

The Detection, Characterization, and Retrieval of Terrestrial Exoplanet Atmospheres

Jacob Lustig-Yaeger

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Victoria Meadows, Chair

David Crisp

Michael Line

Program Authorized to Offer Degree:
Astronomy

©Copyright 2020
Jacob Lustig-Yaeger

University of Washington

Abstract

The Detection, Characterization, and Retrieval of Terrestrial Exoplanet Atmospheres

Jacob Lustig-Yaeger

Chair of the Supervisory Committee:
Professor Victoria Meadows
Astronomy

With the upcoming launch of the James Webb Space Telescope (*JWST*), the first light of thirty-meter class ground based telescopes in the 2020s, and plans in the works for a next-generation space-based telescope designed from the outset to probe exoplanet habitability and biosignatures, we are entering the era of terrestrial exoplanet science. Soon we will have the opportunity to study the presence, composition, habitability, and biosignatures of terrestrial exoplanet atmospheres, enabling new perspectives on comparative planetology, astrobiology, and our place as humans within the cosmos. In this dissertation, we explore optimal astronomical observations and theoretical modeling approaches for constraining terrestrial exoplanet environments.

In the first part of this dissertation, we investigate the extent of terrestrial exoplanet science that may be achievable in the *JWST* era. We begin by studying the feasibility and observational cost of detecting and characterizing the composition of the seven Earth-sized TRAPPIST-1 exoplanets with *JWST*. By positing several terrestrial atmospheric compositions that are plausible given the tumultuous history of atmospheric escape expected for this late M dwarf system, we use *JWST* noise models to determine the optimal observing modes and number of transits/eclipses necessary to detect the planet's atmospheres. We show that the likely common presence of CO₂ in terrestrial atmospheres should drive our ability to detect all seven TRAPPIST-1 planet atmospheres with *JWST*'s NIRSpec Prism, although as many as 30 transits may be required if the planets are cloudy Venus-like worlds. We find that water may be prohibitively difficult to detect in both Venus-like

and habitable atmospheres due to its presence lower in the atmosphere where transit transmission spectra are less sensitive. Although the presence of biogenic O_2 and O_3 will be extremely challenging to detect, abiotically produced oxygen from past ocean loss may be detectable for all of the TRAPPIST-1 planets via O_2 – O_2 collisionally induced absorption features at 1.06 and 1.27 μm , or via NIR O_3 features for the outer three planets. These results constitute a suite of hypotheses on the nature and detectability of highly evolved terrestrial exoplanet atmospheres that may be tested with *JWST*.

Although the presence of clouds in the TRAPPIST-1 planet atmospheres are not expected to prevent the atmospheres from being detected, we use models of cloudy Venus-like exoplanets to demonstrate that their atmospheres may remain significantly misunderstood. We find that *JWST*/Mid-IR Instrument (MIRI) Low Resolution Spectrometer (LRS) secondary eclipse emission spectroscopy in the 6 μm opacity window could probe at least an order of magnitude deeper pressures than transmission spectroscopy, potentially allowing access to the sub-cloud atmosphere for the two hot innermost TRAPPIST-1 planets. In addition, we identify two confounding effects of sulfuric acid aerosols that may carry strong implications for the characterization of terrestrial exoplanets with transmission spectroscopy: (1) there exists an ambiguity between cloud-top and solid surface in producing the observed spectral continuum; and (2) the cloud-forming region drops in altitude with semimajor axis, causing an increase in the observable cloud-top pressure with decreasing stellar insolation. Taken together, these effects could produce a trend of thicker atmospheres observed at lower stellar insolation—a convincing false positive for atmospheric escape and a “mirage” of the cosmic shoreline. These studies emphasize that *JWST* should be able to detect terrestrial exoplanet atmospheres around the smallest of stars, but characterizing the unique nature of secondary atmospheres, including habitability and biosignature detection, may be beyond *JWST*’s capabilities.

In the second part of this dissertation, we investigate terrestrial exoplanet science cases that a future space telescope could undertake. A future coronagraph-equipped, direct-imaging telescope

may offer the best chance to directly search for oceans on the surface of exoplanets. To explore the depths of ocean detection, we introduce a novel approach for exoplanet ocean detection that combines two previously distinct techniques—rotational mapping and ocean glint—into the most rigorous approach to date for determining exoplanet habitability. We demonstrate that time-series observations of Earth-like exoplanets at crescent phase can be used to search for the “blinking” of ocean glint as continents, which interrupt the ocean glint spot, rotate into and out of the illuminated planet crescent. This effect causes large-amplitude, time-varying signals in the reflected light, which can be used to longitudinally map the ocean glint. In our simulations, we find that the retrieved apparent albedo of ocean-bearing longitudes is increased by a factor of 5 at crescent phase due to the contribution from glint, compared to the albedo of the same longitudes at quadrature phase. Meanwhile, the predominately land-bearing longitudes exhibit no significant change in apparent albedo with phase. Thus, by using time-dependent information we can decompose the planet into unique spatial regions that provide greater leverage to identify the scattering properties of liquid water. Although these observations would be difficult to acquire in practice, we calculate that NASA’s LUVOIR mission concept could use this method to detect oceans for the closest Earth-like exoplanet targets (e.g. within 10 pc).

Constraining the prevalence of habitable and inhabited exoplanets is one of the main science objectives of the LUVOIR telescope concept. Accomplishing this ambitious objective will require directly-imaged reflectance spectra of many Earth-like exoplanets that can be analyzed for signs of habitability and biosignatures. We use the LUVOIR-A (15 m) and LUVOIR-B (8 m) coronagraph design specifications provided in the LUVOIR Final Report to perform exposure time calculations and determine the number of potentially habitable Earth-like exoplanets for which LUVOIR can undertake a robust search for habitability and biosignatures. We find that a dedicated one-year characterization program with LUVOIR-A (LUVOIR-B) would be capable of observing the spectra of 24 (11) exo-Earth candidates between $0.3 - 1.5 \mu\text{m}$. We highlight the challenges faced when observing exo-Earths in the farthest UV and farthest NIR bandpasses, and we explore the effect

to our spectral yields of imposing a target prioritization strategy based on preliminary evidence about each target. These spectral yields demonstrate that a LUVOIR-like telescope could be used to characterize the nature of a large enough sample of HZ terrestrial exoplanets to place statistical constraints of the prevalence of planetary habitability and life in the Solar neighborhood.

In the third part of this dissertation, we develop, validate, and apply a novel terrestrial exoplanet atmospheric retrieval model. Exoplanet atmospheric retrieval models have become invaluable tools for the interpretation of observed exoplanet spectra, however surprisingly little work has been done to develop and test retrieval methods for *terrestrial* exoplanets despite the excitement for upcoming spectral observations of terrestrial exoplanets with *JWST*. We describe a novel exoplanet atmospheric retrieval algorithm built specifically for terrestrial exoplanets that may span a broad range of atmospheric temperatures, pressures, and compositions. Our retrieval is built around the physically rigorous line-by-line radiative transfer code, SMART, originally built for Solar System terrestrial remote sensing. Our new SMART Exoplanet Retrieval code (SMARTER) has been developed with a modular design featuring a suite of different forward and inverse models that enable atmospheric inference for a variety of different observing techniques—including transmission, emission, and reflectance spectroscopy—coupled to an assortment of different Bayesian retrieval methods—including optimal estimation, Markov chain Monte Carlo, nested sampling, and a novel machine learning approach. We validate SMARTER against high-resolution transmission spectrum observations of Earth observed in solar occultation, demonstrating key capabilities and challenges for interpreting the spectra of terrestrial exoplanets with complex atmospheres.

Finally, to assess the sensitivity to biosignatures and test common retrieval assumptions, we employ our new SMARTER model to interpret the atmosphere of TRAPPIST-1e as if it were observed by an intensive *JWST* transmission spectroscopy campaign, assuming several different plausible habitable and uninhabitable archetypal planetary states. We find that the CH_4 – CO_2 disequilibrium biosignature pair, previously suggested by [Krissansen-Totton et al.](#) for anoxic biospheres like the Archean Earth, may also be retrieved if the planet possesses a modern Earth-like atmosphere due to

the increased photochemical lifetime of CH_4 in the atmospheres of late M dwarf planets. However, enhanced CH_4 absorption features only worsen the already dim prospects for detecting H_2O , which could complicate the interpretation of planetary habitability. We also find that stacking 25 transits could lead to a weak detection of CO_2 in a cloudy Venus-like atmosphere, or alternatively reveal O_2 as the bulk atmospheric gas via $\text{O}_2\text{--O}_2$ CIA for an oxygen-dominated post-runaway atmosphere. We uncover that adding refraction to our transmission retrievals modifies the known degeneracy between reference radius and pressure allowing for an additional mode of interpretation wherein the spectral continuum is controlled by refraction rather than an opaque medium. We also identify a strong implicit prior on the mean molecular weight caused by the common assumption of a background gas (such as H_2/He or N_2) that can significantly affect the retrieved temperature via scale height degeneracies. Finally, we verify that the TRAPPIST-1e planet mass constraints from Transit Timing Variation (TTV) measurements are sufficient to enable dozens of observed transmission spectra to be stacked without incurring additional posterior uncertainties on atmospheric parameters.

The results of this dissertation both underscore the compelling future and highlight some of the key challenges ahead for terrestrial exoplanet atmospheric characterization. In the near term, observations with *JWST* are well-positioned to begin detecting and characterizing terrestrial exoplanet atmospheres, but are ultimately unlikely to constrain the presence of habitable conditions or biosignatures even for the most optimistic targets. Investing resources in a future space telescope that is capable of directly imaging Earth-like exoplanets is crucial to optimize the ambitious pursuit of habitable and inhabited planets. In advance of these upcoming missions, our new retrieval model can continue to be used, tested, and improved upon to maximize our chances of correctly interpreting the first precise spectra of rocky exoplanets.

TABLE OF CONTENTS

	Page
List of Figures	vi
List of Tables	xxi
Chapter 1: Introduction	1
1.1 At the Dawn of the Terrestrial Exoplanet Characterization Era	2
1.2 Terrestrial Exoplanet Atmospheres	3
1.2.1 Formation & Evolution	3
1.2.2 Planetary Habitability & The Habitable Zone	5
1.2.3 The Search for Life & Biosignatures	6
1.3 Detecting Terrestrial Exoplanets	7
1.3.1 Direct Detection	8
1.3.2 Indirect Detection	9
1.3.3 Targets for Characterization	13
1.4 Characterizing Terrestrial Exoplanets	13
1.4.1 Observing Techniques	14
1.4.2 Modeling Techniques	21
1.5 Next Steps	26
1.5.1 Milestones for Terrestrial Exoplanet Characterization	26
1.5.2 A Timeline for the Characterization of Terrestrial Exoplanets	28
1.6 Purpose & Structure of this Work	32
1.7 Rotational Mapping	34
1.7.1 samurai: Surface Albedo Mapping Using Rotational Inversion	34
1.8 Telescope Noise Modeling	34
1.8.1 coronagraph: A Noise Model for Direct-Imaging Exoplanets	34
1.8.2 PandExo: A Noise Model for Exoplanet Spectroscopy with JWST	34

1.9	Radiative Transfer	34
1.9.1	smart: Spectral Mapping Atmospheric Radiative Transfer	34
Part I:	Developing a Foundation to Characterize Transiting Exoplanets in the Era of JWST	35
Chapter 2:	Detecting & Characterizing the TRAPPIST-1 Exoplanet Atmospheres with JWST	36
2.1	Introduction	38
2.2	Methods	42
2.2.1	JWST Noise Modeling	42
2.2.2	Noise Model Inputs	45
2.2.3	Observing Experiments	46
2.3	Results	54
2.3.1	Detecting Atmospheres	55
2.3.2	Distinguishing Specific Molecules	68
2.4	Discussion	71
2.4.1	Do the TRAPPIST-1 planets have atmospheres?	71
2.4.2	What is the nature of the TRAPPIST-1 planet atmospheres?	77
2.4.3	Further Considerations	80
2.5	Conclusion	82
2.6	Chapter Appendix	84
2.6.1	Signal-to-Noise Ratio on Transit and Eclipse Depths	84
2.6.2	Atmospheric Detectability by Instrument	86
2.6.3	Brightness Temperatures	86
Chapter 3:	A Mirage of the Cosmic Shoreline: Venus-like Clouds as a Statistical False Positive for Exoplanet Atmospheric Erosion	95
3.1	Introduction	97
3.2	Methods	100
3.3	Results	103
3.3.1	Transmission Spectroscopy	103
3.3.2	Emission Spectroscopy	107
3.4	Discussion	110

3.4.1	Accessing the lower atmospheres of Venus-like exoplanets	111
3.4.2	An ambiguity between cloud-top and solid surface	112
3.4.3	Clouds as a statistical false positive for atmospheric loss	113
3.4.4	Mitigation strategies & opportunities	114
3.5	Conclusions	115
Part II:	Developing a Foundation to Characterize the Habitability & Biosignatures of Directly-Imaged Exoplanets	117
Chapter 4:	Detecting Ocean Glint on Exoplanets Using Multiphase Mapping	118
4.1	Introduction	120
4.2	Methods	122
4.2.1	Generating Synthetic Lightcurves	123
4.2.2	Lightcurve Fitting	126
4.3	Results	130
4.3.1	Multi-Phase Longitudinal Mapping	130
4.3.2	Observational Feasibility	137
4.4	Discussion	148
4.4.1	Increasing the Robustness of Exoplanet Ocean Detection	149
4.4.2	Observational Feasibility and Yield Estimates	155
4.5	Conclusions	158
Chapter 5:	Searching for Biosignatures and Confirming Habitability with LUVOIR . . .	160
5.1	Introduction	161
5.2	Methods	164
5.2.1	Observation Overview	164
5.2.2	Targets	166
5.2.3	Coronagraph Model	167
5.2.4	Exposure Times & Overheads	169
5.3	Results	171
5.4	Discussion & Conclusion	176
5.4.1	Targeted versus Agnostic Observing Strategies	179
5.4.2	Future Work	179

Part III:	Retrieving the Atmospheres of Terrestrial Exoplanets	182
Chapter 6:	Development of a Terrestrial Exoplanet Atmospheric Retrieval Model	183
6.1	Introduction	184
6.2	The SMART Exoplanet Retrieval (SMARTER) Model	189
6.2.1	Bayes' Theorem	191
6.2.2	The Spectral Mapping Atmospheric Radiative Transfer (SMART) Model .	196
6.2.3	Forward Model	201
6.2.4	Inverse Model	214
6.3	SMARTER Validation using Earth's Transmission Spectrum	228
6.3.1	Observed Data	229
6.3.2	Retrieval Experimental Setup	231
6.3.3	Retrieval Validation Results	233
6.3.4	Discussion & Conclusion	239
6.4	Discussion & Conclusion	242
6.4.1	Continued Validation using Titan's Transmission Spectrum	242
6.4.2	Developing a Physically Augmented Atmospheric Retrieval Code	244
Chapter 7:	Retrieving TRAPPIST-1e's Atmospheric Composition with JWST	248
7.1	Introduction	250
7.2	Simulated JWST Data	253
7.3	Retrieval Model	255
7.3.1	Forward Model	256
7.3.2	Inverse Model	260
7.4	Results	264
7.4.1	Atmospheric Composition, Habitability & Biosignatures	265
7.4.2	Comparison to Heterogeneous Vertical Profiles	268
7.4.3	Refraction	270
7.4.4	Planet Mass & TTV Constraints	273
7.4.5	Mean Molecular Weight & The Bulk Atmospheric Composition	276
7.4.6	Random Noise Instances	279
7.4.7	APPROXPOSTERIOR Validation	281
7.5	Discussion	284

7.5.1	Characterizing the Atmosphere of TRAPPIST-1e	284
7.5.2	Terrestrial Exoplanet Atmospheric Retrieval Subtleties	286
7.5.3	Active Learning for Atmospheric Retrievals	289
7.6	Conclusion	289
Chapter 8:	Discussion	291
8.1	On the Era of Terrestrial Exoplanet Science	292
8.1.1	Detecting Terrestrial Exoplanet Atmospheres	292
8.1.2	Characterizing Terrestrial Exoplanet Atmospheres	295
8.1.3	Identifying Habitable Exoplanets	297
8.1.4	Confirming Exoplanet Biosignatures	300
8.2	Ongoing & Future Work	302
Chapter 9:	Conclusions	305
Bibliography		310
Appendix A:	Vita	363

LIST OF FIGURES

Figure Number		Page
1.1	Graphic depicting the many methods that exoplanet astronomers can use to characterize exoplanet environments, and the orbital phase at which each method is usually applied. Methods like transits, secondary eclipses, and planet-planet occultations are only viable for transiting exoplanet systems, while other methods, such as phase curves, direct imaging reflectance spectroscopy, and rotational mapping, may be used for transiting and non-transiting exoplanet architectures. This figure was developed by J. Lustig-Yaeger for the VPL CAN-8 proposal, and made beautiful by Ron.	15
2.1	Secondary eclipse spectrum models of TRAPPIST-1b assuming different end-member planet mineralogical compositions. The black line shows a featureless blackbody curve, corresponding to the zero bond albedo planet equilibrium temperature, from which all of the thermal emission curves from rock forming minerals (thick color lines) deviate due to non-unity, wavelength-dependent emissivities. The three line styles for the quartz curves demonstrate the reduction in effective emissivity due to n reflections within the rock (see equation 2.9). The thin purple and teal lines show the expected emission spectrum for TRAPPIST-1b if it were to possess a climatically and photochemically self-consistent thick O_2 - or a CO_2 -dominated atmosphere, respectively (Lincowski et al., 2018).	50
2.2	Signal-to-noise contours on the simulated 10 bar high CO_2 transmission spectrum, observed with JWST/NIRSpec Prism with 6 groups per integration, relative to a featureless spectrum as a function of the number of transits observed and wavelength of the instrument. The bottom panel shows the noiseless spectrum (black) and the best-fitting featureless spectrum (orange), both convolved to the instrument resolution. The right panel shows how the total expected SNR from Equation 2.6 increases with more transits observed.	52

2.3	Signal-to-noise contours on the simulated 10 bar high CO ₂ emission spectrum observed with JWST/MIRI/LRS relative to a featureless spectrum as a function of the number of occultations observed and wavelength of the instrument. The bottom panel shows the noiseless spectrum (black) and the best-fitting featureless spectrum (orange), both convolved to the instrument resolution. The right panel shows how the total expected SNR from Equation 2.6 increases with more eclipses observed.	53
2.4	Signal-to-noise contours on the O ₂ contribution to the transmission spectrum of TRAPPIST-1b if it possesses a 10 bar desiccated high O ₂ atmosphere observed with NIRSpec G140H as a function of the number of occultations observed and wavelength of the instrument. The bottom panel shows the full model spectrum (black) and the model spectrum with O ₂ removed (orange), both convolved to the resolution of NIRSpec G140H. The right panel shows the total expected SNR ($\langle \text{SNR} \rangle$) from Equation 2.6. The O ₂ features at 1.06 and 1.27 μm are due to O ₂ -O ₂ collisionally-induced absorption (CIA), and could lead to $\langle \text{SNR} \rangle = 5$ detection of O ₂ in ~ 7 transits.	54
2.5	Brightness temperatures for different TRAPPIST-1b atmospheric models across the JWST/MIRI imaging wavelength range. The average brightness temperature integrated across each MIRI filter are shown as color points. The zero bond albedo equilibrium temperature of TRAPPIST-1b (black horizontal dashed line) is shown to compare against the model brightness temperatures, which vary as a function of wavelength due to the interplay between atmospheric temperature structure and gaseous opacities. The black lines from the top of the plot correspond to the (atmospheric model averaged) SNR on a single secondary eclipse, shown increasing from the top of the right y-axis to the bottom. These SNR indicators are shown capped with the effective width of each MIRI filter, and help to identify which filters may best offer secondary eclipse detections.	56
2.6	Number of transits (blue) or eclipses (red) needed to detect an atmosphere for all seven known TRAPPIST-1 planets with MIRI photometry (left) and JWST spectroscopy (right). Line and marker styles designate the type of atmosphere assumed. MIRI imaging assumes that the number of occultations are split evenly between the 2-3 photometric filters that are optimal to detect each atmosphere (see §2.3.1). Transmission spectroscopy is shown using NIRSpec Prism* (see §2.3.1) and emission spectroscopy is shown using MIRI LRS (see §2.3.1). Detecting the atmospheres of the TRAPPIST-1 planets is optimal using transmission spectroscopy with NIRSpec Prism across a range of terrestrial atmospheric compositions.	59
2.7	Model transmission spectrum of TRAPPIST-1b with a 10 bar CO ₂ atmosphere. Synthetic data are simulated for two transits observed with NIRSpec Prism* and binned to a resolution of $R = 8$.	60

2.8	Simulated JWST/NIRSpec Prism transmission spectra of TRAPPIST-1c for different CO ₂ -dominated atmospheric states. The blue model shows a clear sky spectrum with large CO ₂ absorption features and the red model includes a self-consistent H ₂ SO ₄ cloud. The number of coadded transits simulated for each NIRSpec observation are set so that the atmosphere is strongly detected ($\langle \text{SNR} \rangle = 10$). That is, to obtain approximately equal constraints on the presence and composition of TRAPPIST-1c's atmosphere will require 16 transits for the clear sky CO ₂ case and 72 transits for the cloudy case.	61
2.9	Simulated JWST/NIRSpec Prism* transmission spectra of TRAPPIST-1e for different possible atmospheric compositions. The teal and blue models show spectra for water-dominated atmospheres, with and without a water cloud, respectively. The orange and red spectra are for CO ₂ -dominated atmospheres, with and without H ₂ SO ₄ aerosols, respectively. The number of coadded transits simulated for each NIRSpec observation are set so that the atmosphere is strongly detected ($\langle \text{SNR} \rangle = 10$). That is, to obtain approximately equal constraints on the presence of TRAPPIST-1e's atmosphere will require ~ 30 transits for a clear sky CO ₂ atmosphere, ~ 50 transits for a clear sky H ₂ O atmosphere, ~ 90 transits for a cloudy H ₂ O atmosphere, and ~ 120 transits for a CO ₂ -dominated Venus-like atmosphere. .	62
2.10	Number of transits for each TRAPPIST-1 planet necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different self-consistent atmospheric compositions using JWST NIRSpec Prism with the optimized readout mode of Batalha et al. (2018).	63
2.11	Estimated detectability of pure quartz surface emissivity features in the secondary eclipse spectrum of TRAPPIST-1b with JWST's MIRI LRS instrument. The top panel shows the MIR effective emissivity of a quartz surface for three different assumed values for the number of surface reflections (see Equation 2.9): 0 (blue), 5 (orange), and 10 (green). The bottom panel shows theoretical secondary eclipse spectra for TRAPPIST-1b assuming the effective emissivities from the top panel. The inset axis shows the number of secondary eclipses that would be required to observe with MIRI LRS to detect features in the emission spectrum with a $\langle \text{SNR} \rangle = 2.0$, for each of the three assumed effective emissivities. Only an extremely polished pure quartz surface could possibly be detected with JWST.	67

2.12	Theoretical transmission spectra of TRAPPIST-1 b assuming three different atmospheric compositions with modeled noise for JWST observations. <i>Top:</i> Transmission spectrum of a 10 bar desiccated O ₂ atmosphere shown with error bars calculated for 12 transits with NIRSpec G140H—sufficient for $\langle \text{SNR} \rangle \sim 5$ on the O ₄ features. <i>Middle:</i> Transmission spectrum of a 10 bar outgassing O ₂ atmosphere shown with error bars calculated for 3 transits with NIRSpec Prism*—sufficient for $\langle \text{SNR} \rangle \sim 5$ on the H ₂ O features. <i>Bottom:</i> Transmission spectrum of a 10 bar CO ₂ atmosphere shown with error bars calculated for 125 transits with NIRSpec Prism*—sufficient for $\langle \text{SNR} \rangle \sim 5$ on the NIR H ₂ O features—and 125 transits with MIRI LRS—sufficient for $\langle \text{SNR} \rangle \sim 5$ on both the 6 μm H ₂ O feature and the 7.3 and 8.7 μm SO ₂ features.	72
2.13	Number of TRAPPIST-1b transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	88
2.14	Number of TRAPPIST-1c transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	89
2.15	Number of TRAPPIST-1d transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	90
2.16	Number of TRAPPIST-1e transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	91
2.17	Number of TRAPPIST-1f transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	92
2.18	Number of TRAPPIST-1g transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	93
2.19	Number of TRAPPIST-1h transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.	94

3.1	Transmission spectrum models of the TRAPPIST-1 planets assuming they possess thick Venus-like atmospheres with H_2SO_4 clouds (top panels) and without clouds (bottom panels) (models from Lincowski et al., 2018). The left panel shows transmission spectra in units of pressure probed into the atmosphere. The right panel shows the atmospheric temperature structure (solid lines; lower axis) and the cumulative vertical cloud/haze optical depth from the top of the atmosphere (dashed lines; upper axis). Thicker line styles indicate the region of the atmosphere that the transmission spectrum is sensitive to, and dots are shown on the temperature profiles to indicate the highest possible pressure that may be probed by the transmission spectrum for such planets. H_2SO_4 clouds effectively prevent the lower atmosphere of Venus-like exoplanets from being remotely sensed by a transmission spectrum. However, for the cooler planets, cloud formation occurs at lower altitudes and enables the transmission spectrum to probe deeper into the atmosphere.	104
3.2	Maximum pressure probed in the transmission spectrum (left panel) and emission spectrum (right panel) of clear (blue line) and cloudy (red line) Venus-like models of the TRAPPIST-1 planets as a function of semi-major axis. The surface pressures of Venus, Earth, and Mars are shown as dashed horizontal grey lines for reference. The maximum transit pressure appears to increase with semi-major axis for models with H_2SO_4 clouds, while it decreases with semi-major axis for clear sky models. Although the maximum pressure probed in emission exceeds that of transmission, it does not approach the simulated Venusian surface pressure.	105
3.3	Analogous to Figure 3.1 but for emission spectrum models of the TRAPPIST-1 planets assuming they possess thick Venus-like atmospheres with H_2SO_4 clouds (top panels) and without clouds (bottom panels). The left panel shows the atmospheric pressure from which the planets thermal emission emanates ($\tau = 1$). Various NIR and MIR atmospheric windows, which enable the planet to thermally cool to space, may allow distant observers to detect the higher pressures and temperatures of the sub-cloud atmosphere, but these windows can be closed by aerosol extinction (e.g. TRAPPIST-1 d) and gaseous absorption (e.g. H_2O at $6\ \mu\text{m}$), but they are infeasible for JWST to detect for planets cooler than TRAPPIST-1 c. . . .	108
3.4	Synthetic secondary eclipse spectra of the TRAPPIST-1 planets assuming they possess either thick Venus-like atmospheres with H_2SO_4 clouds (top panel) or no clouds (bottom panel). Reflected stellar flux dominates at short wavelengths, while thermal emission dominates at long wavelengths.	110

4.1	Multi-phase inverse problem example for the cloudless Earth. <i>Top</i> : Synthetic lightcurves for two wavelength intervals with error bars for the cloudless Earth, and retrieved fits to the observations (lines) at quadrature (solid lines) and crescent (dotted lines) phase. Data were simulated for the cloudless Earth at 5 pc observed with a 15 m telescope with 50% throughput, which yielded a 1-hour photometric precision of 4.5% at quadrature and 6.5% at crescent. <i>Bottom Left</i> : Retrieved surface albedos from a two surface fit to rotational lightcurves of the cloudless Earth at 90° phase and 135° phase. The color of each line identifies the surface and the line style denotes the phase. The envelope about each line shows the 1σ interval about the median posterior PDF derived from MCMC. From quadrature to crescent phase, the average albedo of Surface 1 increases by a factor of ~ 16 with 15σ confidence, while the average albedo of Surface 2 increase by $\sim 15\%$ at $<1\sigma$. <i>Bottom Right</i> : Median retrieved longitudinal slice maps. Each longitudinal slice is colored proportional to the retrieved area covering fraction of the given surface, with white meaning that none of the given surface is found within the slice. Maps are shown in the Robinson projection.	131
4.2	Histograms showing the increase in albedo for Surface 1 (blue) and Surface 2 (orange) between quadrature and crescent phase observations of the cloudless Earth. Surface 1 increased in apparent albedo by 0.47 ± 0.02 from quadrature to crescent phase, indicating that it has a highly non-Lambertian scattering phase function. Surface 2 has a negligible albedo increase, indicating that it is consistent with Lambertian scattering.	133
4.3	<i>Top</i> : Longitudinal slice maps showing the median (left) and standard deviation (right) of the change in area covering fraction between quadrature and crescent phase observations of the cloudless Earth. <i>Bottom</i> : Longitudinal slice-averaged fractional area change of Surface 1 from quadrature to crescent phase. The maps agree best for slices containing majority open ocean, with the marginally statistically significant change in the maps with phase primarily driven by disagreements in land-bearing slices.	134
4.4	Similar to the lower panels of Fig. 4.1, but now for Earth with time-varying clouds and at three different levels of photometric precision. From top to bottom, the data have a 1-hour S/N (photometric precision) of 23 (4.3%), 10 (10%), and 2.4 (42%), at quadrature, which correspond to observations of Earth at 5 pc with a 15 m and 50% throughput telescope, a 15 m and 10% throughput telescope, and a 6 m with 10% throughput telescope, respectively.	136

- 4.5 Coronagraph inner working angle (IWA) requirements for observing Earth-analog exoplanets. The top panel corresponds to a planet at quadrature phase ($\alpha = 90^\circ$) and the bottom panel corresponds to a planet at crescent phase ($\alpha = 135^\circ$). The IWA is given as the factor which multiplies λ/D , at a fixed wavelength $\lambda = 0.8 \mu\text{m}$, as a function of telescope diameter D and the distance to the system d 139
- 4.6 Change in surface albedo (top) and change in surface longitudinal area covering fractions (bottom) as a function of the signal-to-noise of the observations. The lower x-axis corresponds to the 1-hour photometric signal-to-noise for the Earth at quadrature phase, while the upper x-axis corresponds to the 1-hour signal-to-noise at crescent phase for observations of the same exoplanet, at the same distance, using the same model telescope configuration, with the same integration time. The upper panel shows the phase-dependent albedo increase of Surface 1 (blue) and Surface 2 (orange) with $1-\sigma$ error bars derived by randomly sampling the MCMC posterior distributions. The lower panel shows a similar result but for the change in longitudinal area covering fraction. The red shaded region indicates the crescent phase photometric precision where a one surface fit (homogeneous map) is favored by the Bayesian Information Criterion (BIC) over a two surface fit. 140
- 4.7 Distance (color contours) out to which an Earth-like exoplanet could have crescent phase ($\alpha = 135^\circ$) observations fit with a two surface model, as a function of telescope diameter and throughput. A 1-hour photometric precision of $\sim 15\%$ ($S/N=6.7$) is determined by the BIC to be the threshold between detecting one surface (homogeneous planet) and two surfaces (heterogeneous planet) at crescent phase, considering our observational assumptions, particularly the limited bandwidth. Minimally detectable two surface maps at quadrature phase can be performed out to approximately twice the distance shown here for crescent phase. . . . 143
- 4.8 *Left:* Projected separation between star and planet at the inner edge of the habitable zone (moist greenhouse limit) for low mass stars. Quadrature and crescent (135°) phase projected separations are shown as solid and dotted lines. *Right:* Maximum distance (color contours) out to which a planet at the inner edge of the HZ can be directly imaged as a function of stellar mass and telescope diameter. Phase, wavelength, and IWA are fixed at 90° , $0.8 \mu\text{m}$, and $3\lambda/D$. Distances can be readily scaled to the apparent phase $\alpha = 135^\circ$ by multiplying by $\sin(\alpha)$ 145

4.9	Yield estimates for observations of HZ exoplanets at quadrature (left panel) and crescent phase (right panel) as a function of telescope diameter for nearby G, K, and M dwarfs. Systems are counted if a hypothetical planet in the HZ falls outside of the IWA (and is therefore observable), assuming $\theta_{IWA} = 3\lambda/D$, and if the 1-hour S/N at crescent phase exceeds our photometric threshold of 6.7 (see Fig. 4.6), assuming a 20% throughput. Both panels share the same y-axes, where the left y-axis gives the number of stars that would have observable HZ planets. The right y-axis gives the expected number of exoplanets after considering the geometric probability of passing through 135° phase (70%) and assuming a 20% occurrence rate for planets in the HZ. The shaded regions are bounded below and above by the inner and outer edges of the HZ, respectively, and demonstrate a sensitivity in yield depending on whether the observed planet is at the inner or outer edge of the HZ. The dashed lines correspond to the sum of all partially observable HZs (e.g. half a count for a star with half of the HZ observable outside the IWA). Observations of HZ exoplanets are limited by photometric precision for G dwarfs while M dwarfs are significantly more limited by the IWA.	146
5.1	Simulated spectrum of an Earth analog exoplanet observed with LUVOIR in this DRM observing program (SNR = 8.5). The shaded background regions indicate the wavelength ranges of the UV (blue) and visible (green), and NIR (red) channels. The vertical lines show 20% bandpasses in the UV and visible channels and 10% bandpasses in the NIR channel.	165
5.2	LUVOIR’s wavelength-dependent optical throughput used in this study. The throughput minimum is about 5% at 200 nm and reaches a maximum of 52% by 2 μm . . .	170
5.3	Raw contrast and core throughput curves as a function of planet-star separation for three different coronagraph designs considered in the LUVOIR Final Report. The coronagraph contrast curves (dotted lines) are shown relative to the left y-axis and the core throughput curves (solid lines) are shown relative to the right. The left panel shows the Apodized Pupil Lyot Coronagraph (APLC) considered for the LUVOIR-A concept; the central panel shows the Deformable Mirror assisted Vortex Coronagraph (DMVC) considered for LUVOIR-B); and the right panel shows the Phase-Induced Amplitude Apodization (PIAA) coronagraph considered for LUVOIR-B.	170

5.4	Required science exposure time per bandpass (left axis) to reach SNR=8.5 for LUVOIR-A (top panel) and LUVOIR-B (bottom panel). The blue (UV channel), green (visible channel), and red (NIR channel) “boxes and whiskers” show the 50% (box extent) and 95% (whisker extent) confidence intervals about the median (white bar) science time required to observe the spectrum shown in Figure 5.1, for the targets in the biased catalog of exo-Earth candidates. The percentage of stars in the target catalog that can be completely observed in each bandpass—considering IWA and OWA constraints is displayed above each box. Exposure times required to achieve a fixed precision on the spectrum can vary by orders of magnitude. . . .	172
5.5	Number of spectra (color map and contour lines) that may be observed by LUVOIR-A (left panel) and LUVOIR-B DMVC (right panel) as a function of both total observing time and fraction of “interesting” targets in the exo-Earth candidate sample, η_{water} . The two shortest wavelength UV bandpasses and the four longest NIR bandpass are not included, so the spectrum spans $0.29 - 1.46 \mu\text{m}$. A higher fraction of candidate targets increases the probability that nearby, more easily observed, stars are included in the exo-Earth candidate sample, which yields more complete spectra in shorter duration programs. . . .	174
5.6	Number of spectra (SNR=8.5) observable in a 6-month or 12-month LUVOIR-A habitable exoplanet characterization program for different observed wavelength ranges (spectral completeness) and assuming different η_{water} fractions. The quoted “plus-minus” values only represent the uncertainty in which specific stellar targets are randomly selected to host “interesting” planets via η_{water}	176
5.7	Number of spectra (SNR=8.5) observable in a 6-month or 12-month LUVOIR-B habitable exoplanet characterization program for different observed wavelength ranges (spectral completeness) and assuming different η_{water} fractions. The quoted “plus-minus” values only represent the uncertainty in which specific stellar targets are randomly selected to host “interesting” planets via η_{water}	177
5.8	Example reflected light spectra of the modern Earth (top panel), low O ₂ Proterozoic Earth (second panel), hazy Archean Earth (third panel), and Venus (bottom panel) as they would be observed, as exoplanets, in this theoretical LUVOIR observing program. . . .	181

6.1	Diagram of the standard retrieval workflow used by SMARTER. An initial guess for the surface and atmospheric state is passed to the forward model, along with gas and aerosol optical properties. The forward model calculates the spectrum and uses an instrument model to degrade it to the quality of the observed data, before passing the spectrum to the inverse model for comparison with the observed spectrum via Bayes' theorem. The inverse model then determines a new state for the planet surface and atmosphere, and runs it back through the forward model. The retrieval iterates in this way, until the inverse model's specific convergence criteria are met and a best fitting set of parameters (or the full posterior distributions) are obtained.	190
6.2	Example of uniformly (blue) and normally (orange) distributed priors. The solid lines show the analytic probability density functions for each distribution, while the histograms show 10,000 random samples from each.	195
6.3	Temperature Jacobian for the Earth's thermal emission radiance spectrum. The central panel shows the full Radiance Jacobian matrix as a color contour to emphasize positive (blue) and negative (red) departures from zero. The lower panel shows the radiance spectrum in units of $\text{W/m}^2/\mu\text{m/sr}$, which shares the x-axis with the Jacobian. The right panel shows the temperature-pressure structure of Earth's atmosphere, which shares the y-axis with the Jacobian. Positive/Blue (negative/red) regions of the Jacobian indicate that the radiance will increase (decrease) as the temperature at those atmospheric pressure levels increases. This highlights the atmospheric pressures that shape the observable spectrum.	198
6.4	Example transmission spectrum output from the SMARTER transit forward model. This spectrum corresponds to an Earth-size planet transiting a TRAPPIST-1-size star. The planet atmosphere is modeled with a 200 K isothermal TP profile with an evenly mixed atmospheric column with 10 ppm H_2O , 100 ppm CO_2 , and the remaining bulk atmosphere composed of N_2 .	204
6.5	Example emission spectrum output from the SMARTER emission forward model. This spectrum corresponds to an Earth-size planet eclipsed by TRAPPIST-1. The planet atmosphere is modeled with an adiabatic tropospheric TP with an isothermal stratosphere and with an evenly mixed atmospheric column with 1 ppm O_3 , 10 ppm H_2O , 100 ppm CO_2 , and the remaining bulk atmosphere composed of N_2 . The dashed lines correspond to isothermal blackbody spectra at the surface (purple) and stratospheric (cyan) temperatures.	206
6.6	Grey surface albedo prior for the reflectance spectroscopy forward model. Using a uniform prior on the $\log_{10}$ surface albedo, exerts a bias towards lower albedo surfaces.	209

6.7	Earth’s reflectance spectrum from the 3D VPL Earth Model (cyan line) compared to the SMARTER 1D retrieval forward model (purple line). The faint error bars approximate the observational uncertainty on the VPL Earth Model spectrum that may be achieved by the 15m LUVOIR telescope concept. The 1D forward model was calculated using best fitting parameters from a retrieval to the synthetic LUVOIR data that was initialized using the fiducial parameters from Feng et al. (2018, see their Table 1 and Figure 4 for reference).	212
6.8	Optimal estimation retrieval results for the transit forward model example shown in Fig. 6.4. The top panel shows the best fitting spectra from 28 randomly initialized Nelder-Mead optimizations. The bottom panels show the corresponding best fit parameter states, with the “true” state given by the black square, and the global MAP solution is the darkest blue point and is indicated with the white central dot. Both panels are color-coded according to the log-likelihood, which is shown in the bottom right panel.	216
6.9	MCMC retrieval “trace plot” example showing the evolution of the ensemble of walkers (color lines) as a function of iteration for each retrieval parameter (sub-panel rows). The right column of subplots shows the walker positions, collapsed along the step-dimension into histograms of the most-likely parameter values. This type of plot is not typically provided in published papers, but is usually produced to aid in determining whether an MCMC simulation has converged.	218
6.10	Example “corner plot” from the same MCMC retrieval as Fig. 6.9. Subplots along the diagonal show the marginalized posterior distributions for each retrieved parameters, while off-diagonal subplots show the covariances between pairs of parameters. Contours are used to indicate the most dense regions of parameter space and points are used to show posterior samples in lower probability regions.	219
6.11	Distribution in spectra sampled from the posterior for the transit example retrieval case. The dark blue line shows the median model fit to the data and the cyan envelopes show the 1σ and 3σ confidence intervals about the median.	220
6.12	MULTINEST retrieval “trace plot” example showing the reduction in prior volume as the algorithm converges on the higher-likelihood region of parameter space. The left column of plots shows each parameter’s values as a function of the negative prior volume ($-\ln X$), which increases monotonically as the sampling volume decreases and the algorithm progresses forward. The points are colored by their relative posterior weight and represent the location and progression of the live points. The right column of plots shows the marginalized posterior distributions.	223
6.13	Corner plots showing the MULTINEST live points colored by their respective posterior weight (left corner plot; same coloring as Fig. 6.12) and contours of posterior probability (right corner plot; same quantities as 6.10).	224

6.14	Corner plots showing the APPROXPOSTERIOR training set (left corner plot) and MCMC posterior samples from the GP surrogate model (right corner plot) for the transit example case.	226
6.15	Empirical transmission spectrum of the cloudless Earth from Macdonald and Cowan (2019). The top panel shows the full wavelength range in high-resolution (light blue) and binned to a lower resolution ($\Delta\lambda = 0.02 \mu\text{m}$; dark blue). The lower panels zoom into the $4.2 \mu\text{m}$ CO_2 band, showing progressively smaller wavelength ranges and highlighting the fully resolved spectral lines with no apparent observational noise.	230
6.16	Spectral resolution of Earth’s observed transmission spectrum from Macdonald and Cowan (2019). The resolving power is plotted in millions, and is shown decreasing as a function of wavelength. This is a characteristic relationship that occurs because the observed spectrum has a fixed FWHM in units of wavenumbers (cm^{-1}).	231
6.17	Earth’s 1D atmospheric thermal structure and composition. The temperature structure (dashed black) is shown along the top x-axis and the gas mixing ratios are shown along the bottom x-axis, both as a function of pressure.	232
6.18	Earth’s observed cloud-free transit transmission spectrum from Macdonald and Cowan (2019) binned to lower resolution and fit using SMARTER. The data are shown in black, the best-fitting model spectrum is shown in white and blue, and various colors are used to indicate each molecule’s contribution to the model spectrum. Residuals between the data and model are shown in red.	234
6.19	Characteristic thermal profiles of Earth’s atmosphere from the VPL Earth Model. The profiles are derived from satellite data acquired during the EPOXI fly by of Earth in 2008 (see Robinson et al., 2011). Each latitude plotted represents the globally averaged TP for all HEALPix bins with the corresponding north and south latitudes. <i>Click to download these TP profiles</i>	236
6.20	Earth’s observed transit transmission spectrum from Macdonald and Cowan (2019) fit using SMARTER at the native spectral resolution. The data are shown in black, the best-fitting model spectrum is shown in blue, and various colors are used in the bottom panels to indicate each molecule’s contribution to the model spectrum. In the bottom panel, HNO_3 , CH_4 , and H_2O show subtle absorption features above a $\text{O}_2\text{--N}_2$ CIA, $\text{O}_2\text{--O}_2$ CIA, and H_2O continuum.	238
6.21	A comparison between the $6\mu\text{m}$ $\text{O}_2\text{--N}_2$ CIA band from Moreau et al. (2001) and the $6\mu\text{m}$ $\text{O}_2\text{--O}_2$ CIA band from HITRAN 2016 (Gordon et al., 2017). At Earth-like temperatures and pressures, the $\text{O}_2\text{--N}_2$ and $\text{O}_2\text{--O}_2$ CIA bands appear to agree in intensity.	240

6.22	Titan’s observed transit transmission spectrum from Robinson et al. (2014b) fit using SMARTER. The model assumes a fractal aggregate hydrocarbon haze that gives the spectral continuum its characteristic linearly decreasing slope. Molecular absorption features are seen from methane (CH_4), acetylene (C_2H_2), carbon monoxide (CO), and a C–H stretch transition modeled using propane (C_3H_8) as a proxy for all aliphatic hydrocarbons.	243
6.23	Titan’s 1D atmospheric thermal structure and composition. The temperature structure (dashed black) is shown along the top x-axis and the gas mixing ratios are shown along the bottom x-axis, both as a function of pressure.	244
6.24	Diagram of a physically and chemically self-consistent retrieval. An initial guess for the gas fluxes from the surface is passed to the forward model, which first calculates the atmospheric composition from climate and photochemical models, and then produces a spectrum to compare with the observed data.	246
6.25	Diagram of a “retrieval-on-a-retrieval” approach to physically interpreting the retrieved atmospheric composition. An initial data-driven “standard” retrieval is used to infer constraints on the atmospheric composition. These atmospheric constraints are then passed into a second “retrieval” that attempts to fit climate and photochemistry models to the inferred atmospheric vertical profiles.	247
7.1	Simulated transmission spectrum observations of TRAPPIST-1e with the <i>JWST</i> NIRSpec Prism assuming the planet possesses a photochemically self-consistent modern Earth-like atmosphere. The noise calculations assume that TRAPPIST-1e had 25 transits observed and coadded.	255
7.2	Fit to synthetic <i>JWST</i> /NIRSpec Prism data of TRAPPIST-1e assuming the observed planet possesses a clear Earth-like atmosphere. The purple envelopes show the 1σ and 3σ uncertainty about the median retrieved spectral fit.	265
7.3	Corner plot showing 1D and 2D marginalized posterior distributions for the retrieval of the clear modern Earth-like transmission spectrum shown in Fig. 7.2. The upper right panel shows Bayes factors for models excluding the listed molecule to quantify the evidence provided by the spectrum in support of that molecule. From smaller to larger values of $\ln B$, the horizontal dashed lines show the limits for “weak” (2.1σ), “moderate” (2.7σ) and “strong” (3.6σ) detections, respectively. . .	266
7.4	Bayes factors showing the evidence in favor of each molecule from retrievals on 25 transits of TRAPPIST-1e assuming multiple plausible atmospheric compositions. .	267

7.5	Retrieved marginalized posterior distributions compared to the 1D vertical atmospheric profiles used to generate the synthetic <i>JWST</i> spectra. Each horizontal row of plots shows a different atmospheric case. Each subplot shows a normalized histogram of the 1D marginalized posterior, compared to that parameters true vertical profile shown as a function of pressure along the vertical y-axis.	269
7.6	Marginalized posterior distributions for retrievals with (cyan) and without (purple) refraction in the transmission spectrum forward model. Simulations were performed on simulated data assuming 25 transits were observed with <i>JWST</i> /NIRSpec Prism and the planet possesses a clear Earth-like atmosphere (first panel), a 10 bar clear CO ₂ atmosphere (second panel), a 10 bar desiccated O ₂ atmosphere (third panel), or a 10 bar cloudy CO ₂ atmosphere (fourth panel).	271
7.7	Two-dimensional marginalized posterior projections for the fit to the 25 transit data of the 10 bar dry O ₂ atmospheric composition using a model with (cyan) and without (purple) refraction included in the forward model. Refraction modifies the covariance between the reference pressure and other key parameters, particularly the solid-body planet radius, at higher pressures.	273
7.8	Marginalized posterior distributions for retrievals with (cyan) and without (purple) planet mass included in the state vector. Retrievals were performed on simulated data of TRAPPIST-1e with an Earth-like (clear-sky) atmosphere assuming 25 transits (top and bottom panels) and 50 transits (middle panel) were observed with <i>JWST</i> /NIRSpec Prism. Gaussian priors obtained from TTV measurements are used for the planet mass (shown in orange); the bottom panel uses a Gaussian prior with twice the standard deviation.	274
7.9	Two-dimensional marginalized prior (orange) and posterior (cyan) constraints on the planet mass and temperature from the 25 transit data of the Earth-like atmospheric composition. A positive correlation is evident between the mass and temperature.	275
7.10	Marginalized posterior distributions for retrievals with three different assumed VMR filling gases: N ₂ (purple), Ar (cyan), and H ₂ (pink). Retrievals were performed on simulated data of TRAPPIST-1e with an Earth-like (clear-sky) atmosphere assuming 25 transits were observed with <i>JWST</i> /NIRSpec Prism.	276

7.11	Two-dimensional marginalized prior (color points and histograms) and posterior (black points) samples/distributions for the VMR of the filling gas and the atmospheric mean molecular weight, assuming N ₂ (purple), Ar (cyan), and H ₂ (pink) as the filling gas. The central panel shows samples from the prior as color points and equally weighted samples from the posterior as black points, indicating an <i>a priori</i> preference for the mean molecular weight to take on values corresponding to the available end-member bulk atmospheric compositions. The top and right histograms show one-dimensional prior projections for the filling gas VMR and the mean molecular weight, respectively. The choice of filling gas imposes an implicit prior that propagates a substantial preference for the molecular weight of the atmosphere into posterior inferences.	278
7.12	Comparison between retrievals with and without Gaussian noise added to the spectrum. The top panel shows the synthetic <i>JWST</i> spectrum and the median retrieved spectrum with 1σ and 3σ uncertainty envelopes. The corner plot shows the 1D and 2D marginalized posteriors for the mean spectrum retrieval (purple) and the random noise instance spectrum retrieval (orange). The middle-right inset axis shows the detection significance of each molecule from an Bayesian evidence-based model comparison for each noise approach.	280
7.13	Corner plot showing the 1D and 2D marginalized posteriors for a simplified retrieval scenario simulated using both MCMC (purple) and AP (orange). For this low-dimensional case, AP derives statistically identical posteriors as MCMC using $\sim 1000\times$ fewer forward model evaluations.	282
7.14	Posterior comparison between retrievals using MULTINEST (purple) and APPROX-POSTERIOR (orange).	283

LIST OF TABLES

Table Number	Page
2.1 JWST instruments used in this study and their observability of TRAPPIST-1	45
2.2 Detectable molecules with transmission spectroscopy for different plausible TRAPPIST-1 planet atmospheres	69
2.3 Brightness temperatures in different Spitzer and JWST photometric filters for different TRAPPIST-1 planet atmospheric models.	87
4.1 Telescope, instrument, and system model parameters	125
5.1 Parameters for the fit to the number of exo-Earth spectra observable with LUVOIR (see Equation 5.1).	175
6.1 Parameters used in the fit to the VPL Earth Model spectrum shown in Figure 6.7	211
6.2 Retrieved molecule VMR scale factors from MAP fit to Earth’s low-resolution (LR) and high-resolution (HR) transmission spectrum	235
7.1 Retrieval parameters and their respective priors.	261

ACKNOWLEDGMENTS

The work presented in this dissertation represents the culmination of six years of graduate study at the University of Washington that would not have been possible without the support of so many individuals. I would like to start by thanking my graduate advisor, Victoria Meadows, for her academic, professional, and funding support throughout my graduate career. It has been an honor to learn from you and collaborate with you on so many fascinating projects. I also acknowledge the generous financial support of the NASA NExSS Virtual Planetary Laboratory, which has provided me with countless opportunities for inspiring conversations and collaborations over the years. I owe a debt of gratitude to my dissertation committee members, Eric Agol, David Crisp, Michael Line, Dale Winebrenner and Eric D’Asaro, for guiding me through the finish line.

There are countless close collaborators that I have sincerely appreciated working with and whose expertise I am grateful for. I would like to thank Giada Arney, Aki Roberge, and Shawn Domagal-Goldman for advising my contributions to the LUVOIR Mission Concept Study Team; Eddie Schwieterman for establishing the first generation Python interface for SMART, which lives on in the heart of my retrieval model, and for running the VPL Earth Model for my projects; Andrew Lincowski for countless collaborations throughout our concurrent graduate careers; Tyler Robinson for his generous expertise on SMART modeling; and Jonathan Fortney for originally inspiring my love of exoplanet science during my undergraduate education at UC Santa Cruz.

The trials of graduate school would have been too much to bear without the support of my friends and family. I deeply appreciate having had the opportunity to develop strong friendships through competitive foosball games with fellow UW Astronomy graduate students. I am grateful for the close friendships with David Fleming and Rodrigo Luger, whose support both in and out of the Astronomy Department enriched the entire experience. Science has never been as fun as when

collaborating with you both. I thank my parents, Marsha Sue Lustig and Bob Yaeger; my sister, Rebekah; and all of our close family and friends for fostering in me a love of science from a young age. Ultimately, none of my graduate studies would have been possible without the loving support of my wife and best friend, Emily Less. Em, thank you for comforting me in challenging times and for challenging me in comfortable times.

The perspectives, approaches, and results presented here all emanate from collaborative efforts. Throughout this dissertation, I will use the royal “we” to acknowledge these collaborations and to acknowledge the continued collaborative journey between *me*, the author, and *you*, the reader.

DEDICATION

To my wife, Em. Not all the stars in the sky, nor their encircling planets, can bring such a richness to this universe, as you do.

Chapter 1

INTRODUCTION

1.1 At the Dawn of the Terrestrial Exoplanet Characterization Era

The era of terrestrial exoplanet characterization is upon us. Since our humble beginnings as a species, humanity has looked to the sky, pondering our origins and our place within the cosmos. A century of modern astronomy has provided the lens through which we view our own context as “star stuff” within an ever expanding universe. Less than three decades of extrasolar planet studies have transformed the way we view the stars, as not just atomic factories that fuse the elements that make up our planet and bodies, but as hosts to previously unseen planets and their mysteries. We now routinely detect Earth-sized exoplanets, and with the next-generation of ground and space based telescopes planned, we will have the opportunity to characterize their atmospheres and to place the ubiquity or rarity of the Solar system, Earth, and life within the broader context of the universe.

However, despite this dawning of terrestrial exoplanet characterization, many significant questions remain at the intersection between the astronomical observations that telescopes must make and the terrestrial exoplanet models that are required to predict and interpret these observations. In this work, we build a foundation to optimally characterize terrestrial exoplanets in the near-term with the James Webb Space Telescope (JWST) and in the far-term with next-generation telescopes by linking key observing capabilities and telescope requirements to the objectives of detecting terrestrial exoplanet atmospheres, characterizing their composition, assessing their habitability, and searching for biosignatures. Furthermore, we develop, validate, and apply a rigorous new terrestrial exoplanet atmospheric retrieval model that is capable of interpreting the observations of both JWST and next-generation telescopes. This work lays a path into the era of terrestrial exoplanet characterization by establishing specific observations that have a high likelihood of success, as well as taking a crucial first step towards an astrobiological framework for the assessment of exoplanet habitability and biosignatures.

1.2 Terrestrial Exoplanet Atmospheres

Terrestrial atmospheres are the thin volume of gas spanning between the solid or liquid planetary surface and the cold unforgiving vacuum of space. The Solar System terrestrial exoplanets—Venus, Earth, and Mars—provide a touchstone for the field of comparative planetology, habitability, and astrobiology. Earth’s atmosphere bears evidence of the dominant metabolisms at work in our planet’s biosphere, exemplified by the oxygen in the air that we breathe. But Venus’s atmosphere tells a different tale of a world that has lost whatever global water inventory it once had and is now our textbook example of the outcome of the runaway greenhouse process, an uninhabitable planet (Arney and Kane, 2018; Kane et al., 2018), and a warning sign of a possible future for our home planet. Shifting our focus beyond the Solar System, the expected diversity of terrestrial *exoplanet* atmospheres is poised to revolutionize our understanding of planetary climate, chemistry, dynamics, and habitability. Furthermore, terrestrial exoplanet atmospheres are the interface between any potential extraterrestrial life that may exist on an exoplanet’s surface and the observation of that life at interstellar distances, and therefore studying these atmospheres is a key step in the search for life beyond the solar system.

1.2.1 Formation & Evolution

Unlike gas giant exoplanets, terrestrial planets are more likely to lose their primordial atmospheres, and develop secondary atmospheres that reflect the diverse planetary processes that generate and sculpt them over time.

The composition of the solid component of a terrestrial exoplanet is thought to be determined by the composition of, and location within, the protoplanetary disk in which it formed (Bond et al., 2010; Wang et al., 2018). In addition to the initial endowment of volatiles locked within the planet’s interior, volatiles may be delivered via planetary embryos and planetesimals during late stage planet formation (Raymond et al., 2004, 2007). The rocky planet then outgasses volatiles from the interior, forming a secondary atmosphere which can reflect the planetary interior composition (Schaefer and Fegley Jr, 2010). Over time, interactions between a terrestrial planet atmosphere

and its host star can shape, and even erode, the planetary atmosphere (Dong et al., 2018).

