the observed signal, with a predicted peak three times higher than the measured value. The model absorption peak is also blueshifted by 15 km/s, while the measured peak is redshifted by ∼4 km/s. Lastly, the fiducial model outflow contains two components, while the observed signal appears to be unimodal. It is particularly difficult to explain the large discrepancy between the predicted and measured peak locations, as velocity blueshifts are a near-universal feature of all 3D outflow models (Wang & Dai, 2018; Wang & Dai, 2020). If this 15 km/s discrepancy between model and data is due to a non-zero orbital eccentricity, it would be in severe tension with the radial velocity data, which indicates the planet is moving toward us at $7.4^{+9.3}_{-7.0}$ km/s at mid-transit—that is, in the wrong direction. Interestingly, while the amplitude of the simulated signal is much too big, the band-integrated light curve has a similar shape to the observed light curve (see Figure 7.2). Both model and data exhibit a slow increase in absorption from ingress to 0.5 h past the white light mid-transit, followed by a fast decrease.

The poor match between the fiducial model and the data prompted us to try different combinations of metallicity, stellar wind speed, and stellar wind density. Figure 7.14 shows the excess absorption spectrum of the different models we tried, compared to each other and to the data, while Table 7.3 reports their mass loss rates, peak absorption, and equivalent width of the absorption. The qualitative behavior of the model when the metallicity is increased is very similar to the behavior of TPCI: the

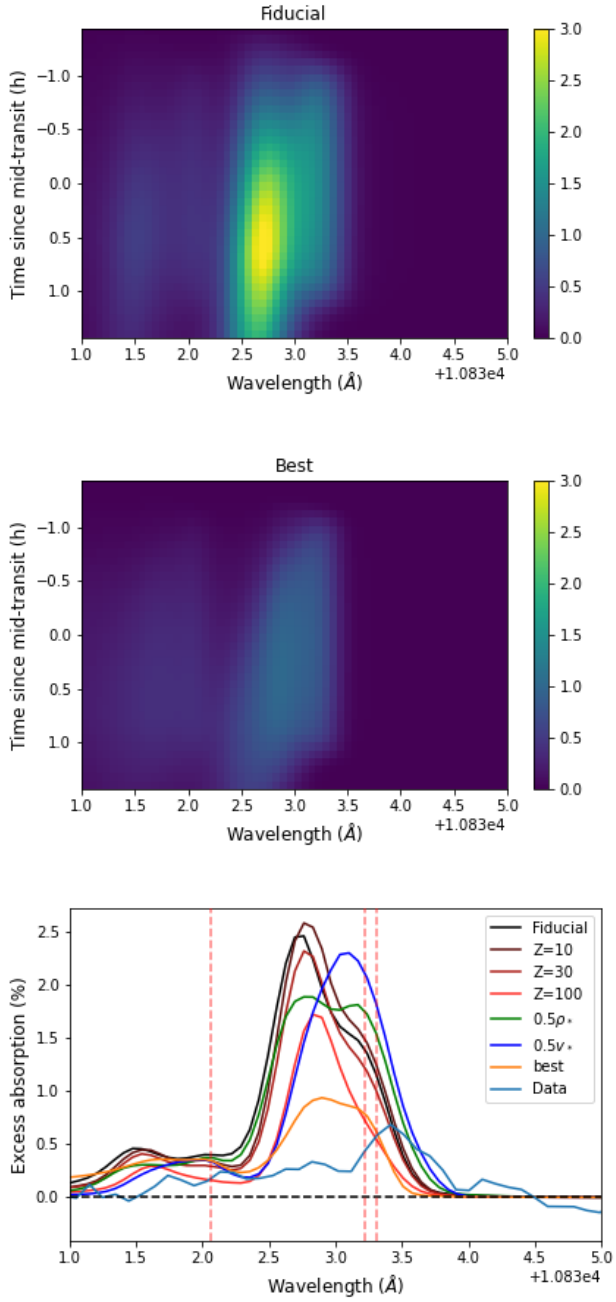


Figure 7.14: Excess absorption as a function of wavelength and time, as predicted by the fiducial model (top) and the best-fit model (middle). The fiducial model uses our best guess for all parameters, while the best-fit model has a higher stellar wind speed of 500 km/s (versus the fiducial 400 km/s), a higher metallicity (10× solar versus the fiducial 1× solar), and 1/3 the fiducial stellar EUV flux. Bottom: average in-transit excess absorption predicted by all Microthena models we tried, compared to data.