Photochemical reactions between stellar UV photons and the upper atmosphere of the planet drive the destruction and formation of molecules. For two exoplanets whose atmospheric compositions would otherwise be identical, differences in their host star’s UV spectral energy distribution can produce significantly different atmospheric compositions, which modify the observable spectra of the planets (Segura et al., 2003, 2005; Domagal-Goldman et al., 2011). For example, in the atmosphere of a modern Earth-like planet orbiting the M dwarf star Proxima Centauri (e.g. Proxima Centauri b; Anglada-Escudé et al., 2016), methane has a longer photochemical lifetime due to the star’s lower near UV emission compared to far UV, which suppresses the generation of CH₄-destroying OH radicals from water vapor, and may allow methane to persist at a higher abundance than seen in Earth’s atmosphere (Segura et al., 2005; Meadows et al., 2018a). On the other hand, O₂ photolysis would be decreased due to lower XUV emission from the M dwarf, which would in turn suppress O₃ production. These photochemical processes are crucial for predicting the type of atmospheres we expect around different types of stars.

In addition to driving photochemical reactions, stellar X-ray and ultraviolet (UV) radiation (together referred to as XUV radiation) and particle fluxes can drive atmospheric escape processes that can strongly influence the volatile content of terrestrial planets, and in some cases lead to complete atmospheric loss (Tian, 2015; Lehmer and Catling, 2017). Dong et al. (2018) found that the inner TRAPPIST-1 planets have endured sufficiently high atmospheric escape rates to call into question whether they possess atmospheres at present. However, even substantial atmospheric escape rates can be offset by vigorous outgassing from the interior (Garcia-Sage et al., 2017), which complicates the ability to predict the atmospheric extent of terrestrial exoplanets *a priori*, and strongly motivates continued and expanded observational reconnaissance.

Stellar evolution can also drive ocean loss and significant planetary atmosphere evolution, especially for planets orbiting M dwarfs. Luger and Barnes (2015) modeled the effects of XUV-driven water photolysis during the extended super-luminous pre-main-sequence phase of M dwarf stars. They found that exoplanets that may currently reside within the habitable zone faced a much harsher irradiation environment in the past as their low-mass parent star contracted on to

the main-sequence, possibly taking as long as one billion years to do so. During this time, XUV radiation would have efficiently split water molecules in the planetary atmosphere, enabling the light hydrogen molecules to escape to space, while also causing the accumulation of oxygen in the atmosphere. Models suggest that this mechanism may produce tens to thousands of bars of oxygen (Luger and Barnes, 2015; Bolmont et al., 2017; Lincowski et al., 2018; Meadows et al., 2018a; Wordsworth et al., 2018), producing a false positive biosignature (Schwieterman et al., 2016b). However, the atmospheric oxygen may sufficiently oxidize the surface and mantle such that final inventory of oxygen in the atmosphere is reduced (Schaefer et al., 2016; Wordsworth et al., 2018). The planetary outcomes of stellar evolution, and the balance between these atmospheric losses and planetary sinks are not fully understood, firmly underscoring the need for observations that can effectively test these theoretical predictions.

1.2.2 Planetary Habitability & The Habitable Zone

Terrestrial exoplanets that can form with water and do not lose their initial water inventory, or entire atmosphere to space, may possess climates and surface temperatures amenable to the presence of liquid water on their surfaces. We will define such a planet with liquid water on its surface as a *habitable* planet. However, we acknowledge that a planet that does not fit this definition of habitable, may maintain a habitable region on it, such as liquid water under an icy surface. We make this distinction for detectability reasons: a planet with liquid surface water will naturally have an ocean-atmosphere exchange of gases that may aid in the remote characterization of the planet.

While planetary habitability is an intrinsic characteristic of a planet, the concept of the *habitable zone* (HZ) is a useful approach for assessing whether a planet may be *potentially* habitable on the basis of the incident stellar flux received by the planet. The HZ is defined as the region around a star where an Earth-like planet (similar to Earth in mass, radius, oceans, atmospheric composition and active carbonate-silicate weathering cycle) could maintain liquid water on its surface (Kasting et al., 1993; Underwood et al., 2003; Kopparapu et al., 2013). The HZ is traditionally bounded at the inner edge by the runaway greenhouse limit, where the oceans evaporate increasing the atmo-

spheric opacity to outgoing thermal radiation and causing runaway heating. The outer edge of the HZ is defined by the maximum greenhouse limit, where the greenhouse effect caused by adding more CO₂ to the atmosphere becomes outweighed by Rayleigh scattering by CO₂, causing runaway cooling (Kopparapu et al., 2013). Whether or not a potentially habitable planet discovered within the HZ of its host star is actually habitable depends on numerous stellar, planetary, and planetary system factors that can act in concert to preclude, maintain, or destroy habitable climate conditions (Meadows and Barnes, 2018). As a result, the HZ is a necessarily simplified approach to increase our chances of discovering habitable planets, and targeting HZ planets are a best-we-can-do-right-now approach to increase our chances of finding life. Indeed, habitable planets may exist outside the HZ and biospheres may exist beyond habitable planets, but until we can use empirical results to refine this Earth-centric framing, we must formulate testable hypotheses with falsifiable predictions from a standpoint of scientific knowledge.

1.2.3 The Search for Life & Biosignatures

Terrestrial exoplanets offer one of the best opportunities to search for life beyond Earth and to answer the culturally significant question “Are we alone?”. This astrobiology cornerstone centers around the search for observational evidence of life’s impact on its environment, or biosignatures. A biosignature can be an object, substance, or pattern that is produced by life, and which accumulates to detectable levels (Marais and Walter, 1999; Des Marais et al., 2008; Seager et al., 2012; Schwieterman et al., 2018). Biosignatures have classically been proposed in three categories: (1) gaseous biosignatures evident in the spectrum of an atmosphere, (2) surface biosignatures evident in the reflected spectrum of a planetary surface, and (3) temporal biosignatures evident in seasonal changes to a planetary atmosphere or surface (Schwieterman et al., 2018). In all categories, a good biosignature to search for is one that is *reliable* and is unlikely to be mistaken as a common planetary process; is *survivable* and avoids normal sinks in a planetary environment; and is *detectable* via accumulation to detectable levels for the planned observing techniques (Meadows, 2017; Meadows et al., 2018b).

In addition to targeted biosignatures, attempts have been made to develop *agnostic* biosigna-

tures that can be recognized by a pattern in the environment without assuming detailed knowledge about how the life operates. One approach is to determine which gases stand out in an atmosphere on the basis of disequilibrium chemistry (Hitchcock and Lovelock, 1967; Sagan et al., 1993; Krissansen-Totton et al., 2016; Krissansen-Totton et al., 2018a). Another promising new agnostic approach to biosignatures is to search for complex chemical reaction network topologies that can indicate biological influence (Sole and Munteanu, 2004; Walker et al., 2018). There is a growing consensus for the need to understand as much as possible about the planetary environment in which a biosignature may be found, both to rule out false positives, in the case of targeted biosignatures searches, and to find the biogenic anomaly, in the case of agnostic searches (Meadows et al., 2018b; Catling et al., 2018). Thus, being able to interpret a biosignatures in the context of its environment requires the characterization of terrestrial exoplanet atmospheres. The first step in this long journey begins with the detection of Earth-sized exoplanets.

1.3 Detecting Terrestrial Exoplanets

Detecting planets outside of our Solar System is an extreme astronomical challenge that has only become possible, and eventually commonplace, over the last three decades. This difficulty stems from the fact that planets are much smaller in mass and radius, and significantly fainter than their stellar hosts. In combination, these factors make exoplanet detection novel to modern astronomy, and they make detecting and studying small, rocky, *terrestrial* exoplanets the hardest challenge yet.

The first planets to be discovered outside the Solar System were found orbiting pulsars (Wolszczan and Frail, 1992; Thorsett et al., 1993), but the first exoplanet discovered around a main-sequence, Sun-like star was the Jupiter-mass planet, 51 Pegasi b, which orbits very close to its star (Mayor and Queloz, 1995). This single discovery of a “hot Jupiter” exoplanet immediately upended the picture of planet formation that suggested that all planetary systems may form like our own (Cenadelli and Bernagozzi, 2018), beginning what has become a revolution in planetary science.

Nearly thirty years after the first exoplanet detection, the exoplanet field has quickly matured into a major cornerstone of modern astronomy and astrophysics research. There are now over

4,100 confirmed exoplanets¹, with thousands more awaiting confirmation. From the high occurrence of exoplanets around main-sequence stars, we now know that planets naturally form in the wake of star formation, coalescing out of the materials that do not become incorporated into the stars (Batalha, 2014; Lissauer et al., 2014). Not only are exoplanets common, but small exoplanets are particularly common (Dressing and Charbonneau, 2013; Batalha, 2014; Dressing and Charbonneau, 2015), and “super-Earth” and “sub-Neptune” exoplanets—for which our Solar System has no direct analog—may be the most common planet type of all (Howard et al., 2010). An emerging paucity of planets with radii near $\sim 1.6 R_{\oplus}$ suggests an empirical transition between gaseous and rocky exoplanets (Weiss and Marcy, 2014; Rogers, 2015; Fulton et al., 2017). These numerous discoveries have been the result of a variety of different techniques that are used to detect exoplanets, which broadly fit into two categories—*direct detection* and *indirect detection*. Each one of these detection techniques is more difficult for small, terrestrial planets than for larger gas giants.

1.3.1 Direct Detection

The direct detection of an exoplanet is conceptually straightforward, but quite difficult in practice. Light emitted or reflected from an exoplanet must be directly imaged. All of the planets in the Solar System have been detected via direct means². However, the faint light of even the largest exoplanets exists within the overwhelming glare of their parent stars. To overcome this stellar glare, astronomers must use clever techniques to suppress the starlight and reveal the exoplanetary system hidden beneath. This is typically done using a *coronagraph* instrument integrated into a telescope’s optics, which can suppress starlight. In addition to coronagraphs, designs are being considered for a *starshade* (Seager et al., 2015; HabEx Study Team, 2019), which is an external occulter that would fly in front of a space-telescope to physically block a star’s light prior to it

¹As of April 5, 2020. This has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program

²With the potential exception of “Planet Nine”, whose possible existence is the subject of current research and debate as a result of *indirect* evidence that a planet may be responsible for unexplained gravitational influences in the outer Solar System (see Batygin et al., 2019, for a recent review).

entering the telescope optics. Optical interferometry can also directly image and detect exoplanets using destructive interference to null stellar light; this method was used for the first time to detect an exoplanet just last year ([Gravity Collaboration et al., 2019](#)). To date, only 50 exoplanets (all gas giants) have been discovered using direct imaging, but as we will discuss throughout this work, it represents one of the most promising avenues for future terrestrial exoplanet atmospheric studies. The vast majority of exoplanet detections have been made using indirect methods.

1.3.2 Indirect Detection

Due to the challenges associated with direct exoplanet detection, most exoplanets have been detected by indirect methods, where the planet's existence must be inferred from its impact on the observable light from the star. The most common techniques for indirect exoplanet detection are the transit, radial velocity, gravitational microlensing, and astrometry methods.

Transits The transit of an exoplanet across the disk of its host star periodically blocks a small amount of light, enabling the existence of the planet to be inferred. The amount of light blocked by a transiting planet, ΔF , is proportional to the size of the planet,

$$\Delta F = \left(\frac{R_p}{R_s} \right)^2 \quad (1.1)$$

$$\approx 84 \text{ ppm} \left(\frac{R_p}{1 R_\oplus} \right)^2 \left(\frac{1 R_\odot}{R_s} \right)^2, \quad (1.2)$$

where R_p and R_s are the radii of the planet and star, respectively ([Winn, 2010](#)). Observing the same planet transit multiple times allows the orbital period to be measured, and using Kepler's third law and estimates of stellar mass allows the semi-major axis of the planet to be calculated. Detecting transiting exoplanets requires high precision time-series stellar photometric monitoring. The dedicated exoplanet hunting mission *Kepler* ([Borucki et al., 2003](#)) and now *TESS* ([Ricker et al., 2010](#)) have been extremely successful at detecting thousands of exoplanets. However, observing an exoplanet transit requires a special alignment between the orbital plane of the planet and the Earth,

such that only a small fraction of total exoplanetary systems are seen to transit. Since planetary formation does not appear to favor any particular inclination relative to other stars, the probability of a random exoplanet having observable transits is

$$\mathcal{P}_{\text{transit}} = \frac{R_s}{a} \quad (1.3)$$

$$\approx 0.5\% \left(\frac{a}{1 \text{ AU}} \right) \left(\frac{1 R_\odot}{R_s} \right) \quad (1.4)$$

$$\approx 2\% \left(\frac{a}{0.029 \text{ AU}} \right) \left(\frac{0.121 R_\odot}{R_s} \right) \quad (1.5)$$

where a is the planet's semi-major axis. This implies that (1) most exoplanets do not transit, (2) planets that are closer to their star have a higher probability of transiting—a detection bias that must be taken into account, and (3) there is a higher chance of a HZ planet transiting a late M dwarf (2% chance for TRAPPIST-1e) than a HZ planet transiting a Sun-like G dwarf (0.5% chance).

Radial Velocity The radial velocity (RV) method of exoplanet detection leverages the fact that a planet does not simply orbit a star, but both bodies orbit their mutual center of gravity. As a result, an orbiting planet will induce a reflex orbital motion in the star. This reflex motion can be observed in the high resolution spectrum of the star as it causes the absorption and emission lines to be Doppler shifted in tune with the orbital period of the exoplanet. The strength of the Doppler shift is dependent upon the velocity of the star along the radial line of sight, and is characterized by the radial velocity semi-amplitude,

$$K = \left(\frac{2\pi G}{P_{\text{orb}}} \right)^{1/3} \frac{M_p \sin i}{(M_p + M_s)^{2/3}} \frac{1}{\sqrt{1 - e^2}}, \quad (1.6)$$

where P_{orb} is the planet's orbital period, M_p and M_s are the respective planet and star masses, i is the planet's inclination, e is the planet's orbital eccentricity, and G is the gravitational constant (Perryman, 2011). Like the transit technique, the radial velocity signal is largest when planetary systems are aligned near “edge-on” so that the star's motion has a large component in the direction of the observer. However, Equation 1.6 contains a degeneracy between planet mass and inclination:

the semi-amplitude can remain fixed by making the planet more massive while inclining the orbit so that less of the stellar reflex motion is in the radial velocity direction. As a result, RV planet detections report the *minimum* mass of the planet by assuming that all of the star’s motion is along the line of sight (unless the inclination can be constrained by observing a transit). The RV method is biased towards detecting massive planets in short orbital periods, and was used to detect the first (traditional) exoplanet, 51 Pegasi b (Mayor and Queloz, 1995). Today, ~ 800 exoplanets have been detected by RV, using, in particular, the HARPS (La Silla Observatory; Pepe et al., 2000; Mayor et al., 2003), HARPS-N (Roque de los Muchachos Observatory; Cosentino et al., 2012), and HIRES (W. M. Keck Observatory; Vogt et al., 1994) instruments. Recently, RV instruments on ground-based telescopes have reached the necessary precision to detect terrestrial exoplanets orbiting M dwarfs—most remarkably, the planet Proxima Centauri b found in the HZ of the nearest star to the Sun (Anglada-Escudé et al., 2016).

Gravitational Microlensing The gravitational microlensing technique makes use of general relativistic distortions in space-time by massive objects to detect exoplanets (Gould and Loeb, 1992). This technique requires a temporary alignment of two stars as seen from observers on Earth. A “lens star” becomes sufficiently aligned with a more distant background “source star” causing the light from the source star to be gravitationally magnified. If the lens star happens to host an exoplanet, then the lightcurve of the magnification event can show a smaller, secondary magnification peak caused by additional lensing from the planet. Gravitational microlensing provides a measurement of the exoplanet’s mass as a fraction of the lens star’s mass and an estimation of its projected separation from the star at the moment of the event. The Wide Field Infrared Survey Telescope (WFIRST; Spergel et al., 2015), scheduled for launch in 2025, will conduct a gravitational microlensing survey pointed towards the galactic center and is predicted to detect ~ 1400 new exoplanets with ~ 200 having masses $\leq 3M_{\oplus}$ (Penny et al., 2019). This survey complements other planet detection surveys by probing the population of long period exoplanets. However, these chance microlensing events do not occur multiple times for any given lensed extrasolar system, and therefore this population of exoplanets does not offer any opportunities for follow up atmospheric

characterization efforts.

Astrometry Like RV, the astrometric method of exoplanet detection leverages the stellar reflex motion induced by an orbiting planet to infer that exoplanet’s existence. In this case, the motion of the star is observed directly in stellar photometry as it revolves around the star-planet center of mass on the sky plane. Observing this effect requires extremely precise measurements of stellar positions in the sky on orbital timescales. The angular semi-amplitude of the astrometric signal is given by

$$\theta_{\star} = \left(\frac{M_p}{M_s} \right) \left(\frac{a}{d} \right) \quad (1.7)$$

$$\approx 0.3 \mu\text{as} \left(\frac{M_p}{1 M_{\oplus}} \right) \left(\frac{1 M_{\odot}}{M_s} \right) \left(\frac{a}{1 \text{ AU}} \right) \left(\frac{10 \text{ pc}}{d} \right) \quad (1.8)$$

where M_p and M_s are the respective planet and star masses, a is the planet semi-major axis, and d is the distance to the exoplanetary system (Shao et al., 2009). For an Earth-Sun analog system observed at 10 pc the astrometric signal is just three tenths of a micro-arcsecond. The dedicated astrometry mission *Gaia* (Prusti et al., 2016) has a precision of about $10 \mu\text{as}$ for bright stars, insufficient to detect Earth-analogs with astrometric signals of just a fraction of a μas . However, *Gaia* is still expected to detect thousands to tens of thousands of larger exoplanets (Robichon, 2002), with recent detection estimates at roughly 21,000 Jupiter-mass long-period planets in the nominal 5-year mission, rising to about 71,000 in the extended 10-year mission (Perryman et al., 2014). Because astrometry preferentially detects exoplanets at long orbital periods as well as those near 0° inclination (i.e. in “face-on” orbits), it is an ideal method to find targets for follow-up atmospheric characterization using direct-imaging. Unfortunately, due to the precision required to detect Earth-like exoplanets, these characterization efforts will be limited to extrasolar gas giants until future telescopes can focus on terrestrial exoplanets.

In summary, while microlensing and astrometry will provide statistical information and more massive targets for direct imaging follow-up, in the near-term, the vast majority of terrestrial exoplanet candidates available for follow-up will be discovered by transit and radial velocity surveys.

Due to the nature of these detection methods, these terrestrial planets are highly likely to reside close to their stars, and to orbit M dwarfs.

1.3.3 Targets for Characterization

Despite the challenges associated with detecting terrestrial exoplanets, transit and RV surveys have detected a handful of extremely promising terrestrial exoplanets that will be the subject of further observations to characterize their atmospheres. In particular, the closest star to the Sun hosts the possibly terrestrial planet, Proxima Centauri b, in its habitable zone (HZ) with an 11.2 day orbit (Anglada-Escudé et al., 2016). The TRAPPIST-1 system hosts at least seven transiting Earth-size planets, with three residing in the HZ (e, f, g), three interior to the HZ (b, c, d), and one exterior to the HZ (h) (Gillon et al., 2016, 2017b; Luger et al., 2017c). The HZ terrestrial exoplanet LHS1140 b is another example of a rocky exoplanet orbiting an ultra cool star (Dittmann et al., 2017). Finally, the highly irradiated rocky planets GJ1132 b (Berta-Thompson et al., 2015) and LHS3844 b (Vanderspek et al., 2019) are examples of hot, rocky planets that may or may not possess atmospheres.

Due to the relative ease of detecting small planets orbiting small stars, all of the above planets orbit small, cool, M dwarf stars. However, M dwarfs are very different from Sun-like G dwarfs, and are expected to alter the atmospheres of their terrestrial planets in complex ways that may prove difficult to interpret. All things being equal, M dwarfs would likely not be the top choice in stellar host for the first rocky exoplanets to study in detail, but alas, they are the only option for the near-term characterization of terrestrial exoplanet atmospheres. So we will gladly indulge in the unfamiliar.

1.4 Characterizing Terrestrial Exoplanets

“Characterization is the sum of all the observable qualities, a combination that makes the character unique: physical appearance coupled with mannerisms, style of speech and gestures, sexuality, age, IQ, occupation, personality, attitudes, values, where he

lives, how he lives.” – McKee (2016), “Story”

Like Robert McKee’s description of human “characterization,” exoplanet characterization is the sum of all the observable qualities, a combination that makes the planet’s character unique: physical appearance coupled with the context of the stellar environment, composition of atmosphere and thermal structure, formation scenario, evolutionary history, surface pressure, temperature, and composition, status of habitability, and presence of biosignature gases. However, after an exoplanet is detected, much about the planet’s character remains a mystery, and additional dedicated telescope observations must be obtained to characterize the nature of the planetary environment. There are numerous observational techniques that astronomers can use to gain sensitivity to exoplanetary atmospheres and surfaces, and thereby characterize their nature.

1.4.1 *Observing Techniques*

Figure 1.1 shows the diverse arsenal of techniques that may be used to characterize terrestrial exoplanets and their atmospheres. We will briefly review the most relevant observational techniques for exoplanet characterization—including transit and eclipse spectroscopy, phase curves, and direct-imaging photometry and spectroscopy—and discuss their prospects for studying terrestrial exoplanet environments.

Transit Transmission Spectroscopy Observing the transit of an exoplanet at multiple wavelengths enables the atmospheric composition to be sensed using transmission spectroscopy (Seager and Sasselov, 2000; Charbonneau et al., 2002). As a planet transits across the face of a star, it blocks light from the star proportional to the planet’s geometric area. However, the radius of the planet with an atmosphere is a function of wavelength due to extinction of stellar light as it filters through the atmosphere and interacts via scattering and absorption with the gases and aerosols that compose the atmosphere. As a result, an exoplanet transmission spectrum may be constructed by measuring the depth of a transit lightcurve (e.g. Equation 1.1) at many wavelengths (Seager and Sasselov, 2000), leading to potentially observable increases in the planet’s absorbing radius at

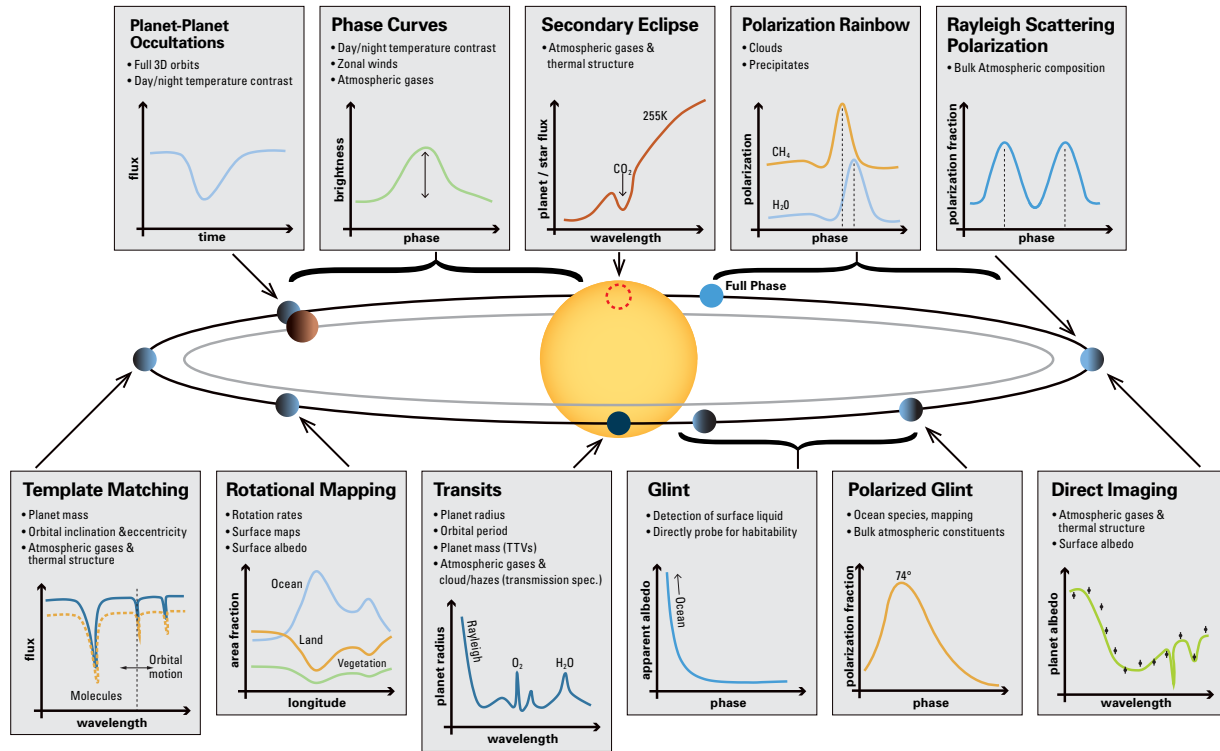


Figure 1.1 Graphic depicting the many methods that exoplanet astronomers can use to characterize exoplanet environments, and the orbital phase at which each method is usually applied. Methods like transits, secondary eclipses, and planet-planet occultations are only viable for transiting exoplanet systems, while other methods, such as phase curves, direct imaging reflectance spectroscopy, and rotational mapping, may be used for transiting and non-transiting exoplanet architectures. This figure was developed by J. Lustig-Yaeger for the VPL CAN-8 proposal, and made beautiful by Ron.

wavelengths where the atmosphere has absorption bands. However, a planet that does not possess an atmosphere would have a featureless, flat line transmission spectrum.

The strength of a molecular absorption feature in a transmission spectrum depends on the scale height of the atmosphere and the optical depth along the slant path traced by the stellar photons through the planet's atmosphere (Fortney, 2005). Absorption features can be strongly muted by the presence of clouds and hazes, as we have seen for hot Jupiters (Sing et al., 2016), but also by Rayleigh scattering (Lecavelier Des Etangs et al., 2008), and a critical refraction limit (Misra et al., 2014; Bétrémieux and Kaltenegger, 2014), which act to raise the continuum atmospheric absorbing

radius. The signal strength for transmission spectroscopy is often expressed simply as

$$\Delta dF_T = \frac{2R_p N_H}{R_s^2} \left(\frac{kT}{\mu g} \right) \quad (1.9)$$

(Winn, 2010), where the term in parenthesis is the pressure scale height of the atmosphere (H) and N_H is the number of scale heights above the continuum that a molecular absorption feature extends. The N_H parameter is the only factor that encodes information about the wavelength-dependent absorption and the vertical extent of an atmospheric gas, and usually varies between 0-10 to truncate features due to the presence of clouds and hazes. For reference, the vertical extent of Earth's $9.6 \mu\text{m}$ O_3 and $15 \mu\text{m}$ CO_2 absorption features are $N_H = 2.6$ and $N_H = 6$, respectively, for an isothermal scale height of 8 km, indicating the relative strength of the CO_2 band compared to the O_3 band. These features roughly correspond to absorption features in a Earth-Sun transmission spectrum of ~ 1 ppm. In practice, and throughout the rest of this work, we will be more rigorous than Equation 1.9, and explicitly simulate the spectra of planetary atmospheres.

Over the past 15 years transit transmission spectroscopy has emerged as a key tool used to detect atoms and molecules in the atmospheres of giant exoplanets (Charbonneau et al., 2002; Redfield et al., 2008). Pioneering observations with the *Hubble Space Telescope's* (*HST*) Wide Field Camera 3 (WFC3), which operates between $1.1\text{--}1.7 \mu\text{m}$, have achieved unprecedentedly precise transmission spectra that have been used to detect H_2O in gaseous exoplanet atmospheres, but with anomalously low abundances compared to solar composition (e.g. Mandell et al., 2013; Wakeford et al., 2013; Deming et al., 2013; Kreidberg et al., 2015). Instead of a rich catalogue of exoplanet atmospheric gases, *HST* transmission spectroscopy has revealed predominantly featureless or near-featureless spectra that hint at the presence of high altitude clouds and/or hazes acting to obscure the column of absorbing gases below for these gas-dominated worlds (Berta et al., 2012; Ehrenreich et al., 2014; Knutson et al., 2014; Kreidberg et al., 2014; Nikolov et al., 2015). Modeling efforts have reproduced the observed flat transmission spectra by considering chemical equilibrium cloud condensation and photochemical production of hydrocarbon hazes (Morley et al., 2013; Charnay et al., 2015a,b). The obstruction of lower exoplanet atmospheres by clouds and hazes remains a

key concern for current and future transmission spectroscopy observations, particularly upcoming terrestrial exoplanet observations with *JWST*, which will be explored in Chapters 2 and 3 of this work.

Secondary Eclipse Spectroscopy As a transiting planet passes behind the host star in a secondary eclipse, astronomers are able to detect and quantify the missing flux emitted or reflected by the planet via comparison with the combined flux of the star and planet prior to the eclipse. Obtaining a planetary spectrum during secondary eclipse was first proposed by Richardson et al. (2003), but was not unambiguously detected until Deming et al. (2007) observed the secondary eclipse of GJ1214b with $8\mu\text{m}$ *Spitzer* photometry, followed by the secondary eclipse spectrum of HD189733b observed by Grillmair et al. (2008). Since then exoplanet brightness temperatures and the presence of absorbing molecules, such as H_2O , CH_4 , CO and CO_2 , have been reported in the literature (e.g. Charbonneau et al., 2008; Christiansen et al., 2010; Fressin et al., 2010; Line et al., 2014), demonstrating a key ability for secondary eclipse to probe deeper into atmospheres than transmission spectroscopy, and be less affected by clouds.

At visible and near-IR wavelengths a secondary eclipse spectrum reveals the planet in reflected light, while at mid- to far-IR wavelengths it reveals the thermal emission spectrum. However, due to the wavelength-dependence of the stellar spectral energy distribution (SED), secondary eclipses are much more favorable to detect in thermal emission. The strength of molecular absorption in a secondary eclipse emission spectrum is controlled by the radiative transfer of thermal photons emitted from the planet's atmosphere. Since an isothermal gas emits a featureless blackbody spectrum, the strength of an absorption feature in an emission spectrum is controlled by the thermal structure of the planetary atmosphere. Typically, the highest flux is emitted from deep, hot layers in the atmosphere and is absorbed by greenhouse gases in higher, cooler layers as it radiates into space. In this way, secondary eclipse observations are sensitive to the thermal structure of planetary atmospheres.

Phase Curves The orbit of exoplanets around their host stars can produce observable thermal and reflected signals as different planetary phases are periodically made visible to distant observers. Phase-dependent lightcurves (i.e. phase curves) can reveal temporal and spatial information about exoplanets (Cowan et al., 2015), which may be used to measure the day-night temperature contrast for tidally locked and synchronously rotating close-in exoplanets (Cowan et al., 2007), and can even be inverted to reveal longitudinal maps of exoplanets (Knutson et al., 2007; Cowan and Agol, 2008). Phase-resolved emission spectroscopy can allow the thermal structure of the atmosphere to be retrieved and mapped as a function of planetary longitude (Stevenson et al., 2014), but require high signal-to-noise data.

The first exoplanet phase curves were observations of the hot Jupiters, *v* Andromedae b by Harrington et al. (2006) and HD189733b by Knutson et al. (2007), using the *Spitzer Space Telescope*. However, more recently, smaller Earth-sized exoplanets have been discovered that are amenable to phase curve characterization, such as the *TESS* discovered exoplanet, LHS 3844b (Vanderspek et al., 2019), which Kreidberg et al. (2019) observed with *Spitzer* and argue must not have a thick atmosphere on the basis of the large detected thermal phase curve amplitude with a hot spot centered at the sub-stellar point. Such a large and longitudinally symmetric difference in emitted flux between the day and night sides of the planet is not expected for a thick atmosphere whose strong equatorial jets would efficiently recirculate heat from the hot dayside to the cooler nightside, thus decreasing the phase curve amplitude and offsetting the hot spot from the sub-stellar point, as has been observed for hot gaseous exoplanets (e.g. Knutson et al., 2014; Demory et al., 2016; Dang et al., 2018). Due to its mid-IR sensitivity, *JWST* promises to significantly advance our understanding of the atmospheric dynamics and thermal structures of highly irradiated gaseous exoplanets (Cowan et al., 2015) and provide a key diagnostic tool for interpreting airless rocky worlds (Maurin et al., 2012; Meadows et al., 2018a).

Direct-Imaging Spectroscopy & Photometry The direct observation of photons emitted from or reflected off an exoplanet is one of the most promising, yet challenging, methods to characterize their atmospheres. Like the direct detection of exoplanets outlined in Section 1.3, characterizing an

exoplanet using direct imaging requires the suppression of the host star's light before a spectrum of the much fainter exoplanet can be obtained. Currently, only a coronagraph can perform such an observation. The planet-to-star flux ratio that a coronagraph must be able to achieve can be expressed as

$$\frac{F_p}{F_s} = \frac{A\Phi(\alpha)F_{\text{TOA}}}{\sigma T_{\text{eff}}^4} \left(\frac{R_p}{R_s}\right)^2 \quad (1.10)$$

where F_{TOA} is the incident flux from the star at the planet's top of atmosphere (TOA), $\Phi(\alpha)$ is the planet's Lambertian phase function, T_{eff} is the stellar effective temperature, σ is the Stefan-Boltzmann constant, and we have assumed (for now) that the planet has a wavelength *independent* albedo, A (Robinson et al., 2016). For an Earth-Sun analog,

$$\frac{F_p}{F_s} \approx 2.2 \times 10^{-10} \left(\frac{A}{0.25}\right) \left(\frac{\Phi(\alpha)}{0.5}\right) \left(\frac{F_{\text{TOA}}}{1370 \text{ W/m}^2}\right) \left(\frac{5778 \text{ K}}{T_{\text{eff}}}\right)^4 \left(\frac{R_p}{1 R_{\oplus}}\right)^2 \left(\frac{1 R_{\odot}}{R_s}\right)^2, \quad (1.11)$$

and thus, to observe an Earth-like exoplanet, a coronagraph must be able to suppress better than two in ten billion photons from the star. However, for an Earth-analog orbiting a late M dwarf like TRAPPIST-1 (Gillon et al., 2016, 2017b), the planet-to-star flux contrast is more favorable,

$$\frac{F_p}{F_s} \approx 4.4 \times 10^{-7} \left(\frac{A}{0.25}\right) \left(\frac{\Phi(\alpha)}{0.5}\right) \left(\frac{F_{\text{TOA}}}{1370 \text{ W/m}^2}\right) \left(\frac{2511 \text{ K}}{T_{\text{eff}}}\right)^4 \left(\frac{R_p}{1 R_{\oplus}}\right)^2 \left(\frac{0.121 R_{\odot}}{R_s}\right)^2, \quad (1.12)$$

due to the lower luminosity and smaller radius of the star, so a coronagraph can afford nearly three orders of magnitude less starlight suppression. However, direct-imaging with a coronagraph must suppress photons from the star without suppressing photons from the planet. The inner working angle (IWA) of a coronagraph is defined as the minimum angular distance from the star at which a directly imaged planet can be observed without facing significant light suppression from the coronagraph. The IWA is wavelength-dependent and is typically given as a factor (N) times the

diffraction limit (λ/D),

$$\theta_{\text{IWA}} = N \frac{\lambda}{D}, \quad (1.13)$$

which must be less than the angular separation between planet and star (θ_{p-s}) for the planet to be observable, $\theta_{\text{IWA}} < \theta_{p-s}$. For an Earth-analog exoplanet orbiting a Sun-like star at 10 pc, the coronagraph IWA must be less than,

$$\theta_{p-s} \approx 100 \text{ mas} \left(\frac{a}{1 \text{ AU}} \right) \left(\frac{10 \text{ pc}}{d} \right), \quad (1.14)$$

but for an Earth-like exoplanet orbiting a cool M dwarf where the HZ is much closer to the star, the coronagraph IWA must be more tightly constrained. For instance, observing Proxima Centauri b (at $a = 0.0485$ AU) if it were at 10 pc from Earth, would require $\theta_{\text{IWA}} < \theta_{p-s} \approx 4.85$ mas—a nearly 100 times stricter constraint on the IWA compared to an Earth-Sun analog (Meadows et al., 2018a). Note that at 1.3 pc, a coronagraph need only archive $\theta_{\text{IWA}} < 37$ mas to directly image Proxima Centauri b. Thus, planet-star flux contrast would indicate that faint M dwarfs are the best stellar hosts for directly imaging HZ Earth-like exoplanets, while planet-star angular separation would indicate that G dwarfs are the ideal stellar hosts, making it non-trivial to determine which systems should be prioritized.

Overcoming these coronagraph challenges will enable reflected light to be observed from terrestrial exoplanets. Terrestrial exoplanet reflectance spectroscopy is sensitive to gas absorption and scattering in the atmosphere (Feng et al., 2018), surface reflectively features (Schwieterman et al., 2015a; Robinson et al., 2014a), and cloud and haze extinction (Rugheimer et al., 2013; Arney et al., 2017, 2018). However, at present, coronagraphs that are sensitive to exoplanets have only been used on ground-based telescopes, such as SPHERE on the VLT (Beuzit et al., 2008) and the Gemini Planet Imager (Macintosh et al., 2008, 2014), which necessarily focus on young, hot (and therefore self-luminous) Jovian exoplanets at large separations from their host star (e.g. Galicher et al., 2014; Macintosh et al., 2015; Zurlo et al., 2016).

Directly-imaged exoplanets can also be observed with photometry as a function of time to characterize dynamic global weather patterns (Luger et al., 2019) and rotational effects (Ford et al., 2001; Pallé et al., 2008; Cowan et al., 2009). As a planet rotates on its spin axis, contrasting surface or atmospheric spatially-distinct features can rotate in and out of view, allowing the planet to be mapped as a function of longitude. Although these techniques have a heritage in the astrophysics community for mapping stars (e.g. Strassmeier, 1994; Cameron, 2001) and brown dwarfs (e.g. Yang et al., 2014; Zhou et al., 2018), the short-cadence precision for exoplanets is currently insufficient to take full advantage of these methods.

1.4.2 Modeling Techniques

Each of the aforementioned observing techniques for the characterization of terrestrial exoplanets have associated modeling techniques that are used to predict observable signatures and to infer physical and chemical information from the observations. In general, predicting the observable signatures of exoplanets is referred to as *forward modeling* and involves running a physical model under a set of assumptions and/or measured constraints about the planetary environment. Interpreting observations to infer or measure the underlying physical and chemical state of an exoplanet is referred to as *inverse modeling* and involves iteratively running a forward model to constrain that model’s input parameters given observed data in a Bayesian framework. Both forward and inverse modeling are crucial requirements to properly predict and interpret future terrestrial exoplanet observations, and form two cornerstones of the modeling work in this thesis.

Forward Modeling — Predictions

Forward models used for exoplanet characterization usually consist of one or more coupled models that simulate the state of a planetary atmosphere, the observables of that planet, and how it may appear through different telescopes. For terrestrial exoplanet atmospheres, a typical forward modeling suite consists of climate and photochemistry models to simulate the planetary surface and atmospheric conditions, radiative transfer models to simulate the spectrum and photometry of

the planet, and telescope and instrument noise models to simulate the degradation of the planetary signal when it is observed as an astronomical object.

Exoplanet climate models come in many different forms, each with a different degree of physical sophistication. One-dimensional climate models predict the temperature-pressure (TP) profile with altitude for a planetary atmosphere. Although lacking spatial-resolution and calculating globally-averaged conditions, 1-D climate models can be nimble when applied to unfamiliar exoplanet conditions (e.g. [Kopparapu et al., 2013](#); [Lincowski et al., 2018](#)). However, while they track the radiative-convective heating in a single atmospheric column, they lack sensitivity to atmospheric dynamics. On the other hand, three-dimensional general circulation models (GCMs) (e.g. [Wordsworth et al., 2011](#); [Shields et al., 2013](#); [Turbet et al., 2016](#); [Way et al., 2017](#); [Wolf et al., 2017](#)) track atmospheric dynamics and can predict 3-D observables, including clouds altitudes and distributions, but they tend to be inflexible to modeling unfamiliar, non-Earth-like conditions, they require very long run times, and their spectral models are typically at insufficient resolution to assess key observables. Exoplanet photochemistry models typically operate in 1-D ([Segura et al., 2005](#); [Arney et al., 2016](#); [Lincowski et al., 2018](#)), and can be most readily coupled to 1-D climate models, although recent efforts have been made to apply 1-D photochemistry models to 3-D output from GCMs ([Fauchez et al., 2019](#); [Pidhorodetska et al., 2020](#)) and to develop full 3-D climate-chemistry models ([Chen et al., 2018, 2019](#)) to capture 3-D gas abundance distributions, particularly for tidally locked and synchronously rotating M dwarf exoplanets.

Given an atmospheric state (TP profile and atmospheric gas mixing ratios as a function of altitude) that is consistent with climate and/or photochemistry modeling, a radiative transfer (RT) model is then used to simulate the transit transmission, thermal emission, or reflected spectrum of the planet. When incident radiation interacts with a particle in an atmosphere, that light can (a) be absorbed by the particle, (b) cause the emission of additional photons from the particle at the same wavelength, or (c) scatter to a new trajectory. Radiative transfer models simulate such interactions between light and matter by numerically solving the radiative transfer equation. Nearly all exoplanet radiative transfer models are 1D models, however there are significant differences in physical rigor and computation expense between them. Line-by-line radiative transfer models

solve the radiative transfer equation at sufficiently high spectral resolution to resolve individual lines from the input atomic and molecular line list, provided from a spectroscopic database (e.g. HITRAN; Rothman et al., 2010, 2013; Gordon et al., 2017). Line-by-line RT models can be quite computationally expensive, so efforts are often made to offload the computational expense by pre-computing fixed resolution molecular cross-sections (e.g. Line et al., 2013; Morley et al., 2017) or using k -tables with the correlated- k method (Goody et al., 1989; Lacis and Oinas, 1991; Fu and Liou, 1992; Mlawer et al., 1997; Irwin et al., 2006). However, these expedient methods can introduce errors in the spectra of atmospheres that vary in composition significantly with altitude.

RT models produce high resolution spectra that must be degraded down to the resolution of any telescope instrument that may make the observation and, in some cases, synthetic observational noise is added. These instrument models include a “window” function with a fixed-width over which to integrate the high resolution model spectrum and a final sampling resolution for the mock observed spectrum. Typical window convolution functions are rectangular (“boxcar” or “top-hat”), triangular, and Gaussian, as these effectively reproduce the behavior of spectrometers. At a bare minimum an instrument noise model will include photon shot noise from Poisson counting statistics, however, it is also common to include instrument throughput, detector quantum efficiency, background thermal and scattered light, detector dark current, read noise, and any other noise sources that may complicate the interpretation of the exoplanet spectrum (Robinson et al., 2016).

The component models used for terrestrial exoplanet forward modeling may be used independently or coupled together to ultimately predict what the planet will look like as an astronomical object. As we will see next, inverse modeling also requires the use of forward models to physically interpret a planetary environment given astronomical observations.

Inverse Modeling — Retrievals

Inferring a state of a physical system from observational data of that system is generally referred to as “solving the inverse problem.” In the context of inferring planetary atmospheric information from an observed spectrum, as was pioneered by the Earth and planetary science remote-sensing

communities, the inverse problem is referred to as *retrieval* modeling (Rodgers, 1976, 2000). The goal of a full atmospheric retrieval is to detect and constrain the abundances of spectrally active gases in the atmosphere, infer the temperature-pressure profile, quantify the presence and scattering properties of aerosols, and infer temperature, pressure, and optical scattering properties of a surface, if applicable. In general, atmospheric retrieval models operate by iteratively evaluating a forward model within a Bayesian statistical framework to determine the forward model parameters that best fit the spectrum and their uncertainties given the data uncertainties.

There are many methods that can be used to solve the inverse problem with varying degrees of statistical rigor. Optimal Estimation (OE) is well established in the retrieval literature (Rodgers, 2000), and can be used to infer atmospheric parameters by determining a best-fit to a spectrum. However, OE makes an assumption that the posterior distribution, or updated information after taking the data into account, for each parameter is Gaussian, which allows the retrieval to converge quickly, but it cannot capture complicated degeneracies between atmospheric parameters that may be important for low SNR, low resolution data (Line et al., 2013). Many retrievals use Markov Chain Monte Carlo (MCMC; e.g. Goodman and Weare, 2010; Foreman-Mackey et al., 2013) methods to numerically sample the posterior distributions for the model parameters, which can handle non-Gaussian posteriors and correlations between parameters. However, MCMC simulations require tens of thousands to millions of forward model evaluations to converge, and so are typically only used with computationally expedient forward models. The nested sampling method (e.g. MultiNest; Feroz and Hobson, 2008; Feroz et al., 2009, 2013) is gaining popularity as a middle-ground method between OE and MCMC, which allows posterior distributions with multiple distinct peaks to be analyzed. There is always a balance between physical rigor in the forward model and statistical rigor in the inverse model. Due to the *modus operandi* of the inverse model to iterate the physical forward model, atmospheric retrievals typically employ forward models with simplified physics, such as correlated- k RT, shifting the balance in favor of statistical rigor.

Numerous spectrum retrieval models have been developed for and applied to exoplanet atmospheric characterization (Irwin et al., 2008; Madhusudhan and Seager, 2009; Line et al., 2013; Benneke and Seager, 2013; Waldmann et al., 2015; Cubillos et al., 2016; Lavie et al., 2016; Howe

et al., 2017; Feng et al., 2018; Kitzmann et al., 2019; Damiano and Hu, 2020). These retrieval models have proven to be key workhorses for exoplanet atmospheric discovery, including the detection of atmospheric water vapor (Kreidberg et al., 2015, 2018) and the diagnosis of spectrally-obscuring aerosols (Kreidberg et al., 2014; Sing et al., 2016; Berta et al., 2012), both in the atmospheres of exoplanet gas giant atmospheres. However, many of the currently available retrieval models were developed specifically for gas giant exoplanets with hydrogen and helium atmospheric compositions, where simplifications about the atmosphere may be afforded, including the use of simplified TP profiles and equilibrium thermochemistry, rather than photochemistry.

Currently available retrieval models may be ill-suited to assess terrestrial exoplanet environments. Motivated by low information content spectra, current retrieval models favor statistical rigor over physical rigor, allowing physically simplified, and highly degenerate retrieved solutions. Simplified physical models may efficiently constrain the possible retrieved solutions, but they run the risk of misinterpreting planetary properties. This was demonstrated by the common assumption of a single-column atmosphere, rather than separate day-night temperature distributions, when retrieving gas abundances from planetary spectra (see Line and Parmentier, 2016; Feng and Robinson, 2017). Similarly, retrieval models designed primarily for gaseous planets with primordial atmospheres (e.g. Line et al., 2013) have not been widely adapted and validated for terrestrial planets with a diversity of relatively thin and high molecular weight secondary, outgassed atmospheres where photochemistry produces additional species not predicted by thermochemistry models. A notable exception includes the recent work of Feng et al. (2018), which crucially demonstrated the ability to detect water vapor, ozone, and oxygen in an Earth-like atmosphere with a next-generation direct-imaging telescope. However, Feng et al. (2018) motivates the need for significantly more work in this area, particularly applied to transiting terrestrial exoplanets observed with JWST. Perhaps the greatest challenge is that many different suites of environmental parameters will fit noisy terrestrial exoplanet spectra, and additional methods need to be developed to properly assess and break this retrieval degeneracy for terrestrial exoplanets to properly interpret their observed spectra with upcoming mission.

1.5 Next Steps

Starting in the next few years and extending into the next decade, powerful new telescopes will ensure that observations of terrestrial exoplanets will become commonplace. Before these telescopes come online, it is important to understand their capabilities and limitations specifically for terrestrial exoplanet characterization so that we can optimally make use of them and further plan to make progress beyond them. In this section, we outline a road map for the characterization of terrestrial exoplanets and then discuss a timeline for the next generation of telescopes that may be capable of accomplishing the major milestones of terrestrial exoplanet characterization.

1.5.1 Milestones for Terrestrial Exoplanet Characterization

Characterizing terrestrial exoplanets, like any science, requires a multi-pronged, multi-step approach that can test increasingly more complex hypotheses with increasingly precise datasets. The discovery of each exoplanet reveals information about the parent star and fundamental properties of the planetary body, such as size and mass or density, and so begins the process of characterizing the nature of that planet. Some planets will be found to contribute little more than statistics to a greater trend, such as the prevalence of airless Mercury-like exoplanets that possess no atmosphere to characterize. Other planets may possess habitable atmospheres that capture the focus of the scientific field and the imagination of the public as years are devoted to an in-depth observational reconnaissance aimed at detecting signs of extant life. In this way, we pose a hierarchy of broad scientific questions that may be used as terrestrial exoplanet characterization milestones:

- 1. Does the planet possess an atmosphere?**
- 2. What is the composition and temperature structure of the atmosphere?**
- 3. Is the planet habitable?**
- 4. Is the planet inhabited?**

Does the planet possess an atmosphere? Whether or not any given exoplanet has an atmosphere may be the key initial factor in the decision to observe that planet with more telescope resources. This is especially important for transiting exoplanets around M dwarfs, where observational biases favor short period planets that face high stellar irradiation and flaring that could ablate and potentially completely erode any primordial atmosphere that the planet may have once had. The presence of an atmosphere can (in most cases) be inferred by simply detecting gas absorption features in a transmission, emission, or direct-imaging spectrum, irrespective of the specific molecule responsible for the absorption.

What is the composition and temperature structure of the atmosphere? If the planet has an atmosphere with detectable spectral features then the next step is to determine (1) what gases are causing the absorption features and attempt to infer their abundances, (2) what is the bulk gaseous constituent of the atmosphere, and (3) what is the temperature structure of the atmosphere. These questions may be tackled using an atmospheric retrieval code to invert spectroscopic observations and quantify the nature of the exoplanet environment (e.g. [Line et al., 2013](#)). Answering these questions is never easy, and for many exoplanets a sufficiently thorough understanding of their atmosphere's nature will never be attained and the characterization hierarchy will end on this step. However, for a select few terrestrial planets the next two questions are essential.

Is the planet habitable? Assessing the habitability of an exoplanet will require the confident detection or refutation of surface liquid water. Gathering precise molecular abundances and temperature structure information may be sufficient to place constraints on surface temperature and stability of any surface liquid water ([Robinson, 2018](#)). However, the possibility of surface water and the detection of surface water are different. The direct detection of surface liquid water—using rotational mapping (e.g. [Cowan et al., 2009](#); [Cowan and Strait, 2013](#)), ocean glint (e.g. [Williams and Gaidos, 2008](#); [Robinson et al., 2010, 2014a](#)), or a combination of the two methods ([Lustig-Yaeger et al., 2018](#))—would require a direct-imaging mission, but would offer a less ambiguous approach than combining spectroscopic inference with climate modeling deduction ([Robinson, 2018](#)).

Is the planet inhabited? The discovery of an inhabited exoplanet would be both scientifically

and culturally transformational as it would answer the fundamental human question: *Are we alone?* At present, numerous studies have outlined approaches for the assessment of exoplanet biosignatures (e.g. [Meadows et al., 2018a](#); [Schwieterman et al., 2018](#); [Catling et al., 2018](#)) with a general consensus that the field should (1) be guided by Earth-based life, but (2) be open to non-Earth-based life, and (3) be prepared to assess a putative biosignature in the context of its environment, including understanding the potential for false positive mechanisms. Detecting biosignatures on exoplanets will likely require retrieval methods, performed on high precision spectroscopic data, and coupled to a probabilistic framework to assess the likelihood that the observed biosignature is in fact due to biology. This step represents the final stop in the characterization hierarchy and one of the ultimate goals of terrestrial exoplanet characterization.

We will continuously revisit these thematic questions throughout this work as we determine what observations and methods are optimal for answering each.

1.5.2 A Timeline for the Characterization of Terrestrial Exoplanets

The ability to successfully progress through the major milestones of terrestrial exoplanet characterization in the coming decades, and what type of exoplanets will be studied, will necessarily track the timeline for the next great astronomical observatories as they come online. Each new telescope is uniquely suited for different types of discoveries (see [Fujii et al., 2018](#), for a relevant recent review). The location of the telescopes—on the ground or in space—the aperture size, and the operating wavelength range, are just a few key factors that will dictate the terrestrial exoplanet science capabilities.

The *James Webb Space Telescope* (JWST) is now set to launch in March 2021 and is likely to capture the first spectrum of a terrestrial exoplanet with sufficient precision to identify the presence of atmospheres ([Morley et al., 2017](#)), and to begin constraining the atmospheric composition ([Barstow and Irwin, 2016](#)). JWST will only be capable of studying the atmospheres of terrestrial exoplanets transiting the nearest late M dwarf systems, such as TRAPPIST-1, and it may only provide compelling datasets for hot planets interior to the habitable zone. Since transit transmission spectroscopy can only probe the upper regions of exoplanet atmospheres, JWST reconnaissance

of terrestrial exoplanets may be limited to the upper atmospheres of hot exoplanets orbiting M dwarfs³. JWST will also be capable of both secondary eclipse spectroscopy and thermal phase curves of terrestrial exoplanets, but these techniques are even more biased towards hot planets orbiting M dwarfs than transmission spectroscopy. As a result, for more temperate, habitable zone planets, transmission spectroscopy is likely the most viable characterization technique. However, it cannot probe the near-surface atmosphere of planets, which will make habitability and biosignature assessment quite challenging. Moreover, a directly imaged spectrum in reflected light can contain significantly more information from the lower atmosphere and surface environment than transmission spectroscopy of transiting planets.

Ground-based 30-meter class telescopes are poised for first light in the mid 2020's. The European Extremely Large Telescope (E-ELT; 39.3 m; [Gilmozzi and Spyromilio, 2007](#)) and the Giant Magellan Telescope (GMT; 25.4 m; [Johns, 2006](#); [Johns et al., 2012](#)) have first light planned for 2024. The Thirty Meter Telescope (TMT; 30 m; [Nelson and Sanders, 2008](#)) has first light planned for 2027⁴. These large ground-based telescopes will be capable of continuing transiting terrestrial exoplanet science beyond JWST's nominal lifetime, and extending these investigations into the visible. The 30-m class ground-based telescopes will also pioneer (1) high-contrast exoplanet direct imaging and (2) high-resolution Doppler spectroscopy, both of which may be used to study non-transiting terrestrial exoplanets.

Exoplanet direct-imaging from the ground requires the use of an internal coronagraph to block starlight and reveal small, faint Earth-sized exoplanets that would otherwise be outshined by a factor of $\sim 10^7$ (M dwarfs systems) to $\sim 10^{10}$ (G dwarf systems). Due to atmospheric extinction and turbulence, which disrupt the wavefront of light entering the telescope, current laboratory experiments show a limit to ground-based coronagraph design contrast of $\sim 10^{-8}$ at best ([Males and Guyon, 2018](#)), making them efficient observatories for lower mass K and M dwarf stars ([Arney, 2019](#)), but not for Earth-Sun analogs in G dwarf systems.

³It is worth noting that the recently launched Transiting Exoplanet Survey Satellite (TESS) is well positioned to find such M dwarf extrasolar systems, providing a key synergy with JWST ([Ricker et al., 2010](#)).

⁴However, this telescope remains extremely controversial due to disputes with Native Hawaiians, who cherish the mountain top as sacred. Delays are likely to push the first light of TMT beyond 2027.

High-resolution Doppler spectroscopy is a relatively new (exoplanet) method that aims to directly resolve exoplanetary spectral lines from stellar and telluric lines by leveraging the fact that the exoplanet lines are uniquely Doppler shifted due to the radial velocity of the planet around the star (Snellen et al., 2010; Brogi et al., 2012). This method provides an independent measure of the planet’s orbital inclination relative to the observer, as well as the presence of gaseous atmospheric absorbers by cross-correlating a pre-computed template spectrum with the observed spectrum (Brogi et al., 2014, 2016). High-resolution spectroscopy is essential for resolving the orbital Doppler shift of the exoplanet spectral lines relative to the stellar and telluric lines. Thirty meter class telescopes should be able to collect enough photons to enable such high dispersion spectroscopy. However, resolving an exoplanet’s orbital Doppler shift is only feasible for planets with rapid orbital velocities. This imparts a characterization bias towards compact M dwarf systems, where planets in the habitable zone orbit faster than habitable zone planets in Solar analog systems. Ultimately, 30-m class ground-based telescopes will be capable of significantly extending terrestrial exoplanet characterization into the lower atmospheres of non-transiting planets, but like JWST, they may only be capable of making these leaps forward for low mass M dwarf systems.