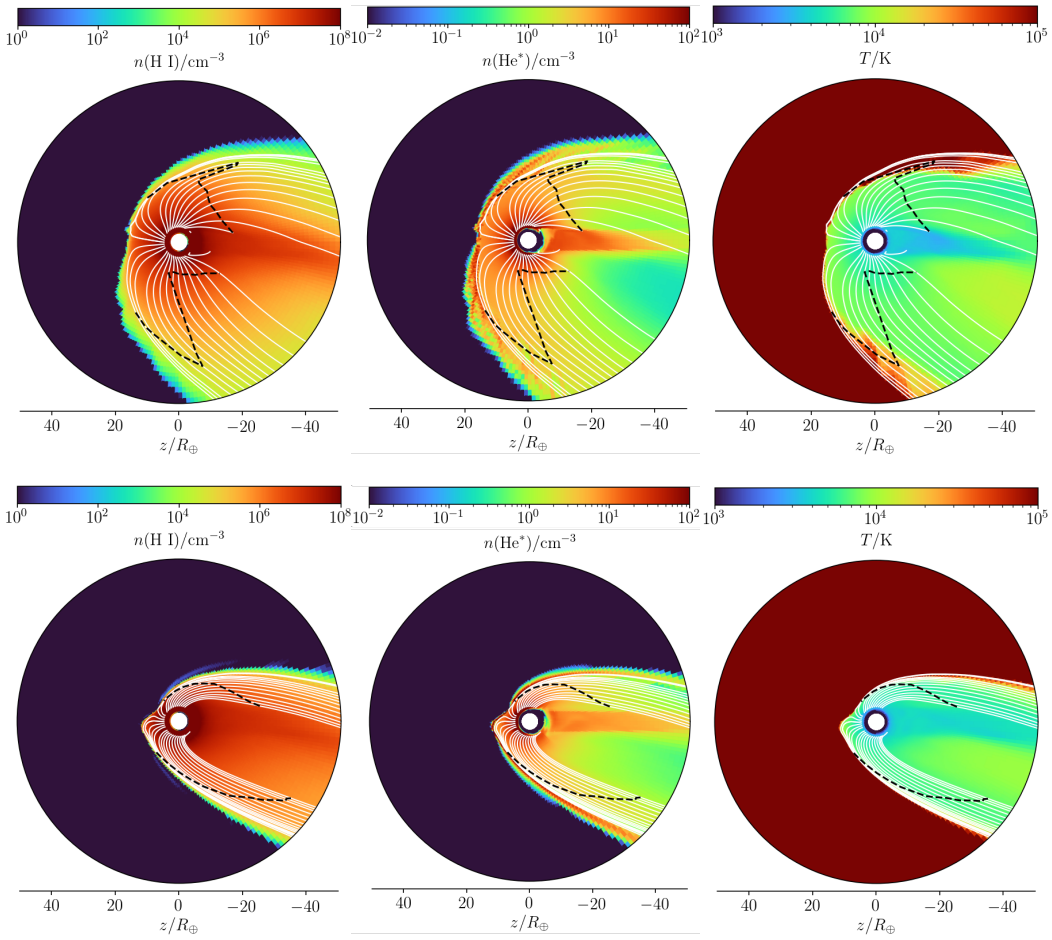


Figure 7.15: Neutral hydrogen number density (left), triplet helium density (middle), and temperature (right) from the time-averaged (over the last $\sim$ 10 kinematic timescales) fiducial (top) and best-fit (bottom) 3D models. The star is toward the left, and orbital motion is upwards. These plots show the profiles in the orbital plane. The white lines are the streamlines, while the dashed black lines represent the inner sonic surface.

strength of helium absorption is similar at Z=1, 10, and 30, but substantially lower at Z=100. In both models, absorption changes non-monotonically with metallicity, reaching its peak at Z=10. The 3D models reveal that metallicity only marginally affects the blueshift. In addition to increasing the metallicity, we tried halving the stellar wind density, which resulted in broader but slightly less blueshifted absorption with a slightly lower (by ∼30%) peak. We also tried halving the stellar wind speed, which marginally decreased the absorption peak and equivalent width, but decreased the blueshift by 10 km/s. For all of these models, the mass loss rate remains remarkably similar: it decreases slightly with metallicity, from 0.12 $M_\oplus$/Gyr at Z=1 to 0.086 $M_\oplus$/Gyr at Z=100, but the fiducial, $0.5\rho_*$, and $0.5v_*$ models have mass loss rates within several percent of each other.

Because all of the models that assume the fiducial stellar spectrum substantially overpredict the observed absorption, we also tried rescaling the input EUV flux. Since the outflow is driven largely by EUV heating, and triplet helium is mostly produced by EUV-driven photoionization followed by recombination of electrons with helium, lowering the EUV flux should lower the corresponding absorption signal. In the end, the model we found that best matches the observations (labelled "Best" in all the plots) has 1/3 the nominal EUV, a stellar wind with a speed of 500 km/s, and a metallicity of 10x. As a result of the low EUV flux, the mass loss rate is 1/3 fiducial (Table 7.3); as a result of the low mass loss rate and the higher stellar wind speed, the outflow is more confined. In the middle and lower panels of Figure 7.14, we show results for this best fitting model. This model is a much better match to the amplitude of the observed absorption signal and the corresponding transit light curve, although the predicted absorption signal is still bluer than the observed, and stronger in both the peak and the equivalent width.

Figure 7.15 shows the neutral hydrogen number density, triplet helium number density, and temperature in the orbital plane for the fiducial and the best-fit models. Both models exhibit a cometary tail, which is caused by the dense and fast stellar wind pushing the outflow away from the star. The tail is angled partly because of the planetary orbital velocity, and partly because of the Coriolis force. In both models, the stellar wind is strong enough to prevent the planetary outflow from becoming supersonic at many locations, making it possible for interplanetary conditions to affect the launching of the wind. As expected, the "best" model has a substantially more confined outflow.

Table 7.3: Summary of Model Predictions

Dim	Model	$\dot{M}$ ($M_\oplus$/Gyr)	Peak (%)	EW (mÅ)
1	Z=1	0.32	4.3	39.6
1	Z=10	0.42	6.1	47.0
1	Z=30	0.37	5.1	35.0
1	Z=100	0.13	1.7	9.6
3	**Fiducial**	0.12	2.5	23.3
3	Z=10	0.11	2.6	23.0
3	Z=30	0.10	2.3	19.7
3	Z=100	0.086	1.7	12.5
3	$0.5\rho_*$	0.12	1.9	22.3
3	$0.5v_*$	0.12	2.3	20.5
3	**Best**	0.041	0.9	11.4
3	Data	?	0.68 ± 0.08	7.0 ± 0.4

Order-of-magnitude empirical estimate of mass loss rate

By making a few assumptions, we can use the equivalent width of the helium signal to obtain a decent estimate for the mass of metastable helium in the outflow and an order-of-magnitude calculation of the mass loss rate. The first assumption we make is that helium absorption from most of the planetary outflow is optically thin. This is supported by the fact that we do not see the secondary peak at 10832 Å, which is 1/8 the height of the primary peak in the optically thin regime, but higher when the primary peak becomes optically thick. The optical thinness assumption tends to cause us to underestimate the mass loss rate.

The optical depth from the star to the observer is:

$$\tau(\lambda) = N\sigma_\lambda P(\lambda), \tag{7.2}$$

where N is the column density, $\sigma_\lambda \equiv (\pi e^2 g_l f_l)/(m_e c^2)$, and $P(\lambda)$ is the line profile ($\int_{-\infty}^{+\infty} P(\lambda)d\lambda = 1$). Integrating $1 - e^{-\tau}$ over λ, and assuming the optically thin limit (where $1 - e^{-\tau} \approx \tau$), we obtain the standard equation for the equivalent width:

$$W_\lambda = N\sigma_\lambda. \tag{7.3}$$

We can perform another integration over the cross-sectional area of the star, which turns the column density N into the total number of metastable helium atoms, N'.