The Wide Field Infrared Survey Telescope (WFIRST; Spergel et al., 2015) is set to launch in 2025 and provide a testbed for space-based coronagraphy (Noecker et al., 2016). WFIRST is a 2.4-meter (same as the *Hubble Space Telescope*) general purpose astrophysics telescope that will primarily perform gravitational microlensing exoplanet detections for its exoplanet science time allocation. However, it will launch with an internal coronagraph as a technology demonstration for future missions. WFIRST could plausibly rendezvous with an external starshade occulter to increase contrast and throughput (Soto et al., 2017), but there is no plan for that currently. The WFIRST coronagraph will be capable of studying gaseous exoplanets at wide orbital separations, and will not have the aperture size or design contrast to observe rocky exoplanets. Funding for WFIRST remains uncertain, and it’s practicality for astrobiologically-relevant science is limited to it being a stepping stone to future-mission observations of terrestrial exoplanets via its coronagraphic observations of larger planets.

More recently, NASA commissioned four Decadal Survey Mission Concept Studies under the

leadership of a community-driven Science and Technology Definition Team (STDT) to provide reports for the NASA 2020 Decadal Survey on possible next-generation flagship telescopes. Two of the concept studies focused on relatively large aperture, space-based, exoplanet direct-imaging telescopes: the Habitable Exoplanet Observatory (HabEx; 4-m; [Mennesson et al., 2016](#); [HabEx Study Team, 2019](#)) and the Large UV Optical Infrared (LUVOR; 8-15 m) Surveyor ([Roberge et al., 2019](#); [LUVOR Mission Concept Study Team, 2019](#)). These two space telescopes would launch no sooner than the mid-2030's and be capable of directly-imaging terrestrial exoplanets orbiting Solar-analog G dwarf planetary systems. They would accomplish this by meeting the coronagraph/starshade design contrast of $<10^{-10}$. The third decadal mission concept is the Origins Space Telescope (OST; 6.5-m [Battersby et al., 2018](#); [Origins Space Telescope Study Team, 2019](#)), which would be a JWST successor operating in the mid- to far-infrared, highly capable of transiting exoplanet science. Like JWST, OST would focus on transit transmission and thermal emission studies of terrestrial planets orbiting small M dwarf stars. The fourth mission concept, Lynx ([Gaskin et al., 2019](#)), is an x-ray telescope that would not be capable of exoplanet science directly. Given the technological challenges required, studying the atmospheres of potentially habitable terrestrial exoplanets orbiting G dwarf stars—the true Earth-analogs—will only be possible with a telescope designed from the outset with that goal in mind. Furthermore, these goals will only be met before ~ 2040 if one of these two direct-imaging telescopes (HabEx or LUVOR) is selected for funding priority in the 2020 Decadal Survey.

Between now and ~ 2040 the field of exoplanets is sure to undergo numerous scientific revolutions as near- and far-term next-generation telescopes come online and provide the data to constrain previously unconstrained problems. JWST will soon begin to confront the existence and composition of terrestrial atmospheres for exoplanets transiting M dwarfs. In the mid/late-2020's, 30-meter ground-based telescopes will extend and advance these ventures by probing the potential habitability and biosignatures of transiting and non-transiting K and M dwarf planetary systems. Extending terrestrial exoplanet characterization to the habitability of Earth-like planets orbiting Sun-like stars, and searching for life as we know it to exist on Earth, will require a space-based, direct-imaging telescope, such as HabEx or LUVOR, which will launch no sooner than the mid-2030's. These

ambitious telescope concepts will only see the light of (another world's) day if progress continues to be made towards motivating their unique science cases.

1.6 Purpose & Structure of this Work

The purpose of my dissertation work has been to provide the theoretical scientific underpinning required to support the observation and engage in the assessment of terrestrial exoplanet environments. To this end, we seek to identify how the major milestones of terrestrial exoplanet characterization may be optimally achieved as we progress into the era of next generation telescopes.

In Part I, we use predictive forward modeling to determine the observational requirements and optimal observing techniques to detect and characterize transiting terrestrial exoplanet atmospheres with *JWST*. This work focuses on imminent observations of the seven known Earth-size exoplanets orbiting the late M dwarf, TRAPPIST-1, to support upcoming *JWST* proposals (see Chapter 2). We also use this best-case-scenario system as an indicator for the extent of information that we should expect to learn about terrestrial exoplanets in the *JWST* era, and what information may remain uncharacterized (see Chapter 3). These works help to indicate that although *JWST* will be capable of detecting terrestrial exoplanet atmospheres, characterizing their nature is likely to fall short of robust habitability and biosignature assessment, which must therefore rely on future missions.

In Part II, we use forward modeling to assess the ability for a next-generation, direct imaging mission to detect signs of habitability and biosignatures for non-transiting terrestrial exoplanets orbiting F, G, K, and M dwarfs. To accomplish this, in Chapter 4, we propose a novel means for a future direct imaging telescope to directly infer the presence of an ocean on a rocky exoplanet by spatially mapping the ocean glint signal. In Chapter 5, we conduct a synthetic observational campaign using designs for the next-generation mission, LUVOIR, to determine the number of HZ Earth-like planets for which the mission would be able to detect habitable surface conditions and search for life. These works help to illuminate the bright future of terrestrial exoplanet characterization and to expose the need to develop optimal strategies to interpret terrestrial exoplanet data.

In Part III, we develop, validate, and apply a physically-robust terrestrial exoplanet retrieval

model that can be used to infer the atmospheric characteristics of terrestrial exoplanets observed in the near-term transiting small M dwarfs with *JWST* and in the far-term orbiting Sun-like stars with a LUVOIR-class mission. In Chapter 6, we describe the methodology of our retrieval model, including our physically rigorous forward model and our multi-pronged suite of inverse models, and we validate the model by retrieving on observed spectra of Earth. In Chapter 7, we apply our new retrieval model to synthetic *JWST* transmission spectra of the TRAPPIST-1 System to demonstrate the usefulness of the model for interpreting real terrestrial exoplanet data. This retrieval model takes a crucial first step towards developing an astrobiological framework to assess exoplanet habitability and biosignatures.

In Chapter 8, we provide a discussion wherein we revisit the major milestones of terrestrial exoplanet characterization and report on how our findings have shaped the future outlook and optimal approaches for terrestrial exoplanet atmospheric detection, characterization, assessment of habitability, and the search of biosignatures. We conclude in Chapter 9.

1.7 Rotational Mapping

1.7.1 samurai: Surface Albedo Mapping Using Rotational Inversion

1.8 Telescope Noise Modeling

1.8.1 coronagraph: A Noise Model for Direct-Imaging Exoplanets

1.8.2 PandExo: A Noise Model for Exoplanet Spectroscopy with JWST

1.9 Radiative Transfer

1.9.1 smart: Spectral Mapping Atmospheric Radiative Transfer

Part I

**DEVELOPING A FOUNDATION TO CHARACTERIZE TRANSITING
EXOPLANETS IN THE ERA OF JWST**

Chapter 2

DETECTING & CHARACTERIZING THE TRAPPIST-1 EXOPLANET ATMOSPHERES WITH JWST

Portions of this chapter were originally published in collaboration with Victoria S. Meadows and Andrew P. Lincowski in the June 2019 edition of the Astronomical Journal (Lustig-Yaeger et al., 2019b, AJ, Vol. 158, 27; 2019 © American Astronomical Society, DOI: 10.3847/1538-3881/ab21e0), and are reproduced below with the permission of the American Astronomical Society.

Summary

The James Webb Space Telescope (JWST) will offer the first opportunity to characterize terrestrial exoplanets with sufficient precision to identify high mean molecular weight atmospheres, and TRAPPIST-1's seven known transiting Earth-sized planets are particularly favorable targets. To assist community preparations for JWST observations, we use simulations of plausible post-ocean-loss and habitable environments for the TRAPPIST-1 exoplanets, and test simulations of all bright object time series spectroscopy modes and all MIRI photometry filters to determine optimal observing strategies for atmospheric detection and characterization using both transmission and emission observations. We find that transmission spectroscopy with NIRSpec Prism is optimal for detecting terrestrial, CO₂ containing atmospheres, potentially in fewer than 10 transits for all seven TRAPPIST-1 planets, if they lack high altitude aerosols. If the TRAPPIST-1 planets possess Venus-like H₂SO₄ aerosols, up to 12 times more transits may be required to detect an atmosphere. We present optimal instruments and observing modes for the detection of individual molecular species in a given terrestrial atmosphere and an observational strategy for discriminating between evolutionary states. We find that water may be prohibitively difficult to detect in both Venus-like and habitable atmospheres due to its presence lower in the atmosphere where transmission spectra are less sensitive. Although the presence of biogenic O₂ and O₃ will be extremely challenging to detect, abiotically produced oxygen from past ocean loss may be detectable for all seven TRAPPIST-1 planets via O₂-O₂ collisionally-induced absorption at 1.06 and 1.27 μm , or via NIR O₃ features for the outer three planets. Our results constitute a suite of hypotheses on the nature and detectability of highly-evolved terrestrial exoplanet atmospheres that may be tested with JWST.

2.1 Introduction

The discovery of Earth-sized planets in temperate orbits around nearby, low-mass stars opens a new door into the era of terrestrial exoplanet atmospheric characterization (Berta-Thompson et al., 2015; Anglada-Escudé et al., 2016; Gillon et al., 2016, 2017b; Luger et al., 2017c; Dittmann et al., 2017). Transmission and emission (secondary eclipse) spectroscopy of transiting rocky worlds with the upcoming James Webb Space Telescope (JWST) may offer a first glimpse into the atmospheres of terrestrial exoplanets (Morley et al., 2017; Kalirai, 2018) and a first opportunity to search for signs of habitability (Lincowski et al., 2018) and biosignatures beyond the Solar System (Cowan et al., 2015).

The TRAPPIST-1 system of seven transiting Earth-sized exoplanets (Gillon et al., 2016, 2017b; Luger et al., 2017c) is observationally favorable for the atmospheric characterization of small exoplanets. TRAPPIST-1 is a late M dwarf (M8V; Liebert and Gizis, 2006) with a small radius ($0.121R_{\odot}$; Van Grootel et al., 2018), which increases planetary transit and eclipse depths; it has a low effective temperature (2511 K; Van Grootel et al., 2018), which increases the eclipse depth; and it is nearby to Earth (12.2 pc; Gillon et al., 2016). These system properties increase sensitivity to atmospheric spectral features in transmission and emission spectroscopy, particularly for small, temperate planets.

Observations of the TRAPPIST-1 planets with the *Kepler*, *Hubble* (HST), and *Spitzer* space telescopes suggest that the innermost 6 planets do not have primordial, low mean molecular weight atmospheres, but whether they have high molecular weight atmospheres or no atmospheres at all requires observations with future facilities. HST Wide Field Camera 3 (WFC3) transmission spectroscopy has ruled out H_2 -dominated atmospheres for most of the TRAPPIST-1 planets (de Wit et al., 2016, 2018), and instead they may have secondary outgassed atmospheres composed of relatively high mean molecular weight gases (Moran et al., 2018). Secondary atmospheres are more difficult to detect—or rule out—than primordial atmospheres, and they may span a large range in temperature, pressure, and composition. Consequently a broad variety of potential atmospheres are consistent with the current modest observational constraints (Delrez et al., 2018; Lincowski et al.,

2018).

Alternatively, the planets could have no atmospheres, although outgassing from the possibly high-volatile-content interiors of the TRAPPIST-1 planets may make that outcome less likely. Simulations suggest that the TRAPPIST-1 planets could have had their atmospheres completely stripped (e.g. [Dong et al., 2018](#); [Airapetian et al., 2017](#); [Roettenbacher and Kane, 2017](#)), although some models of M dwarf planets suggest that atmospheric loss rates may be less than the replenishment rate via outgassing from a planetary interior ([Garcia-Sage et al., 2017](#); [Bolmont et al., 2017](#)), such that atmospheres may be retained. The TRAPPIST-1 planets may also have larger volatile reservoirs than Solar System terrestrials. Masses derived from transit timing variation (TTV) analyses of *Kepler* ([Grimm et al., 2018](#)) and *Spitzer* observations ([Delrez et al., 2018](#)) suggest that the planets have densities between 0.6 to 1.0 $\rho_{\oplus}$, consistent with a rocky composition with a significant fraction of ices ([Grimm et al., 2018](#)). This conclusion is supported by the resonant chain structure of the TRAPPIST-1 system, which suggests inward migration from more volatile-rich formation orbits ([Luger et al., 2017c](#)). Determining whether the interplay of planetary processes over time will allow M dwarf terrestrial planets to maintain high-molecular weight atmospheres and support habitability will be key science questions for JWST.

Previous simulations have demonstrated the plausibility of detecting terrestrial atmospheres with JWST for M dwarf planetary systems in general (e.g. [Belu et al., 2011](#); [Barstow et al., 2016](#)), and for the TRAPPIST-1 planets in particular ([Barstow and Irwin, 2016](#); [Morley et al., 2017](#); [Batalha et al., 2018](#); [Krissansen-Totton et al., 2018a](#)). [Morley et al. \(2017\)](#) considered Venus-like, Earth-like, and Titan-like atmospheres in thermochemical equilibrium and found that CO₂-dominated atmospheres could be detected for 6 of the 7 TRAPPIST-1 planets in fewer than 20 transits with JWST NIRSpec/G235M, and secondary eclipse photometry with JWST/MIRI could readily detect thermal emission from TRAPPIST-1 b and c, and possibly d, e, and f. [Batalha et al. \(2018\)](#) investigated optimal strategies for JWST observations of the TRAPPIST-1 system and found that allowing the NIRSpec Prism to slightly saturate at the peak of the stellar spectral energy distribution (SED) could allow the dominant absorber to be detected in 10 transits of H₂O, CO₂, and N₂ dominated atmospheres. [Krissansen-Totton et al. \(2018a\)](#) investigated the potential

detectability of anoxic biosignatures in the atmosphere of TRAPPIST-1 e and concluded that 10 transits observed with JWST NIRSpec Prism may be sufficient to detect CO₂ and constrain the CH₄ abundance enough to rule out non-biological CH₄ production. Recently, [Wunderlich et al. \(2019\)](#) considered the detectability of photochemically self-consistent Earth-like planets in the habitable zone of various M dwarfs spectral types, including TRAPPIST-1, and found that H₂O, CH₄, and CO₂ may be detectable in ~10 transits with JWST. However, this study did not consider non Earth-like planetary compositions, or the effect of clouds and hazes.

These previous simulations of TRAPPIST-1 planetary atmospheres did not include photochemical forcing by the late M dwarf SED on multiple plausible planetary environments, both habitable and uninhabitable. Photochemistry can have significant impacts on terrestrial atmospheric composition ([Segura et al., 2005](#); [Rugheimer et al., 2015](#)) and haze formation ([Arney et al., 2017, 2018](#)), which can in turn modify the resultant temperature structure ([Lincowski et al., 2018](#)), impacting the predicted spectrum in both transmission and emission. Models that include photochemistry and haze formation for terrestrial atmospheres will better predict the atmospheric composition, and inform the preparation and interpretation of upcoming observations with JWST ([Lincowski et al., 2018](#)).

In this chapter we explore the potential for JWST to detect and characterize the TRAPPIST-1 planetary atmospheres and distinguish between model predictions for evolutionary outcomes and different atmospheric states. As input, we use the climatically and photochemically self-consistent atmospheric and spectral simulations of [Lincowski et al. \(2018\)](#), who used a rigorous line-by-line 1D radiative-convective-equilibrium climate model coupled with a 1D photochemical model to simulate different habitable and post-ocean-loss environments for the TRAPPIST-1 planets. The atmospheric bulk compositions considered by [Lincowski et al. \(2018\)](#) are motivated by the early high luminosity of late M dwarf stars ([Baraffe et al., 2015](#)) like TRAPPIST-1, which may drive early ocean loss and the generation of tens to thousands of bars of O₂ ([Luger and Barnes, 2015](#); [Bolmont et al., 2017](#); [Meadows et al., 2018a](#); [Lincowski et al., 2018](#); [Wordsworth et al., 2018](#)). The final inventory of O₂ may be severely reduced by atmospheric and surface loss processes such that only a few bars of O₂ remain ([Schaefer et al., 2016](#); [Wordsworth et al., 2018](#)). Depending on the

initial water inventory, this process may have exhausted the entire planet's water supply, leaving it desiccated, or the planet may have formed with enough water to endure such vigorous loss. For the ocean loss planets with efficient O_2 sinks and ongoing outgassing of volatiles, large quantities of CO_2 may build up, forming a Venus-like atmosphere (Meadows et al., 2018a). Lincowski et al. (2018) also considered a habitable ocean world for TRAPPIST-1 e, for the scenario where it formed with an appreciable H_2 envelope, which was subsequently stripped to reveal a habitable core (Luger et al., 2015).

To understand the nature and possibly habitability of terrestrial exoplanets, we will need a systematic approach to environmental assessment that starts with the most scientifically significant and least expensive observation, and builds from there. For terrestrial planets orbiting M dwarfs, the first property to be determined is whether or not they have an atmosphere. If an atmosphere can be confirmed, then subsequent studies will focus on determining the nature of that atmosphere, and whether there are atmospheric characteristics that could discriminate between evolutionary outcomes. Finally, a deeper dive to search for signs of habitability, including the presence of an ocean (see Robinson, 2018; Lustig-Yaeger et al., 2018) and biosignatures (see Schwieterman et al., 2018) may be warranted for planets whose initial characterization does not preclude habitable conditions.

Here we determine the feasibility of atmospheric characterization for the seven known TRAPPIST-1 exoplanets by first identifying optimal instrument selection and experiments to test whether or not the TRAPPIST-1 planets have atmospheres. We then determine how to best detect specific molecules as a second step of atmospheric characterization, and discriminate between different plausible climate and photochemically self-consistent atmospheres (e.g. O_2 -dominated, CO_2 -dominated; Lincowski et al., 2018) using JWST transmission and emission photometry and spectroscopy.

The different post-runaway, evolved planet atmospheres considered here are by no means a comprehensive set of evolutionary outcomes, but rather a representative subset of potential physically- and chemically-motivated atmospheres for which we can predict spectra. By understanding the detectability of spectral discriminants for the TRAPPIST-1 planets we develop in-

formed hypotheses on the nature of these planets that may later be tested with JWST observations.

The structure of this chapter is as follows: In §2.2 we present the models and methods used to simulate JWST data and find the optimal JWST modes for detecting and characterizing plausible planet compositions. We describe our results in §2.3, offer a discussion of the significance of those results in §2.4, and conclude in §2.5.

2.2 *Methods*

In the following subsections we present our methods for assessing the detectability of different self-consistent atmospheric compositions for the seven TRAPPIST-1 planets with different instruments and observational modes available to JWST. We first describe the JWST noise models used in this work (§2.2.1), which include a MIRI photometry component and a spectroscopy component using PandExo (Batalha et al., 2017). We then detail our model inputs (§2.2.2), and outline a series of experiments that can be used to successively characterize terrestrial planet environments and determine optimal observing modes (§2.2.3).

2.2.1 *JWST Noise Modeling*

We simulate synthetic exoplanet time-series spectroscopy and photometry with JWST to consider observations during transit (transmission) and secondary eclipse (emission). In the following two subsections we detail our modeling of JWST/MIRI photometry (§2.2.1) and JWST spectroscopy (§2.2.1).

JWST/MIRI Photometry

Filter photometry with JWST’s Mid-Infrared Instrument (MIRI) imager has been suggested to offer an efficient means of performing an initial characterization of Earth-sized planets around low-mass stars (Morley et al., 2017). To assess MIRI photometry, we develop a basic MIRI imaging noise model for exoplanet transit and secondary eclipse observations.

The number of photons from the planet incident upon the detector is

$$N_p = T_{\text{exp}} \frac{F_p \mathcal{T} A \lambda \Delta \lambda}{hc} \quad (2.1)$$

where T_{exp} is the exposure time, F_p is the spectral flux density (e.g. W/m²/um) from the planet, $\mathcal{T}$ is the filter throughput, A is the telescope collecting area (25 m²), λ is wavelength, $\Delta \lambda$ is the width of the wavelength bin, h is Planck's constant, and c is the speed of light. Note that the exposure time used in our photometry model is discretized in terms of the simulated planet's transit duration, but does not explicitly depend on the MIRI integration times or number of groups per integration, as our time-series spectroscopy noise modeling with `PandExo` does. Photon conversion efficiency curves for the MIRI imager were acquired online¹ (Glasse et al., 2015). Photons from the star are calculated analogously using Equation 2.1 by replacing the planet flux with the stellar flux.

The signal-to-noise (SNR) ratio on the transit depth is given by

$$\text{SNR}_T = \frac{N_s (R_p/R_s)^2}{\sqrt{(N_s + N_{bg})/n_{\text{out}} + N_{bg} + N_s [1 - (R_p/R_s)^2]}}. \quad (2.2)$$

where N_s is the number of photons from the star, N_{bg} is the number of photons from background sources, and n_{out} is the number of out-of-transit transit durations observed. Background photon noise is calculated using the seven component grey-body model of Glasse et al. (2015), which includes telescope thermal and scattered zodiacal noise contributions. The SNR on the photons detected from a planet observed in secondary eclipse is given by

$$\text{SNR}_E = \frac{N_p}{\sqrt{(N_p + N_s + N_{bg})/n_{\text{out}} + N_s + N_{bg}}}. \quad (2.3)$$

where N_p is the number of photons from the planet. Derivations of Equations 2.2 and 2.3 are provided in Chapter Appendix 2.6.1.

Saturation must be considered when planning long exposures necessary to characterize small

¹<http://ircamera.as.arizona.edu/MIRI/pces.htm>

exoplanets around nearby stars. Using the bright source limits of [Glasse et al. \(2015\)](#) we find that TRAPPIST-1 will saturate MIRI in the two shortest wavelength filters, F560W and F770W, for the shortest exposures allowed in the standard imaging mode. Although these shorter wavelength filters may saturate, we nonetheless consider MIRI photometry in all nine filters (F560W, F770W, F1000W, F1130W, F1280W, F1500W, F1800W, F2100W, and F2550W) to assess the atmospheric information contained in each.

JWST Spectroscopy

We use the JWST time-series spectroscopy simulator `PandExo`² (version 1.1.2; [Batalha et al., 2017](#); [Batalha et al., 2018](#)) to model different observing modes and their associated noise sources for transmission and emission spectroscopy. `PandExo` leverages the core of the Space Telescope Science Institute’s Exposure Time Calculator, `Pandeia`³ (version 1.2.2; [Pontoppidan et al., 2016](#)), to calculate 3-D data cubes for realistic PSF modeling. We refer the reader to [Batalha et al. \(2017\)](#) for a thorough description of the model and its bench-marking.

We consider a broad variety of JWST instruments and modes that are capable of exoplanet transmission and emission spectroscopy and available using `PandExo`. We include the Near-Infrared Camera (NIRCam; [Greene et al., 2007, 2017](#)) using the grism time-series mode; the Near-Infrared Spectrograph (NIRSpec; [Bagnasco et al., 2007](#); [Ferruit et al., 2014](#)); the Near Infrared Imager and Slitless Spectrograph (NIRISS; [Doyon et al., 2012](#)) using the single object slitless spectroscopy (SOSS) mode; and the Mid-Infrared Instrument (MIRI; [Bouchet et al., 2015](#)) low resolution spectrometer (LRS; [Kendrew et al., 2015](#)).

Table 2.1 summarizes the different JWST instruments and modes used to simulate transmission and emission spectroscopy of the TRAPPIST-1 system. Specifically, Table 2.1 lists the instrument, mode, disperser, filter, subarray, read mode, wavelength range, and nominal spectral resolving power used in our `PandExo` calculations. Table 2.1 also presents the number of groups per integration, the observing efficiency, and the number of saturated pixels at the end of the ramp for

²<https://natashabatalha.github.io/PandExo/>

³<https://jwst.etc.stsci.edu/>

Table 2.1. JWST instruments used in this study and their observability of TRAPPIST-1

Instrument	Mode	Disperser	Filter	Subarray	Read Mode	λ [μm]	R [$\lambda/\Delta\lambda$]	TRAPPIST-1		
								N_{groups}	Efficiency	N_{sat}
NIRCam	ssgrism	grism R	f322w2	subgrism64	rapid	2.42 – 4.15	~ 1600	303	99.34%	0
NIRCam	ssgrism	grism R	f444w	subgrism64	rapid	3.70 – 5.00	~ 1600	342	99.42%	0
NIRISS	SOSS	gr700xd	None	substrip96	nrsrapid	0.6 – 2.8	~ 700	52	96.23%	0
NIRISS	SOSS	gr700xd	None	substrip256	nrsrapid	0.6 – 2.8	~ 700	21	90.91%	0
NIRSpec	fixed slit	g140h	f100lp	sub2048	nrsrapid	0.97 – 1.82	~ 2700	39	95.00%	0
NIRSpec	fixed slit	g140m	f100lp	sub2048	nrsrapid	0.97 – 1.84	~ 1000	14	86.67%	0
NIRSpec	fixed slit	g235h	f170lp	sub2048	nrsrapid	1.66 – 3.05	~ 2700	41	95.24%	0
NIRSpec	fixed slit	g235m	f170lp	sub2048	nrsrapid	1.66 – 3.07	~ 1000	14	86.67%	0
NIRSpec	fixed slit	g395h	f290lp	sub2048	nrsrapid	2.87 – 5.14	~ 2700	82	97.59%	0
NIRSpec	fixed slit	g395m	f290lp	sub2048	nrsrapid	2.87 – 5.10	~ 1000	29	93.33%	0
NIRSpec	fixed slit	prism	clear	sub512	nrsrapid	0.6 – 5.3	~ 100	2	33.33%	0
NIRSpec	fixed slit	prism	clear	sub512s	nrsrapid	0.6 – 5.4	~ 100	3	50.00%	0
NIRSpec	fixed slit	prism	clear	sub512	nrsrapid	0.6 – 5.5	~ 100	6	71.43%	47
NIRSpec	fixed slit	prism	clear	sub512s	nrsrapid	0.6 – 5.6	~ 100	6	71.43%	19
MIRI	LRS	p750l	None	slitlessprism	fast	0.5 – 12.0	~ 100	139	98.57%	0

Note. — N_{groups} is the number of groups per integration and N_{sat} is the number of saturated pixels at the end of the ramp. The right three columns are outputs from the `PandExo` JWST noise model specifically for observations of the TRAPPIST-1 system.

our simulated observations of the TRAPPIST-1 system with each instrument. We use the NIRSpec Prism in three configurations. First, we use the SUB512 subarray (frame time: 0.226 s) with 2 groups per integration set by `PandExo` to avoid pixel saturation. Second, we use the SUB512s subarray (frame time: 0.144 s) with 3 groups per integration, again set by `PandExo`. Finally, we simulate a partial saturation strategy by using the SUB512 and SUB512s subarrays with 6 groups per integration to allow for slight pixel saturation near the peak of the SED (Batalha et al., 2018). This modification improves the duty cycle from 33.3% (50%) to 71.4% for the SUB512 (SUB512s) subarray, as shown in Table 2.1 (see “Efficiency” column). We refer to this partially saturated NIRSpec Prism mode as “NIRSpec Prism*” hereafter.

2.2.2 Noise Model Inputs

For the stellar input to the `PandExo` noise model, we approximate the TRAPPIST-1 stellar spectrum, which has yet to be observed, using a PHOENIX stellar model (Husser et al., 2013) with an effective temperature of $T = 2511$ K, metallicity of $[\text{Fe}/\text{H}] = 0.04$, and surface gravity $\log g = 5.23$

(Delrez et al., 2018), normalized to the K band magnitude of TRAPPIST-1 ($K = 10.30$; Grimm et al., 2018). However, we ignore the effects of stellar opacity in the MIR, stellar variability due to rotation and flaring during periods of observation (Vida et al., 2017; Morris et al., 2018a), and heterogeneous stellar photospheres (Rackham et al., 2018; Zhang et al., 2018).

We use the modeled transmission and emission spectra of Lincowski et al. (2018) as inputs into the noise models to assess the detectability of photochemically and climatically self-consistent TRAPPIST-1 planet’s atmospheres. The climate model developed by Robinson and Crisp (2018) and Lincowski et al. (2018) uses line-by-line radiative transfer computed by the Spectral Mapping and Atmospheric Radiative Transfer (SMART) code (Meadows and Crisp, 1996, developed by D. Crisp), and can generate top-of-atmosphere planetary radiances and transmission spectra (Robinson, 2017a) of its equilibrium climate and photochemical states. See Lincowski et al. (2018) for a thorough description of the climate and photochemical modeling and subsequent climate results. The stellar spectrum used in the optical through the MIR in Lincowski et al. (2018) is identical to the stellar spectrum used here for the JWST noise model input.

For target exposure times per transit we use the median transit durations for the TRAPPIST-1 planets from Grimm et al. (2018), and for photometry and spectroscopy noise calculations we assume that an equal amount of time is spent observing in transit/eclipse versus out of transit/eclipse. For PandExo, we assume saturation to be an exposure level 80% of the full well and we impose no strict noise floor for JWST spectroscopy (c.f. Greene et al., 2016). We compute noise calculations across a grid in the number of transits/eclipses ([1, 100]), and then use these results to derive and report the number of transits/eclipses needed to meet specific atmospheric detection and characterization metrics, as described in the following section.

2.2.3 Observing Experiments

We aim to identify the optimal observing approaches for JWST to (1) detect the presence of the TRAPPIST-1 planet atmospheres, and (2) characterize the composition of the atmospheres, assuming the TRAPPIST-1 planets possess atmospheres similar in nature to the evolved atmospheres modeled in Lincowski et al. (2018). We consider an atmosphere to be detected when sufficient

SNR is achieved on *any* spectral feature in a transmission or emission spectrum. For atmospheric characterization, we consider a specific molecule in the atmosphere to be detected when sufficient SNR is attained on the contribution to the spectrum from that molecule, which may include multiple bands from a given molecule. Next, we detail our SNR approach for ruling out a fiducial spectrum, and then describe how the method is used to quantify the detectability of atmospheres and specific molecules within them.

SNR Approach

We now define a signal-to-noise ratio approach to determine the confidence with which we can rule out that our data match a fiducial transmission or emission spectrum—which is a featureless spectrum for the case of detecting atmospheres.

For each atmospheric model and JWST instrument considered, we employ the following procedure for both transit and secondary eclipse geometries. First, we run the `PandExo` JWST noise model across a grid in number of transits/eclipses (n_{occ}) from 1-100, which is sufficient to establish a simple SNR scaling relationship. Second, we determine the signal-to-noise on the difference between the model spectrum and the fiducial spectrum, and calculate the total expected signal-to-noise $\langle \text{SNR} \rangle$ (defined below) by summing this difference over wavelength. Finally, we solve for number of transits/eclipses, n_{occ} , such that a given $\langle \text{SNR} \rangle$ is achieved.

We define the total expected signal-to-noise using a $\Delta\chi^2$ test formalism, which is common for model selection applications. For many random drawings of synthetic data with Gaussian noise, the expected value for $\Delta\chi^2$ between two models (m_1 and m_2) is simply

$$\langle \Delta\chi^2 \rangle = \sum_{i=1}^{N_\lambda} \left(\frac{m_{1,i} - m_{2,i}}{\sigma_i} \right)^2 \quad (2.4)$$

assuming that the observations with uncertainties σ are truly sampled from one of the models (m_1 in this case), and where the sum is over all N_λ spectral elements for a particular instrument. Gaussian noise is not added to the synthetic spectra both to speed up calculation of the mean result, and to avoid any single random data realization from biasing our results (e.g. [Feng et al., 2018](#)).

The numerator in the sum in Equation 2.4 is the “signal” used to discriminate between the two models, while the denominator is the “noise” on the observations. For convenience, we define

$$\text{SNR}_i = \frac{m_{1,i} - m_{2,i}}{\sigma_i}, \quad (2.5)$$

which is the individual signal-to-noise contribution from each spectral element to model m_2 being ruled out in favor of model m_1 . Equation 2.4 may then be rewritten in terms of SNR_i and what we refer to as the total expected signal-to-noise

$$\langle \text{SNR} \rangle = \sqrt{\langle \Delta\chi^2 \rangle} = \sqrt{\sum_{i=1}^{N_\lambda} \text{SNR}_i^2}. \quad (2.6)$$

Equation 2.6 is particularly useful because the quadrature sum over wavelength allows for the comparison between multiple different JWST instruments that may have different spectral resolutions and wavelength ranges, and for a comparison between transmission and emission spectroscopy for the same hypothesis (e.g. *the planet does not have an atmosphere*).

We caution that Equation 2.6 is equal to the confidence in the detection of model m_1 in units of standard deviations (number of “sigma” n_σ), only in the case that each model has one degree of freedom. To avoid assumptions on the degrees of freedom and degeneracies associated with spectral models, we simply report $\langle \text{SNR} \rangle$ with the understanding that these are upper limits on the confidence in the detection. Furthermore, when we report the number of transits/eclipses to rule out the fiducial model to a given $\langle \text{SNR} \rangle$, these are lower limits; additional sources of uncertainty and/or retrieval model complexity may require that more transits/eclipses be observed.

We now detail the fiducial spectral models that are used to detect the presence of an atmosphere and the individual molecules with them.

Detecting the presence of an atmosphere

The spectra of planets with atmospheres can be discriminated from the featureless spectra of airless worlds by the presence of spectral absorption features. This approach works best for atmospheres

with strong absorption features, but will be challenging for transmission spectroscopy if molecular absorption is suppressed by a high mean molecular weight atmosphere (Miller-Ricci et al., 2009), the presence of clouds and hazes (Berta et al., 2012; Ehrenreich et al., 2014; Knutson et al., 2014; Kreidberg et al., 2014; Nikolov et al., 2015), or atmospheric refraction (B  tr  mieux and Kaltenegger, 2014; Misra et al., 2014). In these cases, higher SNR observations will be required to detect the presence of an atmosphere. Here we assess the detectability of realistic terrestrial atmospheres with clouds and hazes in transmission by comparing their spectra to a baseline featureless spectrum that is modeled as the best-fitting constant planet radius with wavelength. For the corresponding emission spectrum test, we model the featureless spectrum as the ratio of two blackbodies at the stellar effective temperature and best-fitting planetary equilibrium temperature (set by a variable bond albedo).

While there are several “false negative” processes that could suppress the signal from an atmosphere as described above, one “false positive” process is worth considering: the possibility that surface mineralogy could produce wavelength dependent features in emission spectra of airless worlds. However, these features are unlikely to be as prominent as atmospheric features, and this is especially true in realistic cases where multiple reflections occur within the mineral surface. Specifically, by Kirchoff’s law of thermal radiation, the emissivity ϵ is related to the reflectance R by (Nicodemus, 1965),

$$\epsilon = 1 - R. \quad (2.7)$$

The emissivity then regulates the efficiency of thermally radiated flux F_λ relative to a perfect blackbody $B_\lambda(T)$ via

$$F_\lambda = \epsilon_e B_\lambda(T), \quad (2.8)$$

for an object with an equilibrium temperature T . However, Equation 2.8 uses the *effective* emissivity,

$$\epsilon_e = 1 - (1 - \epsilon)^{(n+1)}, \quad (2.9)$$

which accounts for the number of reflections n within the material, and ultimately decreases the

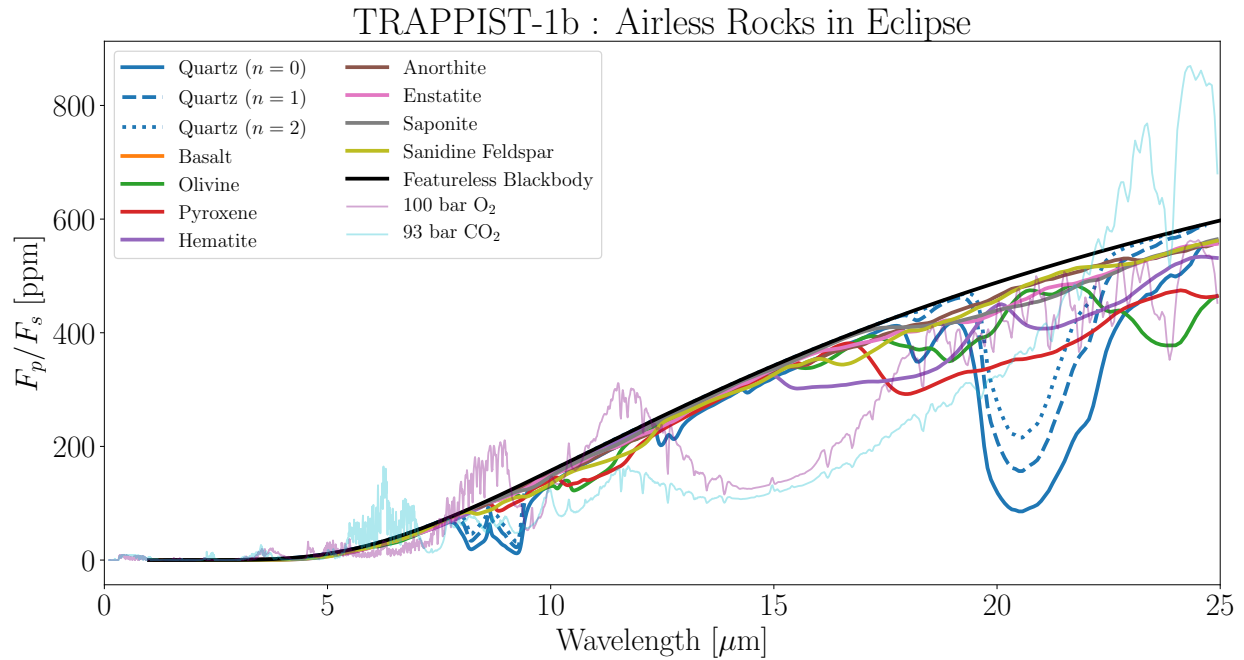


Figure 2.1 Secondary eclipse spectrum models of TRAPPIST-1b assuming different end-member planet mineralogical compositions. The black line shows a featureless blackbody curve, corresponding to the zero bond albedo planet equilibrium temperature, from which all of the thermal emission curves from rock forming minerals (thick color lines) deviate due to non-unity, wavelength-dependent emissivities. The three line styles for the quartz curves demonstrate the reduction in effective emissivity due to n reflections within the rock (see equation 2.9). The thin purple and teal lines show the expected emission spectrum for TRAPPIST-1b if it were to possess a climatically and photochemically self-consistent thick O_2 - or a CO_2 -dominated atmosphere, respectively (Lincowski et al., 2018).

contrast of mineralogical features in the thermal emission spectrum (Kirkland et al., 2003; Hu et al., 2012).

Figure 2.1 compares secondary eclipse spectrum models for different assumed end-member planet mineralogical compositions. Emission spectra are shown for planets at the zero bond albedo equilibrium temperature composed solely of quartz, basalt, olivine, pyroxene, hematite, anorthite, enstatite, saponite, and feldspar—a selection of common rocks and minerals found in terrestrial solar system bodies. For comparison, also shown are emission spectra with strong atmospheric absorption features from the climatically and photochemically self-consistent 100 bar O_2 -dominated

and 92 bar CO₂-dominated atmospheres from [Lincowski et al. \(2018\)](#).

In some cases, mineralogical surface emission features compare in signal contrast to model atmospheric thermal emission features, but under most plausible physical scenarios an airless rock would likely have significantly lower spectral variation than atmospheric features. The single-reflection quartz silicate features between $\sim 8 - 10 \mu\text{m}$ and $\sim 19 - 23 \mu\text{m}$ are strong and rival the strength of the atmospheric features caused by O₃, CO₂, and H₂O over this wavelength range. However, single reflections are unlikely, and both multiple reflections and blends of different minerals would decrease the contrasts on features from any one of the representative end-member cases, producing a relatively featureless thermal emission spectrum.

Throughout the rest of the chapter, unless otherwise stated, we adopt the assumption that detecting deviations from blackbody emission in a secondary eclipse spectrum are evidence of an atmosphere. However, in §2.3 we further explore the detectability of quartz silicate emissivity features in the emission spectrum of TRAPPIST-1b observed with MIRI LRS as an optimistic limiting case on the potential signal from an airless body.

Figures 2.2 and 2.3 provide a specific example of a test for the presence of a 10-bar CO₂ atmosphere, using both transmission (with NIRSPEC Prism*) and emission (with MIRI LRS) spectroscopy, respectively. The bottom panel of each figure shows a direct comparison between the template spectrum and the featureless model that best fits the template spectrum. The color-contours show the magnitude of difference between the two spectra in the bottom panel divided by the noise ($|\text{SNR}_i|$) as a function of wavelength and number of occultations. The right panels show the total expected SNR, $\langle \text{SNR} \rangle$, as a function of number of occultations, which is the total signal of the atmosphere over the wavelength range of the instrument. The quadrature sum over wavelength not only allows for a comparison between different instruments that naturally accounts for the native resolution of and noise incurred by the instrument, but also a comparison between transmission and emission spectroscopy. Comparing Figures 2.2 and 2.3 we see that if TRAPPIST-1b possesses a 10 bar high CO₂ atmosphere, detecting that atmosphere by molecular features in the spectrum will require fewer transits with NIRSpec Prism* than secondary eclipses with MIRI LRS. This example demonstrates how we identify optimal observing modes for atmosphere detections,

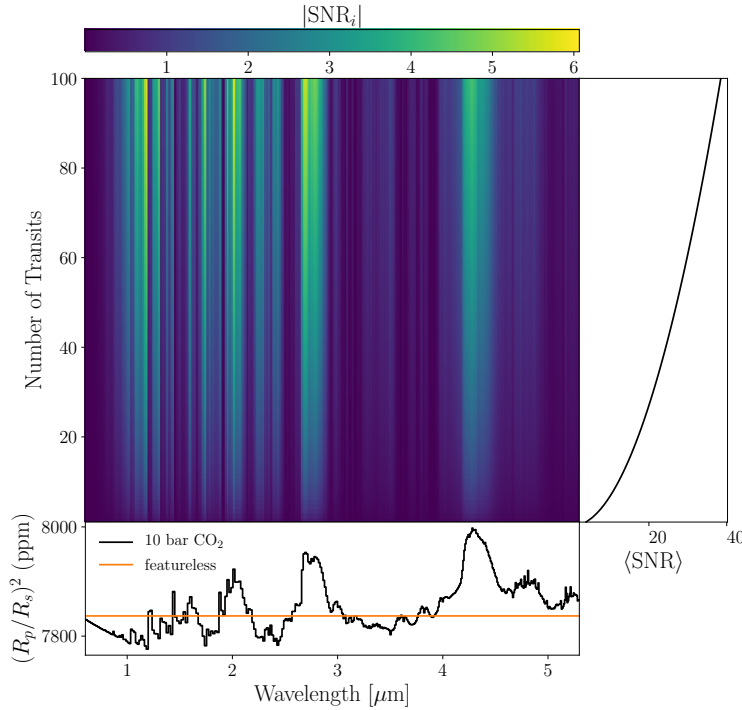


Figure 2.2 Signal-to-noise contours on the simulated 10 bar high CO_2 transmission spectrum, observed with JWST/NIRSpec Prism with 6 groups per integration, relative to a featureless spectrum as a function of the number of transits observed and wavelength of the instrument. The bottom panel shows the noiseless spectrum (black) and the best-fitting featureless spectrum (orange), both convolved to the instrument resolution. The right panel shows how the total expected SNR from Equation 2.6 increases with more transits observed.

and enables a comprehensive study to determine the exposure times needed detect the presence of an atmosphere as a function of observing mode and atmosphere type.

Unless otherwise stated, throughout the rest of the chapter we adopt the convention that an atmosphere is detected if $\langle \text{SNR} \rangle \geq 5$ is achieved on absorption features in the spectrum. We report our results for the detectability of atmospheres at this threshold, but encourage readers to scale our results to their own desired detection thresholds.

Detecting Specific Molecules

To detect individual molecules in the spectrum, we apply the methods described in the previous sections to spectra with and without the absorption features from a given molecular species. To perform these tests we generate additional transmission and emission spectra by running our radiative transfer model for a given atmosphere with each spectrally active molecule removed one at a time. We then use the spectra that are missing contributions from individual molecules as m_2 in Equation 2.4 and 2.6 to calculate the $\langle \text{SNR} \rangle$ on the contribution from each molecule to the

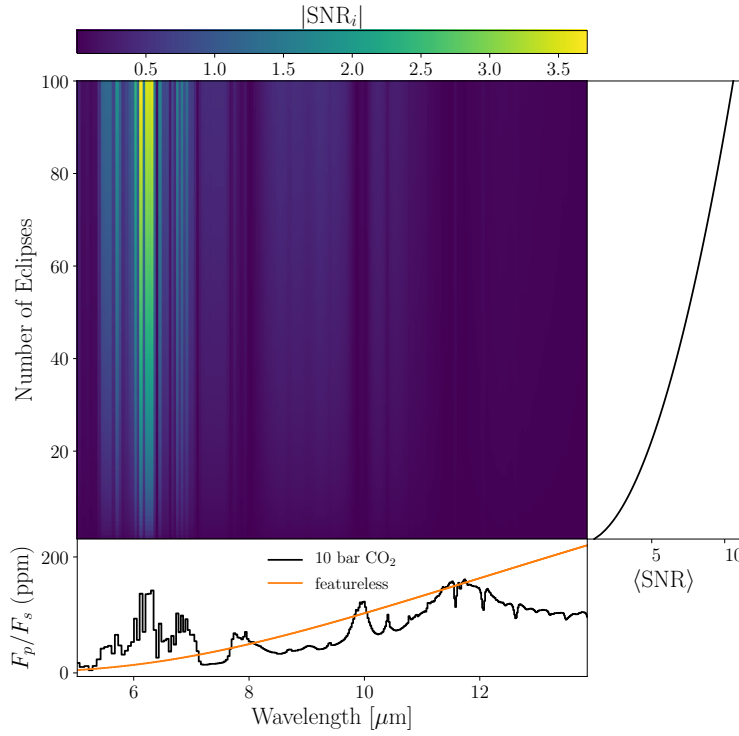


Figure 2.3 Signal-to-noise contours on the simulated 10 bar high CO_2 emission spectrum observed with JWST/MIRI/LRS relative to a featureless spectrum as a function of the number of occultations observed and wavelength of the instrument. The bottom panel shows the noiseless spectrum (black) and the best-fitting featureless spectrum (orange), both convolved to the instrument resolution. The right panel shows how the total expected SNR from Equation 2.6 increases with more eclipses observed.

spectrum. In this case, rather than a “flat line” test, where we attempt to rule out a featureless spectrum and thereby detect the presence of an atmosphere, here we attempt to rule out a spectrum that does not have a particular gas—for instance H_2O —and thereby detect the presence of H_2O . This procedure enables the identification of which JWST instruments and modes are sensitive to detecting individual molecules in an observed spectrum and how much time must be spent on any given target to reduce the noise enough to measure the spectral contributions from each molecule.

Figure 2.4 shows the detectability of O_2 in the atmosphere of TRAPPIST-1b with NIRSpec G140H if the planet possesses a desiccated 10 bar O_2 atmosphere. The strong O_2 - O_2 collisionally-induced absorption (CIA) features at 1.06 and $1.27 \mu\text{m}$ lead to a $\langle \text{SNR} \rangle = 5$ detection of O_2 in 7 transits. Therefore, an oxygen dominated atmosphere for TRAPPIST-1b could be *ruled out* by not detecting these features in 7 transits.

Unless otherwise stated, throughout the rest of the chapter we adopt the convention that molecules in the atmosphere are weakly detected if $\langle \text{SNR} \rangle \geq 3$ is achieved on that molecule’s contribution

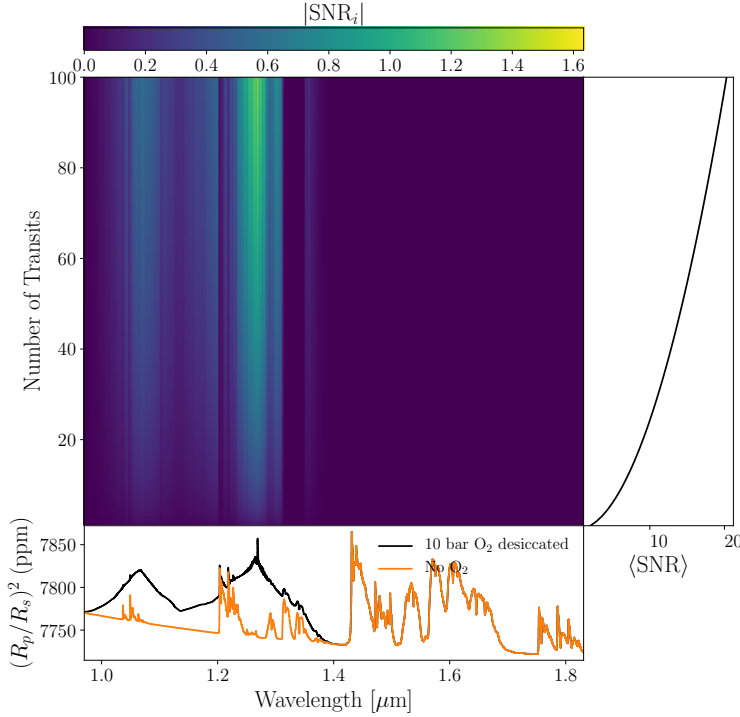


Figure 2.4 Signal-to-noise contours on the O_2 contribution to the transmission spectrum of TRAPPIST-1b if it possesses a 10 bar desiccated high O_2 atmosphere observed with NIRSpec G140H as a function of the number of occultations observed and wavelength of the instrument. The bottom panel shows the full model spectrum (black) and the model spectrum with O_2 removed (orange), both convolved to the resolution of NIRSpec G140H. The right panel shows the total expected SNR ($\langle \text{SNR} \rangle$) from Equation 2.6. The O_2 features at 1.06 and $1.27 \mu\text{m}$ are due to O_2 - O_2 collisionally-induced absorption (CIA), and could lead to $\langle \text{SNR} \rangle = 5$ detection of O_2 in ~ 7 transits.

to the spectrum, and we report our results for the characterization of atmospheres at this weak detection threshold. Keep in mind that we use a weaker threshold $\langle \text{SNR} \rangle$ to report detecting individual molecules than for simply detecting the atmosphere, but we encourage readers to scale our molecular detection results to their own desired thresholds.

2.3 Results

Here we present the full results of our simulations on the detectability and characterization of the TRAPPIST-1 exoplanet atmospheres using JWST. First, we assess the JWST observations needed to detect the presence of an atmosphere for the TRAPPIST-1 planets (§2.3.1). We then address the detectability of individual molecules within the TRAPPIST-1 planet spectra that may be used to distinguish between different atmospheric states and evolutionary scenarios (§2.3.2).

2.3.1 Detecting Atmospheres

We simulate the detectability of the TRAPPIST-1 planetary atmospheres with MIRI photometry (§2.3.1), transmission spectroscopy (§2.3.1), and emission spectroscopy (§2.3.1).

JWST/MIRI Photometry

MIRI photometry may be advantageous for initial assessments prior to the potentially long time commitment necessary to observe the spectrum of the Earth-sized TRAPPIST-1 planets with JWST. We investigate both transit and eclipse photometry with the nine MIRI photometric filters spanning wavelengths from about 5 to 27 μm (Bouchet et al., 2015). We first present results that may help to constrain the presence of an atmosphere from secondary eclipse observations in a *single* MIRI photometric band using brightness temperature arguments. We then present results for transit and eclipse observations in multiple filters, using the tests described in §2.2.3, to assess the observational requirements for ruling out a featureless spectrum.

Single-Band Constraints We calculate brightness temperatures for each of our self-consistent atmosphere models (Lincowski et al., 2018) to help plan and interpret photometric assessments of the TRAPPIST-1 planets in secondary eclipse. Figure 2.5 shows brightness temperature as a function of wavelength for TRAPPIST-1b for different assumed atmospheres. The wavelength-dependent fluxes were also convolved with the nine MIRI filters to calculate the brightness temperature of each atmosphere model, shown as color points in Fig. 2.5. The horizontal dashed line shows the zero bond albedo equilibrium temperature of TRAPPIST-1b—a limit which a planet without additional internal geothermal or atmospheric greenhouse heating would not be expected to exceed. Figure 2.5 also provides the SNR on the depth of a single observed secondary eclipse, averaged over the atmospheric models and displayed increasing from top to bottom on the right y-axis.

A few of the possible TRAPPIST-1b atmospheres, in particular the Venus-like and outgassing O_2 atmospheres, have brightness temperatures that exceed the zero bond albedo equilibrium tem-

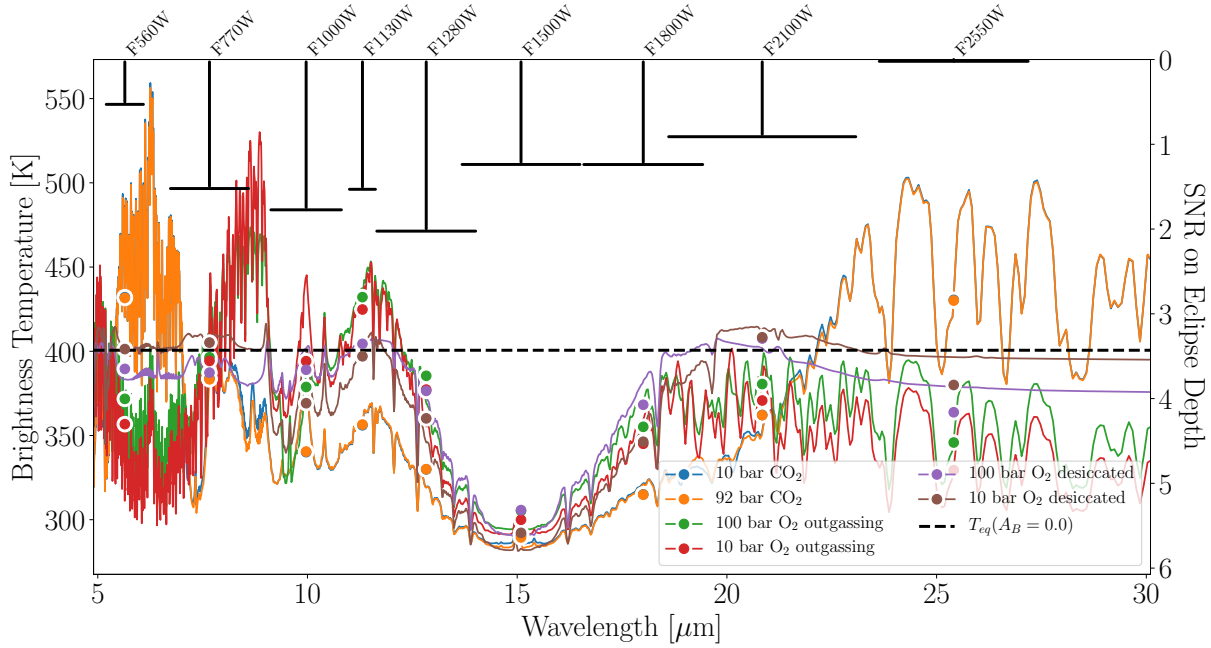


Figure 2.5 Brightness temperatures for different TRAPPIST-1b atmospheric models across the JWST/MIRI imaging wavelength range. The average brightness temperature integrated across each MIRI filter are shown as color points. The zero bond albedo equilibrium temperature of TRAPPIST-1b (black horizontal dashed line) is shown to compare against the model brightness temperatures, which vary as a function of wavelength due to the interplay between atmospheric temperature structure and gaseous opacities. The black lines from the top of the plot correspond to the (atmospheric model averaged) SNR on a single secondary eclipse, shown increasing from the top of the right y-axis to the bottom. These SNR indicators are shown capped with the effective width of each MIRI filter, and help to identify which filters may best offer secondary eclipse detections.

perature in a handful of the MIRI photometric bands. Like Venus, the 10 and 92 bar CO_2 atmospheres have 6 μm windows that provide a glimpse into their hotter, greenhouse heated, lower atmospheres. The F560W MIRI filter could potentially detect this emission. The 10 and 100 bar O_2 outgassing atmospheres also have strong emission windows near 11.5 μm , between the 9.6 μm O_3 band and the 15 μm CO_2 band, which could be detected with the F1130W MIRI filter. The 10 and 92 bar CO_2 atmospheres have sufficiently strong CO_2 absorption to saturate the wings of the 15 μm band and cause a significantly lower brightness temperature at 10-12 μm compared to atmospheres not dominated by CO_2 . The 10 and 92 bar CO_2 atmospheres also exceed the maximum

equilibrium brightness temperature beyond about $22\ \mu\text{m}$, however the F2550W MIRI filter may lack the SNR to provide constraining information.