We then divide each side by the cross-sectional area, so that the left hand side becomes the average equivalent width across the star—the number measured by our observations:

$$\iint W_\lambda dS = \iint N\sigma_\lambda dS \tag{7.4}$$

$$\frac{\iint W_\lambda dS}{\iint dS} = N'\sigma_\lambda \tag{7.5}$$

$$W_{avg} = \frac{N'\sigma_\lambda}{\pi R_*^2} \tag{7.6}$$

$$N' = \frac{R_*^2 m_e c^2}{e^2 g_l f_l \lambda_0^2} W_{avg} \tag{7.7}$$

$$= 8 \times 10^{31} \tag{7.8}$$

$$m_{He^*} = 5 \times 10^8 g, \tag{7.9}$$

where we combined $g_l f_l$ for the two lines at 10833 Å. 5×10^8 g is 500 tons—a remarkably small amount of material to be detectable at interstellar distances. This is the most secure part of our estimate.

Next, we convert the mass of metastable helium to total mass, by assuming that metastable helium comprises 10^{-6} of helium nuclei, and that helium is 25% of total mass. 10^{-6} is roughly the ratio we find in our 3D simulations from 2–5 R_p, as well as the ratio we find from 2–10 R_p for HD 63433 b/c (Zhang et al., 2021b). We obtain $m_{tot} = 2 \times 10^{15}$ g.

Finally, we estimate the replenishment lifetime of the metastable helium atoms: how long do they take to cross the stellar disk and stop being observable? We calculate $\tau = R_*/c_s = 50,000$ s, adopting 10 km/s for the sound speed c_s. The mass loss rate is then $m_{tot}/\tau = 4 \times 10^{10} g/s = 0.22 M_\oplus/Gyr$.

This estimate should not be considered accurate to more than an order of magnitude, but it is reassuring that our empirical estimate is of the same order of magnitude as that of the simulations (Table 7.3).

7.7 Discussion

The fiducial versions of both our TPCI (1D) and Microthena (3D) models predict helium absorption many times stronger than observed. For the same input parameters, the helium absorption predicted by TPCI is much higher than that predicted by Microthena (4.3% vs. 2.5%). The 3D structure, which TPCI cannot model, determines

the shape of the helium absorption signal in both wavelength and time. Though the predicted helium signal is too large, the predicted mass loss rates in Table 5.2 are all reasonable. We assume a 2% hydrogen/helium envelope, which is consistent with the planet's radius, mass, and equilibrium temperature assuming a rocky core (Zeng & Seager, 2008). The mass loss timescales predicted by Microthena would then range from 1.6–2.4 Gyr, and those predicted by TPCI would range from 0.4–1.6 Gyr. All of these timescales are comparable to the planet age, which is high enough that the planet has likely lost a substantial portion of its primordial envelope, but low enough that it is not surprising the planet still has an envelope.

We next consider whether there might be additional modifications to the 3D Microthena models that could better match the amplitude of the observed signal while also reducing the size of the predicted blueshift. We could increase the stellar wind speed by a larger amount, which would eventually succeed in suppressing the mass loss rate (Murray-Clay et al., 2009), but this would also increase the magnitude of the predicted blueshift. If we instead decreased the EUV flux by a larger factor than 3x, it would also decrease the helium signal strength, but our X-ray observations strongly disfavor this scenario. The scaling relations from Linsky et al. (2014) that we used to estimate the EUV spectrum have a dispersion about the fit line of 20–37% for F5–K5 stars, which seems to make a 3x error implausible, but the relations are based on the Lyα flux, which we have not measured for TOI 560 and must infer from the X-ray flux and the stellar rotation rate. In addition, the dispersion might be smaller for young stars only than for all stars. The X-ray and MUV (1230–2588 Å) fluxes are even better known because they were directly measured by XMM-Newton within weeks of both Keck observations. A substantially larger planetary mass (e.g. 16 $M_\oplus$) could also suppress the signal, but would be in tension with both the RV measurements and the mass-radius relation of Chen & Kipping (2017). This mass-radius relation is derived primarily from observations of mature planets; since young planets are expected to be more inflated, it provides a conservative upper limit on the likely mass of this planet. Finally, the redshift would be more consistent with the models if we assumed a much weaker wind, in which case there would be a significant up-orbit stream launched from the dayside, as seen in the 3D hydrodynamic simulations by McCann et al. (2019). This is scenario favored by Czesla et al. (2021) for the redshift they observe from HAT-P-32b; however, their helium signal also shows pre-ingress absorption, which ours does not.

Our 1D TPCI models demonstrate that increasing the metallicity beyond 10× solar gives rise to a rich tapestry of radiative phenomena, which in turn results in significant changes to the structure of the outflow and a reduction in the predicted helium absorption strength. Our 3D models show a qualitatively similar pattern. It is clear, however, that further work is required in order to better address some of the limitations of our models for describing metal-rich atmospheres. For example, both codes are potentially inaccurate at high metallicities because they use the escape probability formalism for radiative transfer, instead of performing radiative transfer completely correctly. The 3D Microthena model only accounts for line cooling from O, S, Si, O+, S+, and Si+, whereas our TPCI models predict that there is a much broader array of elements and ionization states that are important for both heating and cooling at high metallicities (Figure 7.12). In addition, while the 3D model divides all radiation into 7 bins, TPCI uses the incident spectrum at the resolution provided (Figure 6.11); this likely has a significant impact on the predicted thermal structure of the outflow.

In addition to ensuring correct behavior of the model at high metallicities, future studies of metal-rich outflows should also vary the assumed elemental ratios (e.g., C/O) to reflect the diverse array of possible atmospheric compositions for sub-Neptune-sized exoplanets. Condensation of species like NaCl, $MgSiO_4$, and TiO in the deep atmosphere could also further alter the composition of the outflow in ways that may be detectable in future observations.

The parameter space exploration we undertook assumes that the observed outflow is in a steady state. However, there is some tentative evidence for variability in our data: on the second night, the absorption during the portion of transit we could observe was slightly stronger and slightly bluer than on the first night. This effect is only marginally significant (2-3σ), and could be explained by a combination of statistical fluctuations, underestimated errors, stellar variability, and unknown instrumental systematics. However, our 3D simulations show that stellar wind conditions have a significant impact on the absorption signal, so it is by no means implausible that variations in the stellar wind could lead to variations in the outflow properties. During the 2008 solar minimum, for example, the solar wind number density regularly fluctuated by a factor of two over a timescale of days, with dozens of upward spikes corresponding to increases of more than a factor of five (Lei et al., 2011). The solar wind velocity during this period regularly fluctuated between 400 and 600 km/s, again on a timescale of days. TOI 560 is younger, less massive, and

more active than the Sun, and we do not know where it was in its activity cycle during our observations. It is therefore possible that TOI 560's stellar wind may be more variable than that of the Sun, and that the stellar wind conditions during the two transits were appreciably different. Our 3D simulations show that decreasing the stellar wind density by a factor of two decreases the absorption peak by ∼30%, while decreasing the wind speed by a factor of two decreases the blueshift by ∼ 10 km/s. This level of variation is similar to the magnitude of the observed variability between the two nights.

Another effect that neither TPCI nor Microthena accounts for is magnetic fields. Using the same methodology as Zhang et al. (2021a), which is itself based on Owen & Adams (2014), we can estimate the significance of magnetic fields by comparing the magnetic pressure $P_B = B^2/(8\pi)$ to the ram pressure $P_{ram} = \dot{M}v/(8\pi r^2)$. At 3.6 R_p, roughly the position of the helium line photosphere, our Microthena models predict a typical wind speed of 7 km/s. As discussed in Zhang et al. (2021a), the slow rotation of a tidally locked planet (the tidal circulation timescale is 1200 yr, assuming Q=100) gives rise to a weak magnetic field, which does not significantly affect the outflow ($P_{ram}/P_B = 64$ at 3.6 R_p). However, the interplanetary magnetic field can plausibly have a significant effect on the outflow. Based on the scaling relation between surface magnetic field and stellar age in Vidotto et al. (2014), we calculate an interplanetary magnetic field of 0.025 G at the position of TOI 560.01, which leads to $P_{ram}/P_B = 1.1$ at 3.6 R_p. Finally, we emphasize that these are order-of-magnitude calculations with uncertainties of at least a factor of a few in all quantities: interplanetary magnetic fields, stellar activity cycles, and planetary magnetic fields are all very poorly understood for extrasolar systems.

7.8 Conclusion

We observed two transits of TOI 560.01 using Keck/NIRSPEC and detected a strong helium absorption signal during each transit. The observed signal shows a number of intriguing features, including a slow rise in absorption accompanied by a fast decline, a slight redshift, and tentative evidence of variability between the two transits, with the signal being stronger and bluer on the second night at 2-3σ significance. TOI 560.01 is the first mini Neptune with a helium detection. Its size and youth place it in the critical regime where mini Neptunes transition into super Earths, giving rise to the radius gap and dramatically shaping exoplanet demographics.

We applied the 1D code TPCI and the 3D code Microthena to model the outflow. We found it difficult to match the low magnitude of the observed absorption level and the moderate redshift with a solar metallicity planetary atmosphere. Increasing the metallicity to Z=100 suppresses the helium absorption signal in both TPCI and Microthena, but in both models, intermediate metallicities do not. However, neither code is very well suited to modeling high metallicity atmospheres. This problem is more acute for Microthena, which uses a simplified radiative transfer model and includes a limited number of atomic coolants. Further work will be necessary to confirm whether a high metallicity atmosphere is a good explanation for the data, and if so, whether it is the only plausible explanation.

As a low mass mini Neptune around a nearby K star, TOI 560.01 is a favorable target for atmospheric characterization. In this paper, we measure the planet's helium absorption, but its Lyα absorption may also be observable with *HST*/STIS. If so, it will provide complementary insight into the escaping atmosphere: due to the strength of the Lyα line and the strong interstellar absorption, which wipes out the core, Lyα traces energetic hydrogen atoms in the tenuous outer regions of the outflow. *JWST* could also reveal important insights into the outflow by measuring the composition and thermal structure of the deeper atmosphere, which, aside from being scientifically valuable in its own right, will drastically shrink the parameter space of inputs that models can assume.

Finally, TOI 560 is a two-planet system, and TOI 560.02 is also a transiting mini Neptune. This makes the system an excellent test for mass loss models. The two planets share the same contemporary X-ray/EUV environment, and the same irradiation history. In addition, planets of similar size located in adjacent orbits might be expected to have largely similar formation and/or migration histories, and therefore it is reasonable to expect that their primordial atmospheric compositions would be quite similar. This is supported by observational studies of the masses and radii of multi-planet systems in the Kepler sample, which suggest that planets in the same system tend to have similar masses and radii (the 'peas in a pod' theory; Weiss et al. 2018).

Software: numpy (van der Walt et al., 2011), scipy (Virtanen et al., 2020), matplotlib (Hunter, 2007), dynesty (Speagle, 2020), SAS

7.9 Acknowledgements

The helium data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

Based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and NASA.

References

Allart, R., Bourrier, V., Lovis, C., et al. 2019, A&A, 623, A58

Barragán, O., Armstrong, D. J., Gandolfi, D., et al. 2021, arXiv e-prints, arXiv:2110.13069

Bourrier, V., Lecavelier des Etangs, A., Dupuy, H., et al. 2013, A&A, 551, A63

Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., et al. 2018, A&A, 620, A147

Chen, J., & Kipping, D. 2017, ApJ, 834, 17

Czesla, S., Lampón, M., Sanz-Forcada, J., et al. 2021, arXiv e-prints, arXiv:2110.13582

de Jager, C., Namba, O., & Neven, L. 1966, Bull. Astron. Inst. Netherlands, 18, 128

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109

Gaidos, E., Hirano, T., Wilson, D. J., et al. 2020, MNRAS, 498, L119

Gaidos, E., Hirano, T., Mann, A. W., et al. 2020, MNRAS, 495, 650

Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759

Gupta, A., & Schlichting, H. E. 2019, MNRAS, 487, 24

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6

Kasper, D., Bean, J. L., Oklopčić, A., et al. 2020, AJ, 160, 258

Lavie, B., Ehrenreich, D., Bourrier, V., et al. 2017, A&A, 605, L7

Lee, E. J., & Connors, N. J. 2021, ApJ, 908, 32

Lei, J., Thayer, J. P., Wang, W., & McPherron, R. L. 2011, Sol. Phys., 274, 427

Linsky, J. L., Fontenla, J., & France, K. 2014, ApJ, 780, 61

Linsky, J. L., France, K., & Ayres, T. 2013, ApJ, 766, 69

Mamajek, E. E., & Hillenbrand, L. A. 2008, ApJ, 687, 1264

Mazeh, T., Tamuz, O., & Zucker, S. 2007, Astronomical Society of the Pacific Conference Series, Vol. 366, The Sys-Rem Detrending Algorithm: Implementation and Testing, ed. C. Afonso, D. Weldrake, & T. Henning, 119