Despite the few cases with potentially detectable high brightness temperatures, the overwhelming majority of atmospheric models viewed through MIRI filters have brightness temperatures that are consistent with a plausible planetary bond albedos (between 0 and 1), including in the 12-18 μm wavelength range where outgoing thermal radiation is strongly absorbed by the broad 15 μm CO₂ band. In the atmospheres that we considered, no *single* photometric band stands out above the rest as providing a definitive detection of an atmosphere via an emission window. However, using secondary eclipse observations in multiple MIRI filters in and out of the strong 15 μm CO₂ feature may be used to detect the presence of an atmosphere. We discuss this point later in this section.

Although Figure 2.5 only shows brightness temperatures for TRAPPIST-1b, Table 2.3 in the Chapter Appendix 2.6.3 lists the brightness temperature in each MIRI imaging filter for each TRAPPIST-1 planet atmosphere considered here. Table 2.3 also contains calculations for the Warm Spitzer photometric filters. We note, however, that the SNR on secondary eclipses decreases rapidly with planet equilibrium temperature, making precise eclipse photometry beyond TRAPPIST-1c largely infeasible with JWST.

Multi-Band Constraints We now turn to constraints that may be placed on the existence of an atmosphere using a combination of any of the nine MIRI photometric bands to observe either transits or eclipses. For each type of atmosphere, we determine the number and set of MIRI filters that can detect the atmosphere using transit and eclipse photometry in the minimum number of total occultations, assuming an equal number of occultations are observed in each filter. In all cases, 2-3 MIRI filters is optimal for detecting deviations from a featureless spectrum and additional filters are costly given their marginal increase in atmospheric detectability. The F1500W filter is always optimal to include due to the presence of CO₂ in these atmospheres. For transit photometry, the F1500W filter is best combined with F560W for CO₂-dominated atmospheres and F770W for O₂-dominated atmospheres, and typically just 2 filters is optimal. For eclipse photometry, the F1500W filter is best combined with F560W, F770W, and/or F1130W, and typically 3 filters is optimal.

The left panel of Fig. 2.6 shows the total number of transits (blue lines) and eclipses (red lines) needed to detect different atmospheric compositions for each of the seven known TRAPPIST-1 planets using the optimal 2 or 3 MIRI photometric filters. Plotting the number of occultations (transits or eclipses) as a function of the TRAPPIST-1 planets, ordered by semi-major axis, reveals a general trend according to the observation method: emission photometry is comparable with transmission photometry at detecting atmospheres for the innermost/hottest planets (e.g. TRAPPIST-1b and c), but becomes increasingly less efficient as the planets decrease in equilibrium temperature; whereas transmission photometry increases in observational time much more gradually with equilibrium temperature. This strong scaling with temperature occurs because the planet emission, at wavelengths contributing most to the detection of molecular features, is not in the Rayleigh-Jeans limit. The left panel of Fig. 2.6 implies that determining whether or not the potentially habitable TRAPPIST-1 planets (e, f, and g) have atmospheres will be much more efficient with transit photometry than eclipse photometry.

The high CO₂ atmospheres may be surprisingly difficult to distinguish with eclipse photometry because the wings of the 15 μm CO₂ band saturate and extend many microns on either side. This strong absorption effectively mutes the planet’s emitted flux, creating a nearly featureless spectrum. Consequently, MIRI eclipse photometry at 12.8 μm and 18 μm (F1280W and F1800W filters, respectively) may not be sufficiently separated from 15 μm (F1500W) to avoid substantial contamination from the wings of strong CO₂ absorption (see the orange curve in Fig. 2.5), making these filters ineffective for continuum measurement. Instead, the F1130W filter may be better at probing deeper into the atmosphere.

JWST Transmission Spectroscopy

We now present results on the detectability of the TRAPPIST-1 planet atmospheres using transmission spectroscopy with JWST. We begin with a few specific examples of detectable TRAPPIST-1 atmospheres to demonstrate the size of molecular absorption features relative to the expected JWST noise, and then present our full ensemble of results for each TRAPPIST-1 planet, atmosphere, and JWST instrument.

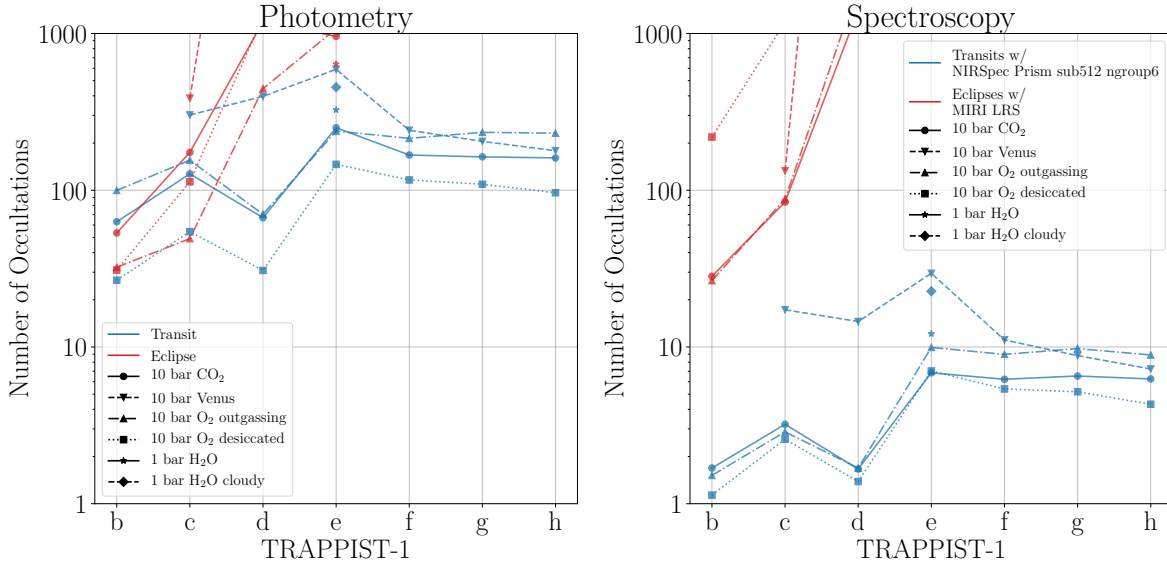


Figure 2.6 Number of transits (blue) or eclipses (red) needed to detect an atmosphere for all seven known TRAPPIST-1 planets with MIRI photometry (left) and JWST spectroscopy (right). Line and marker styles designate the type of atmosphere assumed. MIRI imaging assumes that the number of occultations are split evenly between the 2-3 photometric filters that are optimal to detect each atmosphere (see §2.3.1). Transmission spectroscopy is shown using NIRSpec Prism* (see §2.3.1) and emission spectroscopy is shown using MIRI LRS (see §2.3.1). Detecting the atmospheres of the TRAPPIST-1 planets is optimal using transmission spectroscopy with NIRSpec Prism across a range of terrestrial atmospheric compositions.

Specific Examples of Atmospheric Detection Figure 2.7 shows an example detection of absorption features in the transmission spectrum of TRAPPIST-1 b assuming it possesses a clear 10 bar CO_2 atmosphere. The transmission spectrum is shown with synthetic data simulated for two transits observed with NIRSpec Prism*, which we find to be sufficient to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ (our fiducial detection limit). The synthetic data are shown binned to a resolution of $R = 8$, however the featureless spectrum was ruled out at the native resolution of the NIRSpec Prism ($R \approx 100$). CO_2 absorption features at 1.6, 2.0, 2.8, 4.3 μm drive the detectability of this atmosphere and are apparent in the synthetic data.

Figure 2.8 demonstrates possible transmission spectra of TRAPPIST-1 c for two different aerosol conditions. Both cases are CO_2 -dominated spectral models from [Lincowski et al. \(2018\)](#)

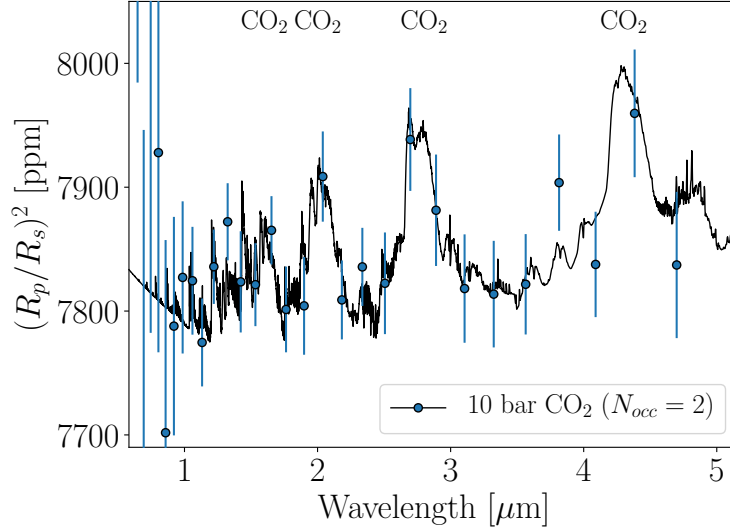


Figure 2.7 Model transmission spectrum of TRAPPIST-1b with a 10 bar CO_2 atmosphere. Synthetic data are simulated for two transits observed with NIRSpec Prism* and binned to a resolution of $R = 8$.

with simulated noise for observations with NIRSpec Prism*. The number of transits observed for each model is calculated such that the atmosphere is strongly detected with $\langle \text{SNR} \rangle = 10$. The blue model shows a clear sky spectrum with large CO_2 absorption features and data uncertainties calculated for 16 observed transits. The red model includes H_2SO_4 clouds at altitudes consistent with H_2SO_4 condensation, with data uncertainties calculated for 72 observed transits.

The aerosol-free atmosphere has strong CO_2 features that can be detected in considerably fewer transits than the case with Venus-like H_2SO_4 aerosols. The self-consistent clouds are located high enough in the atmosphere that over 70 transits are required to detect the CO_2 absorption features with the same confidence as the clear sky case. If TRAPPIST-1 c possess Venus-like aerosols, then about $5\times$ more transits must be observed to achieve comparable constraints on the presence and composition of TRAPPIST-1 c's atmosphere, than our estimates for the clear-sky case.

We find a factor of ~ 4 variation in the number of transits needed to detect the atmosphere of TRAPPIST-1e depending on the type of terrestrial atmosphere it possesses. Figure 2.9 explores the detectability of molecular features in the transmission spectrum of TRAPPIST-1 e. Like Fig. 2.8,

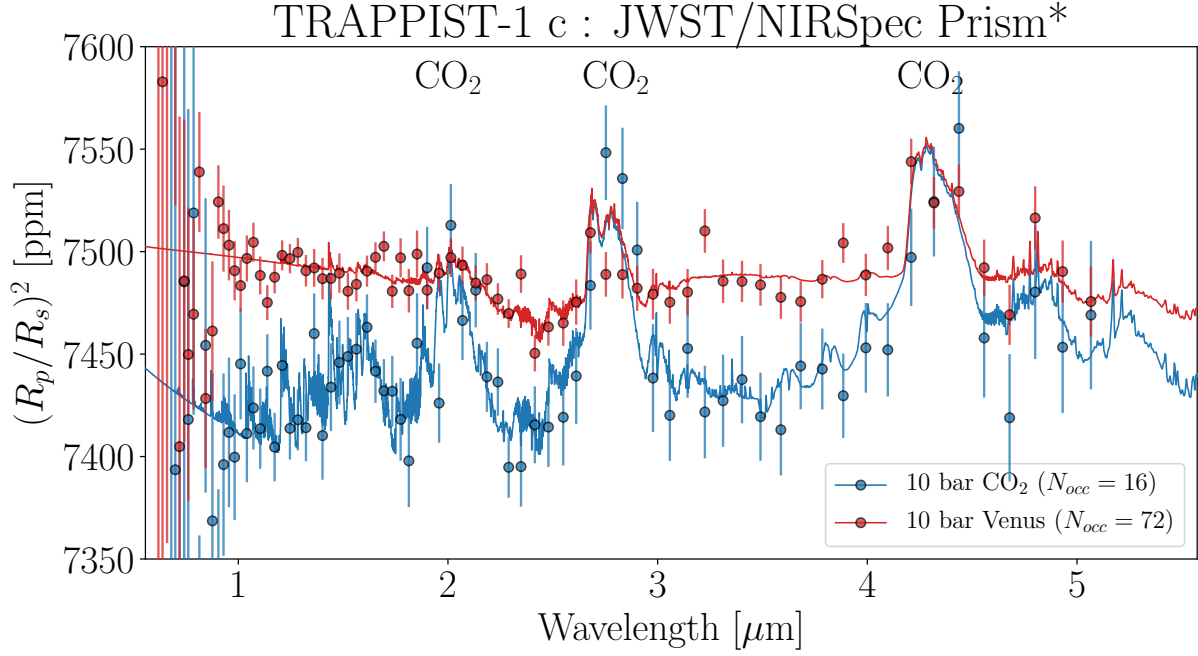


Figure 2.8 Simulated JWST/NIRSpec Prism transmission spectra of TRAPPIST-1c for different CO_2 -dominated atmospheric states. The blue model shows a clear sky spectrum with large CO_2 absorption features and the red model includes a self-consistent H_2SO_4 cloud. The number of coadded transits simulated for each NIRSpec observation are set so that the atmosphere is strongly detected ($\langle \text{SNR} \rangle = 10$). That is, to obtain approximately equal constraints on the presence and composition of TRAPPIST-1c’s atmosphere will require 16 transits for the clear sky CO_2 case and 72 transits for the cloudy case.

Fig. 2.9 shows possible TRAPPIST-1 e model spectra with simulated NIRSpec Prism* observations assuming the number of transits needed to strongly detect features in each atmosphere with $\langle \text{SNR} \rangle \approx 10$. Spectra for water-covered environments with and without a water cloud are shown in teal and blue, respectively, and CO_2 -dominated atmospheres, with and without H_2SO_4 aerosols are shown in orange and red, respectively. To detect features in each spectrum with $\langle \text{SNR} \rangle \approx 10$ ($\langle \text{SNR} \rangle \approx 5$), and thereby obtain approximately equal constraints on the presence of TRAPPIST-1e’s atmosphere, will require ~ 30 (~ 7) transits for a clear sky CO_2 atmosphere, ~ 50 (~ 13) transits for a clear sky aqua planet atmosphere, ~ 90 (~ 22) transits for a cloudy aqua planet atmosphere, and ~ 120 (~ 30) transits for a CO_2 -dominated atmosphere with H_2SO_4 clouds.

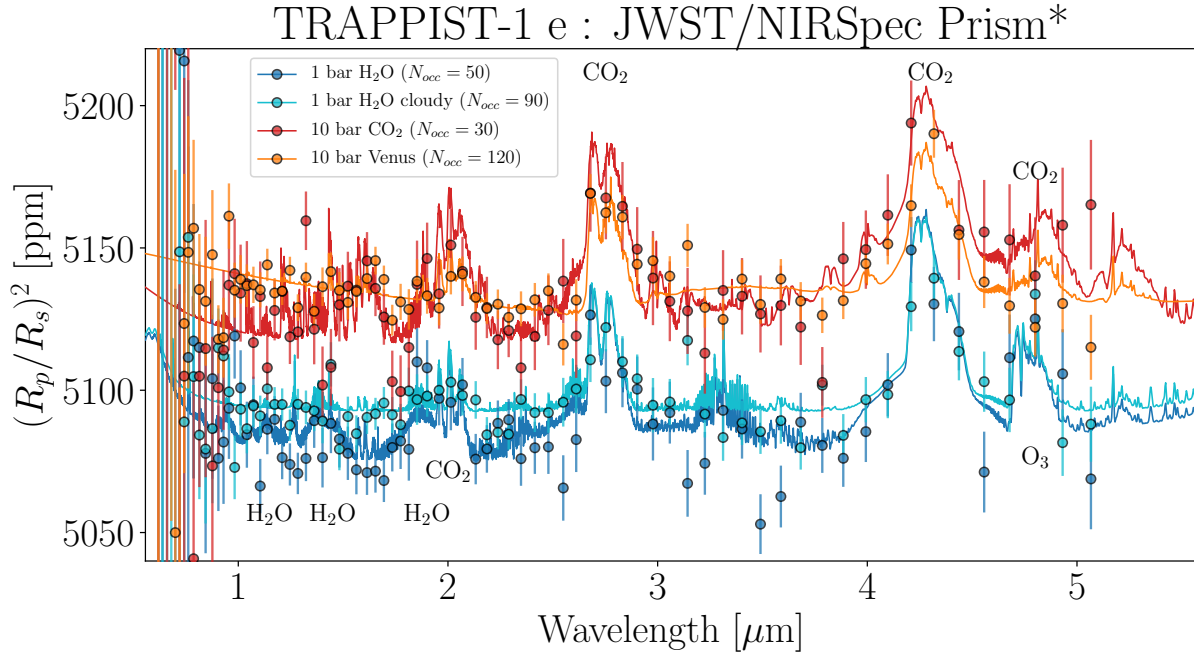


Figure 2.9 Simulated JWST/NIRSpec Prism* transmission spectra of TRAPPIST-1e for different possible atmospheric compositions. The teal and blue models show spectra for water-dominated atmospheres, with and without a water cloud, respectively. The orange and red spectra are for CO₂-dominated atmospheres, with and without H₂SO₄ aerosols, respectively. The number of coadded transits simulated for each NIRSpec observation are set so that the atmosphere is strongly detected ($\langle \text{SNR} \rangle = 10$). That is, to obtain approximately equal constraints on the presence of TRAPPIST-1e's atmosphere will require ~ 30 transits for a clear sky CO₂ atmosphere, ~ 50 transits for a clear sky H₂O atmosphere, ~ 90 transits for a cloudy H₂O atmosphere, and ~ 120 transits for a CO₂-dominated Venus-like atmosphere.

Optimal JWST Observing Modes for Atmospheric Detection Thus far we have only discussed specific cases for the detectability of the TRAPPIST-1 planet atmospheres with JWST transmission spectroscopy; we now report results from our comprehensive study to determine the exposure times needed detect the presence of atmospheres as a function of JWST observing mode and atmosphere type, for each TRAPPIST-1 planet. The number of transits necessary to detect spectral features in the transmission spectrum with $\langle \text{SNR} \rangle = 5$ for TRAPPIST-1 b, c, d, e, f, g, and h are shown in Chapter Appendix 2.6.2 in Figures 2.13, 2.14, 2.15, 2.16, 2.17, 2.18, and 2.19, respectively, as a function of both atmospheric compositions and JWST instrument/mode. Color is used in these

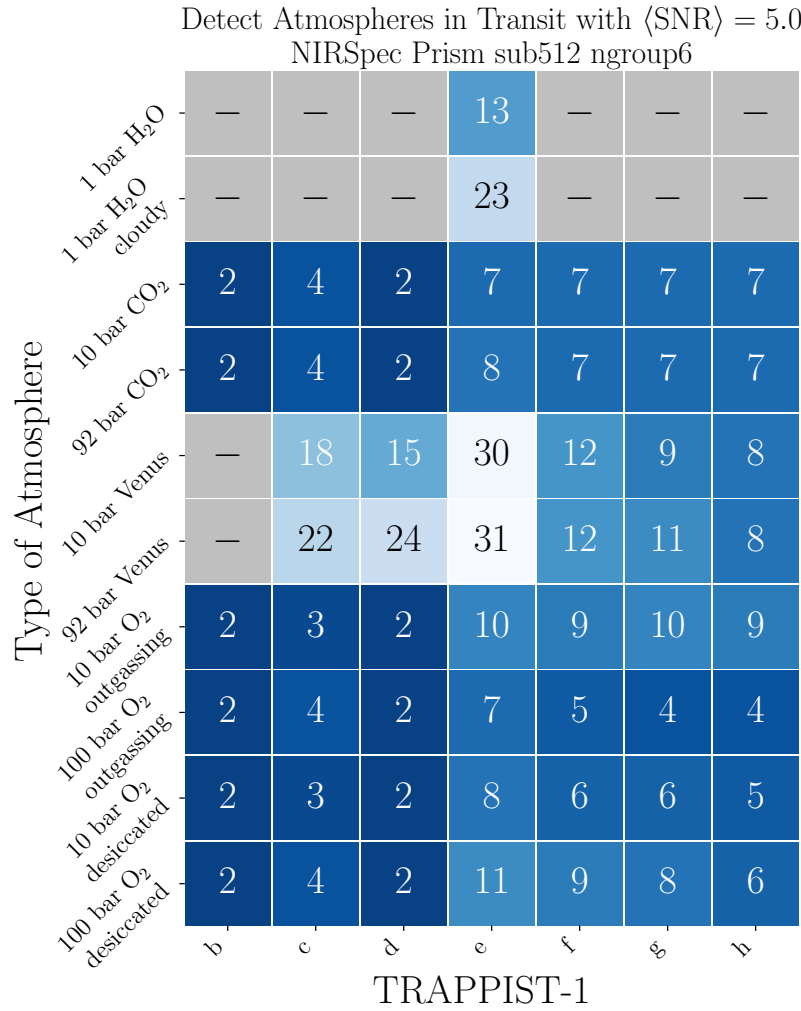


Figure 2.10 Number of transits for each TRAPPIST-1 planet necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different self-consistent atmospheric compositions using JWST NIRSpec Prism with the optimized readout mode of Batalha et al. (2018).

figures to guide the eye to small (more blue) and large (more white) values for the number of transits.

These results suggest that under most circumstances the JWST/NIRSpec Prism is the optimal instrument for detecting the presence of an atmosphere using transmission spectroscopy of the TRAPPIST-1 planets. If a partial saturation strategy is used with the NIRSpec Prism (Batalha et al., 2018), then it will be more capable of detecting atmospheric features than any other JWST instrument or mode. However, if it turns out that the systematics introduced via partial saturation are not beneficial, the SUB512S subarray will offer improved performance over the SUB512 sub-

array due to its shorter readout time. Although we note that SUB512 should be considered the subarray of choice because SUB512s has limited access to important background pixels and it may be difficult to keep the curved trace inside only 16 pixels in the cross-dispersion direction. The NIRSpec G395M/H disperser offers comparable results with the standard (no partial saturation) NIRSpec Prism.

Most of the atmospheres that we consider may require fewer than 12 transits to detect for all of the TRAPPIST-1 planets using the optimized NIRSpec Prism mode. Figure 2.10 displays the number of transits to detect the TRAPPIST-1 planet atmospheres using only the NIRSpec Prism* instrument for each TRAPPIST-1 planet and for each atmosphere considered in this work.

Our results show that the outer TRAPPIST-1 planets require only 2-7 more transits than TRAPPIST-1 b to detect clear-sky atmospheres. The right panel of Fig. 2.6 summarizes the best-case-scenario results for detecting the atmospheres of the TRAPPIST-1 planets with JWST transit spectroscopy (blue lines). The number of transits required to detect the atmosphere at $\langle \text{SNR} \rangle = 5$ for each of the atmospheric models is shown for observations with NIRSpec Prism*. Like our transit photometry results, the number of transits only weakly increases with semi-major axis. However, significantly fewer transits are required for transmission spectroscopy atmosphere detections with NIRSpec Prism* than with MIRI filter photometry. The emission spectroscopy results in Fig. 2.6 (red lines) are discussed in §2.3.1.

The atmospheric composition of the planets has a relatively minimal effect on the number of transits needed to rule out a featureless spectrum, with Venus-like clouds being a significant exception. All atmospheric cases considered for TRAPPIST-1 c could be detected in 3-4 transits with NIRSpec Prism*, except for the 10 and 92 bar Venus atmospheres (with H_2SO_4 aerosols), which would require $4.5\times$ and $5.5\times$ the number of transits to detect, respectively, compared to the 10 and 92 bar clear-sky CO_2 counterparts. In general, more transits tend to be required to detect the atmospheres of the cooler worlds, except for TRAPPIST-1 d, which is more comparable to TRAPPIST-1 b in the detectability of its atmosphere. While the atmosphere of TRAPPIST-1 d may be relatively easy to detect if it is without clouds, with clouds the 10 and 92 bar CO_2 atmospheres require $7.5\times$ and $12\times$ the number of transits to detect, respectively, the largest increase due to

clouds seen in the sample. Beyond TRAPPIST-1 e the effect of Venus-like clouds has a diminished impact on the atmospheric detectability, with fewer than $2\times$ the number of transits required to detect the atmospheres for TRAPPIST-1 f, g, and h if they have clouds.

Note that the numerical values in Figure 2.10 and Figures 2.13 - 2.19 can easily be scaled to higher or lower $\langle \text{SNR} \rangle$ thresholds that more or less confidently rule out a featureless spectrum. Since the SNR on an observation (and $\langle \text{SNR} \rangle$) scales with the square-root of the exposure time, and the number of occultations is a proxy for exposure time, we can obtain a new value for the number of occultations:

$$N'_{\text{occ}} = N_{\text{occ}} \left(\frac{\langle \text{SNR} \rangle'}{\langle \text{SNR} \rangle} \right)^2 \quad (2.10)$$

where N_{occ} is the number of occultations necessary to distinguish features in the spectrum with $\langle \text{SNR} \rangle$ (5 in Figures 2.10, 2.13 - 2.19) and $\langle \text{SNR} \rangle'$ is the new signal-to-noise threshold.

The increased transit duration with semi-major axis make our results for the outer TRAPPIST-1 planets appear more optimistic relative to our results for the inner planets. Note that for JWST planning, our reported number of transits/eclipses leads to different telescope time for each TRAPPIST-1 planet since each has a different transit duration. For each observed transit (eclipse) we assumed one transit duration worth of out-of-transit (out-of-eclipse) observing time. Considering the following median transit durations for the TRAPPIST-1 planets from Grimm et al. (2018): 36.40 mins for b, 42.37 mins for c, 49.13 mins for d, 57.21 mins for e, 62.60 mins for f, 68.40 mins for g, and 76.7 mins for h, one can calculate expected on-target time per transit/eclipse (before overheads) by multiplying the transit duration by two plus any overheads.

JWST is not always able to point at TRAPPIST-1 due to the star's proximity to the ecliptic plane. As a result, the star will only be observable to JWST for ~ 100 days per year⁴. Over the course of JWST's nominal 5 year mission the maximum number of observable transits/eclipses is approximately: 331 for b, 206 for c, 123 for d, 81 for e, 54 for f, 40 for g, and 26 for h. Of course, an extended mission lifetime would allow considerably more observations of the TRAPPIST-1 system.

⁴<https://jwst-docs.stsci.edu/display/JTI/JWST+Target+Viewing+Constraints>

JWST Emission Spectroscopy

Unlike transmission spectroscopy, which can make use of many JWST instruments that span a broad wavelength range (e.g. NIRCam, NIRSpec, NIRISS, and MIRI LRS), secondary eclipse spectroscopy of the TRAPPIST-1 planets will only be viable at the longer wavelengths accessible to JWST. This is primarily driven by the increase in eclipse depths with wavelength as the planet-star contrast ratio becomes more favorable.

We find that only MIRI LRS observations of TRAPPIST-1 b and c may potentially rule out a featureless emission spectrum with $\langle \text{SNR} \rangle = 5$ in fewer than 100 observed secondary eclipses. Using an analogous approach to that shown in Figures 2.13 - 2.19, MIRI LRS observations of TRAPPIST-1 b may rule out a featureless emission spectrum in 27-47 secondary eclipses if the planet possess a (10 or 92 bar) clear-sky CO_2 atmosphere or a (10 or 100 bar) outgassing O_2 atmosphere, with the 10 bar outgassing O_2 atmosphere the most readily detectable. Over 100 secondary eclipses are required if the planet possesses a desiccated O_2 atmosphere (10 or 100 bars), as these emission spectra appear remarkably featureless between 5-9 μm . For TRAPPIST-1 c we find that MIRI LRS observations could rule out a featureless emission spectrum in 85-100 secondary eclipses if the planet possesses a (10 or 92) bar clear CO_2 atmosphere or a 10 bar O_2 atmosphere with outgassing. Over 100 secondary eclipses are required for all other atmospheric compositions considered for TRAPPIST-1 c, including the CO_2 atmospheres with H_2SO_4 clouds, which have emission spectra similar to the clear CO_2 atmospheres, but with reduced temperature contrasts in the absorbing and emitting spectral regions that effectively mute the features and drive the spectrum towards a featureless blackbody. All of the exterior TRAPPIST-1 planets have emission spectra that will appear indistinguishable from cool blackbodies due to insufficient SNR.

The right panel of Fig. 2.6 summarizes the best-case-scenario results for detecting the atmospheres of the TRAPPIST-1 planets with JWST emission spectroscopy (red lines) to compare against our filter photometry and transmission spectroscopy results. The number of eclipses required to detect the atmosphere at $\langle \text{SNR} \rangle = 5$ for each of the atmospheric models is shown for observations with MIRI LRS. Compared with transmission spectroscopy, emission spectroscopy

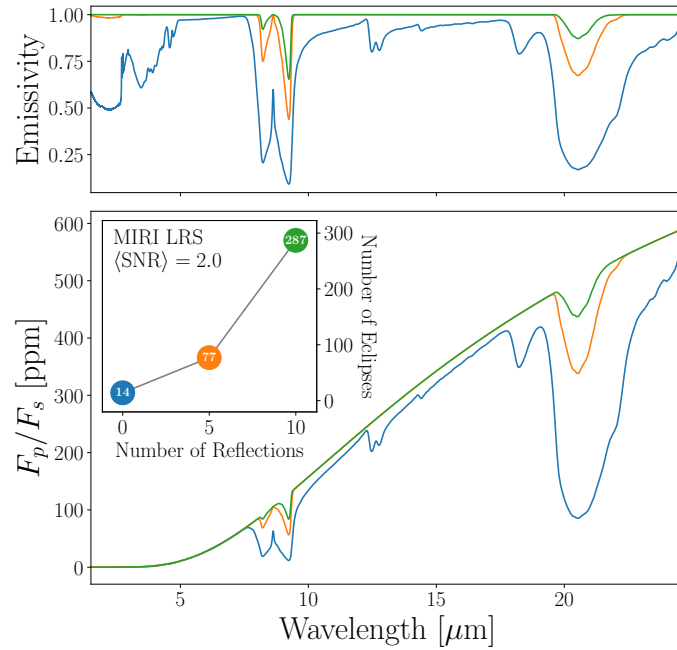


Figure 2.11 Estimated detectability of pure quartz surface emissivity features in the secondary eclipse spectrum of TRAPPIST-1b with JWST’s MIRI LRS instrument. The top panel shows the MIR effective emissivity of a quartz surface for three different assumed values for the number of surface reflections (see Equation 2.9): 0 (blue), 5 (orange), and 10 (green). The bottom panel shows theoretical secondary eclipse spectra for TRAPPIST-1b assuming the effective emissivities from the top panel. The inset axis shows the number of secondary eclipses that would be required to observe with MIRI LRS to detect features in the emission spectrum with a $\langle \text{SNR} \rangle = 2.0$, for each of the three assumed effective emissivities. Only an extremely polished pure quartz surface could possibly be detected with JWST.

is an inefficient method for *detecting* the TRAPPIST-1 planetary atmospheres, particularly for the cooler planets.

If TRAPPIST-1b does not possess an atmosphere, emissivity features from minerals on the surface of the planet are not likely to be detectable in secondary eclipse thermal emission spectra. Figure 2.11 shows the estimated thermal emission spectrum of TRAPPIST-1b if it is airless and possesses a pure quartz surface. Three different effective emissivities are shown in the top panel, corresponding to 0 (blue), 5 (orange), and 10 (green) surface reflections (see Equation 2.9). The bottom panel shows the MIR secondary eclipse spectrum for each surface emissivity, assuming

the surface is at the (highly optimistic) zero bond albedo equilibrium temperature for TRAPPIST-1b (392 K). The inset axis shows the number of secondary eclipses that MIRI LRS would have to observe to detect features in each emission spectrum at a weak $\langle \text{SNR} \rangle = 2.0$. Increasing the number of surface reflections results in surface emissivity features that are not detectable.

2.3.2 *Distinguishing Specific Molecules*

We now present the sensitivity of each JWST instrument to each gas in the [Lincowski et al. \(2018\)](#) TRAPPIST-1 model transmission spectra. Table 2.2 lists the molecules for which JWST could weakly detect ($\langle \text{SNR} \rangle = 3.0$) that molecule's contribution to the spectrum in 100 or fewer transits. For each molecule, the number of transits is listed in parentheses next to each molecular formula along with a footnote identifying which JWST instrument is used for that observation. While in some cases multiple instruments may be sensitive to the same molecule, we list only the instrument that can detect each gas in the minimum number of transits.

The presence of CO₂ dominates the detectability of all the atmospheres simulated in [Lincowski et al. \(2018\)](#) with JWST. Even a relatively small amount of CO₂ (e.g. ~ 290 ppm for the 1 bar H₂O TRAPPIST-1 e) can saturate the strong 2.7, 4.3, 15 μm CO₂ absorption features and lead to the detection of both the atmosphere and CO₂. As a result, the number of transits necessary to detect spectral features in a transmission spectrum, given in Section 2.3.1, are close to the number of transits necessary to detect CO₂. In some cases, the number of transits to detect CO₂ is fewer than that needed to simply detect the atmosphere. This is because the spectral model without CO₂ deviates more significantly from the true spectrum than the best-fitting featureless spectrum. In these cases it is important to defer to the number of transits required to rule out a featureless spectrum, because it provides a more realistic fit to the spectrum.

The inner TRAPPIST-1 planets may have several detectable molecules that can be used to distinguish between different evolutionary scenarios. H₂O and SO₂ may be marginally detectable with MIRI LRS transmission spectra of the inner TRAPPIST-1 planets—b, c, and d— if they possess clear sky CO₂ atmospheres. However, if these planets possess H₂SO₄ clouds then the H₂O and SO₂ may be undetectable. Alternatively, if TRAPPIST-1 b, c, and d have oxygen-dominated

Table 2.2. Detectable molecules with transmission spectroscopy for different plausible TRAPPIST-1 planet atmospheres

Planet	Model	Molecules (Number of transits to $\langle \text{SNR} \rangle = 3.0$)
T-1b	10 bar CO ₂	CO ₂ (1) ^k , CO(47) ^k , H ₂ O(44) ^f , SO ₂ (40) ^f
	92 bar CO ₂	CO ₂ (1) ^k , CO(47) ^k , H ₂ O(41) ^f , SO ₂ (38) ^f
	100 bar O ₂ outgassing	CO ₂ (1) ^k , H ₂ O(2) ^k , O ₃ (11) ^f , O ₂ (3) ^f
	10 bar O ₂ outgassing	CO ₂ (1) ^k , H ₂ O(1) ^k , O ₂ (4) ^f
	100 bar O ₂ desiccated	CO ₂ (1) ^k , O ₂ (3) ^f
	10 bar O ₂ desiccated	CO(13) ^k , CO ₂ (1) ^k , O ₃ (81) ^f , O ₂ (3) ^f
T-1c	10 bar CO ₂	CO ₂ (1) ^k , H ₂ O(45) ^f , SO ₂ (51) ^f
	92 bar CO ₂	CO ₂ (1) ^k , H ₂ O(43) ^f , SO ₂ (52) ^f
	100 bar O ₂ outgassing	CO ₂ (3) ^k , H ₂ O(4) ^k , O ₃ (5) ^f , O ₂ (7) ^f
	10 bar O ₂ outgassing	CO ₂ (2) ^k , H ₂ O(2) ^k , O ₃ (86) ^f , O ₂ (5) ^f
	100 bar O ₂ desiccated	CO ₂ (2) ^k , O ₃ (71) ^f , O ₂ (7) ^f
	10 bar O ₂ desiccated	CO(61) ^k , CO ₂ (1) ^k , O ₃ (64) ^f , O ₂ (6) ^f
	10 bar Venus	CO ₂ (3) ^k
	92 bar Venus	CO ₂ (6) ^k
T-1d	10 bar CO ₂	CO ₂ (1) ^k , CO(48) ^k , H ₂ O(18) ^f , SO ₂ (27) ^f
	92 bar CO ₂	CO ₂ (1) ^k , CO(47) ^k , H ₂ O(17) ^f , SO ₂ (27) ^f
	100 bar O ₂ outgassing	CO ₂ (2) ^k , H ₂ O(2) ^k , O ₃ (2) ^k , O ₂ (3) ^f
	10 bar O ₂ outgassing	CO ₂ (1) ^k , H ₂ O(2) ^k , O ₃ (32) ^f , O ₂ (3) ^f
	100 bar O ₂ desiccated	CO ₂ (1) ^k , O ₃ (26) ^f , O ₂ (3) ^f
	10 bar O ₂ desiccated	CO(49) ^k , CO ₂ (1) ^k , O ₃ (30) ^f , O ₂ (3) ^f
	10 bar Venus	CO ₂ (6) ^k
	92 bar Venus	CO ₂ (7) ^k
T-1e	10 bar CO ₂	CO ₂ (2) ^k , SO ₂ (86) ^f
	92 bar CO ₂	CO ₂ (2) ^k , SO ₂ (88) ^f
	100 bar O ₂ outgassing	CO ₂ (9) ^k , H ₂ O(64) ^k , O ₃ (4) ^k , O ₂ (22) ^f
	10 bar O ₂ outgassing	CO ₂ (4) ^k , H ₂ O(78) ^k , O ₃ (72) ^f , O ₂ (24) ^f
	100 bar O ₂ desiccated	CO ₂ (4) ^k , O ₃ (61) ^f , O ₂ (18) ^f
	10 bar O ₂ desiccated	CO ₂ (3) ^k , O ₃ (40) ^f , O ₂ (19) ^f
	10 bar Venus	CO ₂ (9) ^k
	92 bar Venus	CO ₂ (9) ^k
	1 bar H ₂ O	CO ₂ (9) ^k , H ₂ O(15) ^k , N ₂ (68) ^f
	1 bar H ₂ O cloudy	CO ₂ (13) ^k
T-1f	10 bar CO ₂	CO ₂ (2) ^k , SO ₂ (76) ^f
	92 bar CO ₂	CO ₂ (2) ^k , SO ₂ (79) ^f
	100 bar O ₂ outgassing	CO ₂ (7) ^k , O ₃ (2) ^k , O ₂ (23) ^f
	10 bar O ₂ outgassing	CO ₂ (4) ^k , O ₃ (39) ^f , O ₂ (30) ^f
	100 bar O ₂ desiccated	CO ₂ (4) ^k , O ₃ (10) ^f , O ₂ (20) ^f
	10 bar O ₂ desiccated	CO ₂ (3) ^k , O ₃ (7) ^f , O ₂ (20) ^f
	10 bar Venus	CO ₂ (3) ^k
	92 bar Venus	CO ₂ (3) ^k
T-1g	10 bar CO ₂	CO ₂ (2) ^k , SO ₂ (82) ^f
	92 bar CO ₂	CO ₂ (2) ^k , SO ₂ (89) ^f
	100 bar O ₂ outgassing	CO ₂ (9) ^k , O ₃ (2) ^k , O ₂ (36) ^f
	10 bar O ₂ outgassing	CO ₂ (4) ^k , O ₃ (29) ^f , O ₂ (43) ^f
	100 bar O ₂ desiccated	CO ₂ (5) ^k , O ₃ (6) ^f , O ₂ (28) ^f
	10 bar O ₂ desiccated	CO ₂ (3) ^k , O ₃ (5) ^k , O ₂ (28) ^f
	10 bar Venus	CO ₂ (3) ^k
	92 bar Venus	CO ₂ (3) ^k
T-1h	10 bar CO ₂	CO ₂ (2) ^k , SO ₂ (85) ^f
	92 bar CO ₂	CO ₂ (2) ^k , SO ₂ (86) ^f
	100 bar O ₂ outgassing	CO ₂ (9) ^k , O ₃ (1) ^k , O ₂ (32) ^f
	10 bar O ₂ outgassing	CO ₂ (4) ^k , SO ₂ (54) ^f , O ₃ (23) ^f , O ₂ (31) ^f
	100 bar O ₂ desiccated	CO ₂ (4) ^k , O ₃ (3) ^k , O ₂ (22) ^f
	10 bar O ₂ desiccated	CO ₂ (3) ^k , O ₃ (4) ^k , O ₂ (22) ^f
	10 bar Venus	CO ₂ (2) ^k
	92 bar Venus	CO ₂ (2) ^k

^cNIRSpec G140H

^dNIRSpec G235H

^eNIRSpec G395H

^fMIRI LRS

^kNIRSpec Prism sub512 ngroup6

atmospheres then O_2 should be distinguishable via $\text{O}_2\text{-O}_2$ (O_4) CIA features at 1.06 and 1.27 μm , which are slightly more detectable with NIRSpec G140M/H than the NIRSpec Prism*. An oxygen-dominated planet with outgassing may be distinguished from a completely desiccated world by detecting H_2O , which is only readily detectable in our models of O_2 planets with outgassing.

TRAPPIST-1 e offers an opportunity to characterize a planet in the habitable zone with JWST. Like the inner TRAPPIST-1 planets, O_2 may be detectable in the spectrum of an oxygen-dominated atmosphere via the $\text{O}_2\text{-O}_2$ CIA features. However, modern Earth levels of O_2 and O_3 , and Earth geologic levels of CH_4 in a 1 bar N_2 -dominated atmosphere are not likely to be detectable. If TRAPPIST-1 e possesses such a habitable environment without clouds, tropospheric H_2O may be detectable using the NIRSpec Prism*. However, with full cloud coverage the detectability of H_2O strongly diminishes and becomes unobservable. The broad 4.3 μm $\text{N}_2\text{-N}_2$ CIA feature may be marginally detectable in a 1 bar habitable N_2 -dominated atmosphere with NIRSpec G395M/H. The 9.6 μm O_3 feature would require just over 100 transits with MIRI LRS to detect at $\langle\text{SNR}\rangle = 3.0$ in the clear sky 1 bar N_2 -dominated atmosphere; approximately twice the number of transits would be required if the planet possesses 100% water cloud coverage.

For the outer planets—TRAPPIST-1 f, g, and h—modest atmospheric characterization with transmission spectroscopy may be possible. In particular, oxygen-dominated atmospheres may be distinguished not only by their prominent $\text{O}_2\text{-O}_2$ CIA features, but also by their O_3 features, which become more detectable with the NIRSpec Prism* than the $\text{O}_2\text{-O}_2$ features for the coolest TRAPPIST-1 planets.

The observational difficulty with which individual molecules may be detected in a transmission spectrum—bulk atmospheric constituents or trace gases—varies substantially as a function of atmospheric composition. Consequently, the gases that are both relatively easy to detect and unique to a specific atmosphere make optimal testable hypotheses for that atmospheric composition. Figure 2.12 demonstrates a potential approach for distinguishing between three different atmospheric states for TRAPPIST-1 b using JWST transmission spectroscopy. The models are shown with calculated error bars that correspond to the amount of JWST observing time that would be required to detect specific molecules in the given spectrum.

In the top panel of Fig. 2.12, the spectrum of a 10 bar desiccated oxygen atmosphere displays prominent O_4 features that may be detected with $\langle \text{SNR} \rangle \sim 5$ in 12 transits with NIRSpec G140H. This is a strong discriminant between an O_2 -dominated atmosphere and a CO_2 -dominated atmosphere, both of which have strong and detectable CO_2 features.

The spectrum of a 10 bar oxygen atmosphere with outgassing is shown in the middle panel of Fig. 2.12. This spectrum has similarly detectable O_4 features to the desiccated atmosphere, but also substantial H_2O features (due to Earth levels of geological fluxes) that may be detected with $\langle \text{SNR} \rangle \sim 5$ in just 3 transits with NIRSpec Prism*. Such strong water absorption in an oxygen-dominated atmosphere would indicate incomplete desiccation.

The spectrum of a 10 bar CO_2 atmosphere is shown in the bottom panel of Fig. 2.12. This spectrum contrasts with the outgassing oxygen atmosphere in the detectability of H_2O . Water in a Venus-like atmosphere is scarce, particularly in the upper atmosphere, which would require ~ 125 transits with either NIRSpec Prism* or MIRI LRS to detect with $\langle \text{SNR} \rangle \sim 5$. The added benefit of making this costly observation with MIRI LRS is that SO_2 could also be detected with $\langle \text{SNR} \rangle \sim 5$ from the 7.3 and 8.7 μm features. Nonetheless, detecting any gases other than CO_2 in a CO_2 -dominated atmosphere will be very difficult.

2.4 Discussion

Our results indicate that JWST observations may be able to place strong constraints on the presence of high mean molecular weight terrestrial atmospheres for all of the TRAPPIST-1 planets and, in some cases, detect individual touchstone molecules that may be used to distinguish between different evolutionary histories that the planets may have undergone.

2.4.1 Do the TRAPPIST-1 planets have atmospheres?

For detecting the presence of atmospheres on the TRAPPIST-1 planets, we find that NIRSpec Prism observations are optimal. Transmission spectroscopy with the NIRSpec Prism could lead to a $\langle \text{SNR} \rangle = 5$ detection of atmospheric spectral features in as few as 2-11 transits for TRAPPIST-1

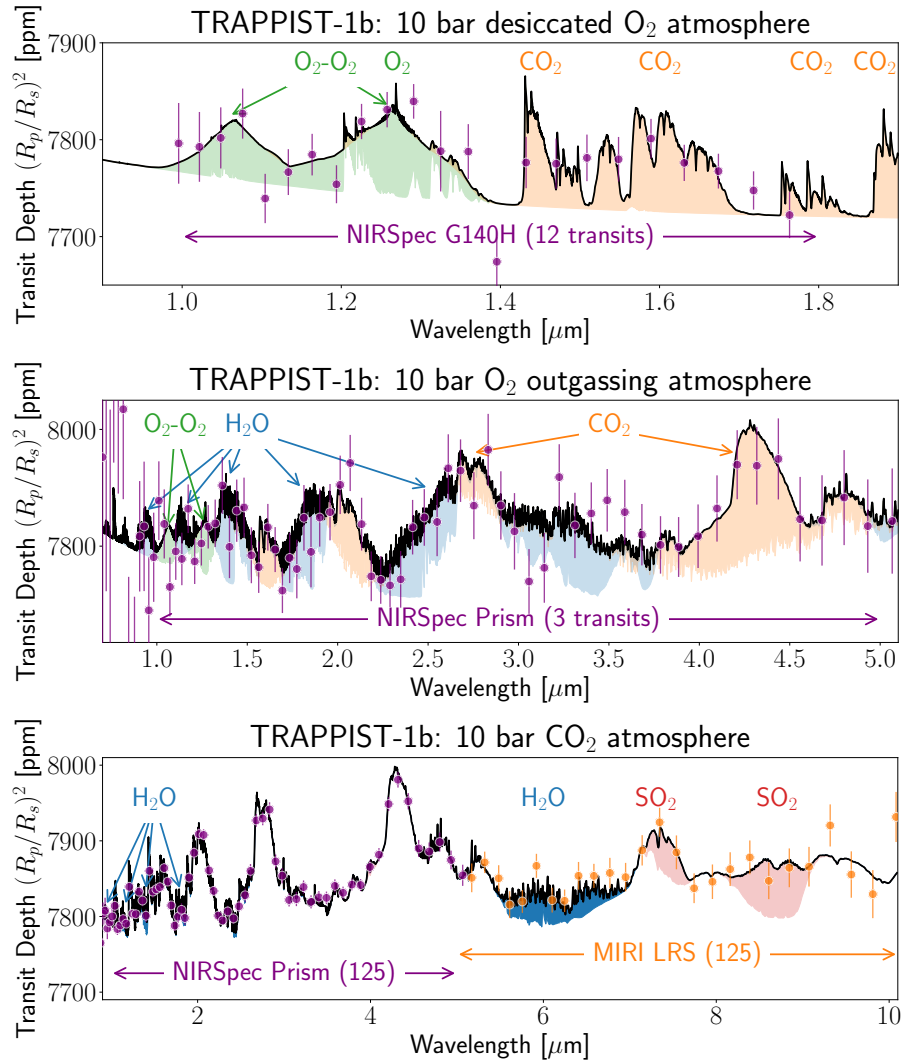


Figure 2.12 Theoretical transmission spectra of TRAPPIST-1 b assuming three different atmospheric compositions with modeled noise for JWST observations. *Top:* Transmission spectrum of a 10 bar desiccated O_2 atmosphere shown with error bars calculated for 12 transits with NIRSpec G140H—sufficient for $\langle\text{SNR}\rangle \sim 5$ on the O_4 features. *Middle:* Transmission spectrum of a 10 bar outgassing O_2 atmosphere shown with error bars calculated for 3 transits with NIRSpec Prism*—sufficient for $\langle\text{SNR}\rangle \sim 5$ on the H_2O features. *Bottom:* Transmission spectrum of a 10 bar CO_2 atmosphere shown with error bars calculated for 125 transits with NIRSpec Prism*—sufficient for $\langle\text{SNR}\rangle \sim 5$ on the NIR H_2O features—and 125 transits with MIRI LRS—sufficient for $\langle\text{SNR}\rangle \sim 5$ on both the 6 μm H_2O feature and the 7.3 and 8.7 μm SO_2 features.

b out to TRAPPIST-1 h if the planets lack high altitude aerosols. However, if the TRAPPIST-1 planets possess Venus-like H_2SO_4 aerosols, reaching the same constraints on the presence of atmospheres may require up to 12 times more transits. CO_2 possesses numerous strong absorption bands from the near- through the mid-IR, such as the 2.0, 2.7, 4.2, and 15 μm bands, which significantly contribute to the ability for JWST to detect the terrestrial atmospheres considered in this work.

Our self-consistent Venus-like H_2SO_4 aerosol modeling reveals trends in the detectability of such atmospheres with semi-major axis. [Lincowski et al. \(2018\)](#) found that TRAPPIST-1 b was too hot for Venus-like aerosols to form in the atmosphere, but that H_2SO_4 aerosols could form in all six exterior planets. These aerosols form at high altitudes in the atmospheres of TRAPPIST-1 c, d, and e and lead to muted CO_2 features that will require about 4-12 times more transits to be detected with JWST. However, TRAPPIST-1 f, g, and h are cool enough for the H_2SO_4 aerosols to form at lower altitudes ([Lincowski et al., 2018](#)), and therefore contribute less to their observable transmission spectra such that their atmospheres may be detected in fewer than 2 times the number of transits compared to the clear sky CO_2 cases for the same planets. Cloudy Venus-like atmospheres for TRAPPIST-1 f, g, and h require fewer transits to detect than cloudy Venus-like atmospheres for TRAPPIST-1 c, d, and e.

Secondary eclipse spectroscopy may require a significantly greater JWST time commitment than transmission spectroscopy to achieve comparable constraints on the detection of the TRAPPIST-1 planet atmospheres. Because of its access to longer wavelengths, MIRI LRS is the only JWST instrument capable of observing eclipse spectra of the TRAPPIST-1 planets with high enough SNR to detect absorption features in the spectra that we considered. However, our estimates for the JWST observing time required to detect *emission* spectrum features with MIRI LRS dwarfed the time required for spectral features to be detected in transmission spectra with NIRSpec, NIRISS, and NIRCам (see Figs. [2.6](#), [2.13-2.19](#)). Furthermore, the disparity between transmission and emission spectroscopy only broadens with semi-major access, making precise MIRI LRS emission spectroscopy beyond TRAPPIST-1c infeasible with JWST. Our self-consistent planet models reveal that atmospheres with high altitude aerosols—that may appear featureless in a transmission

spectrum—may also appear featureless in an emission spectrum as thermal flux is emitted and/or scattered from near the top of the cloud deck.

Initial photometric assessments of the TRAPPIST-1 system with MIRI filter photometry are unlikely to provide more efficient preliminary results than transmission spectroscopy in the NIR with NIRSpec, NIRISS, or NIRCам. Our MIRI transit and eclipse photometry modeling showed that detecting the presence of atmospheres would require approximately an order of magnitude more JWST time than NIRSpec Prism transmission spectroscopy. This is due to (1) the higher SNR on transits afforded to NIRSpec near the peak of the stellar SED in the 1-3 μm range, (2) the usefulness of spectral resolution for atmospheric detections via deviations from a featureless spectrum, and (3) the need to observe transits in each MIRI filter separately to gain any meaningful wavelength resolution.

However, targeted MIRI photometric observations may still provide useful atmospheric constraints. Detecting the presence of atmospheric greenhouse heating may be done by inferring brightness temperatures that exceed the zero bond albedo equilibrium temperature. However, this will depend on the accuracy of MIRI's absolute flux measurements. Specific wavelength bands where the observed atmosphere is optically thin, and therefore emits from hotter depths, provide the key observable. However, these observations may only be feasible for TRAPPIST-1 b and c due to their higher expected thermal emission. The 11.3 μm MIRI filter (F1130W) may be optimal for such measurements because it is sufficiently separated from the strongly absorbing 15 μm CO_2 band. The 5.6 μm MIRI filter (F560W) may also be optimal for detecting high thermal emission from TRAPPIST-1 b and c, if saturation on the star can be avoided in this bandpass. Transit and/or eclipse photometry targeting the 15 μm CO_2 band with the F1500W MIRI filter, and neighboring filters, could be used to detect the atmosphere, but our work shows that targeting the strong 4.2 μm CO_2 feature with NIRSpec Prism transmission spectroscopy will be a much more efficient approach for detecting CO_2 in an atmosphere.

Detecting wavelength-dependent surface emissivity features in the emission spectrum of a planet without an atmosphere will be highly unlikely with JWST. Even for the optimistic case—TRAPPIST-1 b with a pure and smooth quartz surface—we find it would require ~ 90 secondary

eclipses with MIRI LRS to detect the mid-IR silicate absorption (see Fig. 2.11). Consequently, detecting surface features on an airless world would be extremely challenging. This validates our assumption that an airless planet will likely appear featureless in emission, and that detecting spectral features due to surface emissivity variations is an inefficient means to confirm that a planet is airless. Furthermore, this strengthens the case for transmission spectroscopy, which can detect the high mean molecular weight atmosphere of a Venus-like planet enshrouded in H_2SO_4 clouds in just ~ 20 -30 transits (for TRAPPIST-1 c, d, and e). However, actually *confirming* that a TRAPPIST-1 planet does not possess an atmosphere—even if featureless spectra lend favor to that hypothesis—will be a very difficult task, that may require thermal phase curve (Selsis et al., 2011; Maurin et al., 2012; Kreidberg and Loeb, 2016; Meadows et al., 2018a) or planet-planet occultation observations (Luger et al., 2017c) to probe the day-night temperature contrast.

Comparison with Previous Works Our results are in agreement with previous investigations on optimal ways to detect terrestrial exoplanet atmospheres of different compositions. Although emission spectroscopy has been suggested as a means of detecting terrestrial atmospheres (e.g. Belu et al., 2011), we agree with previous works that transmission spectroscopy is more viable (e.g. Hedelt et al., 2013; Barstow et al., 2016; Barstow and Irwin, 2016; Greene et al., 2016), particularly for temperate and cool planets which emit considerably less thermal flux. We agree with B  tr  mieux and Swain (2018) that the JWST integration times needed to detect molecules strongly depend on the composition of the atmosphere—particularly the presence of high altitude aerosols.

Although we considered more atmospheric compositions, our predicted exposure times for the detection of planetary atmospheres agree for the subset of similar atmospheres modeled by Morley et al. (2017), with many of the discrepancies attributable to differences in the assumed planetary masses used in the transmission spectrum models. Our work and that of Morley et al. (2017) has a common focus of distinguishing clear sky Venus-like atmospheres from a featureless spectrum with the NIRSpec/G235 instrument, which offers a basis for comparison. The detectability of these atmospheres is dominated by CO_2 features, which are largely unaffected by photochemistry, which

was not included in [Morley et al. \(2017\)](#). We find that a 10 bar CO₂ atmosphere would require 5, 8, 4, 17, 15, 16, and 15 transits to distinguish TRAPPIST-1 b, c, d, e, f, g, and h, respectively, from a featureless spectrum, and [Morley et al. \(2017\)](#) find that a 1 bar aerosol-free Venus-like atmosphere would require 6, 36, 13, 4, 17, 10, and 4 transits.