McCann, J., Murray-Clay, R. A., Kratter, K., & Krumholz, M. R. 2019, ApJ, 873, 89

Mignone, A., Bodo, G., Massaglia, S., et al. 2007, ApJSupplement Series, 170, 228

Mills, S. M., Howard, A. W., Petigura, E. A., et al. 2019, AJ, 157, 198

Mills, S. M., & Mazeh, T. 2017, ApJ, 839, L8

Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, ApJ, 896, L22

Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, ApJ, 693, 23

Niermann, B., Böke, M., Sadeghi, N., & Winter, J. 2010, European Physical Journal D, 60, 489

Oklopčić, A. 2019, ApJ, 881, 133

Oklopčić, A., & Hirata, C. M. 2018, ApJ, 855, L11

Owen, J. E. 2019, Annual Review of Earth and Planetary Sciences, 47, 67

Owen, J. E., & Adams, F. C. 2014, MNRAS, 444, 3761

Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29

Palle, E., Nortmann, L., Casasayas-Barris, N., et al. 2020, A&A, 638, A61

Pequignot, D., Petitjean, P., & Boisson, C. 1991, A&A, 251, 680

Salz, M., Banerjee, R., Mignone, A., et al. 2015a, A&A, 576, A21

Salz, M., Czesla, S., Schneider, P. C., & Schmitt, J. H. M. M. 2016, A&A, 586, A75

Salz, M., Schneider, P. C., Czesla, S., & Schmitt, J. H. M. M. 2015b, A&A, 576, A42

Salz, M., Czesla, S., Schneider, P. C., et al. 2018, A&A, 620, A97

Schlaufman, K. C. 2010, ApJ, 719, 602

Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77

Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, Nature, 557, 68

Speagle, J. S. 2020, MNRAS, 493, 3132

Stone, J. M., & Proga, D. 2009, ApJ, 694, 205

Stone, J. M., Tomida, K., White, C. J., & Felker, K. G. 2020, ApJS, 249, 4

Turon, C., Creze, M., Egret, D., et al. 1993, Bulletin d'Information du Centre de Donnees Stellaires, 43, 5

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22

Van Eck, S., Neyskens, P., Jorissen, A., et al. 2017, A&A, 601, A10

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J. M., et al. 2008, ApJ, 676, L57

Vidotto, A. A., Gregory, S. G., Jardine, M., et al. 2014, MNRAS, 441, 2361

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Wang, L., & Dai, F. 2018, ApJ, 860, 175

Wang, L., & Dai, F. 2020, arXiv e-prints, arXiv:2101.00045

—. 2021, ApJ, 914, 99

Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, AJ, 155, 48

Wood, B. E., & Linsky, J. L. 2010, ApJ, 717, 1279

Wood, B. E., Müller, H. R., Zank, G. P., Linsky, J. L., & Redfield, S. 2005, ApJ, 628, L143

Wood, B. E., Müller, H.-R., Redfield, S., et al. 2021, ApJ, 915, 37

Zeng, L., & Seager, S. 2008, PASP, 120, 983

Zhang, M., Knutson, H. A., Wang, L., et al. 2021a, AJ, 161, 181

—. 2021b, arXiv e-prints, arXiv:2106.05273

Chapter 8

CONCLUSION

Predictions are hard, especially about the future. The history of exoplanet atmospheric science has been full of surprises, and undoubtedly more are in store. Nevertheless, there are some obvious next steps. In this chapter, I will review potential next steps for the field, with a focus on the research programs that I am personally interested in, have some involvement with, or relate to my thesis.

8.1 James Webb Space Telescope

First, there is the elephant in the room: the *James Webb Space Telescope*, which recently finished deployment. *JWST* will revolutionize the study of exoplanet atmospheres. By far the largest space telescope ever launched, *JWST* has a collecting area 5.6 times that of Hubble, and can observe at much longer infrared wavelengths over a much greater wavelength range. As an example, *HST*/WFC3 can observe between 1.1–1.7 μm with the G141 grism (a commonly used configuration); *JWST*/MIRI LRS, to take just one of many spectrographs, can observe from 5–12 μm simultaneously, and other instruments collectively span 0.6–30 μm. Since exoplanets emit largely in the infrared and molecular absorption features are generally found in the infrared, *JWST* will allow the spectroscopic study of exoplanet atmospheres in unprecedented detail. By my rough count, there are 42 *JWST* General Observer programs in cycle 1 that focus on exoplanet atmospheres, in addition to many more in the Guaranteed Time Observer program and the Early Release Science program. These programs study a wide variety of planets: hot sub-Earths, "super-puffs" with the density of cotton candy, 17 Myr hot Jupiters, highly eccentric planets experiencing intense episodic heating from their stars, white dwarf planets, potentially habitable planets, and much more. They use all techniques (transmission spectroscopy, emission spectroscopy, and phase curves) and a wide variety of instrument/setting combinations.

For the hot Jupiters, such as the poster child HD 189733b, *JWST* will capture exquisite spectra in both transmission and emission. *JWST* GO 1633, for example, will observe five transits of this planet, spanning 2.4–20 μm, covering absorption bands of every major molecule (water, methane, carbon monoxide, and carbon dioxide) as well as strong bands of important minor molecules (HCN, NH_3, H_2S). These