The cases of largest disagreement (e.g. c, e, h) appear consistent with the different masses used in each study. Planet mass affects the atmospheric scale height (via the surface gravity) and therefore the size and detectability of molecular features in a transmission spectrum. [Morley et al. \(2017\)](#) used masses from [Wang et al. \(2017b\)](#), with the exception of TRAPPIST-1 f, for which the mass from [Gillon et al. \(2017b\)](#) was used. In this chapter, and in [Lincowski et al. \(2018\)](#), we use the TRAPPIST-1 planet masses from [Grimm et al. \(2018\)](#). For the case of TRAPPIST-1 c where we use a significantly smaller planet mass than [Morley et al. \(2017\)](#) we find that fewer transits are required to detect spectral features in transit. For the cases of TRAPPIST-1 e and h where we use a significantly larger planet mass than [Morley et al. \(2017\)](#) we find that more transits are required to detect spectral features. This scaling with mass is consistent with the findings of [Morley et al. \(2017\)](#) when they repeated calculations with masses derived from the [Weiss and Marcy \(2014\)](#) mass-radius relationship, and further underscores the need for accurate masses for spectral modeling and fitting.

Our focus on thicker atmospheres (1-100 bar) than [Morley et al. \(2017\)](#) (0.01-1 bar) both explains discrepancies between the studies and further demonstrates that such thick and aerosol-free atmospheres may be more easily detected than thinner atmospheres. Our results for TRAPPIST-1 d, where our masses agree best, show 9 fewer transits required to detect spectral features than [Morley et al. \(2017\)](#), which is consistent with different planet surface pressures used in each study. [Morley et al. \(2017\)](#) demonstrated a trend of increasing atmospheric detectability with increasing surface pressure in simulations of 0.01, 0.1, and 1 bar CO₂ atmospheres. Our results continue this nearly linear trend in log-pressure out to 10 bars. However, our 92 bar CO₂ atmospheres are consistent with the detectability of our 10 bar atmospheres, indicating that this trend in pressure saturates for thick atmospheres as the planet surface drops below the atmospheric regions that are sensed with transmission spectroscopy.

The Krissansen-Totton et al. (2018a) work on the detectability of biogenic gases in an anoxic atmosphere for TRAPPIST-1 e with JWST suggests that relatively few transits are required to constrain atmospheric abundances. Krissansen-Totton et al. (2018a) showed that a retrieval using ~ 10 transits of TRAPPIST-1 e with the NIRSpec Prism was sufficient to begin to constrain the abundances of CO, CO₂, and CH₄. Consequently, the 13 transits of TRAPPIST-1 e that we show may be required to detect a clear 1 bar H₂O atmosphere to a $\langle \text{SNR} \rangle = 5$, may also provide spectra that are sufficient to begin to constrain molecular abundances in an atmospheric retrieval framework.

HST transmission spectroscopy of hot gaseous exoplanets has demonstrated how the presence of clouds and hazes can diminish sensitivity to molecular absorption features (e.g. Berta et al., 2012; Ehrenreich et al., 2014; Knutson et al., 2014; Kreidberg et al., 2014; Nikolov et al., 2015; Sing et al., 2016), but we know *a priori* that planets with such large radii must have atmospheres. However, for terrestrials we will rely on absorption features to test for the presence of atmospheres, and clouds and hazes, if present, will make this more difficult. Batalha et al. (2018) concluded that 10 transits should be sufficient to detect the dominant molecular absorber in the transmission spectrum of TRAPPIST-1f when observed using a partial saturation strategy with the NIRSpec Prism, and that additional observations are unlikely to reveal more information. While our results agree with Batalha et al. (2018) for the case of clear atmospheres, cloudy and/or hazy terrestrial atmospheres may require significantly more observations. Venus-like planets with H₂SO₄ aerosols tend to require more than 10 transits (and up to ~ 30 for TRAPPIST-1e) to reach the same confidence in the detection of the atmosphere. We recommend testing the hypothesis that a TRAPPIST-1 (or similar) planet has a *clear* atmosphere in ≤ 10 transits, and then evaluating the scientific value of additional transits to test the hypothesis that the planet has aerosols.

2.4.2 What is the nature of the TRAPPIST-1 planet atmospheres?

Detecting specific molecules in JWST transmission spectra of the TRAPPIST-1 planets may be possible, and allow for the discrimination between different climate/composition states and evolutionary histories. As previously stated, CO₂ should be the easiest molecule for JWST to detect in

the atmospheres of the TRAPPIST-1 planets. Due to its prevalence in these simulated atmospheres *regardless of evolutionary history*, CO₂ makes for a strong indicator of a terrestrial atmosphere, but a weak discriminant of specific atmospheric state. Furthermore, CO₂ produces nearly the same strength features regardless of atmospheric abundance, particularly the 4.3 and 15 μm bands, which are saturated in a Venus-like atmosphere, an O₂ outgassing atmosphere, and an Earth-like atmosphere, even though the abundance ranges from 90 bars down to 360 ppm (Lincowski et al., 2018). Other molecules, such as O₂, O₃, H₂O, and SO₂, may be detectable with JWST and may help to distinguish between the suite of atmospheres that we considered.

Although oxygen as a biosignature may not be detectable for the potentially habitable TRAPPIST-1 planets, oxygen as a remnant of pre-main-sequence water loss may be easily detected or ruled out. We find that biogenic O₂ in the atmosphere of TRAPPIST-1e may be too difficult to detect with JWST, but the 9.6 μm O₃ feature may be weakly detectable at $\langle\text{SNR}\rangle = 3$ in over 100 transits with MIRI LRS, which is in general agreement with the findings of Wunderlich et al. (2019). However, the 1.06 and 1.27 μm O₂-O₂ CIA features are key discriminants of a planet that has an oxygen abundance greatly exceeding biogenic oxygen production on Earth and may therefore indicate a planet that has undergone vigorous water photolysis and subsequent loss during the protracted super-luminous pre-main-sequence phase faced by late M dwarfs (Luger and Barnes, 2015; Schwieterman et al., 2016b). We find that NIRSpec G140M/H is the optimal JWST instrument for detecting these O₂-O₂ features and could lead to their detection in as few as 7-9, 15, 8, 49-67, 55-82, 79-100, and 62-89 transits of TRAPPIST-1b, c, d, e, f, g, and h, respectively, should they possess such an atmosphere. These quoted number of transits may be sufficient to rule out the existence of oxygen-dominated atmospheres in the TRAPPIST-1 system. Additional evidence of ocean loss could be provided by detection of isotope fractionation, which may also be possible in as few as 11 transits with JWST, for strong isotopologue bands such as HDO (Lincowski et al., 2019).

Detecting ozone absorption may be another strong indicator of a post-runaway, oxygen-dominated atmosphere. For the 10 bar desiccated O₂ atmospheres, ozone absorption features become more detectable for the outer planets. Targeting O₃ at 9.6 μm with MIRI LRS is optimal for the inner

planets (b, c, d, e), although it may require upwards of 100 transits to detect at $\langle \text{SNR} \rangle = 5$. However, targeting the weaker O_3 bands between $3 - 5 \mu\text{m}$ with NIRSpec G395M/H or NIRSpec Prism is optimal for the outer planets (f, g, h), due to their larger ozone column abundances. The $\text{O}_2\text{-O}_2$ CIA bands for these cooler, desiccated planets are more difficult to detect, which may make O_3 a more efficient observational discriminant for such a planet.

The photochemistry underlying the detectable ozone buildup in the cooler planets has implications beyond the TRAPPIST-1 system. [Lincowski et al. \(2018\)](#) noted that the competing effects of the Chapman cycle (with declining photolysis rates with distance from the star), was primarily responsible for the ozone accumulation in the atmospheres of the outer planets, an effect previously noticed by [Grenfell et al. \(2007\)](#). While this is a driving factor, the differences among the planets can more specifically be attributed to catalytic cycles of nitrogen oxides (primarily, N_2O , NO , and NO_2), which drive the destruction of O_3 , as in the stratosphere of Earth ([Seinfeld and Pandis, 2006](#)). Because these atmospheres contain N_2 , O^1D produced from photolysis of oxygen-bearing molecules can react with N_2 to generate nitrogen oxides. The availability of O^1D declines with distance from the star due to lower UV fluxes. Beginning with planet e, the production of nitrogen oxides declines substantially, removing them as a mechanism for the destruction of O_3 , so that O_3 accumulates and becomes well-mixed, generating a large column density. In the atmosphere of an O_2 -dominated atmosphere without N_2 , O_3 levels would likely be higher due to the lack of the nitrogen oxide catalysts.

Detecting water in the atmospheres of the TRAPPIST-1 planets may also help to constrain evolutionary scenarios. The presence of water may be readily detectable for TRAPPIST-1 b, c, and d with the NIRSpec Prism* if they possess high O_2 atmospheres that have not been completely desiccated. High O_2 atmospheres for planets that exited the pre-main-sequence with their atmospheres and interiors completely desiccated, however, will have no water to detect, making water in an oxygen-dominated atmosphere a potentially detectable discriminant of incomplete desiccation or outgassing from the interior. In CO_2 -dominated atmospheres, H_2O may be prohibitively difficult to detect due to its scarcity in the atmospheres of Venus-like worlds, before even considering obscuration by H_2SO_4 clouds.

Detecting water in the atmosphere of one of the potentially habitable TRAPPIST-1 planets could indirectly hint at surface habitability (Robinson, 2018), but this will be challenging for JWST, even with an ideal system like TRAPPIST-1. Tropospheric H₂O may be detectable in the transmission spectrum of TRAPPIST-1 e in ~ 35 transits with NIRSpec Prism* should the planet have a clear-sky 1-bar N₂/O₂ atmosphere with H₂O. However, the cold trap that keeps water vapor concentrated in the lower atmosphere makes habitability difficult to infer with transmission spectroscopy, which cannot readily probe surface environments. Additionally, we find that 100% cloud coverage strongly increases the required JWST time to detect H₂O in such a habitable atmosphere. Given these difficulties for detecting water in the atmosphere of a HZ planet with JWST, robust habitability assessments may ultimately require a future direct imaging telescope that can readily probe rocky planet surfaces to search for more direct evidence of surface liquid water (e.g. Cowan et al., 2009; Robinson et al., 2010; Lustig-Yaeger et al., 2018).

2.4.3 Further Considerations

The results in this chapter lean optimistic, and therefore represent lower limits on the amount of observing time needed to detect and characterize evolved terrestrial atmospheres in the TRAPPIST-1 system. As a result, TRAPPIST-1 observing plans that include fewer transits/eclipses than reported here may require additional observations to make robust inferences on the existence and nature of atmospheres. One source of optimism is that our $\langle \text{SNR} \rangle$ approach (see Equation 2.6) to atmospheric and molecular detection, rather than a full atmospheric retrieval, implicitly assumes that any retrieval will have converged on the true underlying spectrum. As a result, our signal-to-noise metric is not equal to a detection of that specific atmosphere to a given significance (e.g. 3σ). In practice, many different atmospheric compositions will likely be capable of fitting JWST spectra of TRAPPIST-1 planets, potentially leaving large regions of atmospheric parameter space unconstrained. Thus, the confidence of any “one” composition representing the true state of the planet may remain low despite having high confidence in the presence of features in the spectrum that would indicate the presence an atmosphere. Our approach is simply a wavelength range and resolution agnostic method to quantify the detectability of spectral signals emanating from

self-consistent atmospheres above the expected noise of JWST.

Our calculations also lean optimistic because we have neglected astrophysical and systematic sources of noise that may make precise time-series exoplanet observations with JWST more difficult than our estimates suggest. Stellar effects such as limb-darkening (e.g. [Csizmadia et al., 2013](#)) and heterogeneous photospheres (e.g. [Rackham et al., 2016](#); [Rackham et al., 2018](#); [Zhang et al., 2018](#)) are major concerns for precision exoplanet transmission spectroscopy. These effects may be particularly perilous for observations of planets orbiting low mass stars, as the stars may contain large, cool spots and absorbing molecules in their photospheres. We also did not assume a systematic noise floor for any JWST instruments (e.g. [Greene et al., 2016](#)), but this is likely optimistic in terms of the currently unknown on-target performance. Astrophysical and systematic effects must be well understood to detect molecular features at the ~ 10 ppm level, where many terrestrial features reside. However, many of the major absorption features in our TRAPPIST-1 planet spectral models are much larger and more detectable, spanning 50-200 ppm.

During the preparation of this work, a new version of *Pandemia* was released (version 1.3) that increased JWST background noise predictions by 10-40 ppm. For bright targets like TRAPPIST-1 where the noise budget is dominated by stellar photon noise, such an increase in the background has a negligible effect on our estimated number of transits and eclipses to detect and characterize exoplanet atmospheres in the near- through the mid-IR.

We considered a limited number of atmospheres that may not resemble the true state of the TRAPPIST-1 planets, especially since the star's UV spectrum is poorly constrained. Here, as in [Lincowski et al. \(2018\)](#), we used a UV spectrum scaled from measurements of the nearby mid M dwarf Proxima Centauri, but these stars have slightly different spectral types and activity levels. Recently, three synthetic SEDs have been produced by [Peacock et al. \(2019\)](#) for TRAPPIST-1, which are constrained by the Lyman alpha flux and upper limits of GALEX UV photometry. These synthetic spectra differ by having higher NUV and a different distribution of FUV flux than [Lincowski et al. \(2018\)](#), which could slightly change the photochemistry of these planetary atmospheres and the observable molecular absorption in their upper atmospheres. However, only observations—including those outlined in this chapter—will ultimately reveal what planetary pro-

cesses dominate in sculpting the observable atmospheric signatures of terrestrial exoplanets, particularly those orbiting M dwarfs. In the mean time, further study is warranted on the range of possible atmospheric conditions for the TRAPPIST-1 planets, particularly photochemically and climatically self-consistent reducing atmospheres (e.g. [Arney et al., 2017](#)), and UV characterizations of late M dwarfs to improve photochemical predictions and interpretation of observations.

Although the atmospheric states considered in this chapter represent only a selection of possible states, many of their characteristics and molecular features may exist in other atmospheres, making our results apply more broadly. For example, the strength of the H_2O , CO_2 , and O_3 bands in the O_2 -dominated outgassing atmospheres may be similar in other clear atmospheres that may contain more inert gasses, like N_2 . Our JWST detectability calculations for these gases may be useful beyond that particular atmospheric case. The O_2 - O_2 features, however, are specific evidence of high O_2 content. Since all of the atmospheres considered in this work were high mean molecular weight, our results help to elucidate the detectability of such atmospheres and the molecules within, even if the true compositions differ.

Constraining the atmospheres of Earth-size planets transiting even the smallest stars like TRAPPIST-1 will require pushing the limits of JWST. As we demonstrated with the NIRSpec Prism partial saturation, alternate JWST modes that can improve observations of transiting exoplanets can enhance the science return, and dramatically decrease the time investment required to detect and characterize terrestrial atmospheres. Although modes such as NIRCams Dispersed Hartmann Sensor (DHS; [Schlawin et al., 2017](#)) and high efficiency readout patterns for NIRSpec Prism ([Batalha et al., 2018](#)) have not been officially approved, they represent promising avenues towards improved JWST observations that may ultimately make the difference in enabling JWST to constrain terrestrial exoplanet atmospheric compositions.

2.5 Conclusion

We investigated the potential to detect and characterize the atmospheres of all seven known TRAPPIST-1 exoplanets with JWST. Although the planets are small and likely possess high mean molecular weight atmospheres with relatively low scale heights, we found that many molecular absorption

features may be detectable with JWST in ~ 2 -15 transits. These observations may be used to diagnose the presence of atmospheres and, in some cases, discriminate between different plausible atmospheric compositions.

However, we find that an initial photometric assessment of the TRAPPIST-1 planets with MIRI is, perhaps non-intuitively, not as efficient as spectroscopic atmosphere detection in the NIR. To achieve comparable constraints on the detection of atmospheres approximately an order of magnitude more transits or eclipses will need to be observed with MIRI, compared to transits observed with spectrometers in the NIR.

Transmission spectroscopy with NIRSpec Prism may be the most efficient path to detect the presence of atmospheres for the TRAPPIST-1 planets, via detecting CO_2 bands between 1-5 μm . Venus-like atmospheres with high altitude H_2SO_4 aerosols will be more difficult to detect in transmission and emission spectra, however, these aerosols will form at lower altitudes for the temperate and cooler planets such that they obscure less molecular absorption in transmission spectra. Furthermore, post-runaway oxygen-dominated worlds may be identified in transmission using (1) O_2 - O_2 CIA observed with NIRSpec G140M/H, NIRSpec Prism, or NIRISS SOSS, or (2) O_3 absorption observed with MIRI LRS (for b, c, d, and e) or NIRSpec G395M/H (for f, g, and h). If TRAPPIST-1 e is habitable and cloud-free, water could be detected in the troposphere with NIRSpec Prism in about 35 transits, but the presence of water clouds could completely obscure the water vapor absorption features.

We outlined a particular path for characterising the TRAPPIST-1 planets with JWST that narrows down the possible evolutionary histories that the planets may have trod to exist in their current observable states. We recommend using transmission spectroscopy to

1. detect the planet atmospheres via CO_2 absorption;
2. detect or rule out a post-runaway oxygen-dominated atmospheres via O_2 - O_2 CIA or O_3 absorption;
3. constrain the extent of (atmosphere and interior) desiccation, and potentially the habitability, via the H_2O abundance.

However, our results may be used to construct countless additional observing strategies that best augment existing projects and proposals with our testable hypotheses on the nature of TRAPPIST-1 system of Earth-sized exoplanets.

2.6 Chapter Appendix

2.6.1 Signal-to-Noise Ratio on Transit and Eclipse Depths

Transit and eclipse depth measurements are both relative to the out-of-occultation measurement, and thus require accounting for photon fluence in- and out-of the the event of interest. Below we derive the signal-to-noise on measurements of transit and eclipse depths by considering the signal as the occultation depth, which must be calculated from the difference between in- and out-of-event observations, and the noise by propagating individual measurement errors through the signal calculation.

For primary transit, the signal of interest is the number of photons missing when the star is occulted by the planet. Neglecting photons emitted and reflected by the planet prior to and during transit, the out-of-transit photon fluence is measured as

$$N_{\text{out}} = (N_s + N_{bg})n_{\text{out}}, \quad (2.11)$$

where n_{out} is the time of measurement outside of transit in units of the transit duration, and N_s and N_{bg} are the total number of photons counted over a transit duration. The in-transit photon fluence measured is

$$N_{tr} = N_s \left(1 - \left(\frac{R_p}{R_s} \right)^2 \right) + N_{bg}. \quad (2.12)$$

The number of stellar photons blocked by the planet can be estimated from

$$N_{sp} = N_{\text{out}}/n_{\text{out}} - N_{tr}, \quad (2.13)$$

this is the signal we seek to measure. The noise on N_{sp} can be calculated by considering the

variance given standard uncorrelated error propagation:

$$\sigma^2 = \left(\frac{\partial N_{sp}}{\partial N_{out}} \right)^2 \sigma N_{out}^2 + \left(\frac{\partial N_{sp}}{\partial N_{tr}} \right)^2 \sigma N_{tr}^2 \quad (2.14)$$

$$= \left(\frac{1}{n_{out}} \right)^2 N_{out} + (-1)^2 N_{tr} \quad (2.15)$$

$$= \frac{(N_s + N_{bg})}{n_{out}} + N_s \left(1 - \left(\frac{R_p}{R_s} \right)^2 \right) + N_{bg} \quad (2.16)$$

Finally, the signal-to-noise ratio on the transit depth can be constructed by dividing the blocked photons N_{sp} by the standard deviation on that estimate:

$$\text{SNR}_T = \frac{N_s (R_p/R_s)^2}{\sqrt{(N_s + N_{bg})/n_{out} + N_s (1 - (R_p/R_s)^2) + N_{bg}}}. \quad (2.17)$$

For secondary eclipse, the signal of interest is the number of photons missing when the planet is occulted by the star, which is measured assuming that the star is not varying. In this case, the out-of-eclipse photon fluence is measured as

$$N_{out} = (N_p + N_s + N_{bg})n_{out}, \quad (2.18)$$

where n_{out} is the time of measurement outside of eclipse in units of the eclipse duration, N_p is the total number of planet photons over an eclipse duration, and the sum of N_p , N_s and N_{bg} is the total number of photons counted over an eclipse duration. The in-eclipse photon fluence measured is

$$N_{ec} = N_s + N_{bg}, \quad (2.19)$$

which allows the planet photon counts to be estimated as

$$N_p = N_{out}/n_{out} - N_{ec}. \quad (2.20)$$

Again, analogous to the transit calculation, the noise term can be calculated by considering the variance on N_p given standard error propagation:

$$\sigma^2 = \left(\frac{\partial N_p}{\partial N_{\text{out}}} \right)^2 \sigma N_{\text{out}}^2 + \left(\frac{\partial N_p}{\partial N_{ec}} \right)^2 \sigma N_{ec}^2 \quad (2.21)$$

$$= \left(\frac{1}{n_{\text{out}}} \right)^2 N_{\text{out}} + (-1)^2 N_{ec} \quad (2.22)$$

where we have used the fact that N_{out} has a variance of N_{out} . Substituting in for N_{out} and N_{ec} gives:

$$\sigma^2 = \frac{(N_p + N_s + N_{bg})}{n_{\text{out}}} + N_s + N_{bg}. \quad (2.23)$$

Finally, the signal-to-noise ratio on the eclipse depth can be constructed by dividing the estimated planet photons by the standard deviation on that estimate:

$$\text{SNR}_E = \frac{N_p}{\sqrt{(N_p + N_s + N_{bg})/n_{\text{out}} + N_s + N_{bg}}}. \quad (2.24)$$

2.6.2 Atmospheric Detectability by Instrument

Figures 2.13, 2.14, 2.15, 2.16, 2.17, 2.18, and 2.19, show the number of transits needed to detect the atmospheres of TRAPPIST-1 b, c, d, e, f, g, and h, respectively, as a function of both atmospheric compositions and JWST instrument/mode. The detection of an atmosphere requires $\langle \text{SNR} \rangle = 5$ on spectral features in the transmission spectrum.

2.6.3 Brightness Temperatures

Table 2.3 shows the brightness temperature of each [Lincowski et al. \(2018\)](#) TRAPPIST-1 atmospheric model in each of the JWST/MIRI photometric filters and the two warm *Spitzer* bands.

Table 2.3. Brightness temperatures in different Spitzer and JWST photometric filters for different TRAPPIST-1 planet atmospheric models.

Planet	Model	IRAC1 (K)	IRAC2 (K)	F560W (K)	F770W (K)	F1000W (K)	F1130W (K)	F1280W (K)	F1500W (K)	F1800W (K)	F2100W (K)	F2550W (K)
T-1b	10 bar CO ₂	498	328	432	384	340	356	330	291	315	362	430
	92 bar CO ₂	495	328	432	383	340	357	330	289	315	362	430
	100 bar O ₂ outgassing	500	388	372	396	379	432	385	305	355	381	346
	10 bar O ₂ outgassing	514	391	357	394	394	425	377	300	345	371	329
	100 bar O ₂ desiccated	425	375	390	387	389	404	376	306	368	408	364
	10 bar O ₂ desiccated	436	357	401	405	369	397	360	292	346	408	380
T-1c	10 bar CO ₂	445	292	385	328	304	326	295	250	275	315	386
	92 bar CO ₂	437	290	380	324	303	326	293	248	273	313	382
	100 bar O ₂ outgassing	434	337	325	346	312	390	343	266	325	345	308
	10 bar O ₂ outgassing	424	334	313	350	337	378	330	252	304	332	294
	100 bar O ₂ desiccated	387	325	330	328	333	339	327	272	325	342	310
	10 bar O ₂ desiccated	395	313	339	341	322	345	318	257	310	347	322
T-1d	10 bar Venus	427	292	373	318	300	320	296	272	282	313	376
	92 bar Venus	435	287	382	317	297	324	291	250	275	303	383
	10 bar CO ₂	384	256	337	274	269	295	260	219	239	272	340
	92 bar CO ₂	380	254	333	271	268	295	259	215	237	269	337
	100 bar O ₂ outgassing	370	290	280	301	264	352	295	227	290	305	270
	10 bar O ₂ outgassing	361	283	270	299	287	324	285	215	267	291	260
T-1e	100 bar O ₂ desiccated	356	281	278	274	280	279	280	239	282	284	260
	10 bar O ₂ desiccated	362	275	286	286	279	293	277	225	275	292	269
	10 bar Venus	326	258	240	230	230	233	233	227	231	239	243
	92 bar Venus	375	251	332	264	263	288	256	223	245	271	335
	10 bar CO ₂	339	230	297	235	247	272	232	193	208	237	301
	92 bar CO ₂	335	226	292	229	243	267	228	189	206	233	294
T-1f	100 bar O ₂ outgassing	305	236	244	259	231	301	245	203	253	265	245
	10 bar O ₂ outgassing	346	243	243	256	243	267	244	191	239	259	238
	100 bar O ₂ desiccated	341	250	244	238	242	239	246	214	251	243	226
	10 bar O ₂ desiccated	346	246	250	246	242	249	244	200	247	251	232
	10 bar Venus	303	217	213	203	204	204	202	192	204	206	211
	92 bar Venus	311	223	287	228	235	254	225	195	220	242	285
T-1g	1 bar H ₂ O	296	267	243	259	272	277	267	213	258	260	236
	1 bar H ₂ O cloudy	287	253	241	247	250	253	248	212	247	249	243
	10 bar CO ₂	277	201	260	200	221	235	200	172	185	206	263
	92 bar CO ₂	271	196	262	195	216	225	194	171	183	202	260
	100 bar O ₂ outgassing	229	209	218	219	202	234	208	192	221	224	222
	10 bar O ₂ outgassing	326	212	218	216	206	223	210	178	209	218	213
T-1h	100 bar O ₂ desiccated	268	225	228	216	210	209	219	192	227	213	194
	10 bar O ₂ desiccated	298	222	229	223	210	216	218	184	227	219	200
	10 bar Venus	272	189	210	188	193	193	188	171	184	189	199
	92 bar Venus	247	197	239	199	211	218	196	177	195	210	237
	10 bar CO ₂	231	181	236	180	203	205	179	160	170	185	232
	92 bar CO ₂	229	177	240	175	200	197	174	158	167	181	231
T-1i	100 bar O ₂ outgassing	209	191	199	199	188	202	192	180	205	204	194
	10 bar O ₂ outgassing	312	195	202	194	187	199	191	166	188	192	191
	100 bar O ₂ desiccated	224	207	210	200	192	187	201	180	208	189	176
	10 bar O ₂ desiccated	276	205	211	205	192	192	200	174	209	195	178
	10 bar Venus	242	173	210	173	183	181	173	158	168	174	190
	92 bar Venus	226	182	222	181	200	199	177	165	176	191	217
T-1j	10 bar CO ₂	174	157	217	155	176	167	154	145	150	157	186
	92 bar CO ₂	175	154	206	151	174	162	151	143	147	153	182
	100 bar O ₂ outgassing	195	168	171	170	166	169	167	163	174	172	164
	10 bar O ₂ outgassing	296	181	184	170	165	167	170	150	165	164	160
	100 bar O ₂ desiccated	199	183	185	179	170	161	177	165	181	162	149
	10 bar O ₂ desiccated	253	184	188	183	171	164	179	160	186	167	150
T-1k	10 bar Venus	209	155	185	153	162	161	153	145	149	154	162
	92 bar Venus	189	159	195	156	185	169	146	145	144	161	187

T-1b Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	14	16	10	5	5	76	7	6	6	4	2
92 bar CO ₂	14	17	10	5	5	73	6	6	6	4	2
10 bar O ₂ outgassing	9	9	10	4	4	55	6	6	5	3	2
100 bar O ₂ outgassing	10	9	10	5	4	44	7	7	6	4	2
10 bar O ₂ desiccated	7	10	7	3	3	21	5	5	4	3	2
100 bar O ₂ desiccated	11	10	10	5	3	64	8	7	6	4	2
NIRCam F322W2											
NIRCam F444W											
NIRSpec G140H											
NIRSpec G235H											
NIRSpec G395H											
MIRI LRS											
NIRISS SOSS substrip256											
NIRISS SOSS substrip96											
NIRSpec Prism sub512											
NIRSpec Prism sub512s											
NIRSpec Prism sub512 ngroup6											

Figure 2.13 Number of TRAPPIST-1b transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1c Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	27	32	20	8	9	>100	12	12	10	7	4
92 bar CO ₂	28	34	20	8	9	>100	12	12	10	7	4
10 bar Venus	>100	59	>100	40	30	>100	86	84	56	34	18
92 bar Venus	>100	51	>100	60	26	>100	>100	>100	76	46	22
10 bar O ₂ outgassing	18	18	18	6	7	100	11	10	9	6	3
100 bar O ₂ outgassing	19	17	23	9	7	64	16	15	11	7	4
10 bar O ₂ desiccated	16	19	17	6	5	51	11	11	9	5	3
100 bar O ₂ desiccated	26	18	22	11	7	>100	17	16	13	8	4
	NIRCam F322W2	NIRCam F444W	NIRSpec G140H	NIRSpec G235H	NIRSpec G395H	MIRI LRS	NIRISS SOSS substrip256	NIRISS SOSS substrip96	NIRSpec Prism sub512	NIRSpec Prism sub512s	NIRSpec Prism sub512 ngroup6

Figure 2.14 Number of TRAPPIST-1c transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1d Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	15	17	10	4	5	87	7	6	6	4	2
92 bar CO ₂	14	17	10	4	5	84	6	6	5	3	2
10 bar Venus	>100	38	>100	51	21	>100	72	67	49	30	15
92 bar Venus	>100	59	>100	67	33	>100	>100	>100	79	48	24
10 bar O ₂ outgassing	10	10	10	4	4	51	7	6	6	4	2
100 bar O ₂ outgassing	8	8	8	4	4	28	6	6	5	3	2
10 bar O ₂ desiccated	9	10	9	4	3	31	6	6	5	3	2
100 bar O ₂ desiccated	13	9	10	5	4	48	8	7	6	4	2
NIRCam F322W2											
NIRCam F444W											
NIRSpec G140H											
NIRSpec G235H											
NIRSpec G395H											
MIRI LRS											
NIRISS SOSS substrip256											
NIRISS SOSS substrip96											
NIRSpec Prism sub512											
NIRSpec Prism sub512s											
NIRSpec Prism sub512 ngroup6											

Figure 2.15 Number of TRAPPIST-1d transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1e Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

1 bar H ₂ O	88	46	>100	36	20	>100	85	80	41	25	13
1 bar H ₂ O cloudy	>100	60	>100	73	29	>100	>100	>100	81	49	23
10 bar CO ₂	53	60	54	17	17	>100	30	28	22	14	7
92 bar CO ₂	58	72	54	17	20	>100	30	28	23	14	8
10 bar Venus	>100	93	>100	69	45	>100	>100	>100	99	60	30
92 bar Venus	>100	>100	>100	69	50	>100	>100	>100	>100	62	31
10 bar O ₂ outgassing	59	38	87	25	16	>100	53	50	33	20	10
100 bar O ₂ outgassing	36	35	75	27	14	73	45	42	24	15	7
10 bar O ₂ desiccated	42	36	58	17	13	>100	35	33	23	14	8
100 bar O ₂ desiccated	67	36	64	29	16	>100	49	46	34	21	11
	NIRCam F322W2	NIRCam F444W	NIRSpec G140H	NIRSpec G235H	NIRSpec G395H	MIRI LRS	NIRISS SOSS substrip256	NIRISS SOSS substrip96	NIRSpec Prism sub512	NIRSpec Prism sub512s	NIRSpec Prism sub512 ngroup6

Figure 2.16 Number of TRAPPIST-1e transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1f Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	45	52	58	15	14	>100	29	27	20	12	7
92 bar CO ₂	45	50	58	15	14	>100	28	27	20	12	7
10 bar Venus	73	59	>100	22	22	>100	60	57	36	22	12
92 bar Venus	78	61	>100	24	23	>100	64	60	38	23	12
10 bar O ₂ outgassing	54	31	98	24	14	>100	54	51	30	19	9
100 bar O ₂ outgassing	25	27	56	19	10	34	29	28	16	10	5
10 bar O ₂ desiccated	32	27	58	14	11	72	30	29	18	11	6
100 bar O ₂ desiccated	48	31	72	27	13	93	51	48	29	18	9
	NIRCam F322W2	NIRCam F444W	NIRSpec G140H	NIRSpec G235H	NIRSpec G395H	MIRI LRS	NIRISS SOSS substrip256	NIRISS SOSS substrip96	NIRSpec Prism sub512	NIRSpec Prism sub512s	NIRSpec Prism sub512 ngroup6

Figure 2.17 Number of TRAPPIST-1f transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1g Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	46	49	72	16	14	>100	33	31	21	13	7
92 bar CO ₂	49	48	74	17	14	>100	34	32	22	14	7
10 bar Venus	60	54	>100	19	18	>100	49	46	29	18	9
92 bar Venus	70	58	>100	21	20	>100	56	52	33	20	11
10 bar O ₂ outgassing	58	34	>100	27	14	>100	69	66	34	21	10
100 bar O ₂ outgassing	24	29	41	17	11	31	21	20	13	8	4
10 bar O ₂ desiccated	31	26	73	15	11	60	33	31	18	11	6
100 bar O ₂ desiccated	42	30	93	28	13	69	56	53	27	17	8
	NIRCam F322W2	NIRCam F444W	NIRSpec G140H	NIRSpec G235H	NIRSpec G395H	MIRI LRS	NIRISS SOSS substrip256	NIRISS SOSS substrip96	NIRSpec Prism sub512	NIRSpec Prism sub512s	NIRSpec Prism sub512 ngroup6

Figure 2.18 Number of TRAPPIST-1g transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

T-1h Transit : Detect Features with $\langle \text{SNR} \rangle = 5.0$

10 bar CO ₂	44	48	70	15	13	>100	31	30	20	13	7
92 bar CO ₂	46	46	69	16	13	>100	31	30	20	13	7
10 bar Venus	51	50	91	17	15	>100	38	36	24	15	8
92 bar Venus	57	52	93	18	15	>100	40	38	25	15	8
10 bar O ₂ outgassing	50	34	>100	25	13	79	58	55	30	19	9
100 bar O ₂ outgassing	21	28	30	15	10	26	17	16	11	7	4
10 bar O ₂ desiccated	26	24	56	13	9	48	26	25	15	9	5
100 bar O ₂ desiccated	29	26	63	20	10	44	34	33	18	11	6
	NIRCam F322W2	NIRCam F444W	NIRSpec G140H	NIRSpec G235H	NIRSpec G395H	MIRI LRS	NIRISS SOSS substrip256	NIRISS SOSS substrip96	NIRSpec Prism sub512	NIRSpec Prism sub512s	NIRSpec Prism sub512 ngroup6

Figure 2.19 Number of TRAPPIST-1h transits necessary to rule out a featureless spectrum with $\langle \text{SNR} \rangle = 5$ for different atmospheric compositions and using different JWST instruments and modes.

Chapter 3

A MIRAGE OF THE COSMIC SHORELINE: VENUS-LIKE CLOUDS AS A STATISTICAL FALSE POSITIVE FOR EXOPLANET ATMOSPHERIC EROSION

Portions of this chapter were originally published in collaboration with Victoria S. Meadows and Andrew P. Lincowski in the December 2019 edition of The Astrophysical Journal Letters (Lustig-Yaeger et al., 2019a, ApJL, Vol. 887, 1; 2019 © American Astronomical Society, DOI: 10.3847/2041-8213/ab5965), and are reproduced below with the permission of the American Astronomical Society.

Summary

Near-term studies of Venus-like atmospheres with JWST promise to advance our knowledge of terrestrial planet evolution. However, the remote study of Venus in the Solar System and the ongoing efforts to characterize gaseous exoplanets both suggest that high altitude aerosols could limit observational studies of lower atmospheres, and potentially make it challenging to recognize exoplanets as “Venus-like”. To support practical approaches for exo-Venus characterization with JWST, we use Venus-like atmospheric models with self-consistent cloud formation of the seven TRAPPIST-1 exoplanets to investigate the atmospheric depth that can be probed using both transmission and emission spectroscopy. We find that JWST/MIRI LRS secondary eclipse emission spectroscopy in the 6 μm opacity window could probe at least an order of magnitude deeper pressures than transmission spectroscopy, potentially allowing access to the subcloud atmosphere for the two hot innermost TRAPPIST-1 planets. In addition, we identify two confounding effects of sulfuric acid aerosols that may carry strong implications for the characterization of terrestrial exoplanets with transmission spectroscopy: (1) there exists an ambiguity between cloud-top and solid surface in producing the observed spectral continuum; and (2) the cloud-forming region drops in altitude with semi-major axis, causing an increase in the observable cloud-top pressure with decreasing stellar insolation. Taken together, these effects could produce a trend of thicker atmospheres observed at lower stellar insolation—a convincing false positive for atmospheric escape and an empirical “cosmic shoreline”. However, developing observational and theoretical techniques to identify Venus-like exoplanets and discriminate them from stellar windswept worlds will enable advances in the emerging field of terrestrial comparative planetology.

3.1 Introduction

Venus-like exoplanets pose unique opportunities and challenges for the near-term characterization of terrestrial exoplanet atmospheres (Arney and Kane, 2018). Exo-Venuses are key near-term observational targets due to the transit bias in favor of finding and characterizing planets at short orbital period (e.g. high transit probability, high transit frequency, high equilibrium temperature), particularly in the *TESS* era (Ostberg and Kane, 2019). Planets at similar insolation and with similar bulk properties to Venus are also favorable laboratories to empirically test runaway greenhouse theory, identify the location of the inner edge of the habitable zone (HZ), and probe the impact of atmospheric escape on an ensemble of terrestrial planets (Kane et al., 2014). Additionally, the comparative study of Venus and Venus-like exoplanets are mutually beneficial research avenues (Arney and Kane, 2018). Within the exoplanet population, if exo-Venuses are found to be common, they would point to a common end-state of terrestrial exoplanet evolution that Venus exemplifies. However, if true Venus analogs are rare, that may point to a more specific origin for Venus. Within the solar system, future orbiters and descent probes could provide detailed, in situ measurements to help answer outstanding questions about evolutionary processes and the current state of Venus, which will provide crucial context for the population of exo-Venuses (see recent white papers: Kane et al., 2018, 2019; Wilson and Widemann, 2019). However, these exciting opportunities are contingent upon our ability to properly recognize and accurately characterize a Venus-like exoplanet when we see one.

Remote sensing observations have been used to understand and probe beneath the optically thick and global sulfuric acid clouds and hazes, which extend from 48 to 90 km altitude, and obscure the lower atmosphere and surface of Venus at most wavelengths. Although clouds were suspected early on due to Venus's high albedo and UV markings (Hunten et al., 1983), their composition was unknown until optical phase curves ruled out water clouds (Arking and Potter, 1968; Hansen and Arking, 1971), multi-band polarization phase curves matched the real index of refraction for a concentrated solution of sulfuric acid (Hansen and Hovenier, 1971), and NIR absorption features confirmed H_2SO_4 (Pollack et al., 1974). These clouds thoroughly obscure the hot lower

atmosphere at visible wavelengths, but the first clue to the extremely hot nature of the surface environment was a radio brightness temperature measurement of ~ 560 K at 3.15 cm by [Mayer et al. \(1958\)](#), which was later confirmed by spacecraft observations (e.g. [Barath et al., 1963](#)) and descent probes (e.g. [Marov et al., 1973](#)). Despite these challenges, peering beneath the clouds into the hot lower atmosphere has been possible with spectroscopy targeting near-infrared windows on the Venus night side through which thermal emission from below the clouds escapes (e.g. [Allen and Crawford, 1984](#); [Allen, 1987](#); [Carlson et al., 1991](#); [Crisp et al., 1991](#)), enabling remote studies of the Venus lower atmosphere and surface (e.g. [Drossart et al., 1993](#); [de Bergh et al., 1995](#); [Meadows and Crisp, 1996](#); [Barstow et al., 2012](#); [Arney et al., 2014](#)).

Extending the lessons learned from Venus remote sensing to the characterization of potential exo-Venuses may be challenging as the historically most informative Venus observations lack feasible exoplanet analogs, either because they were made from orbiters or descent probes, or used radio brightness or precise optical and polarization phase curves.

After launch, the *James Webb Space Telescope* (JWST) will likely be used to attempt characterization of Venus-like exoplanets ([Barstow et al., 2016](#); [Morley et al., 2017](#); [Lincowski et al., 2018](#)), but these observations may be limited by how transmission and emission spectra are both significantly impacted by Venus-like clouds ([Lustig-Yaeger et al., 2019b](#)). It has been well established by theory and observation that transmission spectroscopy is sensitive to obscuration by high altitude aerosols (e.g. [Fortney, 2005](#); [Berta et al., 2012](#); [Ehrenreich et al., 2014](#); [Knutson et al., 2014](#); [Kreidberg et al., 2014](#); [Nikolov et al., 2015](#); [Morley et al., 2013](#); [Charnay et al., 2015a,b](#)). Additionally, although dayside thermal emission, which is sensitive to the cloud deck temperatures, may be observed for exoplanets via secondary eclipse, nightside NIR thermal emission windows, which are sensitive to the lower atmosphere, may be significantly more challenging to observe for exoplanets when the far brighter dayside portion is included in the disk average. Despite these challenges, modeling efforts in advance of JWST indicate that the presence of cloudy Venus-like atmospheres could be detected for all seven planets in the TRAPPIST-1 system using JWST transmission spectroscopy to identify CO₂ absorption features in the thin atmosphere above the clouds ([Lustig-Yaeger et al., 2019b](#)).

Another potential complication in identifying and interpreting spectra of Venus-like planets comes from the behavior of sulfuric acid cloud formation as a function of semi-major axis. The super-luminous pre-main-sequence phase of late M dwarfs, like TRAPPIST-1, could produce a string of Venuses, extending from interior to, through, and beyond the HZ (Lincowski et al., 2018). During the pre-main-sequence phase, each of the TRAPPIST-1 planets may have been subjected to runaway greenhouse driven water loss and subsequent O₂ buildup (Luger and Barnes, 2015; Bolmont et al., 2017; Lincowski et al., 2018), even for planets well beyond the HZ. The subsequent sequestration of O₂ and the outgassing of volatiles over time (Schaefer et al., 2016; Garcia-Sage et al., 2017) may have allowed high-CO₂ Venus-like atmospheres to develop.

Lincowski et al. (2018) conducted a systematic study of the seven TRAPPIST-1 planets assuming they possess Venus-like atmospheres using a self-consistent 1D photochemical and climate model, which included sulfuric acid cloud formation. Interestingly, these models demonstrated that sulfuric acid clouds form high in the atmospheres of hot Venus-like planets, but drop to lower altitudes for cooler Venus-like planets at lower incident stellar fluxes (Lincowski et al., 2018). Since high altitude clouds can obscure molecular features in a transmission spectrum, the hottest cloudy exo-Venus atmospheres may actually be more difficult to detect than cooler cloudy exo-Venuses (Lustig-Yaeger et al., 2019b)—a practical manifestation of only probing the atmosphere above the clouds. However, if only the upper, above-cloud, region of the atmosphere is readily probed, we may remain ignorant to the existence of a lower atmosphere, unable to distinguish cloud-top from solid surface. Furthermore, the predicted increases in cloud top pressure with semi-major axis occur across a stellar insolation range that could also completely erode planetary atmospheres (Dong et al., 2018). Thus, observing a trend of thicker cloud-truncated atmospheres at lower stellar insolation may produce a statistical false positive for atmospheric escape across a population of terrestrial exoplanets (e.g. Bean et al., 2017; Checlair et al., 2019) and a mirage of the “cosmic shoreline”—an empirical dividing line between planets with and without atmospheres (Zahnle and Catling, 2017).

In this chapter we explore two fundamental questions on the characterization of Venus-like exoplanets and their potential contribution to our understanding of terrestrial exoplanet atmospheric

evolution: (1) *how do we infer the presence of and study sub-cloud atmospheres, and* (2) *what consequences and misinterpretations may arise if we cannot?* In particular, we demonstrate how the presence of sulfuric acid clouds in *thick* Venus-like atmospheres can mimic *thin* cloud-free atmospheres in a transmission spectrum. We then explore how an observed decrease in cloud top altitude as a function of orbital distance could be misinterpreted as a surface pressure trend. Furthermore, such a trend with incident stellar flux could arise due to (1) atmospheric erosion via photoevaporation/thermal escape if the spectral continuum is assumed to be a solid surface, or (2) cloud top altitude variations due to condensation temperature if the continuum is assumed to be a cloud top. Finally, we offer observational and theoretical research avenues that may help to resolve this potential statistical false positive.

In Section 3.2 we describe the TRAPPIST-1 Venus-like atmospheric models used in this chapter. In Section 3.3 we investigate the atmospheric regions probed by the transmission and emission spectra of Venus-like exoplanets applicable to JWST observations. In Section 3.4 we discuss the optimal paths towards inferring the presence of lower atmospheres for Venus-like exoplanets and we also expand on the hypothesis that, if we cannot detect lower atmospheres, atmospheric erosion could be invoked to explain mistakenly thin atmospheres, particularly in statistical characterization populations if Venus-like exoplanets are intrinsically common. We conclude in Section 3.5.

3.2 Methods

We use the clear CO₂ and cloudy Venus-like TRAPPIST-1 planet atmospheric models from [Lincowski et al. \(2018\)](#) as a foundation for the investigations in this chapter. Briefly, [Lincowski et al. \(2018\)](#) used the VPL Climate model, a 1D radiative-convective equilibrium climate model applicable to terrestrial planet atmospheres ([Meadows et al., 2018a](#); [Robinson and Crisp, 2018](#)). The climate model is coupled to a 1D atmospheric photochemistry model originally developed by [Kasting et al. \(1979\)](#) and significantly improved upon by [Zahnle et al. \(2006\)](#); this code is described in detail in [Meadows et al. \(2018a\)](#) and has been used extensively for terrestrial exoplanet photochemical modeling across a broad range of redox states (e.g. [Segura et al., 2005](#); [Arney et al., 2016, 2017](#); [Schwieterman et al., 2016b](#); [Arney, 2019](#)). In particular, the photochemical code was

specifically updated and validated for modeling Venus-like atmospheres (Lincowski et al., 2018).

Lincowski et al. (2018) used radiatively-active, photochemically-self-consistent sulfuric acid aerosols in their climate and spectral calculations. These calculations considered the photochemical production of H_2SO_4 vapor given the forcing from the late M dwarf SED, the temperature-dependent condensation of H_2SO_4 , the sedimentation of H_2SO_4 condensates, and their thermal decomposition at high temperatures in the lower atmosphere. Together these effects determined the aerosol effective radii and H_2SO_4 concentration in each layer. Lincowski et al. (2018) used refractive indices for sulfuric acid solutions from Palmer and Williams (1975) ranging in concentration from 25-100%, which were calculated for each atmosphere using the vapor pressure equilibrium between the H_2O and H_2SO_4 gases and the condensed H_2SO_4 solution. Mass-conserving log-normal aerosol particle distributions were computed (with geometric standard deviation equal to 0.25, as used in Crisp (1986) for Venus), from which mie-scattering phase functions and optical depths were calculated. We note that Lincowski et al. (2018) found that sulfuric acid clouds did not form for a Venus-like TRAPPIST-1 b because the atmosphere was too hot for them to condense, so TRAPPIST-1 b is omitted from our cloudy Venus-like cases. We discuss the observational implications of this result in Section 3.4.4.

Transmission and emission spectra of the TRAPPIST-1 Venus-like planets were produced in Lincowski et al. (2018) using the Spectral Mapping Atmospheric Radiative Transfer (SMART) code (developed by D. Crisp; Meadows and Crisp, 1996). SMART is a line-by-line, multi-stream, multi-scattering radiative transfer code that includes layer dependent gaseous and aerosol absorption and scattering, and treats both stellar and thermal source functions. SMART also calculates transmission spectra for transiting exoplanets using a ray tracing algorithm that includes the refraction of stellar light passing through the atmosphere (Misra et al., 2014; Robinson, 2018). Gaseous rotational-vibrational line absorption coefficients were calculated using the LBLABC model (Meadows and Crisp, 1996) with the HITEMP2010 and HITRAN2012 line lists (Rothman et al., 2010, 2013).

To understand the maximum possible depth that can be probed into clear and cloudy Venus-like atmospheres, we processed the atmospheric structure and spectra of the Venus-like TRAPPIST-1

models from [Lincowski et al. \(2018\)](#) to reveal the average pressure into each atmosphere that can be probed with transmission and emission spectroscopy. For transmission spectroscopy, we used the relationship between altitude and pressure for atmospheres in hydrostatic equilibrium to interpolate the effective transit height to an effective transit pressure. Although this so called “transit pressure” is not a direct observable, it is tied to the observable transit depth, $(R_p/R_s)^2$, and approximates the depth into the atmosphere at which it becomes optically thick in the slant transit geometry. While this is similar to the effective transit height, the transit height increases radially out of the atmosphere and must assume a zero-point altitude that presumes knowledge of the planet’s solid body radius—a key point of interest—but which is unknown *a priori*. Alternatively, the transit pressure increases into the atmosphere from space—a known zero point pressure boundary condition—to the maximum pressure probed, and is therefore a good measure of how deep into the atmosphere the transmission spectrum probes. The maximum transit pressure across a given wavelength range also provides forward modeling insight into the cloud top pressure or reference pressure that would be inferred by atmospheric retrievals (e.g. [Benneke and Seager, 2013](#); [Kreidberg et al., 2015](#); [Line and Parmentier, 2016](#); [Benneke, 2015](#)).

To assess the depth probed into Venus-like atmospheres with emission spectroscopy, we re-computed the radiative transfer with SMART for the [Lincowski et al. \(2018\)](#) Venus-like models to solve for the atmospheric pressure at which the total optical depth is unity at normal incidence. SMART calculates the pressure of optical depth unity in terms of Rayleigh scattering, gaseous absorption, and aerosol extinction. We combine these three terms into a single “emission pressure” or “brightness pressure” spectrum, which reflects the dominating process at each wavelength and approximates the pressure from which thermal emission emerges the atmosphere. Note that unlike the transit pressure, which resembles the observable transmission spectrum $(R_p/R_s)^2$, the “emission pressure” is significantly different from the observable eclipse depths of an emission spectrum (F_p/F_s) , which have a strong wavelength dependence in accordance with the planet and star fluxes. Rather, the emission pressure more closely resembles the transit pressure, allowing for qualitative and quantitative comparisons between the two observational techniques.

3.3 Results

3.3.1 Transmission Spectroscopy

To explore the impact of clouds on the range of pressures probed by transmission spectroscopy, Figure 3.1 shows model transmission spectra for each TRAPPIST-1 planet with and without H_2SO_4 clouds in the upper and lower panels, respectively. The transmission spectrum is shown between about 1-5 μm —the range of the JWST/NIRSpec Prism instrument, which is optimal for detecting the atmospheres of the TRAPPIST-1 planets (Batalha et al., 2018; Lustig-Yaeger et al., 2019b)—in units of the pressure into the atmosphere that is probed. The right panels of Fig. 3.1 show the thermal structure of each TRAPPIST-1 planet atmosphere on the same pressure y-axis as the transmission spectrum. Thicker line styles indicate the vertical region of the atmosphere that the transmission spectrum is sensitive to, and the circular points denote the maximum atmospheric pressure that is probed. The top-right panel also shows the cumulative vertical optical depth for the H_2SO_4 aerosols on the top x-axis, which increases going down into the atmosphere.

Although the clear and cloudy transmission spectra appear similar due to the common presence of CO_2 absorption bands, their respective continua vary by up to 3 orders of magnitude in pressure, which affects the strength and detectability of CO_2 absorption (see Lustig-Yaeger et al., 2019b) and the depth into the atmosphere that may be probed by the spectrum. In units of pressure, the transmission spectra of different planets with atmospheres of similar compositions look quite similar, despite having different radii, masses, and temperatures. At wavelengths where the atmosphere is optically thick due to strong CO_2 absorption (e.g. 2.7, 4.3, and 15.0 μm), the opacity is sufficiently high above the clouds that the peak absorption in the bands occurs at the same pressure in the upper atmosphere (1-10 Pa) for all of the TRAPPIST-1 planets, uninfluenced by the clouds and hazes at higher pressures below. At wavelengths where the atmosphere is optically thin, the presence of H_2SO_4 aerosols significantly raises the spectral continuum altitude to lower atmospheric pressures. For instance, the continuum pressure at 2.5 μm is 10^5 Pa and 10^2 Pa for clear and cloudy TRAPPIST-1d models, respectively. These cloudy results show both a significant departure from the clear atmosphere cases for each TRAPPIST-1 planet ($>100\times$ lower continuum pressures for

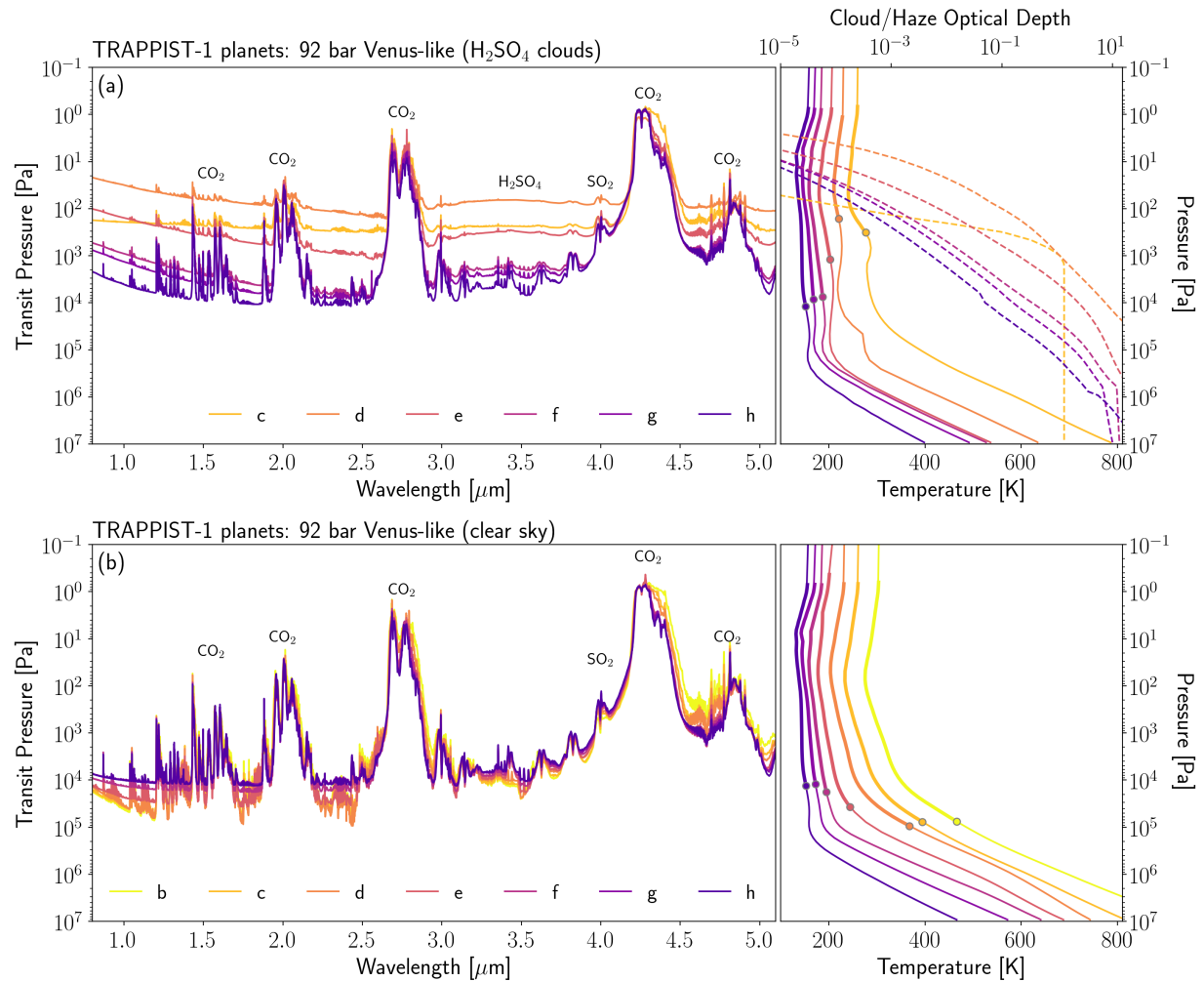


Figure 3.1 Transmission spectrum models of the TRAPPIST-1 planets assuming they possess thick Venus-like atmospheres with H_2SO_4 clouds (top panels) and without clouds (bottom panels) (models from [Lincowski et al., 2018](#)). The left panel shows transmission spectra in units of pressure probed into the atmosphere. The right panel shows the atmospheric temperature structure (solid lines; lower axis) and the cumulative vertical cloud/haze optical depth from the top of the atmosphere (dashed lines; upper axis). Thicker line styles indicate the region of the atmosphere that the transmission spectrum is sensitive to, and dots are shown on the temperature profiles to indicate the highest possible pressure that may be probed by the transmission spectrum for such planets. H_2SO_4 clouds effectively prevent the lower atmosphere of Venus-like exoplanets from being remotely sensed by a transmission spectrum. However, for the cooler planets, cloud formation occurs at lower altitudes and enables the transmission spectrum to probe deeper into the atmosphere.

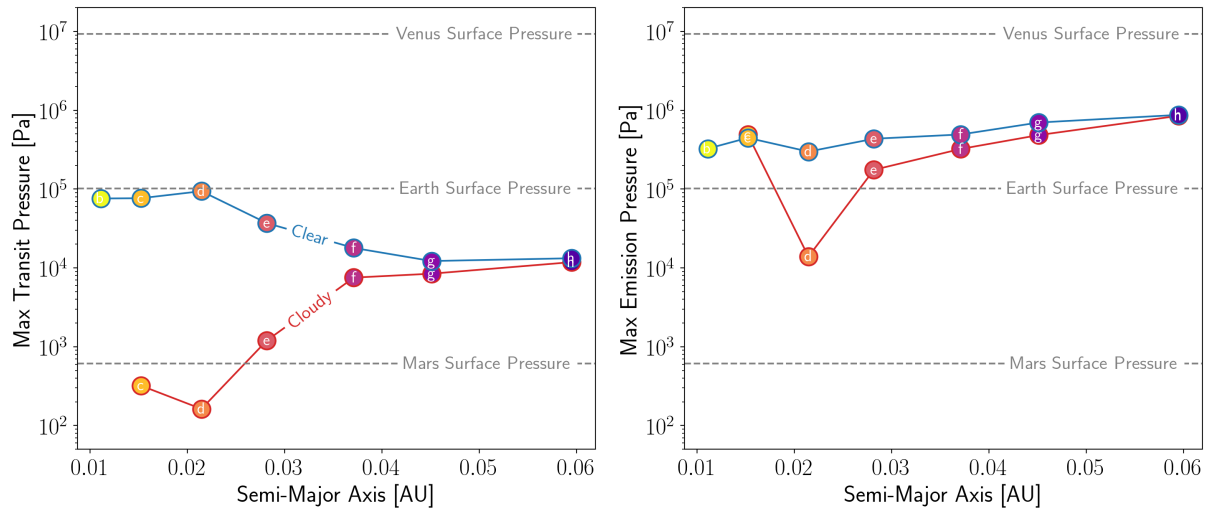


Figure 3.2 Maximum pressure probed in the transmission spectrum (left panel) and emission spectrum (right panel) of clear (blue line) and cloudy (red line) Venus-like models of the TRAPPIST-1 planets as a function of semi-major axis. The surface pressures of Venus, Earth, and Mars are shown as dashed horizontal grey lines for reference. The maximum transit pressure appears to increase with semi-major axis for models with H_2SO_4 clouds, while it decreases with semi-major axis for clear sky models. Although the maximum pressure probed in emission exceeds that of transmission, it does not approach the simulated Venusian surface pressure.

the cloudy inner planets compared to clearsky) and a significant variance in the pressure of the spectral continuum from one cloudy planet to the next ($\sim 100\times$ lower continuum pressure for the cloudy inner planets compared to the cloudy outer planets).

The left panel of Figure 3.2 shows the maximum pressure probed by the transmission spectra (shown in Fig. 3.1) as a function of semi-major axes for the TRAPPIST-1 planets. Clear Venus-like atmospheres (blue lines) have their highest pressures accessible in the $1\text{-}3\ \mu\text{m}$ range between the CO_2 bands (e.g. at about 1.7 , 2.4 , and $3.1\ \mu\text{m}$), while Venus-like atmospheres with H_2SO_4 clouds (red lines) have their highest pressures accessible in the $2.4\text{-}2.6\ \mu\text{m}$ range where H_2SO_4 aerosol scattering is weakest and just short of the $2.7\ \mu\text{m}$ CO_2 band, which has notably broadened wings when not obscured by aerosols. These wavelengths offer the best opportunity to probe deepest into Venus-like atmospheres in transmission, and offer observation leverage for retrieving a cloud-top or surface pressure.

There are no wavelengths at which the transmission spectra of Venus-like TRAPPIST-1 planets access their lower atmospheres. For the inner TRAPPIST-1 planets, if they have Venus-like atmospheres, then the transmission spectrum will only probe down to about the Martian surface pressure (610 Pa). If they are not cloudy, then they may be probed down to about the surface pressure of Earth (101 kPa). For the outer TRAPPIST-1 planets, the presence of clouds minimally affects the maximum transit pressure, and yet they still cannot be probed to higher pressures than about 10^4 Pa. At least two, and up to five, orders of magnitude in pressure exist between the maximum transit pressure and the unseen Venusian surface pressure, which these models share.

Despite the inaccessibility of lower atmospheres, clear and cloudy atmospheres exhibit distinctly opposing trends in the depth into their atmospheres that may be probed as a function of semi-major axis. Clear CO₂ atmospheres gently slope from higher pressures accessible for the inner planets to lower pressures for the outer planets. However, Venus-like atmospheres with H₂SO₄ clouds generally increase from lower pressures accessible for the inner planets to higher pressures accessible for the outer planets.

The divergent scaling with semi-major axis seen between clear and cloudy atmospheres is a result of the underlying physics that controls the transmission spectrum continuum. For clear atmospheres, refraction places a fundamental limit on the depth into the atmosphere that can be accessed by a transmission spectrum (B  tr  mieux and Kaltenegger, 2014; Misra et al., 2014). Clear atmospheres with larger semi-major axes cannot be probed as deeply as those closer to the star due to the dependence of the critical refraction pressure on (1) the angular size of the host star as seen from the transiting planet and (2) the scale height of the planetary atmosphere, both of which decrease with semi-major axis for similar composition planets within the same planetary system.

For cloudy atmospheres, the optically thick cloud deck limits the depth into the atmosphere that can be probed by the transmission spectrum. However, since the fixed temperature of sulfuric acid cloud condensation occurs lower in the atmospheres (at higher pressures) for cooler planets, [Lincowski et al. \(2018\)](#) demonstrate a steady transition from high altitude clouds at low pressures for the inner TRAPPIST-1 planets to low altitude clouds at higher pressures for the outer TRAPPIST-1 planets. This cloud formation trend is manifested in the observable transmission spectrum contin-

uum, and enables increasingly higher pressures to be accessed for planets with increasing distance from their parent star. The clear and cloudy trends in Figure 3.2 converge for planets at sufficiently large orbital separations (e.g. TRAPPIST-1 f, g and h) as the cloud tops drop below the critical refraction pressure.

3.3.2 Emission Spectroscopy

We also consider the limits of atmospheric study for Venus-like planets using emission spectroscopy. Figure 3.3 shows the depth into the atmosphere that is probed by a secondary eclipse thermal emission spectrum for Venus-like models of the TRAPPIST-1 planets with and without H_2SO_4 clouds. Similar to Figure 3.1, the left panels of Figure 3.3 show the “emission pressure”, or the average pressure level in the atmosphere from which most of the thermal emission emerges at normal incidence. We show the emission pressure over a wavelength range that it is applicable to JWST’s Mid-IR Instrument (MIRI) Low Resolution Spectrometer (LRS), which is optimal for observing thermal emission from the TRAPPIST-1 and similar exoplanets during secondary eclipse (Lustig-Yaeger et al., 2019b). The right panels of Figure 3.3 show the atmospheric thermal structure on the same pressure y-axis as the emission pressure for comparison, with line thickness highlighting the thermally emitting region of the atmosphere.

Numerous transparent windows in the near- and mid-IR offer glimpses into the deeper atmosphere of Venus-like planets. For instance, at $6\text{ }\mu\text{m}$ there is a prominent window where thermal emission can be seen coming from pressures of about $10^5 - 10^6\text{ Pa}$. There are also windows in the NIR that probe even deeper into the atmosphere, but which are not shown here due to the insensitivity of JWST to thermal emission in the NIR. Note, however, that these NIR windows have been used extensively to study the surface and near-surface of Venus (see Meadows and Crisp, 1996; de Bergh et al., 2006).

The right panel of Figure 3.2 shows the maximum pressure reached by an emission spectrum over the MIRI LRS bandpass as a function of semi-major axis for the TRAPPIST-1 planets. In all cases considered here, emission spectroscopy probes higher pressures than transmission spectroscopy. In general, cloudy atmospheres emit from higher altitudes and lower pressures than clear

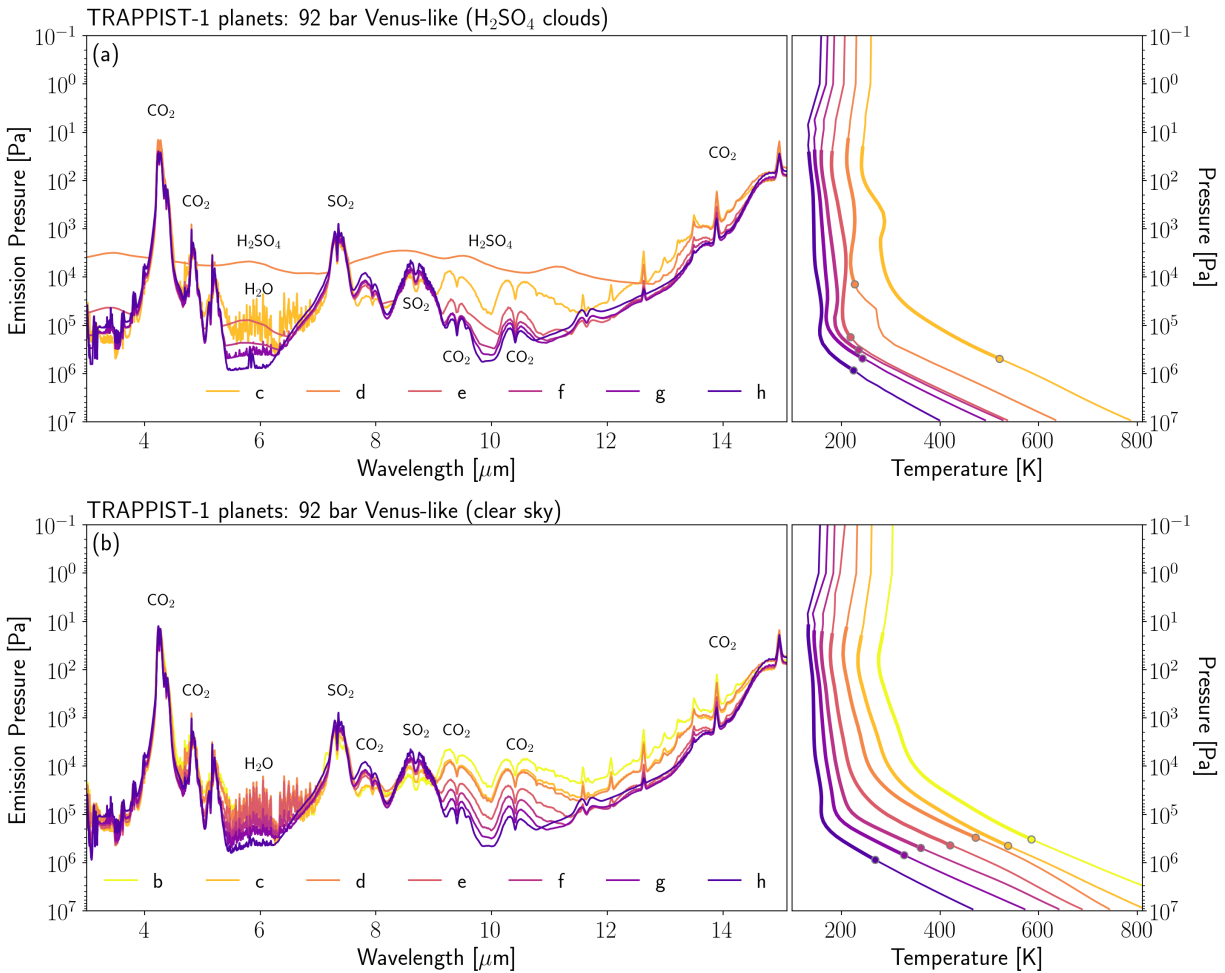


Figure 3.3 Analogous to Figure 3.1 but for emission spectrum models of the TRAPPIST-1 planets assuming they possess thick Venus-like atmospheres with H_2SO_4 clouds (top panels) and without clouds (bottom panels). The left panel shows the atmospheric pressure from which the planets thermal emission emanates ($\tau = 1$). Various NIR and MIR atmospheric windows, which enable the planet to thermally cool to space, may allow distant observers to detect the higher pressures and temperatures of the sub-cloud atmosphere, but these windows can be closed by aerosol extinction (e.g. TRAPPIST-1 d) and gaseous absorption (e.g. H_2O at $6 \mu\text{m}$), but they are infeasible for JWST to detect for planets cooler than TRAPPIST-1 c.

atmospheres. However, even thermal emission from the clear atmospheres is coming from over an order of magnitude lower pressures than the surface.

Whereas transmission spectroscopy is more sensitive to the location of the cloud top due to

the slant optical depth (Fortney, 2005), emission spectroscopy is more sensitive to the total optical depth of the clouds in the atmospheric column. The maximum emission pressure trend with semi-major axis for the cloudy Venus-like models notably tracks the total extinction optical depth of the aerosols, seen in Figure 7¹ of Lincowski et al. (2018). TRAPPIST-1 c and d effectively bracket the small particle haze and thick cloud regimes, respectively, which both exist in the Venus atmosphere (Crisp, 1986). That is, TRAPPIST-1 c's total H₂SO₄ aerosol extinction optical depth is of order unity, which does not substantially modify the thermal emission spectrum from the clear sky case. However, TRAPPIST-1 d's H₂SO₄ aerosol total optical depth peaks among the TRAPPIST-1 planets at $\tau \sim 30$, due to the strong formation rate, cooler temperature, and lower gravity, which allows larger particles to be sustained and suspended. The resulting extended haze and cloud layer significantly mutes spectral features in the thermal emission spectrum and restricts remote sensitivity to the lower atmosphere. Beyond TRAPPIST-1 d, the total aerosol optical depth decreases with semi-major axis, revealing higher atmospheric pressures and explaining the convergence of the clear and cloudy lines in the right panel of Figure 3.2. However, observing thermal emission spectra from cool terrestrial exoplanets is not feasible with JWST, which we discuss next.

Figure 3.4 shows simulated secondary eclipse spectra for our Venus-like TRAPPIST-1 models. At short wavelengths each eclipse spectrum is dominated by reflected light, while at long wavelengths they are dominated by thermal emission. Both radiative source functions decrease with semi-major axis, making eclipse spectroscopy of temperate and cool planets require at least an order of magnitude higher precision observations. Note that thick clouds can mute the thermal emission spectrum features, as best exemplified by TRAPPIST-1 d in our models. The aforementioned 6 μ m spectral window shows enhanced flux of thermal radiation, particularly in the clear sky models, extending nearly 100 ppm above the thermal continuum for TRAPPIST-1 b.

¹[Link to figure on the Astronomy Image Explorer](#)

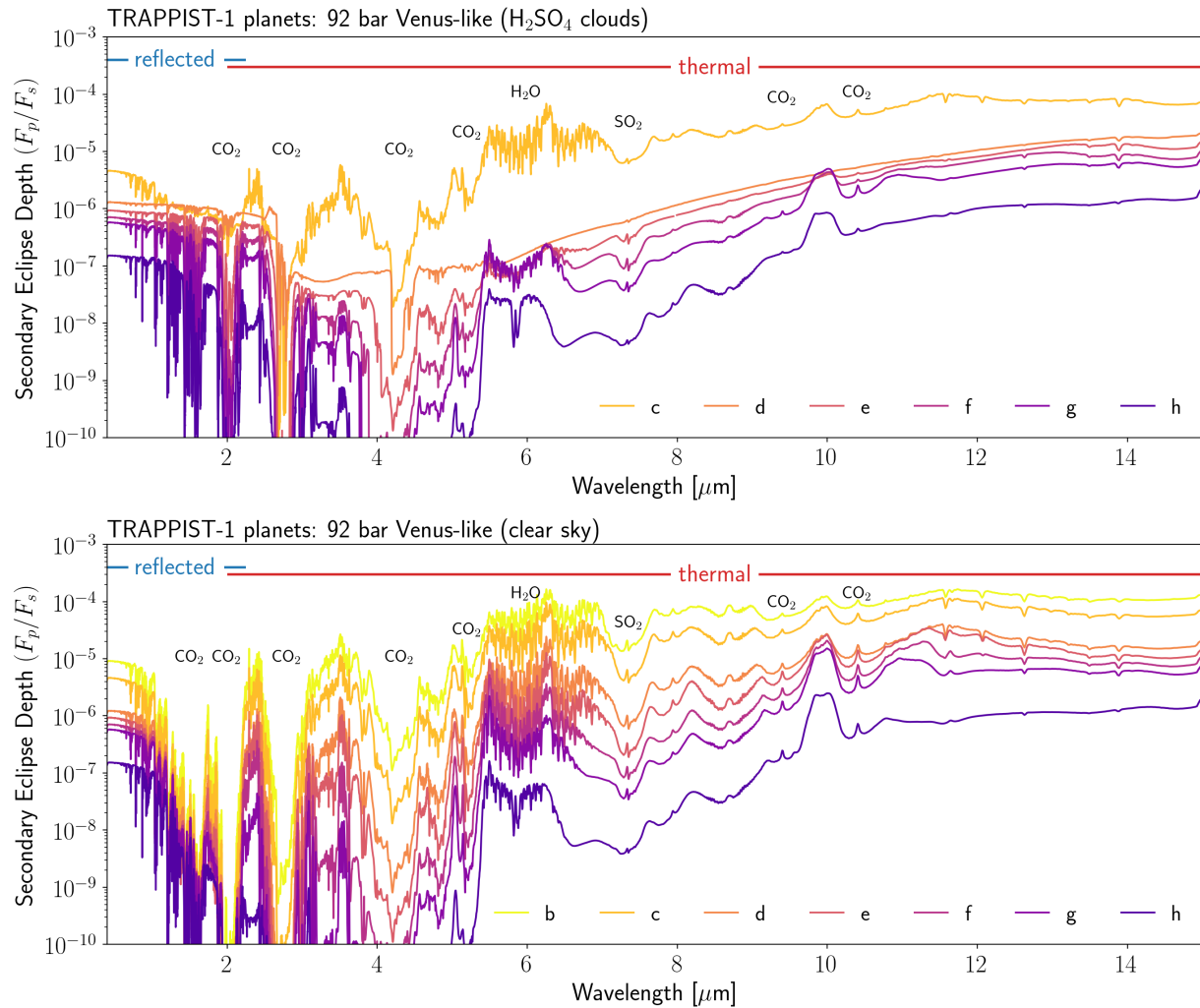


Figure 3.4 Synthetic secondary eclipse spectra of the TRAPPIST-1 planets assuming they possess either thick Venus-like atmospheres with H_2SO_4 clouds (top panel) or no clouds (bottom panel). Reflected stellar flux dominates at short wavelengths, while thermal emission dominates at long wavelengths.

3.4 Discussion

We used the TRAPPIST-1 planets to demonstrate the difficulty intrinsic to identifying and studying the lower atmospheres of Venus-like exoplanets with transmission and emission spectroscopy, which is applicable to near-term efforts with JWST. In the case of true Venus analog exoplanets

with sulfuric acid clouds and 92 bar surface pressures, transmission spectroscopy will only be sensitive to pressures between $10^2 - 10^4$ Pa (0.001 - 0.1 bar). Although the tenuous above-cloud atmosphere could still be detected for all of the TRAPPIST-1 planets with JWST (Lustig-Yaeger et al., 2019b), inferring the presence of the lower atmosphere—which perhaps best defines the very nature of Venus—will be a significant challenge, likely exceeding the scope of transmission spectroscopy. In the following discussion we will present observational approaches that may best constrain lower atmospheres for Venus-like exoplanets (§3.4.1). We then discuss how when lower atmospheres cannot be observationally constrained, there exists a potential ambiguity between cloud-top and solid surface (§3.4.2) that may pose clouds as a false positive for atmospheric erosion (§3.4.3). We will finish with a discussion of possible strategies and opportunities to mitigate these challenges (§3.4.4).

3.4.1 *Accessing the lower atmospheres of Venus-like exoplanets*

Although emission spectroscopy is able to probe deeper than transmission, optically thick clouds may still impede lower atmosphere studies. Additionally, the NIR transparent windows that are used to probe down to the surface of Venus for spatially-resolved remote-sensing studies in our Solar System, are out of observational reach for exoplanets in secondary eclipse because dayside reflected light will overwhelm photons emerging from the lower atmosphere (see Fig. 3.4), analogous to how the illuminated crescent of Venus must be spatially avoided when observing nightside thermal windows (Meadows and Crisp, 1996). Instead, thermal emission measurements that are sensitive to the lower atmosphere must push to longer wavelengths where the reflected stellar SED is naturally dimmed in the Rayleigh-Jeans tail. In particular, there is a $6\ \mu\text{m}$ opacity window that is optimally located for observations with MIRI LRS. However, this MIR window does not probe within an order of magnitude of the surface pressure and falls short of the surface temperature by over 200 K in our TRAPPIST-1 models. Additionally, water vapor has the potential to close the $6\ \mu\text{m}$ opacity window, so atmospheres with more atmospheric water than the Lincowski et al. (2018) Venus-like models may not have this observable window into the lower atmosphere.

Promising observational approaches have been proposed to efficiently identify the presence, or

lack, of hot terrestrial atmospheres using photometry, and in some cases these methods may immediately favor the existence of thick atmospheres. Thermal phase curves with large day-night contrasts can rule out thick atmospheres that would otherwise redistribute heat to the nightside (Seager and Deming, 2009; Selsis et al., 2011; Kreidberg and Loeb, 2016; Koll and Abbot, 2016; Kreidberg et al., 2019), while an offset hot spot from the substellar point could favor a thick atmosphere (Demory et al., 2016). Similarly, secondary eclipse photometry could indicate a low dayside temperature due to atmospheric heat redistribution (Koll et al., 2019) or a high bond albedo due to subsolar clouds (Mansfield et al., 2019). Although promising for warm to hot planets ($T_{eq} = 300 - 880$ K), constraints from secondary eclipses, and thermal studies in general, become quickly infeasible with JWST for cooler planets into and beyond the HZ (Lustig-Yaeger et al., 2019b; Koll et al., 2019). For these cooler planets, transmission spectroscopy is especially favorable because the strength of spectral features scales with the planet’s atmospheric scale height, $H = kT/\mu g \propto T_{eq}$, rather than the planet’s thermal emission which scales much more strongly with temperature for temperate planets not in the Rayleigh-Jeans limit² (Winn, 2010), as shown for the TRAPPIST-1 planets in Fig. 3.4.

3.4.2 *An ambiguity between cloud-top and solid surface*

However, because of a lack of thermal emission data possible for the temperate and cooler planets, the interpretation of their transmission spectra is paramount in the era of JWST, but it may be complicated by an ambiguity between cloud-top and solid surface.

For any single exoplanet the presence of aerosols may be quite difficult to diagnose with transmission spectroscopy. In principle, scattering slopes and/or absorption features from aerosols may be used to break the cloud-surface degeneracy. However, high S/N observations would be needed to detect these features as they are 10-20 ppm in Venus-like TRAPPIST-1 models (Lincowski et al., 2018), which is much smaller than any of the potentially detectable spectral features with JWST (Lustig-Yaeger et al., 2019b). As a result, this cloud-top–solid-surface ambiguity is more likely

²For instance, the blackbody flux scales approximately as T^8 near 300 K and at $15 \mu\text{m}$ (for $< 1\%$ errors incurred by a Taylor series expansion of the exponential).

to emerge for spectra with low S/N either due to prohibitively long exposure times or observing strategies that seek a large sample of spectra at low to moderate S/N for statistical comparative planetology (e.g. [Bean et al., 2017](#); [Checlair et al., 2019](#)).

3.4.3 *Clouds as a statistical false positive for atmospheric loss*

Across a population of exoplanets—either within a single planetary system, as in the case of TRAPPIST-1, or for an ensemble of planets from many systems—inferred trends in cloud-top pressure with stellar irradiation (as seen in Figure 3.2) for similar composition atmospheres may erroneously appear as a surface pressure trend due to atmospheric loss processes. Specifically, the left panel of Figure 3.2 clearly shows that higher pressures are probed for cloudy exo-Venuses with larger semi-major axes. Without our prior knowledge on the inclusion of clouds in our models, and under the veil of the cloud-surface ambiguity, these trends could readily appear as trends in surface pressure. That is, *are we seeing thicker atmospheres as stellar irradiation decreases, or lower cloud decks, or both?*

This potential statistical false positive may be particularly nefarious because atmospheric loss is predicted to play a major role in sculpting the atmospheres of small rocky planets orbiting late M dwarfs. Models indicate that the TRAPPIST-1, and similar close-in, planets may have had their atmospheres completely eroded by x-ray and extreme ultraviolet radiation (XUV; [Airapetian et al., 2017](#); [Garcia-Sage et al., 2017](#); [Roettenbacher and Kane, 2017](#); [Zahnle and Catling, 2017](#); [Dong et al., 2018](#); [Fleming et al., 2020](#)), although sufficient volatile outgassing could help maintain atmospheres ([Bolmont et al., 2017](#); [Garcia-Sage et al., 2017](#); [Bourrier et al., 2017](#)). Furthermore, [Dong et al. \(2018\)](#) found that the outer TRAPPIST-1 planets are capable of retaining their atmospheres over billions of years, while the inner planets may not be able to. Thus, observing a trend of thin atmospheres for the inner TRAPPIST-1 planets to thick atmospheres for the outer TRAPPIST-1 planets may appear consistent with a “cosmic shoreline”—an empirical division between planets with and without atmospheres based on the relationship between total incident stellar radiation and planetary escape velocities ([Zahnle and Catling, 2013, 2017](#)). Testing the cosmic shoreline hypothesis on exoplanet data will require a statistical comparative planetology approach, as outlined

in [Bean et al. \(2017\)](#) and [Checlair et al. \(2019\)](#), but care must be taken to understand and mitigate degenerate exoplanet population trends.

3.4.4 *Mitigation strategies & opportunities*

Ambiguous trends in the maximum pressure seen across a population of planetary transmission spectra can also be used to implicate clouds and potentially expose their composition. First, [Lincowski et al. \(2018\)](#) found that sulfuric acid clouds did not condense in the Venus-like model atmospheres of TRAPPIST-1 b, and so our analysis did not include a cloudy TRAPPIST-1 b. However, as noted by [Lincowski et al. \(2019\)](#), detecting the atmosphere of the innermost planet in multi-planet systems could strongly increase the likelihood that similar size planets at longer orbits have atmospheres, because the loss of volatiles due to escape over a planet’s history is expected to decrease with increasing distance from the star. Second, by understanding what physical and chemical conditions may produce continuum pressure trends in terrestrial transmission spectra, it may be possible to rule out false positives scenarios, in the same way that false positive biosignatures may be identified and mitigated using additional environmental context from the atmosphere and stellar environment ([Meadows, 2017](#); [Meadows et al., 2018b](#); [Catling et al., 2018](#)). If distinct population trends seen with stellar insolation are consistent with predictions from cloud condensation modeling, then clouds could potentially be revealed by statistical characterization where they were unidentifiable in any single planet spectrum. This is a terrestrial exoplanet analog to cloud condensation trends observed in brown dwarf atmospheres across the L/T transition (e.g. [Ackerman and Marley, 2001](#); [Morley et al., 2012](#)).

Furthermore, the distinctly opposing trends that we found between clear and cloudy atmospheres as a function of semi-major axis (see Figure 3.2) highlights the potential to use the transmission spectrum continuum pressure to group similar populations of terrestrial exoplanets. The maximum pressure probed in clear atmospheres is set by the critical refraction pressure which scales with the angular size of the star as seen from the planet ([B  tr  mieux and Kaltenegger, 2014](#); [Misra et al., 2014](#)), and prevents access to higher pressures at larger semi-major axes (for planets orbiting similar sized stars, e.g. late M dwarfs). Conversely, sulfuric acid cloud condensation

allows access to higher pressures at larger semi-major axes. These contrasting trends could be detected by retrieving cloud top or reference pressures for multiple planets within the same system and may enable thick clear and cloudy atmospheres to be distinguished.

These distinguishing characteristics extend into the terrestrial domain the concepts presented in [Sing et al. \(2016\)](#) of an observable distinction between clear and cloudy atmospheres in an ensemble of exoplanet spectra. [Sing et al. \(2016\)](#) found that the strength of water absorption features in the spectra of hot Jupiter exoplanets is correlated with cloud and haze scattering slopes in the spectra, indicating that clouds/hazes may be obscuring the water column, rather than seeing an intrinsic trend in water vapor abundance. Similarly, we have explored how the strength of gaseous absorption features relative to the spectral continuum could potentially be used to discriminate between populations of thin, thick/clear, and thick/cloudy atmospheres, even if the transmission spectra are individually difficult to diagnose.

The transition between planets with and without atmospheres and the transition between terrestrial and gaseous planets are two bookends of the high mean molecular weight, terrestrial atmosphere regime. The emerging paucity of planets with radii $\sim 1.6 R_{\oplus}$ orbiting Sun-like stars ([Rogers, 2015](#); [Fulton et al., 2017](#)) likely constrains the presence of terrestrial atmospheres on the large planet boundary. JWST will offer a first opportunity to investigate this boundary on the small planet end, as we continue to explore the effects of atmospheric escape (e.g. [Lehmer and Catling, 2017](#)), and attempt to map the cosmic shoreline.

3.5 Conclusions

The lower atmospheres of Venus-like exoplanets may elude our characterization efforts with JWST due to the presence of sulfuric acid clouds, which both dictate remote studies of Venus and constitute a potential terrestrial exoplanet analog to the high altitude clouds and hazes that currently limit the characterization of gaseous exoplanets with transmission spectroscopy. For hot exo-Venuses, MIR opacity windows observed during secondary eclipse may offer glimpses of thermal emission from the atmosphere just below the clouds, potentially allowing for high-pressure, greenhouse-heated lower atmospheres to be directly inferred.

However, for temperate to cold exo-Venuses observed with transmission spectroscopy, a sulfuric acid cloud deck may appear indistinguishable from a solid surface at low to moderate S/N as both cause the spectral continuum to be flat. In these cases, Venus-like atmospheres should still be detectable via CO₂ absorption features, but appear like tenuous low pressure atmospheres due to a lack of observational constraint from the lower atmosphere. For Venus-like atmospheric models of the TRAPPIST-1 planets, we demonstrated that the sulfuric acid clouds drop in altitude to higher pressures with semi-major axis. This effect has the potential to be misinterpreted as a trend of increasing surface pressure with decreasing stellar insolation and may appear suspiciously consistent with atmospheric escape. Looking ahead, the prospect of different populations of terrestrial exoplanets—cloudy exo-Venuses and stellar windswept worlds—presenting similar observables motivates the need for additional climate, photochemical, cloud formation, and atmospheric escape modeling to uncover observable characteristics that effectively discriminate between different populations of exoplanets, and observing strategies tailored to test these hypotheses.

Part II

**DEVELOPING A FOUNDATION TO CHARACTERIZE THE
HABITABILITY & BIOSIGNATURES OF DIRECTLY-IMAGED
EXOPLANETS**

Chapter 4

DETECTING OCEAN GLINT ON EXOPLANETS USING MULTIPHASE MAPPING

Portions of this chapter were originally published in collaboration with Victoria S. Meadows, Guadalupe Tovar Mendoza, Edward W. Schwieterman, Yuka Fujii, Rodrigo Luger and Tyler D. Robinson in the December 2018 edition of The Astronomical Journal (Lustig-Yaeger et al., 2018, AJ, Vol. 156, 6; 2018 © American Astronomical Society, DOI: 10.3847/1538-3881/aaed3a), and are reproduced below with the permission of the American Astronomical Society.

Summary

Rotational mapping and specular reflection (glint) are two proposed methods to directly detect liquid water on the surface of habitable exoplanets. However, false positives for both methods may prevent the unambiguous detection of exoplanet oceans. We use simulations of Earth as an exoplanet to introduce a combination of multiwavelength, multiphase, time-series direct-imaging observations and accompanying analyses that may improve the robustness of exoplanet ocean detection by spatially mapping the ocean glint signal. As the planet rotates, the glint spot appears to “blink” as Lambertian scattering continents interrupt the specular reflection from the ocean. This manifests itself as a strong source of periodic variability in crescent-phase disk-integrated reflected light curves. We invert these light curves to constrain the longitudinal slice maps and apparent albedo of multiple surfaces at both quadrature and crescent phase. At crescent phase, the retrieved apparent albedo of ocean-bearing longitudinal slices is increased by a factor of 5, compared to the albedo at quadrature phase, due to the contribution from glint. The land-bearing slices exhibit no significant change in apparent albedo with phase. The presence of forward-scattering clouds in our simulated observation increases the overall reflectivity toward crescent, but we find that clouds do not correlate with any specific surfaces, thereby allowing for the phase-dependent glint effect to be interpreted as distinct from cloud scattering. Retrieving the same longitudinal map at quadrature and crescent phases may be used to tie changes in the apparent albedo with phase back to specific geographic surfaces (or longstanding atmospheric features), although this requires ideal geometries. We estimate that crescent-phase time-dependent glint measurements are feasible for between 1 and 10 habitable zone exoplanets orbiting the nearest G, K, and M dwarfs using a space-based, high-contrast, direct-imaging telescope with a diameter between 6 and 15 m.

4.1 Introduction

Finding and characterizing habitable exoplanets is an emerging frontier, with the detection of liquid surface water as a fundamental goal on the path towards discovering life beyond the Solar System. A promising few small exoplanets have been discovered in the habitable zones of their parent stars (Anglada-Escudé et al., 2016; Gillon et al., 2016, 2017b; Dittmann et al., 2017). These potentially habitable exoplanets are likely high priorities for upcoming observations to characterize their environments and search for signs of habitability. However, characterizing the nature of a planet's habitability will likely require many different lines of evidence synthesized to assess the likelihood of surface liquid water (for a recent review, see Robinson, 2018; Meadows and Barnes, 2018).

Detecting an ocean on the surface of an exoplanet is the most straightforward way to confirm habitability, and high-contrast direct imaging is the most promising type of observation to accomplish this. Unlike transmission spectroscopy, which cannot probe the surface and near-surface atmosphere (García Muñoz et al., 2012; Bétrémieux and Kaltenegger, 2014; Misra et al., 2014), direct imaging can probe all the way to the planetary surface, and is less susceptible to obscuration by atmospheric extinction (Fortney, 2005), such as by aerosols (Kreidberg et al., 2014). As a result, only direct imaging offers a feasible path to exoplanet ocean discovery. There are two ways exoplanet oceans could be directly detected using direct imaging of planetary reflected light: spatial mapping using the planet's rotational lightcurves, and ocean glint using phase-dependent and/or polarization observations (Zugger et al., 2010).

Although any given snapshot of a directly-imaged exoplanet is disk-integrated and contains no spatial information, as the planet rotates on its axis, time-series observations permit spatial mapping of heterogeneities in the atmosphere and on the surface (Ford et al., 2001; Pallé et al., 2008; Oakley and Cash, 2009). Multi-wavelength, time-series observations can be used to infer the longitudinal map and color of multiple surfaces on a planet (Cowan et al., 2009; Cowan and Strait, 2013), and multi-epoch observations can leverage the planet's obliquity and/or inclination to map latitudes and construct two-dimensional color surface maps of exoplanets (Kawahara and Fujii, 2010, 2011; Fujii and Kawahara, 2012; Berdyugina and Kuhn, 2017). Such color maps could

be used to identify large oceans on exoplanets.

An alternative approach to detecting exoplanet oceans is to observe glint—the specular reflection typical of a flat surface or liquid when viewed at an oblique angle relative to the illumination source—a remotely detectable characteristic of the Earth ocean (Sagan et al., 1993). Glint has been suggested as a promising habitability marker (Williams and Gaidos, 2008; Robinson et al., 2010), and produces a substantial relative increase in the brightness of Earth when viewed at crescent phases (Robinson et al., 2014a). Targeted crescent phase observations could be used to identify the reflected light glint signature of an ocean-bearing exoplanet.

Although spatial mapping and glint are promising, both have degeneracies in interpretation that may limit the confidence with which each technique may be used to directly detect oceans. Lightcurve inversion techniques to identify the albedo and geography of multiple surfaces, including oceans (e.g. Cowan et al., 2009), are under-constrained inference problems that suffer from mathematical degeneracies (Fujii et al., 2017), and have traditionally required *a priori* assumptions about the surface composition to find unique solutions (e.g. Fujii et al., 2010). More recently, Fujii et al. (2017) imposed the simple physical prior that the surface albedo be bounded between zero and one. This was insufficient to fully break this “rotational unmixing” degeneracy, but showed promise for achieving simultaneous constraints on both the color and geography of Earth-like oceans due to their low albedo and vast global distribution. Furthermore, simply inferring a dark blue surface is not a definitive ocean detection, as Rayleigh scattering above a grey surface can masquerade as “blue” in color (Krissansen-Totton et al., 2016). Glint also has potential false positives. The non-Lambertian increase in reflectivity towards crescent phase that is characteristic of ocean glint can occur independent of a glinting ocean due to forward scattering from clouds and hazes, as observed for Venus (e.g. Mallama et al., 2006) and Earth (e.g. Robinson et al., 2010). Furthermore, high latitudes, which are more-likely to be ice-covered, have an increased contribution to the disk-integrated flux at crescent phase relative to more illuminated phases, which can naturally cause an increase in reflectivity towards crescent phase (Cowan et al., 2012). Additional strategies and techniques are needed to help increase the robustness of ocean detection.

Here we introduce a sequence of direct imaging observations that combine longitudinal map-

ping with phase-dependent glint measurements to increase the robustness of ocean detection. Specifically, multi-wavelength, time-series observations at multiple phase angles will enable spectral unmixing as a function of illumination. An exoplanet in a nearly edge-on inclination orbit, with a negligible obliquity will have rotational lightcurves that probe the same latitudes at all orbital phases, and therefore the two-dimensional mapping problem (e.g. [Fujii and Kawahara, 2012](#)) will be intractable as longitudinal maps will appear the same at all observed phases. However, recognizing the same longitudinal map at another phase angle would offer a second look at the same surface, thereby allowing observers to trace the reflective properties of a given geographical surface as a function of illumination angle. Thus, the reoccurring map would lend confidence to any assumption that the same surface is being viewed, while the surface albedo necessary to reproduce the observed lightcurves would reveal its scattering phase function. In this way, it may be possible to infer that the same surface that appears dark blue at gibbous phases is significantly more reflective at crescent phase, thereby proving the non-isotropic scattering behaviour of the surface.

In this chapter we demonstrate the potential utility of multi-phase longitudinal mapping as it may be applied to detect oceans on directly-imaged exoplanets. In §4.2 we describe our methods used to generate synthetic lightcurves of Earth and to invert them to map the planet. In §4.3 we present results on multi-phase longitudinal mapping and observational feasibility tests. We offer a discussion in §4.4 and conclude in §4.5.

4.2 *Methods*

To investigate and develop a more robust technique for exoplanet ocean detection, we forward model multi-wavelength, time-series, photometry of the disk-integrated Earth, apply realistic noise for direct-imaging exoplanet observations, and then invert the problem to solve for the geographical maps and colors of dominant surface types. The methods to generate synthetic lightcurves of Earth and fit maps to the lightcurves are described in the following subsections.

4.2.1 Generating Synthetic Lightcurves

We use the VPL Earth Model to create high spectral resolution, time-dependent, disk-averaged observations of Earth at multiple phases (Robinson et al., 2011). The VPL Earth Model gathers observed cloud and ice coverage, surface wind speed, and atmospheric thermal structure and gas mixing ratio profiles from Earth observing satellites. These Earth science data products are used as inputs to the VPL Earth Model, which maps them to a HEALPix¹ (Górski et al., 2005) projection of Earth. Multi-stream, multi-scattering, line-by-line radiative transfer calculations are then performed across the grid of cloud coverages and illumination angles, for each HEALPix pixel, using the Spectral Mapping and Atmospheric Radiative Transfer (SMART) model (Meadows and Crisp, 1996, developed by D. Crisp). SMART, and by extension the VPL Earth Model, include the spectral reflection of the Sun/star off the ocean using the Cox-Munk glint model (Cox and Munk, 1954). Performing SMART calculations yields a database of radiances that contain contributions emerging from different regions of Earth at any given time. Spatially- and spectrally-resolved data cubes of Earth are constructed according to user specified sub-solar and sub-observer longitudes and latitudes (Robinson et al., 2011). An example VPL Earth model image of the crescent-phase Earth with a clearly visible glint spot is provided in Figure 1 of Robinson et al. (2010). These data cubes are then integrated over the visible disk to give a single composite disk-integrated spectrum.

The VPL Earth Model has been validated against spacecraft observations of Earth, including spatially resolved and time-dependent observations made by the EPOXI mission (Robinson et al., 2011), and *phase-dependent* measurements of Earth showing the glint effect, that were taken by the Lunar Crater Observing and Sensing Satellite (LCROSS) (Robinson et al., 2014a). We refer the reader to Robinson et al. (2011, 2014a) and Schwieterman et al. (2015a, 2016a)² for a more complete description of the VPL Earth Model and its validation.

The VPL Earth Model is used to generate multiple different Earth and Earth-like spectral models that enable longitudinal mapping investigations at multiple phases. We make a simplifying

¹<http://healpix.sf.net/>

²Earth spectral database used for this work is available upon request. Previous similar datasets are available at: <http://depts.washington.edu/naivpl/content/spectral-databases-and-tools>

assumption that the sub-solar and sub-observer latitudes are fixed at the equator to simulate a planet with zero obliquity in an edge-on orbit. We compute time-dependent spectra at three phase angles: 45° , 90° , and 135° , which we refer to as gibbous, quadrature, and crescent, hereafter. At each phase we simulate 120 consecutive hours (5 days) of high-resolution, disk-integrated spectra at a 1 hour cadence. We neglect the small ($<5^\circ$) phase change of the planet over the 120 hours of Earth modeling. The computational expense of the VPL Earth Model places limits on the observational duration and observational cadence that we may consider in this work, and in §4.3.1, we focus on experiments with 100 hours of data collection at a 1 hour cadence. We simulate Earth during northern spring. Datasets are generated both with true time-varying cloud coverage, and without clouds.

To simulate realistic observations of our VPL Earth Model spectra as an analog for a directly-imaged exoplanet, we use an open-source Python coronagraph noise model³, originally developed by Robinson et al. (2016) for the WFIRST exoplanet direct imaging mission and expanded upon in subsequent studies (Luger et al., 2017b; Meadows et al., 2018a). The model simulates planetary and background photon count rates as a function of wavelength using analytic parameterizations for telescope, instrumental, and astrophysical noise sources. These sources include Poisson noise, read noise and dark current, coronagraph speckles, scattering off of zodiacal and exo-zodiacal dust, and thermal emission from the mirror. The telescope is assumed to perform a roll maneuver to allow for background subtraction (c.f. Brown, 2005), which increases the background by a factor of two (Robinson et al., 2016).

We perform coronagraph simulations for various telescope designs that encompass telescope mission concepts currently under study by NASA Science & Technology Definition Teams (STDTs) in advance of the 2020 decadal survey. Many of the telescope, instrument, and system parameters used here are the same as those used in Robinson et al. (2016, see their Table 2 and Table 3), with the exceptions listed in Table 4.1. We adopt a raw contrast of $C = 10^{-10}$ to pursue Earth-like exoplanets orbiting Sun-like stars. We consider telescope diameters between 4 - 15 meters, and

³<https://github.com/jlustigy/coronagraph>

Table 4.1. Telescope, instrument, and system model parameters

Parameter	Description	Value(s)
d	distance to system	5 pc
R_p	planetary radius	$1 R_{\oplus}$
α	planet phase angle	$45^\circ, 90^\circ, 135^\circ$
C	coronagraph design contrast	10^{-10}
D_{e-}	dark current	10^{-4} s^{-1}
D	telescope diameter	(4, 15) m
$\mathcal{R}$	instrument spectral resolution	R200
$\mathcal{T}$	end-to-end throughput	(0.02, 0.5)
θ_{IWA}	coronagraph inner working angle	$> 1\lambda/D$
θ_{OWA}	coronagraph outer working angle	$< 60\lambda/D$

Note. — See [Robinson et al. \(2016\)](#) for a complete list of coronagraph parameters and their nominal values. Comma separated values are all considered and comma separated values in parentheses represent explored ranges.

end-to-end throughputs between 0.02 - 0.5. These ranges generally encapsulate those relevant to the Large UV/Optical/IR (LUVOIR⁴; e.g. [Dalcanton et al., 2015](#); [Bolcar et al., 2017](#)) Surveyor and Habitable Exoplanet Imaging Mission (HabEx⁵; e.g. [Mennesson et al., 2016](#)) telescope mission concepts. We adopt a highly optimistic inner working angle (IWA) of $1\lambda/D$ and outer working angle (OWA) of $60\lambda/D$ so that our simulated lightcurves solely represent the photon counting statistics. However, we separately consider constraints on the IWA in §4.3.2. We note that, like [Robinson et al. \(2016\)](#), the IWA and OWA are modeled as strict cutoffs—a step function in throughput from 100% to 0%—whereas formally they are defined as the 50% off-axis throughput point. As a result, our coronagraph modeling neglects throughput effects as the IWA and OWA are approached and exceeded.

⁴<https://asd.gsfc.nasa.gov/luvoir/>

⁵<http://www.jpl.nasa.gov/habex/>

We construct simultaneous, multi-wavelength lightcurves by simulating coronagraph spectroscopy with a resolving power of $\mathcal{R} = \lambda/\Delta\lambda = 200$ at a one hour cadence over a modest 18% bandpass. This gives a low signal-to-noise spectrum each hour, and from hour-to-hour the spectrophotometry samples Earth’s rotational variability. While multiple hours of spectra can be co-added into a higher signal-to-noise spectrum that may be used to retrieve atmospheric gas abundances (Feng et al., 2018), here we instead co-add neighboring spectral resolution elements to increase the signal-to-noise on each hour of photometry. Since current coronagraph designs can only null over relatively modest spectral bandpasses (typically $\Delta\lambda/\lambda \approx 10\% - 20\%$) in a single pointing, we focused on synthetic observations in an 18% bandwidth wavelength range between 0.67 and 0.8 μm . In this wavelength range the Earth exhibits large amplitude rotational variability owing to surface features, such as ocean, soil, and vegetation. This sensitivity is due in large part to the Earth’s atmosphere being optically thin, with very few molecular absorption features and reduced Rayleigh scattering, at most of these wavelengths. The one major exception is the opacity due to the 0.76 μm O_2 absorption band, which is likely to be one of the most sought-after atmospheric diagnostics for Earth-like planets (Meadows, 2017; Meadows et al., 2018b). Consequently, the same multi-wavelength, time-resolved data that is used to map the planet and search for oceans can be coadded to characterize the planet’s atmosphere and search for signs of life. We bin the spectrophotometry into two photometric bands, centered at $\lambda_1 = 0.70 \mu\text{m}$ and $\lambda_2 = 0.77 \mu\text{m}$ with widths $\Delta\lambda_1 = 0.08 \mu\text{m}$ and $\Delta\lambda_2 = 0.07 \mu\text{m}$, sampled over many consecutive hours and with corresponding noise applied, which yield lightcurves for studying the time-variability of the exoplanet. We assume no read time, no reset time, and no down time between consecutive exposures thereby simulating a 100% efficient duty cycle.

4.2.2 *Lightcurve Fitting*

Time-resolved, multi-wavelength planetary lightcurves can be used to constrain the longitudinal map of a planetary surface, using techniques of “spectral-rotational unmixing”, which solve for the longitudinal map of N different unique surface types and their geometric albedo spectrum (Cowan et al., 2009, 2012; Fujii et al., 2017). As described in detail in Fujii et al. (2017) and summarized

here for completeness, the observed apparent albedo of the planet, A_{ij} , at wavelength j and at time i can be decomposed into

$$A_{ij} = \sum_{i,k} \tilde{f}_{ik} a_{kj} \quad (4.1)$$

where $\tilde{f}_{ik}$ is the apparent covering fraction of the k -th surface at the i -th time and a_{kj} is the apparent albedo spectrum of the k -th surface at the j -th wavelength.

Similarly, due to the relationship between time variability and surface heterogeneity for a rotating planet, the observed apparent albedo can be written in terms of a longitudinal map. In this case, $\tilde{f}_{ik}$ may be approximated by

$$\tilde{f}_{ik} \approx \sum_l W_{il} f_{lk} \quad (4.2)$$

where f_{lk} is the area fraction of the k -th surface in the l -th longitudinal slice and W_{il} is the weight of the l -th longitudinal slice at the i -th time, which is a function of the observational geometry. The approximation in Equation 4.2 is only strictly valid when the true area fraction within each longitudinal slice does not change or changes little across the slice for all surfaces. Plugging Equation 4.2 into Equation 4.1, we reach the following expression in terms of longitudinal slices:

$$A_{ij} \approx \sum_{l,k} W_{il} f_{lk} a_{kj} \quad (4.3)$$

Both Equations 4.1 and 4.3 form the two modular forward models in the Surface Albedo Mapping Using Rotational Inference (SAMURAI⁶) Python code developed in this work. SAMURAI attempts to solve the inverse problem for the map ($\tilde{f}_{ik}$ or f_{lk}) and albedo (a_{kj}) of N different surfaces using multi-band $\{j\}$, time-series $\{i\}$ photometric observations $y_i(\lambda_j)$ and their observational

⁶<https://github.com/jlustigy/samurai>

uncertainties $\sigma_i(\lambda_j)$, subject to the following constraints on the model A_{ij} :

$$\begin{cases} \tilde{f}_{ik}, f_{lk} \geq 0 & \text{for any } i, l, k \\ \sum_k \tilde{f}_{ik} = \sum_k f_{lk} = 1 & \text{for any } i, l \\ 0 \leq a_{kj} \leq 1 & \text{for any } k, j \end{cases} \quad \begin{matrix} (4.4a) \\ (4.4b) \\ (4.4c) \end{matrix}$$

Equation 4.4a ensures that the area fraction of any surface k at any time i or in any longitudinal slice l is not negative, while Equation 4.4b enforces that the area fractions for all surface types at any time or in any longitudinal slice sum to unity. Equation 4.4c imposes that the albedo of a surface k at any wavelength j does not exceed 1.

The hard boundaries in 4.4a, 4.4b, and 4.4c make it difficult to efficiently sample the likelihood space using Markov Chain Monte Carlo (MCMC). One way around this is to reparametrize so that the domain of our parameters extends from $-\infty$ to $+\infty$. This can be done by introducing the transformed parameters

$$\begin{cases} b_{kj} = \ln \left(\frac{a_{kj}}{1 - a_{kj}} \right) & \text{for any } k, j \\ g_{lk} = \ln \left(\frac{f_{lk}}{1 - \sum_{m=1}^k f_{lm}} \right) & \text{for } k = 1, \dots, N - 1 \end{cases} \quad \begin{matrix} (4.5a) \\ (4.5b) \end{matrix}$$

where N is the number of surface types included in the model. Note that conditions 4.4a, 4.4b, and 4.4c are implicitly satisfied by this reparameterization, and there is a one-to-one correspondence between albedo matrices a and b , and between area covering fraction matrices f and g . Similar reparameterizations have been applied to gas mixing ratios in exoplanet atmospheric retrieval problems (c.f. Benneke and Seager, 2012).

It is important to note that when reparameterizing our model, we must also transform the prior probability of each model parameter. Since we have no *a priori* information on the albedos and area fractions of individual surfaces, we enforce a flat prior for these variables. This can be obtained by

imposing the prior probabilities

$$\mathcal{P}(b_{kj}(a_{kj})) \propto \frac{e^{b_{kj}}}{(1 + e^{b_{kj}})^2} \quad (4.6)$$

$$\mathcal{P}(g_{lk}(f_{lk})) \propto \frac{e^{g_{lk}}}{(1 + e^{g_{lk}})^2} \quad (4.7)$$

on our transformed variables b and g , where $\mathcal{P}$ is a probability distribution function. Note that Equation 4.7 does not contain the covariance term that arises due to the sum in Equation 4.5b. As a result, Equation 4.7 does not guarantee a flat prior for three or more surfaces, and thus all investigations in this chapter are performed with only two surfaces. We impose one additional prior $\mathcal{O}(g_{lk})$ to ensure that surface covering fractions are ordered from largest to smallest to break a surface labeling degeneracy. Considering all elements in b and g , and $\mathcal{O}$, our full prior is

$$\pi(b_{kj}, g_{lk}) = \mathcal{P}(b_{kj})\mathcal{P}(g_{lk})\mathcal{O}(g_{lk}) \quad (4.8)$$

$$\propto \left(\prod_{k,j} \frac{e^{b_{kj}}}{(1 + e^{b_{kj}})^2} \right) \left(\prod_{l,k} \frac{e^{g_{lk}}}{(1 + e^{g_{lk}})^2} \right) \mathcal{O}(g_{lk}). \quad (4.9)$$

The aim of our parameter estimation is to evaluate the probability of the albedo and map parameters—the posterior probability distribution function (PDF)—given multi-wavelength, time-series data. By Bayes’ Theorem, the posterior PDF is

$$\mathcal{P}(b_{kj}, g_{lk}|y_{ij}) \propto \mathcal{L}(y_{ij}|b_{kj}, g_{lk}) \times \pi(b_{kj}, g_{lk}) \quad (4.10)$$

where the likelihood, $\mathcal{L}(y_{ij}|b_{kj}, g_{lk})$ is easily calculated from the model and data,

$$\ln \mathcal{L}(y_{ij}|b_{kj}, g_{lk}) = -\frac{1}{2} \sum_{i,j} \left(\frac{y_{ij} - W_i^l f_l^k a_{kj}}{\sigma_{ij}} \right)^2. \quad (4.11)$$

Unless otherwise stated, we sample the log-probability (log of Equation 4.10) using the EM-

CEE⁷ implementation of MCMC (Foreman-Mackey et al., 2013). We use a linear combination of 15% affine-invariant (Goodman and Weare, 2010), 70% differential evolution (Nelson et al., 2014), and 15% differential evolution snooker (ter Braak and Vrugt, 2008) MCMC proposal steps to improve convergence in a likelihood space that is, at times, multi-modal, with different sets of surface spectra and geography that can reproduce the observed lightcurves. We transform our sampling parameters b and g back to physical parameters a and f after the MCMC. This analysis results in posterior PDFs for the wavelength-dependent albedo and longitudinal area fraction of each surface included in the fit. Deriving PDFs for our model parameters allows us to quantify uncertainties on our inferences so that we may make robust statements about any changes seen when the experiment is repeated at another phase.

4.3 Results

We performed spectral unmixing experiments at quadrature, crescent, and gibbous phase to determine if a combination of phase-dependent surface mapping and glint observations could provide more robust constraints on the presence of an ocean for an Earth analog exoplanet. We then quantified the observational feasibility and telescope requirements for this technique with respect to coronagraph inner working angle and observed photometric signal-to-noise.

4.3.1 Multi-Phase Longitudinal Mapping

Our first multi-phase mapping experiment focused on models without clouds to provide an ideal end-member to isolate the signal from the ocean. This also allowed us to assess the unique impact of clouds once they were included.