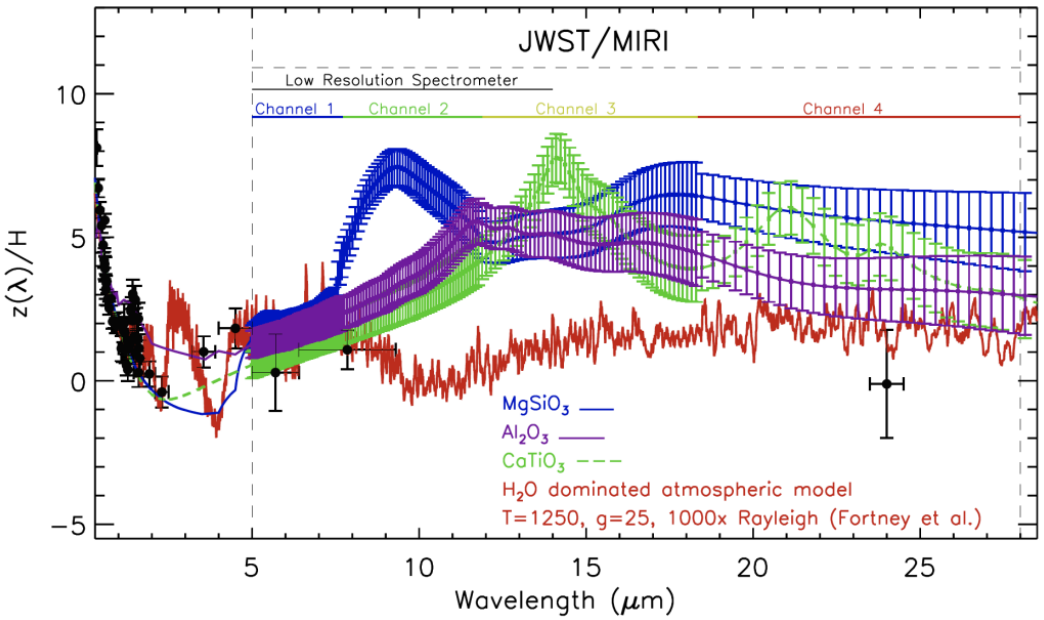


Figure 8.1: The power of *JWST*. Black: observed transit depths; red: a PLATON model of a water-dominated atmosphere; blue, purple, green: simulated *JWST* observations, with error bars, assuming cloudy atmospheres with three different cloud species. Note that all the existing observations redward of 2 μm are no longer obtainable because they are from *Spitzer*, which has been deactivated. Credit: Figure 2 of *JWST* GO 1633 proposal, reproduced with permission.

measurements will give unprecedented insight into the atmospheric composition, potentially enabling the metallicity and C/O ratio to be measured to the percent level. As Figure 8.1 shows, *JWST* has such a broad wavelength coverage that it will be able to detect the broad absorption features from common cloud species. Clouds are ubiquitous in exoplanet atmospheres, and ascertaining their composition will, in itself, be a major advance for the field.

For the sub-Neptunes, not counting the mass loss detections in this thesis, we currently have only a single definitive observation of an atmospheric feature around a single planet—the detection of water in K2-18b. So little is known about their atmospheres that virtually any result would be exciting. Many cycle 1 *JWST* programs are dedicated to observing the atmospheres of these planets. This includes GO 2347, which will obtain a MIRI/LRS phase curve of the lava planet K2-141b; GO 1952, which will use emission spectra to measure the atmospheric composition of the poster child super Earth 55 Cnc e; GO 2372, which will go after the transmission spectrum of the aforementioned K2-18b; and GO 1935, which will take a 0.83–5 μm transmission spectrum of the temperate mini Neptune TOI 421b to robustly

constrain composition and test theories of aerosol formation. These programs will give us the first firm data-driven understanding of the composition, structure, and circulation patterns of sub-Neptune atmospheres.

The ground-breaking nature of *JWST* will hopefully be even more evident for sub-Earths, which, from the planet occurrence rates, appear to be a distinct population from the super Earths (Qian & Wu, 2021). There are currently no observations of their atmospheres. In cycle 1, I will use *JWST* to obtain a spectroscopic (5–12 um) phase curve of GJ 367b, a dense 0.72 $R_\oplus$, 0.53 $M_\oplus$ sub-Earth. Because TESS has already surveyed the whole sky, GJ 367b is likely to forever remain by far the most observationally favorable sub-Earth outside the solar system, and the only one accessible to study by *JWST*. The phase curve will constrain the planet's dayside and nightside temperatures, phase offset, Bond albedo, and heat recirculation efficiency. A significant atmosphere would manifest as a diminished day-night temperature contrast, non-zero phase offset, and if there are clouds, high albedo. If there is no atmosphere, observations may be capable of distinguishing between different surface compositions. Our observations will shed light on the properties and origins of a new population of exoplanets, and serve as a pathfinder to future studies of the planet.

The search for life

The holy grail of exoplanet atmospheric science is to find life elsewhere in the universe by detecting biosignatures in an exoplanet. Unfortunately, scientific progress in the field has only made this goal seem further out of reach, not closer. Every year brings more featureless or ambiguous transmission spectra of small planets, more discoveries of non-biological synthesis pathways of candidate biosignatures, and more appreciation of the complexities of interpreting exoplanet atmosphere observations. For example, Sousa-Silva et al. (2020) proposed phosphine (PH_3) as a biosignature gas because its only known sources under terrestrial conditions are biological and industrial. Less than a year later, Greaves et al. (2021b) reported discovering phosphine on Venus with two different telescopes. There followed a debate (e.g. Villanueva et al. 2020; Greaves et al. 2021a) about whether the data was reduced correctly, what order of polynomial is appropriate to fit the systematics, whether the observed line is from phosphine or sulfur dioxide, and whether any natural processes could account for the phosphine. The debate remains unresolved, although the original authors have substantially weakened their claim (Greaves et al., 2020). Notably, even though Greaves et al. (2021b) were confident in their phosphine

detection, and even though phosphine was considered one of the most promising biosignature gases, they did not claim that their detection was evidence of life on Venus.

A similar story played out on Mars. Given the difficulty of finding a single magic molecule that absolutely cannot be produced non-biologically, many have suggested (e.g. Krissansen-Totton et al. 2016) that a severely out-of-equilibrium atmospheric composition would be evidence of life, even if there are many geological sources of all individual molecules. For example, in the heavily oxidizing environment of Mars, methane might be a biosignature. For decades, multiple groups using multiple instruments have detected methane on Mars (e.g. Krasnopolsky et al. 2004). Webster et al. (2015) detected methane in situ with the Curiosity rover, and showed that it was seasonal. The presence of methane has been suggested as evidence for life (Krasnopolsky et al., 2004). To this day, the mystery has not been resolved, even though a wide variety of methane production mechanisms have been proposed.

Venus and Mars are the closest planets to Earth, and Mars has been monitored in-situ by multiple spacecraft for decades. If discovering candidate biosignatures on these planets does not prove the existence of life, discovering them on exoplanets—even to very high confidence—would mean much less. Absent a truly spectacular discovery, such as the detection of an unambiguous radio transmission, it is unlikely that even detecting an atmosphere identical to Earth's would constitute proof of alien life in the forseeable future.