The upper panel of Fig. 4.1 shows the quadrature ($\alpha = 90^\circ$; solid lines) and crescent ($\alpha = 135^\circ$; dotted lines) phase multi-wavelength lightcurves and median model fits to these data. The lightcurves were simulated for an optimistic case of Earth at 5 pc, observed for 100 hours at each phase, at a 1 hour cadence, with a 15 m, 50% throughput space telescope. The spectrophotometry

⁷EMCEE <https://github.com/dfm/emcee>

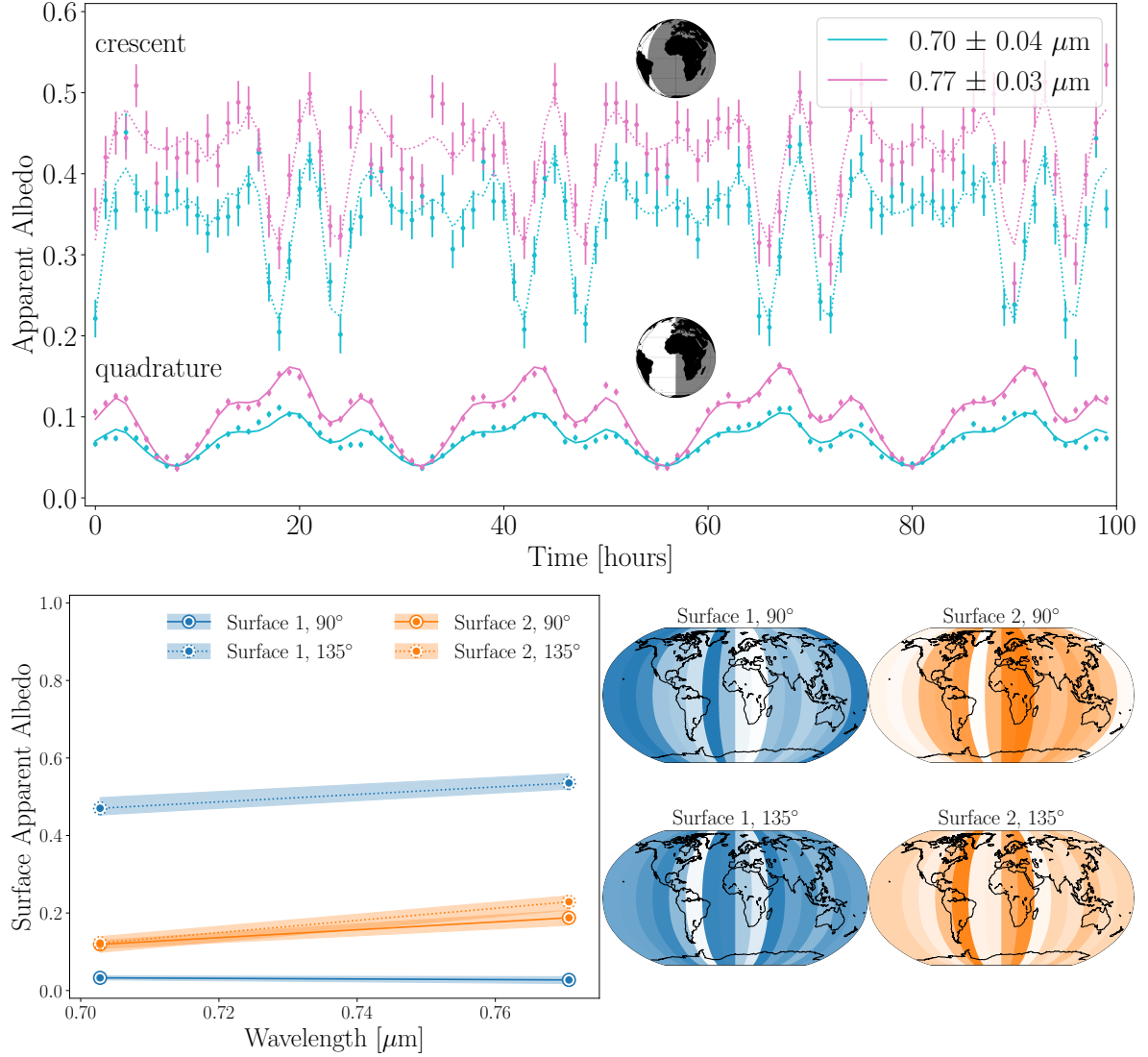


Figure 4.1 Multi-phase inverse problem example for the cloudless Earth. *Top:* Synthetic lightcurves for two wavelength intervals with error bars for the cloudless Earth, and retrieved fits to the observations (lines) at quadrature (solid lines) and crescent (dotted lines) phase. Data were simulated for the cloudless Earth at 5 pc observed with a 15 m telescope with 50% throughput, which yielded a 1-hour photometric precision of 4.5% at quadrature and 6.5% at crescent. *Bottom Left:* Retrieved surface albedos from a two surface fit to rotational lightcurves of the cloudless Earth at 90° phase and 135° phase. The color of each line identifies the surface and the line style denotes the phase. The envelope about each line shows the 1 σ interval about the median posterior PDF derived from MCMC. From quadrature to crescent phase, the average albedo of Surface 1 increases by a factor of ~ 16 with 15 σ confidence, while the average albedo of Surface 2 increase by $\sim 15\%$ at $<1\sigma$. *Bottom Right:* Median retrieved longitudinal slice maps. Each longitudinal slice is colored proportional to the retrieved area covering fraction of the given surface, with white meaning that none of the given surface is found within the slice. Maps are shown in the Robinson projection.

were binned into two wavelength intervals to allow for slopes in the albedo of each surface. This yielded lightcurves at $0.70 \pm 0.04 \mu\text{m}$ and $0.77 \pm 0.03 \mu\text{m}$, with an average 1-hour photometric uncertainty, respectively, of 5.0% (S/N=20.1) and 4.0% (S/N=25.4) at quadrature, and 6.8% (S/N=14.8) and 6.1% (S/N=16.5) at crescent. The decrease in photometric precision seen for the crescent lightcurves in Fig. 4.1 is due to the decrease in the illuminated portion of the planet's disk, for observations using the same telescope.

The lower panels of Fig. 4.1 show the retrieved apparent albedo and longitudinal slice maps of each surface, for a two surface, 16 slice fit to the synthetic multi-wavelength lightcurves shown in the upper panel. The choice of 16 slices was selected so that the angular slice width (22.5°) divides the change in phase (45°) by an integer, thereby allowing the quadrature and crescent phase maps to be shifted and aligned. The slice maps are displayed on a Robinson map projection, and use color intensity to indicate the median retrieved area covering fraction in each slice. The retrieved surface albedos are shown with a 1σ uncertainty envelope about the median posterior PDFs derived from MCMC. The apparent albedo of Surface 1 (blue) exhibited a strong increase from quadrature to crescent phase, becoming $\sim 16\times$ brighter, while the albedo of Surface 2 (orange) increased by only $\sim 15\%$, which was within the derived uncertainties.

Figure 4.2 explores the statistical significance of the increase in apparent albedo of Surface 1 and Surface 2 between the retrievals at quadrature and crescent phase. Histograms showing the change in apparent albedo of each surface were produced by taking 5000 random samples from the wavelength-averaged posterior PDFs at each phase and then subtracting the quadrature phase result from the crescent phase result. Surface 1's apparent albedo increased by 0.47 ± 0.02 , while Surface 2's apparent albedo remained constant with phase (0.02 ± 0.02).

Figure 4.3 examines the difference in the retrieved longitudinal maps with phase. A qualitative comparison shows that the maps at quadrature and crescent phase in Fig. 4.1 are very similar, but a direct comparison of the maps and their uncertainties at each phase reveals both regions of agreement and regions of disagreement. The two longitudinal maps in Fig. 4.3 show the median (left) and standard deviation (right) of the difference between the maps at each phase, in each slice. This is simply the difference between the two Surface 1 slice maps in Fig. 4.1, with uncertainties

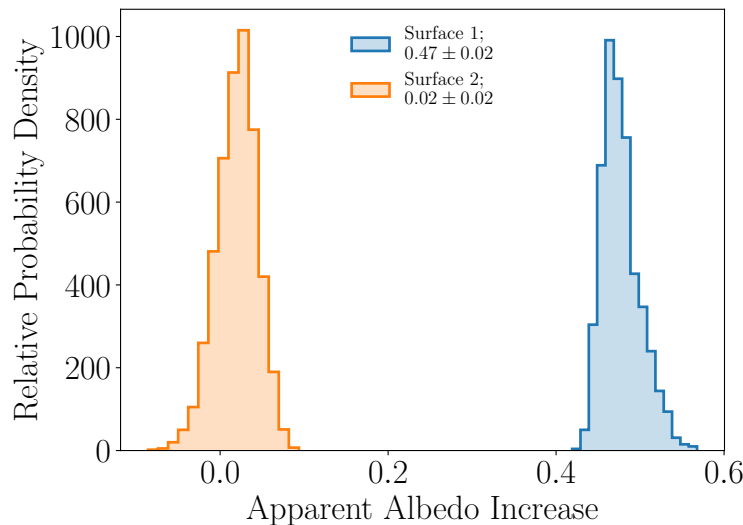


Figure 4.2 Histograms showing the increase in albedo for Surface 1 (blue) and Surface 2 (orange) between quadrature and crescent phase observations of the cloudless Earth. Surface 1 increased in apparent albedo by 0.47 ± 0.02 from quadrature to crescent phase, indicating that it has a highly non-Lambertian scattering phase function. Surface 2 has a negligible albedo increase, indicating that it is consistent with Lambertian scattering.

propagated by randomly sampling the MCMC posterior distributions. Figure 4.3 is shown with respect to Surface 1, although for a two surface longitudinal model Surface 2 is simply one minus the map of Surface 1 (see condition 4.4b). The median difference between the maps is less than 0.5 for all slices. The histogram in the bottom panel of Fig. 4.3 shows the slice-averaged (or total) fractional area increase with the change in phase. The crescent phase results include slightly more contribution from Surface 1 (0.10 ± 0.06) than the quadrature results. However, the longitudinal maps agree best for slices containing majority open ocean, with the marginally significant change in the maps with phase primarily driven by disagreements in the land-bearing slices. This suggests that even despite the ideal system geometry, telescope noise properties, and lack of clouds, the inferred maps have intrinsic differences.

Simulations performed at gibbous phase (45°) show similar results to those at quadrature phase, and are therefore not shown. At gibbous phase the recovered apparent surface albedos are consistent with the recovered albedos at quadrature, with one very dark surface and another more

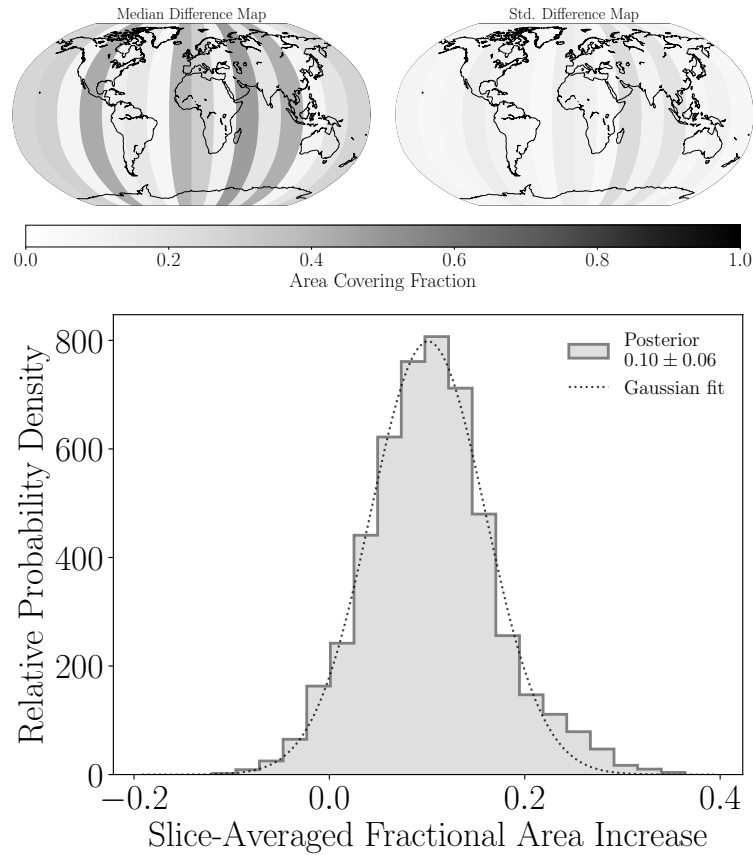


Figure 4.3 *Top*: Longitudinal slice maps showing the median (left) and standard deviation (right) of the change in area covering fraction between quadrature and crescent phase observations of the cloudless Earth. *Bottom*: Longitudinal slice-averaged fractional area change of Surface 1 from quadrature to crescent phase. The maps agree best for slices containing majority open ocean, with the marginally statistically significant change in the maps with phase primarily driven by disagreements in land-bearing slices.

reflective surface that increases in albedo from about 0.1 to 0.2 over the wavelength range of interest. However, the retrieved maps at gibbous are less precise than both the quadrature and crescent phase maps despite having a higher 1-hour photometric signal-to-noise. This difference is due to the larger phase which reflects more total light to the observer, but that light is integrated over a larger portion of the planetary disk which lowers the observer’s sensitivity to each individual longitude, thereby lowering the resolution of the recovered maps (Fujii et al., 2017).

We performed the same experiment on simulations of Earth with real time-varying cloud coverage. Figure 4.4 shows results from multiple surface mapping retrievals performed for Earth at 90° (quadrature; solid lines) and 135° (crescent; dotted lines), where each row of plots corresponds to data generated at a different signal-to-noise ratio. The upper panel of Fig. 4.4 shows longitudinal mapping results from lightcurves simulated for the optimistic case of Earth at 5 pc, observed for 100 hours at each phase, at a 1 hour cadence, with a 15 m, 50% throughput space telescope (same setup at Fig. 4.1). The middle and lower panels of Fig. 4.4 show similar results to the upper panel, except they correspond to data simulated for a 15 m with 10% throughput telescope, and a 6 m with 10% throughput telescope, respectively. These three datasets have 1-hour photometric signal-to-noise at quadrature phase of 23 (4.3%), 10 (10%), and 2.4 (42%), from the top to bottom panels. As in the cloudless case, the spectrophotometry were binned into two wavelength intervals, and the maps were fit with a two-surface, 16 longitudinal slice model. The envelope about each line in Fig. 4.4 shows the 1σ uncertainty about the median retrieved albedo spectrum parameters derived from MCMC.

Comparing results at each phase for the highest S/N data we find that the albedo of Surface 1 exhibits the same large increase at crescent as seen for the cloudless case, even in the presence of realistic clouds, and the longitudinal maps have reasonably good agreement, particularly at longitudes between -180° and 0° . By the same test shown in Fig. 4.2, Surface 1’s apparent albedo increased by 0.57 ± 0.02 with the change in phase, while Surface 2’s apparent albedo increased by a slight margin of 0.08 ± 0.02 . Similarly, the crescent phase results include more contribution from Surface 1 (0.14 ± 0.06) than the quadrature results, and again, the longitudinal maps agree best for slices containing majority open ocean, with the change in the maps with phase driven primarily by

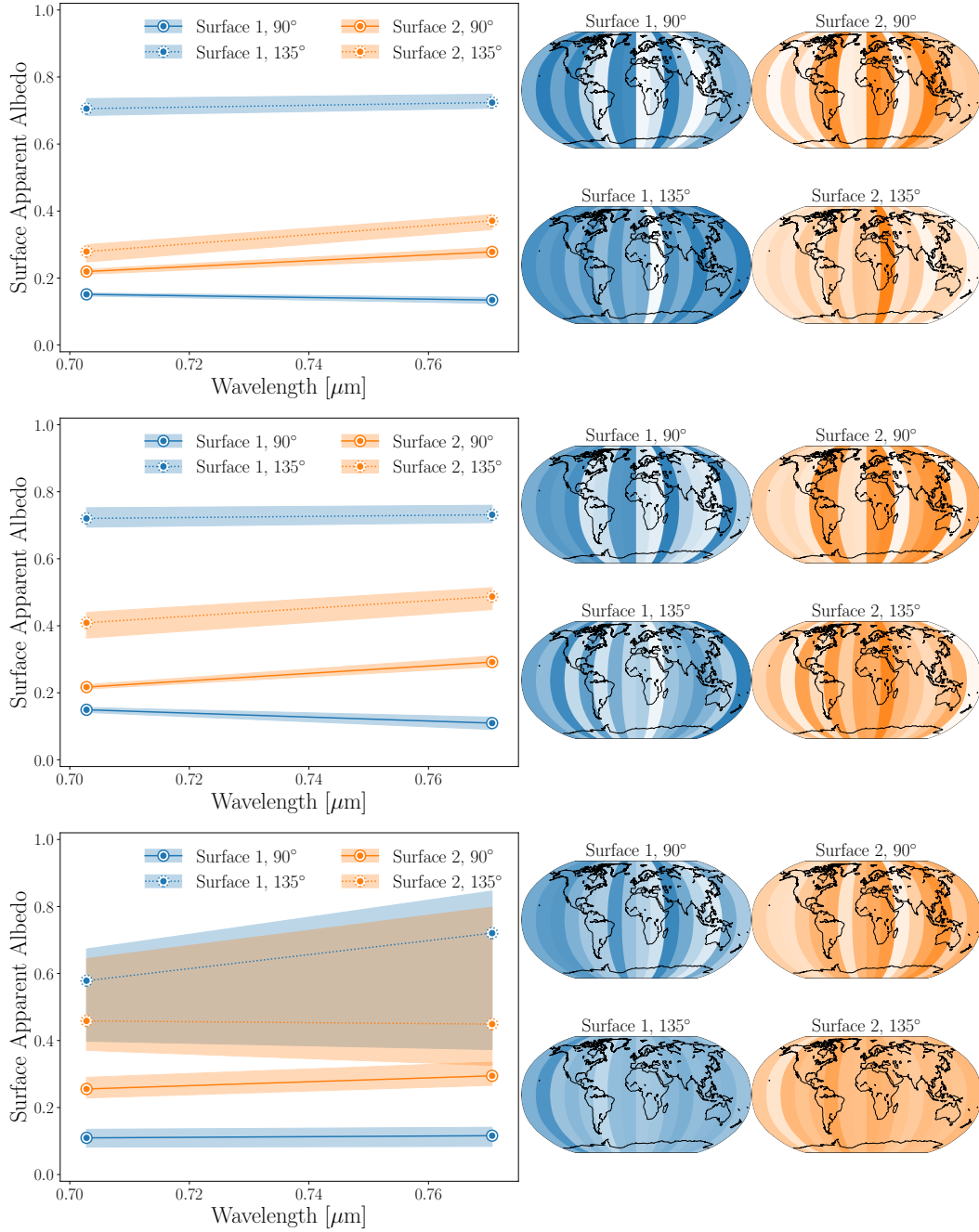


Figure 4.4 Similar to the lower panels of Fig. 4.1, but now for Earth with time-varying clouds and at three different levels of photometric precision. From top to bottom, the data have a 1-hour S/N (photometric precision) of 23 (4.3%), 10 (10%), and 2.4 (42%), at quadrature, which correspond to observations of Earth at 5 pc with a 15 m and 50% throughput telescope, a 15 m and 10% throughput telescope, and a 6 m with 10% throughput telescope, respectively.

disagreements in the land-bearing slices.

The longitudinal mapping results from the lower S/N datasets in Fig. 4.4 (lower panels) demonstrate the photometric limitations of multi-phase mapping. The middle panel with S/N=10 is quite similar to the S/N=23 case, but with slightly larger uncertainties on the crescent phase surface albedos and longitudinal maps. The lower panel with S/N=2.4 no longer captures the variability at crescent phase: the albedo and longitudinal map of both surfaces are unconstrained and significantly overlap. The convergence of the two surfaces implies that a single surface model (mean observed albedo and no map) is warranted at crescent phase. At quadrature phase, the S/N=2.4 result still shows success in extracting both surface's albedos and mapping them to the ocean and land. In the following subsection we further explore the limitations and feasibility of multi-phase longitudinal mapping, particularly with respect to difficult crescent phase observations.

4.3.2 *Observational Feasibility*

We now assess the feasibility of using this phase-dependent technique to detect oceans on extra-solar planets, considering angular separation, photometric sensitivity, and target yields. To obtain direct imaging observations of an Earth-analog exoplanet at a crescent phase of 135° the planet must fall outside the coronagraph inner working angle (IWA). Additionally, the crescent phase lightcurves must be observed with sufficient photometric precision for the relatively short exposure times needed for mapping to be able to retrieve multiple surfaces. Below, we describe these observing considerations, and estimate the number of nearby planetary systems where this technique could be applied to search for oceans with upcoming telescopes.

Inner Working Angle Constraints

The IWA is the minimum angular separation, θ_{IWA} , for which a coronagraph can occult a star without blocking the light of an exoplanet in orbit about it. Since $\theta_{\text{IWA}} \propto \lambda/D$, it is convenient to define

$$\theta_{\text{IWA}} = N\lambda/D \quad (4.12)$$

where N is a factor set by the specific coronagraph design. Smaller values of N enable planets at smaller angular separations from their host star to be imaged. In the limit that the distance to the observed system, d , is much farther away than the projected planet-star separation, r , then

$$\theta_{\text{IWA}} \approx r/d. \quad (4.13)$$

We can solve for the coronagraph specific IWA parameter N in terms of key telescope and system parameters,

$$N \approx 4.8 \left(\frac{r}{1\text{AU}} \right) \left(\frac{D}{10\text{m}} \right) \left(\frac{10\text{pc}}{d} \right) \left(\frac{1\mu\text{m}}{\lambda} \right). \quad (4.14)$$

We initially consider observations of habitable Earth-like exoplanets orbiting Sun-like stars for which $r \approx 1$ AU, at wavelengths near $\lambda \approx 0.8 \mu\text{m}$, but using a telescope with a currently unspecified mirror diameter, and for systems that have yet to be discovered.

To better understand the coronagraphic requirements needed for ocean detection, we explored contours of constant N while varying the telescope diameter and system distance. Figure 4.5 shows coronagraph IWA requirements to directly image an Earth-analog exoplanet at quadrature phase (top panel) and at crescent phase (bottom panel). At quadrature phase the projected planet-star separation is the semi-major axis, $a = 1$ AU, but at crescent phase with $\alpha = 135^\circ$, the projected separation for an edge-on inclination system is $a \sin(\alpha) \approx 0.707$ AU. This tighter planet-star separation imposes a stricter constraint on the coronagraph IWA that ultimately reduces the distance out to which exoplanet systems may be observed at 135° phase relative to quadrature phase for any given coronagraph design. For instance, a 10 meter telescope with $\theta_{\text{IWA}} = 3\lambda/D$ can image Earth analog exoplanets out to ~ 20 pc at quadrature, but only ~ 14 pc at 135° degrees.

Photometric Constraints

We also considered the feasibility of detecting large changes in albedo of a particular surface with phase as a function of observational precision, and the effects of telescope design and target distance on scientific yields. We repeated the same multi-phase longitudinal mapping experiment

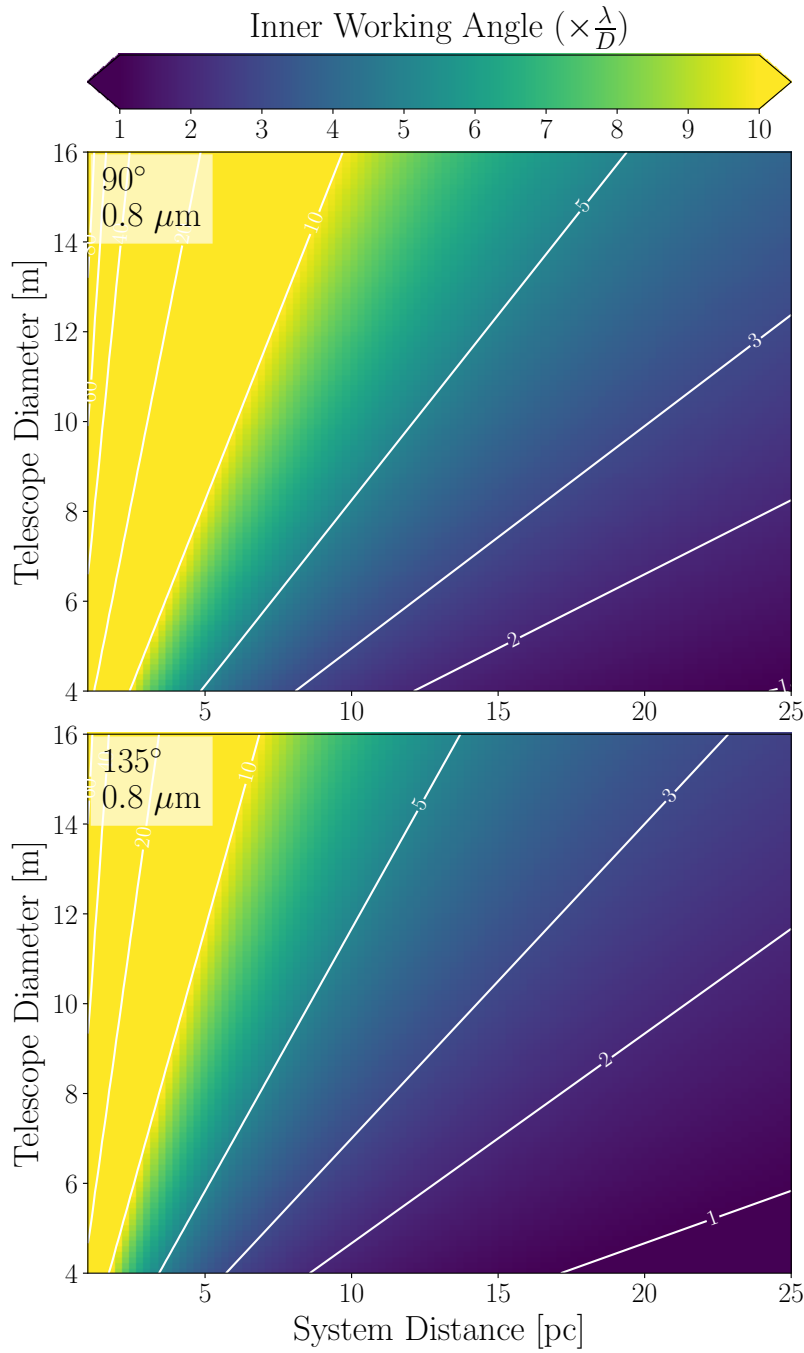


Figure 4.5 Coronagraph inner working angle (IWA) requirements for observing Earth-analog exoplanets. The top panel corresponds to a planet at quadrature phase ($\alpha = 90^\circ$) and the bottom panel corresponds to a planet at crescent phase ($\alpha = 135^\circ$). The IWA is given as the factor which multiplies λ/D , at a fixed wavelength $\lambda = 0.8 \mu\text{m}$, as a function of telescope diameter D and the distance to the system d .

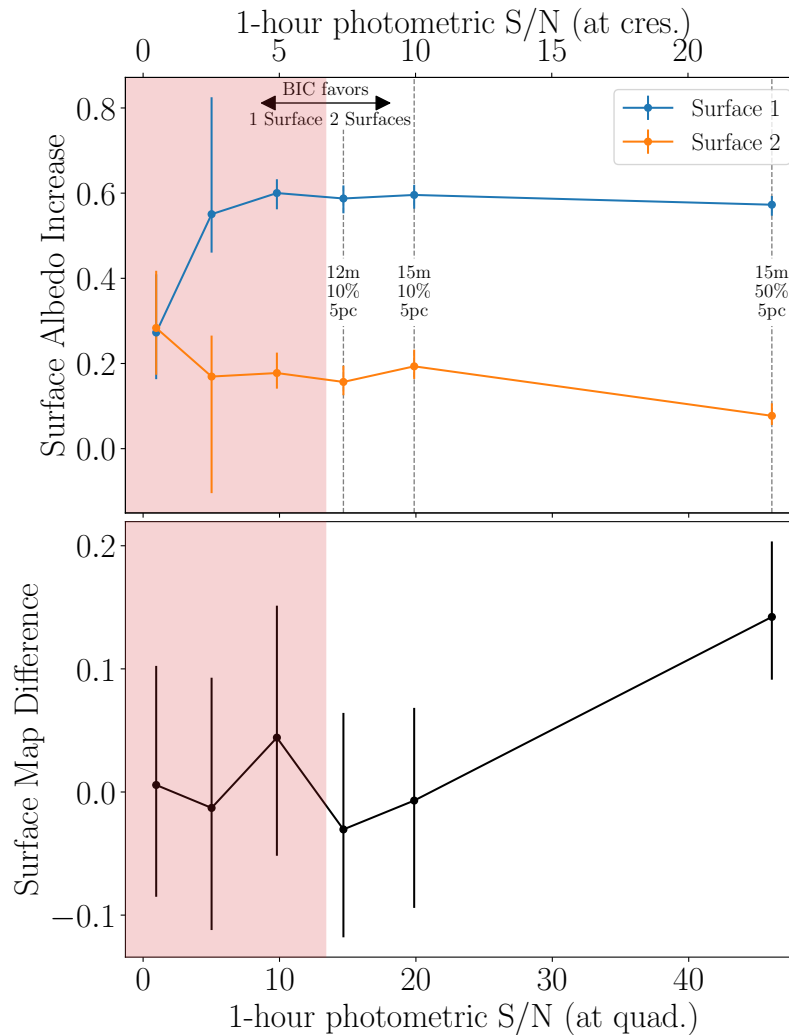


Figure 4.6 Change in surface albedo (top) and change in surface longitudinal area covering fractions (bottom) as a function of the signal-to-noise of the observations. The lower x-axis corresponds to the 1-hour photometric signal-to-noise for the Earth at quadrature phase, while the upper x-axis corresponds to the 1-hour signal-to-noise at crescent phase for observations of the same exoplanet, at the same distance, using the same model telescope configuration, with the same integration time. The upper panel shows the phase-dependent albedo increase of Surface 1 (blue) and Surface 2 (orange) with $1\text{-}\sigma$ error bars derived by randomly sampling the MCMC posterior distributions. The lower panel shows a similar result but for the change in longitudinal area covering fraction. The red shaded region indicates the crescent phase photometric precision where a one surface fit (homogeneous map) is favored by the Bayesian Information Criterion (BIC) over a two surface fit.

as in §4.3.1 for the true cloudy Earth observed for 100 hours with a 1-hour cadence, but now for varying levels of photometric precision. However, instead of fixing the signal-to-noise on each lightcurve, we fixed the underlying telescope architectural parameters used to simulate the lightcurves to enforce self-consistency between observations at quadrature and crescent phase. We then determine for what data quality, or telescope design, statistically significant increases in Surface 1’s albedo may be retrieved concurrent with statistically similar longitudinal maps.

Figure 4.6 shows both changes in retrieved apparent albedo of the two surfaces and changes in retrieved longitudinal maps as a function of data quality. The lower x-axis corresponds to the 1-hour photometric signal-to-noise for the Earth at quadrature phase, while the upper x-axis corresponds to the 1-hour signal-to-noise at crescent phase for observations of the same exoplanet, at the same distance, with the same model telescope configuration, and identical integration time. Thus, fixing all parameters except phase, crescent phase observations are approximately half the signal-to-noise of quadrature phase observations. The upper panel of Fig. 4.6 shows the phase-dependent albedo increase of Surface 1 (blue) and Surface 2 (orange) with $1-\sigma$ error bars derived by randomly sampling the MCMC posterior distributions. This visualization presents posterior PDFs, analogous to those shown in Fig. 4.2, as points with errors along an x-axis of increasing data quality. The lower panel shows a similarly derived result, but for changes in longitudinal area covering fraction, which is analogous to results from Fig. 4.3, but now as a function of S/N.

Figure 4.6 indicates that the phase-dependent apparent albedo change of Surface 1 can be robustly detected and discriminated from the lack of albedo change in Surface 2 across a wide range in photometric precision, although recognizing the same map at multiple phases may prove difficult. For a quadrature (crescent) $S/N > 10$ (5), the increase in apparent albedo with phase for Surface 1 is consistent with 0.6, with uncertainties smaller than 0.04. Over the same range, the albedo of Surface 2 slightly increases. The decrease in albedo change of Surface 2 with photometric precision suggests that greater precision aides in the recognition of Lambertian scattering surfaces. However, the opposite behavior is found for the change in maps with phase. For low S/N on the observed lightcurves the inferred longitudinal maps are consistent with no change—the maps are the same within the propagated observational uncertainties. As photometric S/N increases

the maps at quadrature and crescent phase appear more unique from one another, with an average covering fraction change of 0.14 ± 0.06 for the highest S/N case. Similar to our findings in Fig. 4.3, this scaling behaviour suggests that the inferred maps are actually distinct despite having the same continental distribution and sub-observer latitude. We propose that this discrepancy is due to the effect of a multi-component scattering phase function on the inferred longitudinal maps, which we discuss further in §4.4.1.

The asymptotic behavior of Surface 1’s apparent albedo increase for $S/N > 10$ suggests a natural precision lower limit below which the difference in apparent albedo with phase will be lost in the noise. To further investigate this limit and determine the photometric limits of mapping—particularly at crescent phase—we find the minimum number of surfaces that are justified in fits to different S/N data. We use the Bayesian Information Criterion (BIC; Schwarz et al., 1978) to penalize overly complicated model fits. We take the “optimal” number of surfaces to be the model with the minimum BIC. We define a convention that the optimal number of surfaces determined by the BIC is the number of *detected* surfaces, where a one surface fit is a homogeneous map and a two surface fit is the bare minimum to perform any meaningful mapping. This is similar to the use of principal component analysis to quantify detected surfaces (Cowan et al., 2009; Fujii et al., 2017), except we find it to be significantly more robust against observational uncertainties.

We used the BIC to find the number of detected surfaces across a finely sampled grid in photometric S/N at both quadrature and crescent phase. For each S/N, we calculate the BIC for a one surface model by finding the best-fit from 100 randomly initialized optimizations using the L-BFGS-B algorithm (Zhu et al., 1997). We then increment the number of surfaces and repeat the procedure until we discover the minimum BIC. Both panels of Figure 4.6 are shaded red to denote the crescent phase photometric precision for which only one surface is detected. The boundary occurs at a photometric precision of $\sim 15\%$ ($S/N=6.7$) for observations at crescent phase, which we take to be the minimum precision capable of performing phase-dependent mapping on a partially cloudy Earth-like exoplanet. We also find that the boundary between one and two surfaces requires a slightly improved photometric precision of $\sim 10\%$ ($S/N=10$) for observations at quadrature phase. This indicates that the increased color contrast between the two surfaces at crescent

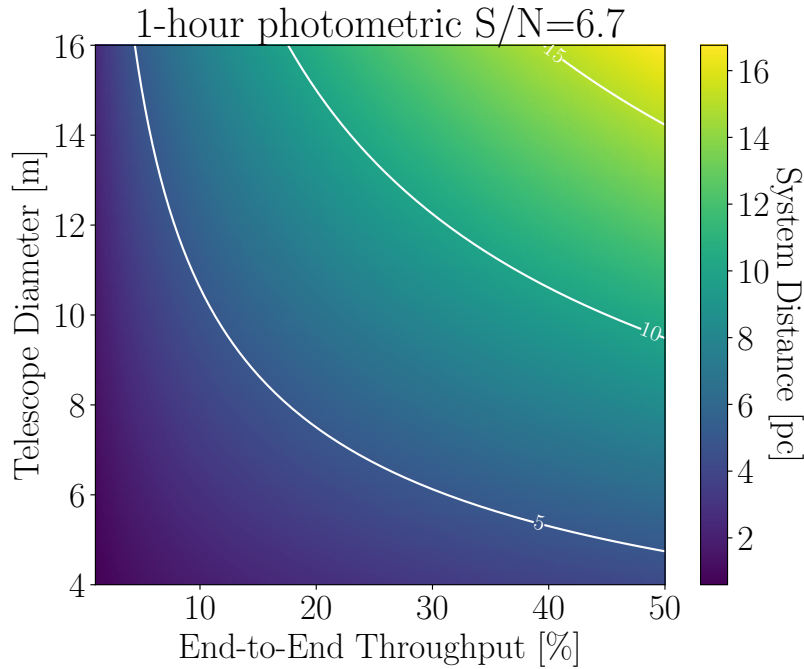


Figure 4.7 Distance (color contours) out to which an Earth-like exoplanet could have crescent phase ($\alpha = 135^\circ$) observations fit with a two surface model, as a function of telescope diameter and throughput. A 1-hour photometric precision of $\sim 15\%$ ($S/N=6.7$) is determined by the BIC to be the threshold between detecting one surface (homogeneous planet) and two surfaces (heterogeneous planet) at crescent phase, considering our observational assumptions, particularly the limited bandwidth. Minimally detectable two surface maps at quadrature phase can be performed out to approximately twice the distance shown here for crescent phase.

aids in multi-surface mapping, although higher precision measurements are easier to acquire at quadrature phase. We also note that three surfaces are detected at photometric precisions $> 2.8\%$ ($S/N=35$) at quadrature.

Adopting the two-surface threshold at crescent phase as the lower limit to the photometric precision, we determine the ability for different possible future telescopes to meet this requirement with respect to scientific yields. Figure 4.7 shows the distance out to which an Earth-like exoplanet could have crescent phase ($\alpha = 135^\circ$) observations meet the S/N threshold for two-surface longitudinal mapping as a function of telescope diameter and end-to-end telescope throughput. For example, a 15 m with 20% throughput could perform crescent phase mapping out to 10 pc, while

an 8 m with 20% throughput could only go out to 5.4 pc, and a 4 m out to only 2.7 pc.

Extending to Low Mass Stars

The forward models used here were developed for Earth-Sun analog systems. However, we can broaden our scope to consider observations of planets with surface oceans orbiting later type stars, using the Earth as a proxy for these planets. Cooler stars have compact habitable zones and therefore smaller angular separations between planet and star compared to G dwarfs. This places a stricter constraint on the IWA to directly image exoplanets in such orbits. Using the [Kopparapu et al. \(2013\)](#) moist/max greenhouse HZ with the 5 Gyr [Baraffe et al. \(1998\)](#) stellar models, we calculate the semi-major axis of the inner and outer edge of the HZ and relate this to the IWA.

Figure 4.8 shows the distance out to which a planet at the inner edge of the HZ (and therefore the entire HZ) may be directly imaged as a function of telescope diameter and host star mass, for 90° phase, $0.8 \mu\text{m}$, and IWA of $3\lambda/D$. A 15 m telescope can directly image planets at the inner edge of the HZ of a $0.6 M_\odot$ K dwarf at 90° and 135° out to 8.3 pc and 5.8 pc, respectively. A 9-meter coronagraph observing the same system would only be able to go out to 5.0 pc and 3.5 pc for phases of 90° and 135° , respectively; a 6-meter out to 3.3 pc and 2.2 pc, respectively; and a 4-meter out to 2.2 pc and 1.6 pc, respectively. Note that the stellar distances for accessing the inner edge of the HZ given in Fig. 4.8 are conservative compared with outer edge calculations, which we also account for next.

Yield Estimates

We now provide initial estimates for how many planets might be accessible to ocean detection, as a function of telescope diameter. We use the Nearby Stellar Systems Catalog for Exoplanet Imaging Missions (ExoCat; [Turnbull, 2015](#)) to estimate target yields for G, K, and M dwarfs in the stellar neighborhood. For each star in ExoCat we use the stellar effective temperature and luminosity to calculate the moist greenhouse inner edge of the HZ and maximum greenhouse outer edge using relations from [Kopparapu et al. \(2013\)](#). We then relate the projected separation of hypothetical

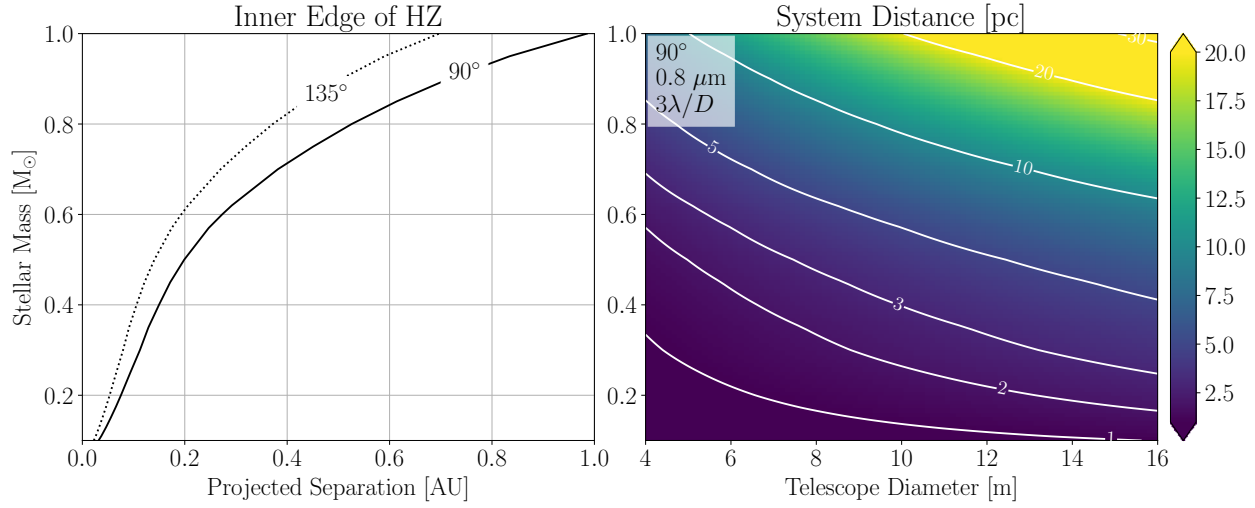


Figure 4.8 *Left*: Projected separation between star and planet at the inner edge of the habitable zone (moist greenhouse limit) for low mass stars. Quadrature and crescent (135°) phase projected separations are shown as solid and dotted lines. *Right*: Maximum distance (color contours) out to which a planet at the inner edge of the HZ can be directly imaged as a function of stellar mass and telescope diameter. Phase, wavelength, and IWA are fixed at 90° , $0.8 \mu\text{m}$, and $3\lambda/D$. Distances can be readily scaled to the apparent phase $\alpha = 135^{\circ}$ by multiplying by $\sin(\alpha)$.

exoplanets at the inner and outer edges of the HZ to the IWA of telescopes of different diameters, assuming $\theta_{IWA} = 3\lambda/D$, to determine if a planet in such an orbit is observable. Additionally, to meet the photometric requirements for mapping, we also check to see if an Earth-size and albedo planet at the inner edge of the HZ exceeds our 1-hour photometric S/N threshold of 6.7, assuming the same telescope and instrumental specifications provided above and a 20% throughput. If a given stellar system satisfies both the IWA and the photometric conditions then we count that system as observable.

However, not all stars will have a HZ planet, and not all HZ planets will be optimally inclined so that they pass through a crescent phase of 135° . These are questions on the occurrence rate of HZ exoplanets and the random distribution of system inclinations, both of which are fundamental to exoplanetary target yield estimates. Numerous studies have attempted to infer the occurrence rate of Earth-like planets in the HZ of different stellar types (e.g. [Dressing and Charbonneau, 2013](#); [Kopparapu, 2013](#); [Foreman-Mackey et al., 2014](#); [Burke et al., 2015](#)), but large uncertainties persist

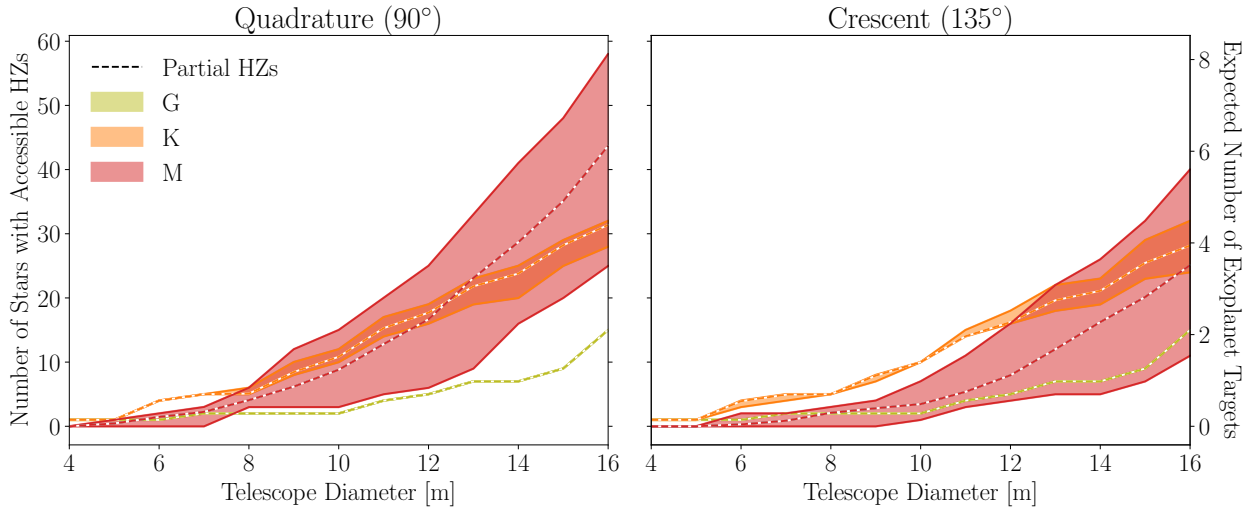


Figure 4.9 Yield estimates for observations of HZ exoplanets at quadrature (left panel) and crescent phase (right panel) as a function of telescope diameter for nearby G, K, and M dwarfs. Systems are counted if a hypothetical planet in the HZ falls outside of the IWA (and is therefore observable), assuming $\theta_{IWA} = 3\lambda/D$, and if the 1-hour S/N at crescent phase exceeds our photometric threshold of 6.7 (see Fig. 4.6), assuming a 20% throughput. Both panels share the same y-axes, where the left y-axis gives the number of stars that would have observable HZ planets. The right y-axis gives the expected number of exoplanets after considering the geometric probability of passing through 135° phase (70%) and assuming a 20% occurrence rate for planets in the HZ. The shaded regions are bounded below and above by the inner and outer edges of the HZ, respectively, and demonstrate a sensitivity in yield depending on whether the observed planet is at the inner or outer edge of the HZ. The dashed lines correspond to the sum of all partially observable HZs (e.g. half a count for a star with half of the HZ observable outside the IWA). Observations of HZ exoplanets are limited by photometric precision for G dwarfs while M dwarfs are significantly more limited by the IWA.

for temperate rocky worlds. We assume an occurrence rate of 20% for G, K, and M dwarfs, which is consistent with published values. Although trends in occurrence rate with stellar type may exist (Mulders et al., 2015), we elect to keep our calculations simple so that they may be easily scaled as occurrence rates are refined. The prior probability of randomly oriented system inclinations is $\sin(i)di$ (Luger et al., 2017b). This means that 70% of randomly oriented systems will have planets tracing orbital paths that pass through a phase of 135° ($i > 45^\circ$). We use the 20% occurrence rate and the 70% random inclination prior to scale our potential stellar targets into an expected number of exoplanets for which oceans may be searched for.

The left y-axis of Fig. 4.9 shows the number of stars for which an assumed optimally-aligned planet ($i = 45^\circ - 90^\circ$) in the HZ of G (yellow), K (orange), and M (red) dwarfs could be observed at phase angles of 90° (left panel) and 135° (right panel) as a function of telescope diameter. The right y-axis of Fig. 4.9 shows the expected number of HZ exoplanets that could potentially be searched of oceans using our crescent phase glint mapping technique. These observations are assumed to be at $0.8 \mu\text{m}$, using a coronagraph with an IWA of $3\lambda/D$. The shaded regions in Fig. 4.9 are bounded below by the inner edge of the HZ and above by the outer edge. The dashed lines correspond to the sum of all partially observable HZs, where a partial HZ is calculated as the radial fraction of the HZ that is observable beyond the IWA. Thus the dashed lines fall between the corresponding shaded region because 100% of the HZ is observable if the inner edge falls outside the IWA, while 0% of the HZ is observable if the outer edge is equal to the IWA.

Figure 4.9 suggests that K and M dwarfs will make up the bulk of the stellar hosts for surface ocean assessments. Low mass K and particularly M dwarfs make up the bulk of the nearby stellar population. However, these stars have compact HZs that are difficult to observe in direct imaging. Additionally, the planet-to-star contrast ratio increases towards lower mass, fainter stars. This decreases the noise due to unocculted stellar light (speckles), particularly for systems with contrast ratios that are lower than the design contrast of the direct imaging mission (10^{-10} assumed here). Fig. 4.9 suggests that K dwarfs strike a balance between these effects for the smaller telescope diameters considered, and may be optimal stellar hosts for direct imaging studies of HZ planets. G dwarfs have appreciably separated HZs, so that if the systems are near enough to achieve a

sufficient photometric precision, then the IWA will not be a restricting factor for crescent phase observations. This is in agreement with our earlier findings. M dwarfs occupy a different domain: their favorable planet-to-star contrast ratios more easily satisfy S/N requirements, so that IWA is instead more limiting on both the stellar distance out to which M dwarfs may have their HZs observed, and crescent/gibbous phase observations of any HZ planet.

Considering the IWA and S/N constraints, we find that 1 to nearly 10 HZ planets could be assessed for signs of time-varying ocean glint signals with telescopes in the 6 to 15 m diameter range. The nearby single stars Tau Ceti, Epsilon Eridani, and Lacaille 8760 (AX Microscopii) are the respective G, K, and M dwarfs that have the largest angular separation to the inner edge of the HZ, making them quite optimistic targets, considering IWA constraints. Telescopes smaller than 12 m may have the potential for crescent phase observations around more K dwarfs than either G or M dwarfs. Telescopes larger than 12 m may have more M dwarfs targets than K dwarf targets. A 15 m with an IWA of $3\lambda/D$ could perform crescent phase $0.8\ \mu\text{m}$ direct imaging in the HZ (including the sum of partially observable HZs) of approximately 9 G, 25 K, and 20 M stars. A 12 m could do the same for 5 G, 16 K, and 8 M stars; a 9 m for 2 G, 8 K, and 3 M stars; a 6 m for 1 G, 4 K, and 0 M stars; and a 4 m for 1 G, 1 K, and 0 M stars. Considering random system geometries and assuming an occurrence rate, as discussed above, we find that a 15 m, 12 m, 9 m, 6 m, and 4 m will have approximately 8, 4, 2, 1, and 0 expected planet(s), respectively, in the HZ amenable to crescent phase mapping investigations.

4.4 Discussion

We investigated the unique time- and phase-dependent observables of Earth oceans in the context of direct habitability assessments for terrestrial exoplanets. In the following subsections we discuss how our findings may be used to increase the robustness of ocean detection on exoplanets (§4.4.1) and assess the potential yield and required observational conditions for ocean searches on extrasolar planets (§4.4.2).

4.4.1 *Increasing the Robustness of Exoplanet Ocean Detection*

We find that the robustness of exo-ocean detection can be increased via solutions to the apparent albedo and longitudinal map of multiple surfaces observed at multiple orbital phase angles. We retrieved the longitudinal map and apparent albedo of two unique surfaces from simulated lightcurves of Earth at quadrature, crescent, and gibbous phase. At quadrature and gibbous phase, one low albedo surface is mapped to longitudinal slices that we know from Earth to be ocean bearing, while the other more reflective surface is mapped to the continents. At crescent phase, the maps of the dominant surfaces appeared quite similar to the quadrature phase solutions, and the albedo of the land-containing slices was nearly the same as at quadrature, but the apparent albedo of the ocean-dominated slices exhibited a large and discernible increase in albedo of ~ 0.5 . This sizable increase in apparent albedo of the ocean-bearing longitudinal slices seen at crescent phase comes from the ocean glint spot. Since our inverse model assumes surfaces are Lambertian scatterers, deviations from Lambertian are encoded in the inferred apparent albedo of the surface, or the albedo that the surface must have (when integrated over the longitudinal slice) to reproduce the observed flux. The fact that the longitudinal maps of each surface appear quite similar at quadrature and crescent phase provides a means to identify that the same surfaces are mapped. This helps to attribute changes in surface reflectivity to specific surface components. The planetary surface that had appeared dark at quadrature, must reflect light more efficiently at crescent phase. Such a defined deviation from Lambertian behavior is strongly suggestive of ocean glint.

Although our cloudless Earth tests demonstrated that glint effects can be extracted from crescent phase longitudinal mapping, our tests with realistic time-varying clouds revealed limitations and caveats for the technique. Clouds made it much more difficult to extract the same map at both quadrature and crescent phases. Crescent phase observations are more sensitive to clouds than more nadir viewing geometries. Clouds therefore inhibit, but do not fully prevent, accurate surface longitudinal mapping at crescent phase. We were still able to detect the presence of large landmasses, like Africa, which effectively interrupted the glint spot.

However, clouds had minimal impact on the surface albedo results with phase. Despite the

added difficulty for accurately detecting the same map, we did confidently detect the increase in albedo of the ocean-bearing slices, along with the lack of albedo increase of the land-bearing slices. Both retrieved surfaces exhibited a slight increase in albedo going from quadrature to crescent phase when clouds were included. Without clouds Surface 2 (land) did not increase in inferred albedo (0.02 ± 0.02) with phase, while with clouds it increased by about 0.1 (0.08 ± 0.02), a 30% increase. Surface 1 (ocean) without clouds increased in albedo by 0.5 (0.47 ± 0.02) with phase, a factor of ~ 20 increase, while with clouds it increased by 0.6 (0.57 ± 0.02), a factor of ~ 5 increase with phase. This added increase of 0.1 in the apparent albedo of both surfaces towards crescent is due to forward scattering by clouds, which we find imparts a significantly smaller phase-dependent effect than ocean glint (c.f. [Robinson et al., 2010](#)), and the effect of glint strongly correlates with only one of the surface components. [Cowan et al. \(2009\)](#) also noted that clouds appeared to increase the albedo of both principle components used in their study, rather than appearing as a unique principle component (or surface). As a result of decomposing the lightcurves into two surfaces, we see a much stronger effect from glint ($\sim 5\times$) than that noted by [Robinson et al. \(2010\)](#) ($\sim 2\times$), which could make this technique a more sensitive tool for ocean detection than attempting to observe phase-dependent brightness in the disk-average.

Longitudinal mapping at crescent phase naturally breaks the latitude-albedo false positive for ocean glint [Cowan et al. \(2012\)](#) by attributing the global albedo increase observed at crescent phases to a specific surface on the planet. High S/N, high cadence, reflected light observations at crescent and quadrature (or gibbous) phase enable the longitudinal map and albedo to be retrieved. The time-variability of a rotating planet's disk-integrated light is used to deconvolve the relative albedo increase seen towards crescent phase into specific surfaces that are responsible for the relative increase. This may increase the confidence that high latitude ice is not responsible for the relative albedo increase. Furthermore, the boosted strength of the glint signal found after the deconvolution ($\sim 5\times$) exceeds the likely increase in disk-integrated albedo that is expected from the latitude-albedo effect for an Earth-like planet with polar ice caps ($\sim 2\times$; [Cowan et al., 2012](#)).

We find that phase-dependent deviations in the longitudinal maps of, our ideal case, the cloudless Earth further implicate non-Lambertian surface scattering behavior. The difference between

the quadrature and crescent phase maps in Fig. 4.1 can be attributed to the footprint on the planet that contributes the most flux at each phase, and therefore drives our inferences. Cowan et al. (2013) refer to this as the kernel of convolution, which has a longitudinal and latitudinal extent determined by the system geometry and phase function of the reflecting surface (Schwartz et al., 2016). At quadrature phase, the planet’s surface is well described by a Lambertian surface with a kernel that extends well above and below the equator (for our case here with a 90° inclination and 0° obliquity). However, at crescent phase the ocean deviates from diffuse Lambertian scattering as light reflected off the ocean is dominated by the glint spot. The effective size of the glint spot is significantly more concentrated in both longitude and latitude compared with the Lambertian kernel of convolution. Consequently, the ocean glint probes a relatively small latitudinal extent. Thus, inferred glint maps are highly sensitive to interruptions (e.g. land, ice, and clouds) along the direct path that the glint spot takes across the surface of the planet. In our test case of Earth with 90° inclination and 0° obliquity, the dominant latitude and path of the glint spot is the equator.