This pessimism does not mean, however, that no progress is being made toward detecting life on other planets. In fact, *JWST* will be able to probe the atmospheres of several terrestrial exoplanets in the habitable zone. The most promising targets are in the truly astounding TRAPPIST-1 system: a system of 7 transiting planets, all of them with Earth-like radii and masses, and 3–4 of which are in the habitable zone, all orbiting an M dwarf that is one of the smallest stars in existence. It is highly unlikely that a similarly favorable system will ever be found. There are eight cycle 1 *JWST* programs targeting the TRAPPIST-1 system, probing the planets' atmospheres in both transmission and emission. Depending on the precise nature of the atmosphere, *JWST* may be able to detect biosignatures in the atmospheres of these planets. For example, a CO_2/CH_4 atmosphere (which would be in chemical disequilibrium) on TRAPPIST-1e could be robustly detectable in 14 transits (Meadows et al., 2021). Such a find would be thrilling, but would not constitute strong evidence of life,

as this is precisely the combination of gases that has been detected on Mars. On the other hand, it is possible that dozens of transits and eclipses will yield nothing but a flat transit spectrum and a featureless eclipse spectrum, indicating that the planets may have already lost their atmospheres. Such a result would be severely disappointing, as planets around M dwarfs are by far the easiest to study, and such a result would suggest that M dwarfs' high activity tends to strip the atmospheres of terrestrial planets.

The need for better models

Along with much better data from *JWST* comes the need for better models. It would be shocking if any existing model could perfectly explain *JWST* observations, considering that they cannot perfectly explain existing observations. *JWST* will give modellers new information about what their models are missing, motivating them to add new physics, improve their parameterizations of existing physics, or even pursue entirely new research directions. For example, existing observations reveal that clouds are ubiquitous, but are silent on the composition of the clouds. *JWST*'s broad infrared wavelength coverage should enable it to distinguish between candidate cloud species (Figure 8.1), providing guidance on a crucial part of any atmosphere model. As another example, the assumption of chemical equilibrium allows for simple parameterizations of atmospheric composition (e.g. metallicity and C/O ratio), but it is unclear under what conditions chemical equilibrium holds, as the vertical mixing strength (K_{zz}) is uncertain to many orders of magnitude. If *JWST* detects molecular features that should not exist based on chemical equilibrium assumptions, we would be able to refine our photochemistry and/or vertical circulation models to account for the observations.

Existing Bayesian retrieval models, including PLATON, adopt many simplifications that might be of dubious validity with *JWST*-quality data. Aside from the assumption of chemical equilibrium (or in some other codes, vertically constant abundances) and the simple treatment of clouds, these codes usually assume a homogenous terminator (Powell et al., 2019), no nightside emission (Kipping & Tinetti, 2010), no refraction (Misra et al., 2014), parameterized temperature-pressure profiles, and Voigt molecular line profiles with broadening parameters that are a mixture of guesswork and extremely spotty empirical data. Improving these models will require physically motivated parameterizations of the underlying physics and better understanding of said physics, which in turn will require a combination of self-consistent simulations, cloud formation models, better molecular line list computations, and laboratory ex-

periments of molecular absorption and of cloud particle characteristics. It will of course also require additions to the retrieval codes themselves, some of which are simple (e.g. a two-terminator approximation), and some of which will be challenging (e.g. a physically realistic cloud parameterization). A 3D retrieval may end up being necessary (e.g. Nixon & Madhusudhan 2022).

The wealth of data from *JWST* will advance our understanding of exoplanet atmospheres by an extraordinary amount, but deeper understanding will come hand in hand with new mysteries. The atmospheres of brown dwarfs and solar system planets are still active areas of research, even though brown dwarfs are routinely observed in much greater detail than exoplanets, and solar system planets are observed with infinitely higher SNR than brown dwarfs. In particular, brown dwarfs exhibit a dramatic L-T transition around 1400 K (Vos et al., 2019), where cooler brown dwarfs become dramatically bluer. At the transition point itself, brown dwarfs exhibit high variability, which is commonly attributed to patchy clouds. The physical mechanism behind the transition is still unclear. Will clear transition points also be found for exoplanets, and will variability also be an important tool for studying exoplanet atmospheres? No variability has been definitively detected for any exoplanet, although 55 Cnc e exhibits signs of variability in its *Spitzer* eclipse depths (Demory et al., 2016). Variability may turn out to be a whole new subfield domain of exoplanet atmospheres that *JWST* will enable.

8.2 Mass loss

In the last two papers of this thesis, I reported an escaping hydrogen or helium atmosphere from two out of the three young mini Neptunes I observed. The logical next step is to expand the sample to multiple planets spanning a range of radii, ages, and densities. I am in the process of doing this, and have so far detected escaping atmospheres from 4 out of the 5 young mini Neptunes I have observed. It appears that, consistent with the canonical picture, primordial atmospheres are common around these planets, and they evaporate quickly enough to lose a significant fraction of their envelopes within their lifetimes.

However, much work remains to be done, because the 3D hydrodynamic models we ran to predict the outflows do not match the data. Worse, my work and those of other similar studies have shown that it might be unrealistic for any model to match the data in detail. The enormous uncertainties surrounding the stellar wind velocity and density (which can vary by a factor of a few on timescales of

days), the interplanetary and planetary magnetic fields, the planetary mass, and the atmospheric composition may make it unrealistic for any forward model to match the data to better than a factor of a few. Some of these uncertainties can be reduced in the near future. For example, favorable targets for mass loss observations are usually also favorable targets for *JWST* atmospheric characterization, which could measure the atmospheric composition. However, the stellar wind conditions and magnetic fields are not observable, making it very hard to see how these uncertainties can be addressed in the near future. If this is the case, the importance of a large sample size is all the more important, as it would average out uncertainties and allow statistical comparisons between model and data. In addition, a large sample size could reveal patterns that shed light on the underlying physics. By my rough estimate, there are optimistically ~20 mini Neptunes (mostly TESS discoveries) where a helium detection could be possible with 10 m class telescopes in one transit; there are certainly not hundreds.

Yet another source of uncertainty, and of potential future work, is variability. Since the absorption signal depends on the high-energy flux of the star and the stellar wind conditions, it should change when they change—which should be all the time, especially for young and active stars. Spurious variability is also expected when the stellar surface is inhomogenous, with different regions having dramatically different depths for the stellar He I 1083 nm line. Inhomogenous stellar surfaces are not exactly unexpected for young stars. Guilluy et al. (2020) observed 5 transits of HD 189733b in the He I line and found variability at the 50% level; using simultaneous Hα data, they were able to determine that two transits were affected by stellar inhomogeneities, and that the remaining 3 had consistent He I depths. Similar studies ought to be performed for young mini Neptunes with detected helium signals to see whether they exhibit any variability, and if so, whether the variability is likely planetary or stellar.

On the model side, there is also much work to be done. In particular, magnetic fields need to be studied in much greater detail. For example, most simulations show that magnetic fields should decrease the outflow rate and the signal strength (Owen & Adams, 2014; Khodachenko et al., 2015; Arakcheev et al., 2017), but Carolan et al. (2021) predicts an increase in the outflow rate. Also, there has been little work on 3D modelling photoevaporation from high metallicity atmospheres, or H_2O/CO_2 atmospheres. These atmospheres are plausible around mini Neptunes, and it is worthwhile to study how their observational signatures differ from hydrogen/helium

atmospheres. Finally, despite 20 years of Lyα mass loss observations, there is still no solid answer to this simple question: what does a Lyα observation tell us, other than that hydrogen is escaping? The inability to observe the core of the line, corresponding to the inner regions of the outflow, makes the problem of interpretation daunting. Recently, Owen et al. (2021) used an analytic model to explore Lyα transits and found that they are linearly sensitive to the velocity of the outflow, but only logarithmically sensitive to the mass loss rate. Detailed comparisons with 3D simulations will be needed to see if the analytic model is accurate, and more Lyα data will need to be taken to test its predictions.

To a lesser extent, the problem of interpretation extends to helium observations as well. It would be useful to have a model that is intermediate in complexity between 3D hydrodynamic codes which take days to run, and the 1D Parker wind model of Oklopčić & Hirata (2018), which assumes a spherically symmetric isothermal outflow with simplified physics. Such a parameterized model could be used in a Bayesian retrieval code akin to those for static atmospheres, and could be used to infer the mass loss rate, the stellar wind conditions, and other parameters. The problem is that mass loss is fundamentally not a spherically symmetric process, and it is an open question whether modelling it as spherically symmetric plus a perturbation while abandoning physical consistency for the temperature profile will yield any valid results.

Lastly, hydrogen and helium absorption lines are exceptional for studying mass loss from planets with H/He atmospheres, but it is possible that mini Neptunes have high metallicity, and probable that even smaller planets have no H/He atmosphere at all. It is my hope that mass loss studies will detect the panoply of other elements that surely make up the exotic atmospheres of smaller planets. Far UV detections of OI, CII, MgII, and FeII have been claimed (Vidal-Madjar et al., 2004; Sing et al., 2019), but these detections are all for hot Jupiters and have a relatively low signal-to-noise ratio. The discovery of a metastable line similar to He I 1083 nm, but for a metal, would allow ground telescopes to study the escape of metals, opening an entire new subfield in exoplanet atmospheres. This subfield could provide new insights into the atmospheric evolution of smaller, rocky planets with high mean molecular weight atmospheres. This question, of which small planets keep their atmospheres and which lose them, will be crucial for understanding the potential habitability of exoplanets.

References

Arakcheev, A. S., Zhilkin, A. G., Kaigorodov, P. V., Bisikalo, D. V., & Kosovichev, A. G. 2017, Astronomy Reports, 61, 932

Carolan, S., Vidotto, A. A., Hazra, G., Villarreal D'Angelo, C., & Kubyshkina, D. 2021, MNRAS, 508, 6001

Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2016, MNRAS, 455, 2018

Greaves, J. S., Richards, A. M. S., Bains, W., et al. 2020, arXiv e-prints, arXiv:2011.08176

Greaves, J. S., Rimmer, P. B., Richards, A. M. S., et al. 2021a, arXiv e-prints, arXiv:2108.08393

Greaves, J. S., Richards, A. M. S., Bains, W., et al. 2021b, Nature Astronomy, 5, 655

Guilluy, G., Andretta, V., Borsa, F., et al. 2020, A&A, 639, A49

Khodachenko, M. L., Shaikhislamov, I. F., Lammer, H., & Prokopov, P. A. 2015, ApJ, 813, 50

Kipping, D. M., & Tinetti, G. 2010, MNRAS, 407, 2589

Krasnopolsky, V. A., Maillard, J. P., & Owen, T. C. 2004, Icarus, 172, 537

Krissansen-Totton, J., Bergsman, D. S., & Catling, D. C. 2016, Astrobiology, 16, 39, pMID: 26789355

Meadows, V., Lincowski, A., Lustig-Yaeger, J., & Crisp, D. 2021, in Bulletin of the American Astronomical Society, Vol. 53, 0302

Misra, A., Meadows, V., & Crisp, D. 2014, ApJ, 792, 61

Nixon, M. C., & Madhusudhan, N. 2022, arXiv e-prints, arXiv:2201.03532

Oklopčić, A., & Hirata, C. M. 2018, ApJ, 855, L11

Owen, J. E., & Adams, F. C. 2014, MNRAS, 444, 3761

Owen, J. E., Murray-Clay, R. A., Schreyer, E., et al. 2021, arXiv e-prints, arXiv:2111.06094

Powell, D., Louden, T., Kreidberg, L., et al. 2019, ApJ, 887, 170

Qian, Y., & Wu, Y. 2021, AJ, 161, 201

Sing, D. K., Lavvas, P., Ballester, G. E., et al. 2019, AJ, 158, 91

Sousa-Silva, C., Seager, S., Ranjan, S., et al. 2020, Astrobiology, 20, 235

Vidal-Madjar, A., Désert, J. M., Lecavelier des Etangs, A., et al. 2004, ApJ, 604, L69

Villanueva, G., Cordiner, M., Irwin, P., et al. 2020, arXiv e-prints, arXiv:2010.14305

Vos, J. M., Allers, K., Apai, D., et al. 2019, arXiv e-prints, arXiv:1903.06691

Webster, C. R., Mahaffy, P. R., Atreya, S. K., et al. 2015, Science, 347, 415