The effect of kernel size helps to explain why our cloud-free longitudinal maps at quadrature phase (Fig. 4.1) show a sensitivity to continents in the mid-latitudes, such as North America and Asia, which does not appear in the crescent phase maps. These continents never intersect the path of the latitudinally compact glint spot, and are therefore not mapped at crescent phase. Thus, the latitudinal heterogeneity of the planet leads the inferred maps to slightly disagree between quadrature and crescent phase. This does not occur strictly due to the system geometry as explored in Schwartz et al. (2016), but instead because the scattering phase function of the planetary surface(s) affects the width of the convolution kernel. Conversely, the differences found between the quadrature and gibbous phase maps are driven more by the illumination geometry, which favors smoother longitudinal maps at gibbous phase because more illuminated longitudes are integrated together, thus decreasing spatial resolution in the maps (Fujii et al., 2017).

For demonstrative simplicity we have considered planets with zero obliquity and in an edge-on inclined orbit. Since the sub-observer latitude does not change with phase, this is the optimal case for detecting the same longitudinal map at multiple phases. However, even for this optimistic case with the same latitude dominating the peak reflected flux, we had difficulty recognizing the same

map at multiple phases due to latitudinal heterogeneity, as discussed above.

This presents the possibility of false negatives that may arise due to the difficulty of inferring the same longitudinal map at multiple phases. If the observed planet has a prominent north-south continental dichotomy and a non-zero obliquity, like Earth, or a non-edge-on inclination, then time-series observations at multiple phases will constrain the two-dimensional map of the planet (e.g. Kawahara and Fujii, 2010, 2011; Fujii and Kawahara, 2012; Berdyugina and Kuhn, 2017) and the obliquity (Schwartz et al., 2016; Farr et al., 2018). For these cases it will be more difficult to distinguish between changes with phase and changes in surface type. As a result, multi-phase longitudinal mapping will perform best for system geometries that are not amenable to two-dimensional spin-orbit mapping. However, exoplanets with non-zero obliquity may still exhibit strong lightcurve variations at crescent phase that are best explained by the combination of a Lambertian and a non-Lambertian component. It may be that congruence of the multi-phase maps is not necessary to place strong constraints on the detection of glint. Indeed, our finding that the glint signal is found in only one inferred surface, and that the glint effect far exceeds forward scattering by clouds, may favor an ocean glint interpretation, albeit with less confidence than if the multi-phase maps matched.

Even the detection of glint is not uniquely indicative of surface liquid water. Surface liquid reservoirs of any composition may cause a specular reflection due to their mirror-like smoothness. This is perhaps best exemplified by the glint detection from Titan’s hydrocarbon lakes (Stephan et al., 2010). Freshly-formed smooth ice may also deviate from Lambertian scattering (Williams and Gaidos, 2008). Furthermore, terrestrial glint has been identified from light scattering off oriented ice crystals in high altitude cirrus clouds (Kolokolova et al., 2010; Marshak et al., 2017). Such signals could be distinguished from surface glint due to their lack of reproducibility on subsequent rotations, or by their anti-correlation with water vapor absorption features in the spectrum that would indicate a cloud-truncated path length. These cases exemplify why additional context is necessary for conclusions to be drawn on the surface liquid composition, such as surface pressure and temperature constraints (Robinson, 2018), or the phase-dependent relative change in the surface albedo as discussed in this work.

Because the VPL Earth Model only accurately reproduces Earth, our study only considered Earth's true geography, rotation rate and atmospheric composition. Future studies should generalize the technique of glint mapping with randomly sampled rotation rates and geographies. Furthermore, we assumed that the rotation rate was precisely known. However, a simple periodogram analysis indicates that our time-series data used for longitudinal mapping with a one-hour cadence and 100-hour duration is more than sufficient to confidently infer the rotation rate of Earth. In general, when the S/N is sufficient for longitudinal mapping to favor two or more surfaces and the observational duration exceeds the rotation rate, then the rotation rate can reliably be retrieved. This agrees with the findings of [Pallé et al. \(2008\)](#) and [Fujii et al. \(2014\)](#). However, the rigorous treatment of directly imaged terrestrial exoplanet rotation rates is beyond the scope of this chapter, and we leave it for follow-up work.

Additionally, closer-in planets are more likely to be in tidally locked rotation states, for which our assumption that the rotational period is much shorter than the orbital period breaks down. Further work should investigate the potential for glint detection in reflected light phase curve measurements of tidally locked exoplanets in the habitable zones of late K and M dwarfs. In these cases the interruption of a glinting ocean by land may manifest as a strong asymmetry in the phase curve between the waxing and waning crescent phases.

Using the apparent albedo assumes that the illumination phase angle and planet radius are known, but in practice, both of these quantities will be difficult to precisely constrain for directly imaged exoplanets. This is due to uncertainties in orbit determinations and the radius-albedo-phase degeneracy ([Nayak et al., 2017](#)). Planet size cannot be measured unless the planet is transiting, and it is highly unlikely that the few planets amenable to glint mapping observations have transiting orbits. Alternatively, constraints on the planet's radius may be derived for non-transiting planets from a combination of photometric measurements and constraints on the planet's albedo, radius, and atmospheric composition determined via spectrum retrieval (e.g. [Nayak et al., 2017](#); [Feng et al., 2018](#)). Multi-epoch observations can be used to derive the planet's orbit, either prior to, or using the multi-epoch observations studied in this chapter. The radius-albedo degeneracy could potentially be broken by relating the inferred surface composition to its laboratory measured re-

flectance spectrum, particularly if phase-dependent glint were to provide a strong indication on the composition of the reflecting surface. Similarly, identifying the slight increase in reflectivity towards crescent for both surfaces could indicate forward scattering by clouds and may be used to rule out a low planet albedo, thereby helping to constrain the planet radius. In the event that the planet’s radius is unconstrained or incorrectly assumed, multi-phase observations may still reveal *relative* increases in the apparent albedo of multiple surfaces with the caveat that the apparent albedo and longitudinal area covering fraction of each surface are not precisely known.

A number of improvements to the fitting procedures used in this work could be made in the future. We fit the light curves at each phase independently, but multiple epochs can instead be fit simultaneously (e.g. [Farr et al., 2018](#)), while allowing for the surface albedo to change with phase. A stronger prior could be enforced that only one of the surfaces is allowed to change albedo as a function of phase. This would turn the physical prior that land does not glint into a model constraint, thereby reducing the number of fitting parameters and helping with lower S/N cases. Additionally, a Gaussian process could be used to enforce smoothness in the longitudinal maps and allow for the characteristic angular length scale to be inferred (e.g. [Farr et al., 2018](#)). This could provide enough flexibility in the mapping procedure to more accurately assess the congruence of maps at multiple phases.

Our work has extended rotational mapping to multiple observed planetary phases to combine specific observables of exoplanetary habitability. Previous studies have demonstrated the utility of multi-wavelength lightcurve inversions to map directly imaged exoplanets in one-dimension ([Cowan et al., 2009](#); [Fujii et al., 2010, 2011](#); [Cowan and Strait, 2013](#)), and two-dimensions ([Kawahara and Fujii, 2010](#); [Fujii and Kawahara, 2012](#); [Berdyugina and Kuhn, 2017](#); [Farr et al., 2018](#)). While previous two-dimensional exoplanet mapping concepts have leveraged the planet’s obliquity and/or inclination to scan latitudes using observations at multiple orbital epochs ([Schwartz et al., 2016](#)), we have instead considered phase-dependent effects for observations at multiple epochs for cases where negligible obliquity and a nearly edge-on planet orbit would prevent two-dimensional mapping. Our crescent phase simulations of Earth demonstrate a clear increase in reflectivity relative to more illuminated phases due to specular reflection off the oceans, as seen in previous studies

(Williams and Gaidos, 2008; Robinson et al., 2010, 2014a). The variability of Earth increases towards crescent phase as the illuminated crescent may be more completely dominated by a single end-member surface spectrum (Oakley and Cash, 2009; Robinson et al., 2010). We demonstrated for the first time that crescent phase variability can be leveraged to map ocean glint.

4.4.2 *Observational Feasibility and Yield Estimates*

The observations discussed above will require a telescope that can achieve simultaneous multi-wavelength, time-series, direct imaging observations of an Earth-like exoplanet at crescent phase. This will require a next-generation, high contrast, space-based direct imaging mission equipped with a coronagraph or a starshade. For glint-mapping observations to be feasible, light from the planet must fall outside the coronagraph or starshade inner working angle and at sufficient signal-to-noise to justify spectral unmixing inverse methods, without overfitting.

We find that the IWA constraints that must be overcome to directly image an Earth-analog at a crescent phase of 135° are less stringent than the constraints on the photometric precision. A 9 m, 12 m, and 15 m space-based coronagraph with an $IWA < 3\lambda/D$ would be capable of observing an Earth-Sun analog at a 135° phase angle out to ~ 13 , 17, and 21 pc, respectively. However, a 9 m, 12 m, and 15 m telescope with a 20% end-to-end throughput can only reach the requisite $\sim 15\%$ 1-hour photometric precision on multi-wavelength lightcurves of Earth at 135° phase out to ~ 6 , 8, and 10 pc, respectively.

We estimate that there will be potentially habitable exoplanets orbiting nearby stars that will be amenable to glint searches using a next-generation, space-based telescope. The IWA, which is less problematic than photometric precision for G dwarfs, becomes a more serious concern for later type stars where the HZ is more compact. On the other hand, the larger population of nearby K and M dwarfs compared to G dwarfs, and the more favorable planet-star contrast ratios for these later type stars both help to counteract the IWA concern. This allows for the possibility of quadrature and even crescent phase observations of G, K and M dwarf HZ planets using coronagraph-equipped telescopes with diameters in the 6-15 m range. Telescopes with diameters approximately ≥ 8 m may have more K and M dwarf targets than G dwarf, depending on how exoplanet occurrence

rates scale with stellar type.

Since performing longitudinal mapping at multiple phases requires relatively high S/N observations, then larger mirror size, and/or improvements in optical throughput and coronagraph spectral bandwidth could increase the potential number of oceans detected. Our derived 1-hour S/N requirement assumed 100 hours of consecutive observations at each phase, however, lower S/N with more exposures or higher S/N with fewer exposures could still achieve qualitatively similar results. This could potentially open the door to glint mapping observations for smaller telescopes, such as a 4 m, but would require more than 100 hours of observation at each phase, and can likely only search for oceans on planets orbiting in the HZ of Alpha Centauri A and/or B, due to their proximity.

Although our study specifically modeled coronagraph-equipped telescopes, our results are also applicable to external occultors, such as a starshade. The main difference between coronagraph modeling and starshade modeling for the purposes of this study are the IWA, throughput, and spectral bandwidth. Although starshades can attain λ/D IWA, they are currently being considered for smaller telescopes (≤ 4 m; [Seager et al., 2015](#); [Spergel et al., 2015](#); [Mennesson et al., 2016](#)) than those using only an internal coronagraph. Therefore, a 4m-class starshade-enabled telescope with an IWA of λ/D would be comparable to an 8m with internal coronagraph and an IWA of $2\lambda/D$. For equivalent IWAs the higher yield of possible ocean observations will go to larger aperture and/or enhanced throughput, and external occultors more readily achieve higher throughput than internal occultors due to fewer optical elements ([Mennesson et al., 2016](#)). Additionally, starshade technology permits larger spectral bandpasses to be simultaneously observed. This will allow more photons to be collected, effectively increasing the throughput, and will also provide a broader wavelength range over which to identify unique planetary surface albedo components. Figures 4.5 and 4.7 can easily be interpreted for a telescope with a starshade by considering IWAs, diameters, and throughputs that are more applicable to such an observatory. However, the multi-epoch approach discussed in this work will require two pointings at the same target, possibly months apart, which is more constrained by the formation flying operations of the starshade.

Due to contrast considerations, large aperture ground-based telescopes will be unlikely to ob-

serve habitable planets around G dwarfs in reflected light, but may excel for K and M dwarfs. Ground-based 30 meter class telescopes have an estimated coronagraph raw contrast of $10^{-7} - 10^{-8}$ (Guyon, 2005; Pueyo and Norman, 2013), compared with estimates for space-based observatories of $\sim 10^{-10}$. Since Earth-Sun analogs have a flux contrast in the visible of $\sim 10^{-10}$, ground-based observations will face $10^2 - 10^3$ times more noise from an inability to fully block stellar light, which will more than erase gains from larger diameter mirrors. However, ground-based direct imaging of nearby habitable zone planets orbiting K and M dwarfs may be optimal for such 30-m class facilities due to more favorable planet-star contrast ratios, and the larger aperture helps to ameliorate some IWA concerns for these compact systems.

While similar previous studies have focused on broad wavelength coverage (e.g. Cowan et al., 2009), we focused on a narrow wavelength region (0.68 - 0.8 μm) to remain relevant to current coronagraph design that is unable to effectively control the wavefront over broad wavelength ranges (Pueyo and Norman, 2013). As a result of the narrow spectral coverage and the intrinsic faintness of target habitable exoplanets, we simulated multi-wavelength, time-series photometry by rebinning hourly spectroscopy. This allows for simultaneous multi-wavelength, time-series without the need to change photometric filters. This also means that the entire duration of the time-series may be coadded into one spectrum that contains the disk- and time-averaged spectrum of the planet, which would be an excellent data product for atmospheric retrieval studies (e.g. Feng et al., 2018). Thus, we argue that time-series spectroscopy is more favorable than traditional filter photometry for the time intensive observations discussed in this chapter. Low noise detector technology will enable this type of time-dependent spectroscopic observation. Microwave Kinetic Inductance Detectors (MKIDs) are a promising new technology that directly measure photon energy as they excite the detector, allowing for simultaneous photon counting and low-resolution spectroscopy (Meeker et al., 2018). Such advances would naturally allow for simultaneous multi-wavelength time-series observations. Furthermore, the wavelength range that we considered (0.68 - 0.8 μm) contains the 0.76 μm O₂ A-band and the vegetation red-edge—two potential biosignatures—and thus our IWA results are also directly applicable to observing biosignatures on potentially habitable nearby exoplanets. Ultimately, the atmospheric context necessary to more confidently character-

ize the habitability of an exoplanet target may be contained within the same dataset used for the mapping results presented in this chapter, but along the spectral-, rather than temporal-dimension.

4.5 Conclusions

We successfully mapped Earth as an exoplanet using simulated multi-wavelength, time-series observations at multiple orbital phases, while including ocean glint and forward scattering by clouds. We confirmed the findings of [Cowan et al. \(2009\)](#) that planet mapping in the specular reflection regime can identify two surface components corresponding to oceans and continents. However, our results extend this work to show that the specularly reflecting glint spot dominates the inferred crescent phase maps, just as a dark blue surface dominates our quadrature maps for ocean-dominated worlds. Interruptions in the crescent phase glint signal as Lambertian scattering continents rotate into view leads to a strong blinking effect, which allows for separate longitudinal maps to be constructed of Lambertian and non-Lambertian surfaces. We identified that the partitioning of Lambertian and non-Lambertian surfaces between retrieved surfaces is in stark contrast to the effects of forward scattering by clouds, which introduces a global apparent albedo increase for all surfaces towards crescent phase. Such multi-epoch studies may be used to detect the strong deviations from Lambertian scattering that oceans exhibit and thereby increase the robustness of ocean detection on directly imaged exoplanets.

We found that recognizing the same longitudinal map at multiple phases will be difficult for a variety of reasons. Clouds introduce uncertainty into the surface maps; non-zero obliquities and inclinations $< 90^\circ$ cause latitudinal sensitivities with phase; and the effective latitudinal extent of the mapped longitudes is a function of phase. However, even cases with distinct longitudinal maps at each phase may exhibit a time-dependent rotationally-induced glint signal, which differs significantly from forward scattering by clouds.

We estimate that space-based, high-contrast telescopes with diameters larger than 6-m will have at least one potentially habitable target amenable to glint mapping via crescent phase observations, with a 15-m telescope offering nearly 10 targets. Although the technique of glint mapping discussed here is not applicable to all potentially habitable planets, it may be a key tool for in-

creasing the robustness of habitability assessment for the most promising direct imaging targets. The discovery of an ocean on an exoplanet will transform exoplanet science by unveiling a truly habitable planet beyond Earth, bringing us one step closer to answering fundamental questions on the prevalence of life in the universe.

Chapter 5

SEARCHING FOR BIOSIGNATURES AND CONFIRMING HABITABILITY WITH LUVOIR

Portions of this chapter were originally written in collaboration with Giada Arney and the entire LUVOIR Mission Concept Study Team as a Design Reference Mission in the LUVOIR Final Report Appendices (B.3 Signature Science Case #2: Searching for biosignatures and confirming habitability), and are reproduced below with the permission of NASA.

Summary

Constraining the prevalence of habitable and inhabited exoplanets is one of the main science objectives of the LUVOIR telescope concept. Accomplishing this ambitious objective will require directly-imaged reflectance spectra of many Earth-like exoplanets that can be analyzed for signs of habitability and biosignatures. We use the LUVOIR-A (15 m) and LUVOIR-B (8 m) coronagraph design specifications reported in the LUVOIR Final Report to perform exposure time calculations and determine the number of potentially habitable Earth-like exoplanets for which LUVOIR can undertake a robust search for habitability and biosignatures. We find that a dedicated one-year characterization program with LUVOIR-A (LUVOIR-B) would be capable of observing the spectra of 24 (11) exo-Earth candidates between $0.3 - 1.5 \mu\text{m}$. We highlight the challenges faced when observing exo-Earths in the farthest UV and farthest NIR bandpasses, and we explore the effect to our spectral yields of imposing a target prioritization strategy based on preliminary evidence about each target. These spectral yields demonstrate that the LUVOIR telescope could characterize the nature of a large enough sample of HZ terrestrial exoplanets to place statistical constraints of the prevalence of planetary habitability and life in the Solar neighborhood.

5.1 Introduction

Determining whether a potentially habitable exoplanet is actually habitable and whether it hosts life will be one of the greatest astronomical challenges in the coming decades. The Large Ultraviolet / Optical / Infrared Surveyor (LUVOIR) has been designed with access to a broad swath of UV-visible-NIR wavelengths to specifically enable searches for a diverse range of habitability indicators and biosignature features (Dalcanton et al., 2015; Roberge et al., 2019; LUVOIR Mission Concept Study Team, 2019). This ambitious science objective has driven many design choices for

LUVOIR, however at present it remains unclear exactly how capable the LUVOIR mission would be at confirming habitability and searching for biosignatures.

As detailed in the LUVOIR Final Report (Chapter 3; [LUVOIR Mission Concept Study Team, 2019](#)), LUVOIR’s proposed strategy for studying habitable and possibly inhabited exoplanets employs a multilevel targeted approach to test whether habitable conditions are possible on planets orbiting nearby stars and to then determine whether any of the habitable exoplanets exhibit remotely detectable signs of life. To test these hypotheses, LUVOIR will first perform an initial two-year habitable exoplanet survey. This initial survey will use precursor information to establish a ranked list of promising targets based on known information about nearby stars available in the 2030s. Then multi-color photometry and proper motion will be used to determine which faint sources of light around target stars are actually planets. Next, observations spaced out over each planet’s orbit will be required to constrain the orbital motion and determine which planets are in the HZ. Finally, the initial survey will conclude with observations of low resolution, low S/N spectra to search for the H₂O feature at 0.94 μm . This will produce a catalogue of Earth-like exoplanet (exo-Earth) candidates that are likely to be both rocky and in their star’s HZ.

Following the exoplanet yield estimation methods of [Stark et al. \(2014\)](#), the LUVOIR team anticipate detecting approximately 54 and 28 exo-Earth candidates with LUVOIR-A and -B, respectively (Signature Science Case #1; [LUVOIR Mission Concept Study Team, 2019](#)). However, not all habitable planet candidates will be actually habitable. Habitability depends on the availability of surface liquid water, which in turn depends on volatile delivery, planetary evolution, and many additional factors (see [Meadows and Barnes, 2018](#)). Furthermore, only a fraction of these candidates may show signs of habitability via atmospheric water absorption features in the initial exo-Earth survey (we refer to this fraction as η_{water}). Therefore, significantly more information will be needed to follow-up on the exo-Earth candidates.

The LUVOIR habitable planet characterization program would employ a similar multilevel approach to combine and synthesize multiple observations into a comprehensive picture on the presence of habitability and biosignatures. First, UV observations of the star will be used to understand the photochemical environment of the planet. Second, astrometric observations will be

used to constrain the planet masses. Third, relatively high S/N spectra will be acquired to constrain the H₂O abundance, search for biosignatures, such as O₂, O₃, and CH₄ (see [Schwieterman et al., 2018](#)), and check that any biosignatures are not false positives ([Meadows, 2017](#); [Meadows et al., 2018b](#)). Finally, additional observations may be required for the most promising targets to extend the spectrum to find additional features and/or revisit the planet to search for seasonal variations.

In this chapter we focus on answering a single question about LUVOIR’s habitable exoplanet detection and characterization strategy that directly ties the field of Astrobiology with the engineering challenges of mission design optimization. That is,

How many potentially habitable Earth-like exoplanets will LUVOIR be able to rigorously search for habitable conditions and life?

This is ultimately a question of scientific yield that links characteristics about the mission—including exposure times and mission duration, specific telescope design specifications, instrument sensitivities, and observing strategies—to many crucial aspects of stellar, planetary and Earth science—including the demographics of nearby stars, exoplanetary occurrence rates, and strategies to detect surface oceans and biosignatures.

To determine the anticipated yield of Earth-like exoplanet spectra with NASA’s LUVOIR mission, we use the design specifications for the two LUVOIR concepts—LUVOIR-A (15 m diameter) and LUVOIR-B (8 m diameter)—from the LUVOIR Final Report to perform exposure time calculations for the list of habitable exoplanet candidates produced by the LUVOIR Team (see Signature Science Case #1; [LUVOIR Mission Concept Study Team, 2019](#)). We thus determine the amount of time required on each target to observe a spectrum that has sufficient S/N to probe habitability and biosignatures. Using this information we determine the number of exo-Earth candidates that could be observed in a dedicated characterization survey. Finally, we examine the sensitivity of these calculations on the spectral completeness (fraction of LUVOIR’s entire wavelength range that is observed) and the fraction of candidates that actually have H₂O features in their preliminary spectrum. This approach and the results presented here represent a modified version of a Design Reference Mission (DRM) conducted for the LUVOIR Final Report and originally appeared

in the The LUVOIR Final Report Appendices as *B.3 Signature Science Case #2: Searching for biosignatures and confirming habitability* (LUVOIR Mission Concept Study Team, 2019). The organization of this chapter is as follows: we describe our methods in Section 5.2; we present our results in Section 5.3; and we offer a discussion and conclusion in Section 5.4.

5.2 Methods

We begin the description of our methods next with an overview of our simulated observations (Section 5.2.1), followed by details about the stellar and planetary targets (Section 5.2.2). We then present the coronagraph model used for exposure time calculations (Section 5.2.3), and last we discuss our exposure times and overheads (Section 5.2.4).

5.2.1 Observation Overview

Our primary goal is to obtain a spectrum of each exo-Earth candidate that has sufficient wavelength coverage and SNR to detect signs of habitability and life, and to rule out potential false positives for any observed spectroscopic biosignatures. Based on the work of Feng et al. (2018), we believe that obtaining a spectrum across the bulk of LUVOIR’s ECLIPS coronagraph wavelength range ($0.2 - 2 \mu\text{m}$) is sufficient to accomplish this goal for modern Earth-like planets. Figure 5.1 displays a characteristic modern Earth spectrum from Robinson et al. (2011) with the expected noise for the observations described in this program. This wavelength range is also sufficient to characterize and detect life on an Archean or Proterozoic Earth analog, as described in Chapter 3 of the LUVOIR Final Report (LUVOIR Mission Concept Study Team, 2019).

We perform our nominal exposure time calculations assuming a planet with a wavelength-independent 20% albedo (a “gray” planet with the Earth average albedo). This gray spectrum is agnostic to specific molecular features, allowing us to estimate the integration time required to obtain a given SNR on the mean spectral continuum at any wavelength. This is done to ensure that the low planetary albedo at the bottom of absorption bands does not bias our observations towards unnecessarily long exposure time estimates. However, we also consider the detectability of key

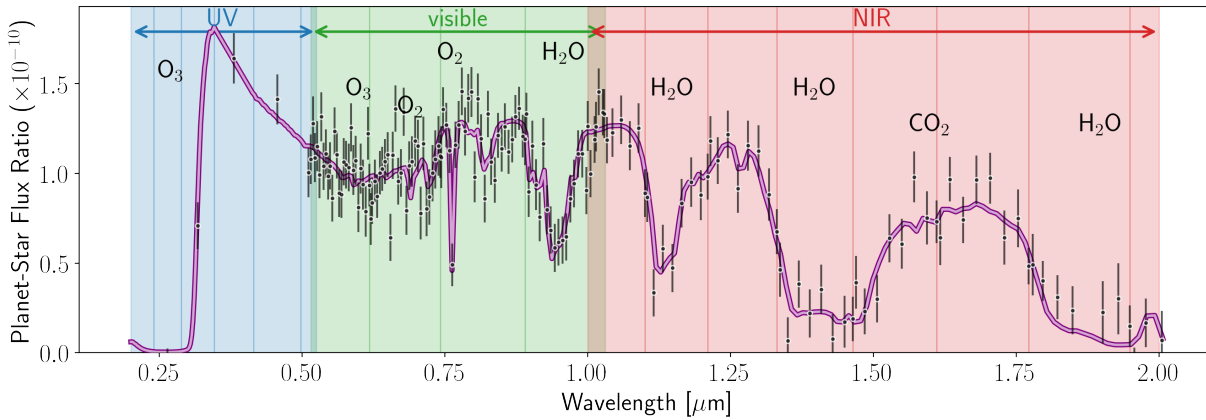


Figure 5.1 Simulated spectrum of an Earth analog exoplanet observed with LUVOIR in this DRM observing program (SNR = 8.5). The shaded background regions indicate the wavelength ranges of the UV (blue) and visible (green), and NIR (red) channels. The vertical lines show 20% bandpasses in the UV and visible channels and 10% bandpasses in the NIR channel.

individual molecular bands on modern Earth, and for this we use a partially cloudy Earth spectrum (Robinson et al., 2011).

In our calculations, we require SNR=8.5 on the continuum across the complete LUVOIR spectral range (0.2 – 2.0 μm) to detect and place abundance constraints on atmospheric gases, based on work by Feng et al. (2018). Feng et al. (2018) performed atmospheric retrievals from 0.4-1.0 μm at a spectral resolution of $R=140$ (the same as LUVOIR in the visible channel) for a range of SNR values. However, a subtlety of the Feng et al. (2018) results is that their quoted SNR values are calculated at a standard wavelength of 0.55 μm , not at the continuum near any particular spectral feature of interest. Therefore, the equivalent SNR on the continuum at other wavelengths will be different from the Feng et al. (2018) SNR values at 0.55 μm due to wavelength-dependent differences in the SNR between the LUVOIR design used here and the telescope noise model used in Feng et al. (2018). For instance, Feng et al. (2018) find that a spectrum with SNR=10 at 0.55 μm can place order-of-magnitude constraints on the O_2 abundance. We find that SNR=10 at 0.55 μm corresponds to SNR=8.5 near the O_2 A band at 0.76 μm . Based on the assumption that the O_2 A band is one of the more challenging and important spectral features that will retrieve, we apply this

SNR=8.5 requirement for O₂ across our spectrum, where our SNR=8.5 is equivalent to [Feng et al. \(2018\)](#)’s SNR=10 at 0.55 μm for O₂ retrieval. Note that [Feng et al. \(2018\)](#) show that constraints on O₂ are weaker than those on H₂O and O₃ at a given SNR, so we can assume that our spectrum will also produce as strong or stronger abundance constraints on H₂O and O₃.

Additionally, LUVOIR’s broad wavelength coverage extends farther into the UV and farther into the NIR than the spectrum from the [Feng et al. \(2018\)](#) study, enabling access to a range of biosignatures and biosignature false positives. Specifically, this enables LUVOIR to access features like the strong Hartley-Huggins ozone band (0.2 – 0.3 μm), which is a critical biosignature for planets with low oxygen abundances, like the mid-Proterozoic Earth ([Planavsky et al., 2014](#); [Reinhard et al., 2017](#)). LUVOIR can also access the 1.6 μm CO₂ band which, together with CH₄ and/or organic haze (with spectral features across the visible and NIR), form a biosignature for Archean Earth-like planets ([Arney et al., 2018](#); [Krissansen-Totton et al., 2018a](#)).

5.2.2 Targets

LUVOIR’s initial two-year exo-Earth detection survey provides the list of candidates from which targets for our intensive spectroscopic follow-up will be selected. Based on the exo-Earth yields anticipated for LUVOIR-A and B (54 planets for LUVOIR-A; 28 planets for LUVOIR-B), we perform a biased draw of targets from the total list of stars observed in the exo-Earth imaging surveys (LUVOIR Final Report Appendices Section B.2), weighted towards stars with higher habitable zone completeness. This becomes our biased catalog of systems with exo-Earth candidates.

We make a few simplifying assumptions for the target exoplanets. We assume that the exoplanet being observed is one Earth-radius and positioned at quadrature phase. We will not know the radius of directly imaged exoplanets at this stage and this stands to be a major characterization hurdle that is beyond the scope of this work. However, our assumption of quadrature phase is reasonable based on the constrained orbits of the targets at this stage of their characterization. We further assume that each planet is located at the inner edge of the habitable zone (moist greenhouse; [Kopparapu et al., 2013](#)) for the star about which it orbits. This is a conservative assumption from the perspective of the IWA, but optimistic from the perspective of planet-star contrast (due to the $1/r^2$ reflected flux

dependence). Blackbodies are assumed for each stellar spectrum, set by the known stellar effective temperature; 90% of our stellar catalog is comprised of FGK stars for which a blackbody is a reasonable assumption. We caution that this may change our calculations in the UV where stellar emission is significantly non-thermal. The exozodi level in each system is set to the same value drawn from the LBTI exozodi distribution for each target system used in the initial detection survey (see LUVOIR Signature Science Case #1 or [Stark, 2019](#)), which ranges from about 0.1 – 100 zodis.

5.2.3 Coronagraph Model

We make extensive use of the open-source Python coronagraph telescope noise model, `coronagraph`, developed for exoplanet science and significantly improved upon in this work. The Python `coronagraph` package is based on Interactive Data Language (IDL) code originally developed by T. Robinson¹ and is described in detail with science applications in [Robinson et al. \(2016\)](#). The Python code is published in the Journal of Open Source Software (JOSS; [Lustig-Yaeger et al., 2019](#)), it is continuously managed in an open-source GitHub repository², and it is available for download from the Python Package Index (PyPI) by typing “`pip install coronagraph`” at a standard linux/unix terminal.

Briefly, `coronagraph` uses analytic relations to calculate the photon count rates for a variety of astrophysical, telescope, and instrumental noise sources. These include photons from coronagraph speckles, zodiacal and exozodiacal dust, telescope thermal emission, dark current, read noise, and clock-induced charge. The model also includes Earth’s atmospheric transmission and emission spectrum from the UV through the near IR for ground-based telescope modeling ([Meadows et al., 2018a](#)). Photons from a user-provided exoplanet source spectrum are compared against the aforementioned sources of noise to provide wavelength-dependent signal-to-noise (S/N) ratios and synthetic observations. The model also features “reverse” S/N calculations, which provide the requisite exposure times per wavelength interval such that a user-specified S/N is achieved on the exoplanet spectrum. We will make use of these reverse S/N calculations in this work.

¹https://github.com/tdrobinson/coronagraph_noise

²<https://github.com/jlustigy/coronagraph>

The coronagraph model was designed to assess exoplanet science capabilities for near- and far-term future telescopes. Since the computations are all relatively simple and analytic, `coronagraph` is ideal for fast sweeps across a range of possible telescope, instrumental, astrophysical, and planetary parameters to investigate primary noise sources and the feasibility of different science cases. However, `coronagraph` may be impractical for existing telescopes or telescope concepts that have mature designs due to the “first order” nature of the model. Additionally, systematic errors (e.g. from speckle subtraction), stray light from companion stars and/or background objects, and time-dependent (red) noise terms are not considered. More specific or detailed use cases than those outlined above may be appropriate for the Exoplanet Open-Source Imaging Mission Simulator³ (EXOSIMS; Delacroix et al., 2016). However, for the purposes here the `coronagraph` model is ideal due to the large number of star-planet targets that we will simulate observations for, and the flexibility and speed of the model allow for a rapid response to changes in telescope design parameters, which occurred numerous times over the course of this work.

`coronagraph` has already been used in numerous practical science applications. This includes peer-reviewed work on the potential for direct-imaging Proxima Centauri b with ground- and space-based telescopes (Meadows et al., 2018a), the detectability of exoplanet aurorae on Proxima Centauri b (Luger et al., 2017b), and for assessing methods to directly detect exoplanet oceans (Chapter 4; Lustig-Yaeger et al., 2018). The coronagraph model is also used within NASA Goddard’s interactive tool *Coronagraphic Spectra of Exoplanets*⁴, and continues to be used in work such as this to motivate science cases for next-generation, space-based, direct-imaging mission concepts (Mennesson et al., 2016; Bolcar et al., 2017).

The coronagraph package may also be used to simulate signal-to-noise ratios and synthetic spectra for transiting exoplanets observed with transmission and emission spectroscopy. The transiting exoplanet modules extend the potential science applications of `coronagraph` to non-coronagraph-equipped telescopes, enabling studies relevant to a far-infrared surveyor mission concept (Cooray and Origins Space Telescope Study Team, 2017). This capability was used in the

³<https://github.com/dsavransky/EXOSIMS>

⁴http://luvoir.stsci.edu/coron_model

LUVOIR Final Report to assess the mission’s transit capabilities, but this work is not included in the nominal direct-imaging observing strategy under investigation here.

5.2.4 *Exposure Times & Overheads*

We adopt the nominal LUVOIR telescope and ECLIPS coronagraph instrument parameters for the LUVOIR-A and LUVOIR-B architectures provided in the LUVOIR Final Report (LUVOIR Mission Concept Study Team, 2019). One particularly strong constraint on LUVOIR’s ability to observe broad wavelength spectra is that LUVOIR can only observe spectra within one bandpass at a time per channel. We assume 20% instantaneous bandpasses across the LUVOIR UV and visible spectral ranges and 10% bandpasses across the NIR (shown by the vertical shaded regions in Fig. 5.1). Simultaneous observations in two channels at a time are accounted for in our reported total exposure times. It is assumed the third channel will be used for wavefront sensing, although observations in all three channels simultaneously may become possible in the future with technological improvements.

A few crucial modifications were made to the original coronagraph model during this project that are all now available in the online version of the model. First, the model now allows the user to provide a wavelength-dependent throughput curve to replace the single-valued panchromatic throughput. Figure 5.2 shows the optical throughput used for LUVOIR as a function of wavelength. Note that it peaks near 50% at $2\ \mu\text{m}$ and reaches a minimum of 5% at 200 nm in the UV. The coronagraph model was also upgraded to include the dependence of the coronagraph contrast and throughput on the separation between the planet and star as a function of λ/D . Figure 5.3 shows the raw coronagraph contrast and core throughput curves used for each type of coronagraph considered for the LUVOIR mission. The behavior is different for each coronagraph design as a function of the targets planet’s separation from the star. We account for these relationships to capture the subtle effects induced in each system observed. We use these specific relationships shown in Fig. 5.3 for the Apodized Pupil Lyot Coronagraph (APLC) on LUVOIR-A, and the Deformable Mirror assisted Vortex Coronagraph (DMVC) on LUVOIR-B. We also studied the Phase-Induced Amplitude Apodization (PIAA) coronagraph for LUVOIR-B. The APPLC has multiple masks for

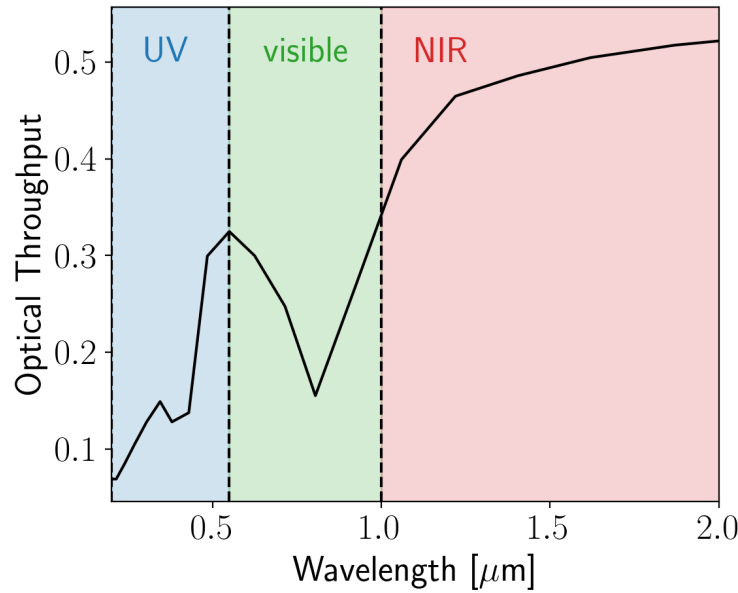


Figure 5.2 LUVOIR's wavelength-dependent optical throughput used in this study. The throughput minimum is about 5% at 200 nm and reaches a maximum of 52% by 2 μm .

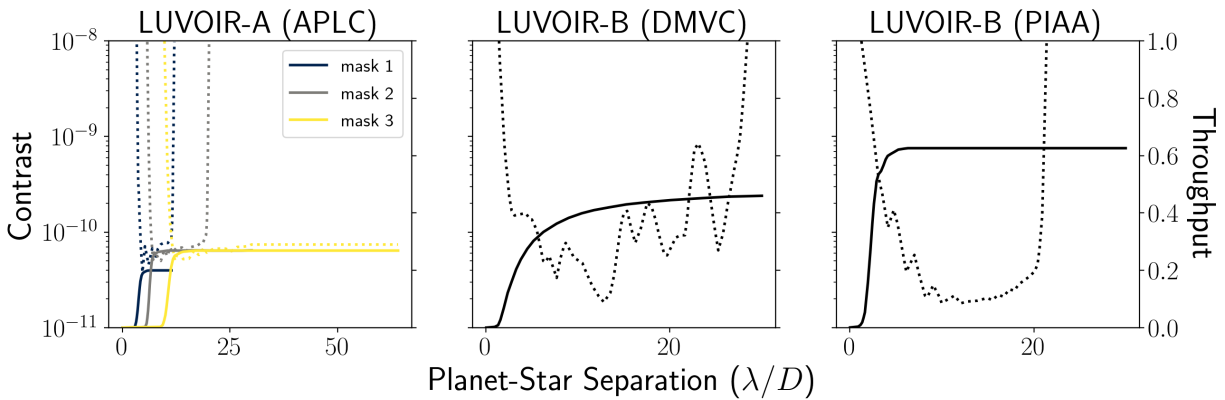


Figure 5.3 Raw contrast and core throughput curves as a function of planet-star separation for three different coronagraph designs considered in the LUVOIR Final Report. The coronagraph contrast curves (dotted lines) are shown relative to the left y-axis and the core throughput curves (solid lines) are shown relative to the right. The left panel shows the Apodized Pupil Lyot Coronagraph (APLC) considered for the LUVOIR-A concept; the central panel shows the Deformable Mirror assisted Vortex Coronagraph (DMVC) considered for LUVOIR-B); and the right panel shows the Phase-Induced Amplitude Apodization (PIAA) coronagraph considered for LUVOIR-B.

different planet-star separations. We perform exposure time calculations using all three masks and assume the use of the mask that gives the minimum exposure time for each individual star and for each bandpass.

We also account for overheads on our observations. For each target, we assume 1-hour for the combination of slew, dynamic settle, and thermal settle time. We then add 0.6 (1.25) hours for architecture A (B) to dig the dark hole down to the desired contrast for each bandpass observed. Finally, we impose a 10% overhead on the total science time for one assumed iteration of the wavefront control system.

For each star in the biased catalog of exo-Earth candidates, we calculate the exposure time in each bandpass necessary to observe an exo-Earth at $\text{SNR}=8.5$, as discussed above. This produces a lookup table of requisite exposure times for each target in each bandpass. Bandpasses that are not entirely observable when considering the IWA and OWA are not included in exposure time calculations. However, this approach overestimates decreases in spectral completeness, which we quantify as the percentage of the total LUVOIR wavelength range ($0.2 - 2.0 \mu\text{m}$) that is observable for any given exo-Earth target. To address this, we also consider science program cases that omit bandpasses, and necessarily decrease the spectral completeness to increase the spectral yield.

To examine how our spectral yields are impacted by observations that target only a fraction of the full sample of exo-Earth candidates due to a prioritization metric, such as the presence of H_2O , we randomly sample N_{draw} exo-Earth candidates from the biased catalog, where N_{draw} is set by the fraction of water-vapor-bearing targets in the biased catalog, η_{water} . We use the precomputed exposure times per bandpass to calculate the total exposure time to observe the full spectrum for each planet. As we will see, the total integration time to characterize the complete set of N_{draw} planets is exorbitantly large, so we also calculate the number of exo-Earth spectra that can be obtained in a time-limited program, for which we prioritize targets with shorter exposure times.

5.3 Results

We first examine the spread in exposure times among the exo-Earth candidates. We find that bandpass exposure times vary significantly, both from one wavelength to another, and at a given

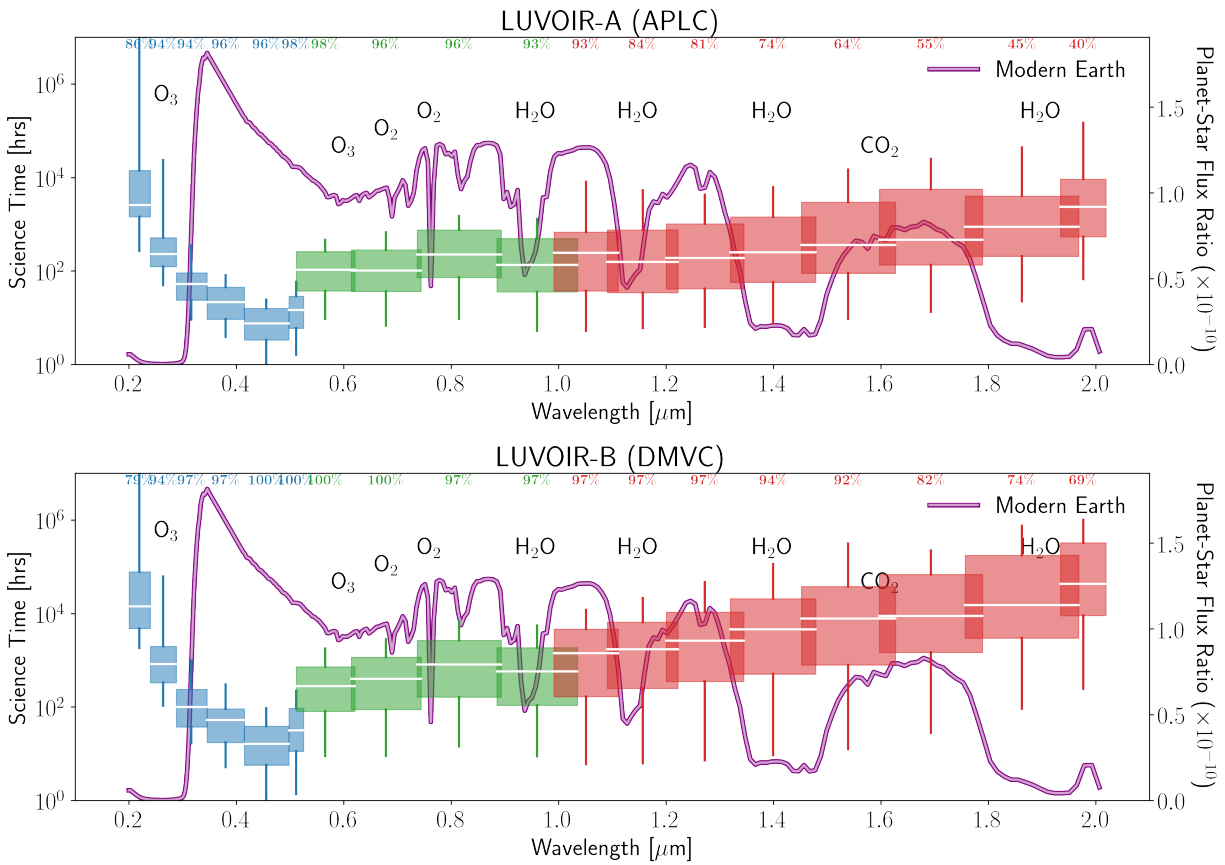


Figure 5.4 Required science exposure time per bandpass (left axis) to reach SNR=8.5 for LUVUOIR-A (top panel) and LUVUOIR-B (bottom panel). The blue (UV channel), green (visible channel), and red (NIR channel) “boxes and whiskers” show the 50% (box extent) and 95% (whisker extent) confidence intervals about the median (white bar) science time required to observe the spectrum shown in Figure 5.1, for the targets in the biased catalog of exo-Earth candidates. The percentage of stars in the target catalog that can be completely observed in each bandpass—considering IWA and OWA constraints is displayed above each box. Exposure times required to achieve a fixed precision on the spectrum can vary by orders of magnitude.

wavelength depending on the specific stars in the sample that ultimately host the target exo-Earth candidates. Figure 5.4 shows the range of exposure times required in each bandpass to observe the LUVUOIR-A (top panel) and LUVUOIR-B (bottom panel) exo-Earth candidates. The percentage of exo-Earth candidates in the biased draw catalog for which spectra in each bandpass may be obtained given IWA and OWA constraints is also shown along the top of each plot.

LUVOIR-A and LUVOIR-B offer nearly complete coverage at UV and visible wavelengths for our set of target stars, but the IWA leads to decreasing spectral completeness towards longer NIR wavelengths. The longest NIR bandpass requires relatively long exposure times, but only $\sim 40 - 70\%$ of the stars in the biased catalog may have access to this entire bandpass given the coronagraph IWA. The large tails seen in the 95% confidence intervals (Figure 5.4) that extend to lower exposure times is caused by the small handful of very nearby targets in the sample that are only a few parsecs away, which require only a few hours to observe in this bandpass.

The shortest UV bandpass is particularly costly to observe for exo-Earth planets, primarily due to the lack of reflected photons available and the low coronagraph throughput here. The large tail seen in the 95% confidence interval of the bluest UV bandpass that extends roughly two orders of magnitude higher in exposure time for LUVOIR-B is caused by the latest type stars (late K and M dwarfs) in the sample which have a low UV flux. These tails may be decreased by the use of better UV SEDs in future calculations.

Omitting the most time intensive bandpass(es) offers an observing strategy for obtaining spectra of more targets in significantly less observing time, without sacrificing significant scientific yield. These expensive bandpasses could be observed later as part of a follow-on program for the most interesting planets that exhibit signs of habitability and life. Hundreds of days of observing time can be saved if these costly bandpasses are not included in the nominal science program. Moreover, the two bluest UV bandpasses may not be immediately necessary for ozone detection on a modern Earth-like planet, which has the ozone Chappuis band between $0.5 - 0.7 \mu\text{m}$ (Feng et al., 2018) and can also access the long wavelength shoulder of the UV ozone Hartley-Huggins band in the third shortest UV bandpass at $\sim 0.3 \mu\text{m}$. For a Proterozoic Earth-like planet where the Chappuis band is not detectable, the third shortest UV band should be able to detect the spectral darkening from ozone. Meanwhile, the reddest NIR bandpasses may provide leverage for the detection of water and/or methane, but both of these molecules have bands at shorter wavelengths that may be more easily detected. Archean Earth in particular may have had a methane-rich atmosphere able to produce CH_4 spectral features even in the visible channel. These reddest bandpasses are most valuable for detecting methane at low modern Earth-like abundances, but this type of expensive

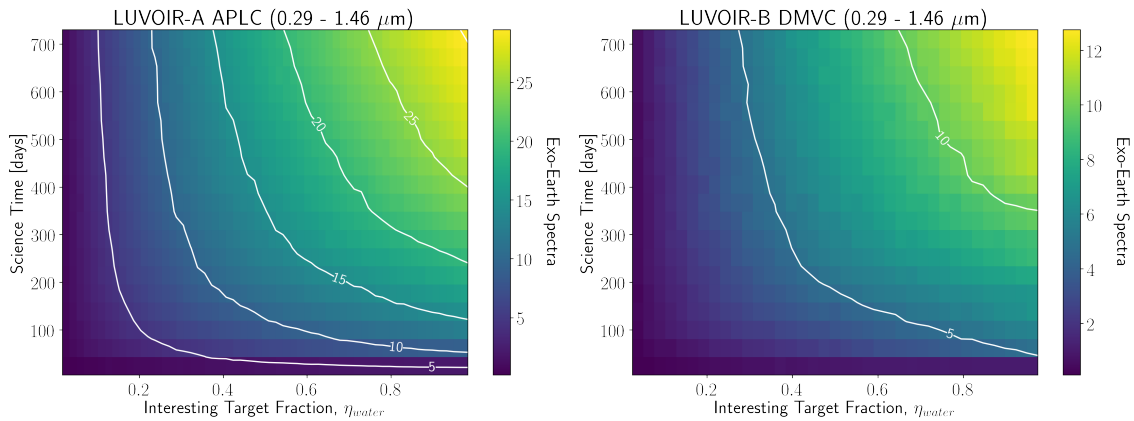


Figure 5.5 Number of spectra (color map and contour lines) that may be observed by LUVUOIR-A (left panel) and LUVUOIR-B DMVC (right panel) as a function of both total observing time and fraction of “interesting” targets in the exo-Earth candidate sample, η_{water} . The two shortest wavelength UV bandpasses and the four longest NIR bandpass are not included, so the spectrum spans $0.29 - 1.46 \mu\text{m}$. A higher fraction of candidate targets increases the probability that nearby, more easily observed, stars are included in the exo-Earth candidate sample, which yields more complete spectra in shorter duration programs.

observation would likely only be sought once other signs of biosignatures were detected at shorter wavelengths. These bandpasses also allow access to CO_2 near $1.6 \mu\text{m}$, which could be targeted as part of the nominal program for planets with shorter integration times, or part of an extended program if integration times are lengthy.

We also find that, if only a subset of “interesting” targets according to η_{water} are observed, the value of η_{water} strongly affects the number of spectra that may be obtained in a time-limited science program. Figure 5.5 shows contours of the average number of spectra that can be acquired as a function of the total observing time and the interesting target fraction, η_{water} . Low values for η_{water} (e.g., 10%) yield only a few exo-Earth targets slated for follow-up in this program, which, on average, require long exposure times due to the lower probability of them being nearby easy targets. On the other hand, optimistic estimates for η_{water} yield many interesting targets for follow-up, which can be prioritized by distance such that many spectra are able to be observed even for a relatively modest observing program duration.

Considering the aforementioned dependence of the spectral yields on both the *a priori* unknown

Mission Concept	N_0	α	β
LUVOIR-A	24	0.71	0.37
LUVOIR-B	11	0.76	0.39

Table 5.1 Parameters for the fit to the number of exo-Earth spectra observable with LUVOIR (see Equation 5.1).

interesting target fraction and the desired spectral completeness, we report spectral yields for a broad selection of different contingencies. Figure 5.6 and Figure 5.7 report the number of exo-Earth spectra that may be obtained for LUVOIR-A and LUVOIR-B, respectively, for different assumed interesting target fractions η_{water} and for different wavelength coverage cases. For our nominal observing program, all exo-Earth candidates are prioritized in order of shortest to longest observing time and the reported spectral yield is the number of spectra that fit into the assumed program duration after overheads are added. Results are presented for both 12-month and 6-month habitable exoplanet characterization science program durations. In addition to the information shown in Figure 5.7, we find that the yield values for LUVOIR-B may increase substantially (by up to 50%) with new PIAA coronagraph designs. For the nominal time allocation of these tests (6 months), we choose to report yields for programs designed to acquire spectra from $0.29 - 1.46 \mu\text{m}$, which we find strikes a balance between spectral completeness, observing time, and scientific yield.

We fit a simple power law scaling relation to the LUVOIR spectral yield results to more easily approximate our solutions. The number of exo-Earth spectra LUVOIR could obtain, N_{spec} , is given by

$$N_{\text{spec}} \approx N_0 \left(\frac{\eta_{\text{water}}}{100\%} \right)^\alpha \left(\frac{t_{\text{exp}}}{365 \text{ days}} \right)^\beta \quad (5.1)$$

where N_0 is the fiducial number of spectra in the nominal reported wavelength range ($0.29 - 1.46 \mu\text{m}$), t_{exp} is the total exo-Earth characterization program duration (exposure times plus overheads), and α and β are fitted powerlaw exponents. The parameters N_0 , α and β are reported for LUVOIR-A and LUVOIR-B in Table 5.1.

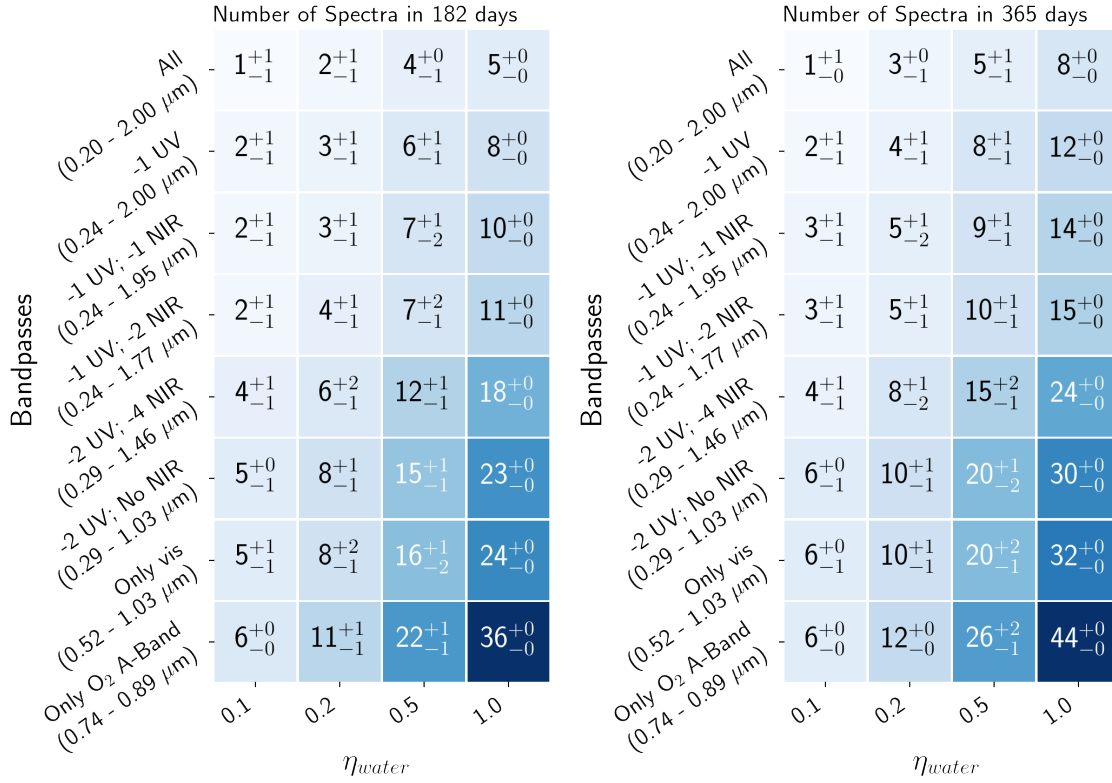


Figure 5.6 Number of spectra (SNR=8.5) observable in a 6-month or 12-month LUVOIR-A habitable exoplanet characterization program for different observed wavelength ranges (spectral completeness) and assuming different η_{water} fractions. The quoted “plus-minus” values only represent the uncertainty in which specific stellar targets are randomly selected to host “interesting” planets via η_{water} .

5.4 Discussion & Conclusion

We calculated the total time required to obtain direct spectra of sufficient quality to confirm habitable conditions and detect biosignatures on Earth-like exoplanets with LUVOIR. The number of Earth-like exoplanets for which these direct spectra can be obtained in a time-limited program vary substantially based on the number of planets chosen for spectral characterization and the fraction of the total wavelength range included in the nominal program (among other crucial factors). For instance, in a one-year observing program, LUVOIR-A (LUVOIR-B) would only be able to observe full spectra of 8 (2) exo-Earths from 0.2 – 2 μm , but if we instead focus on the 0.3 – 1.5 μm range

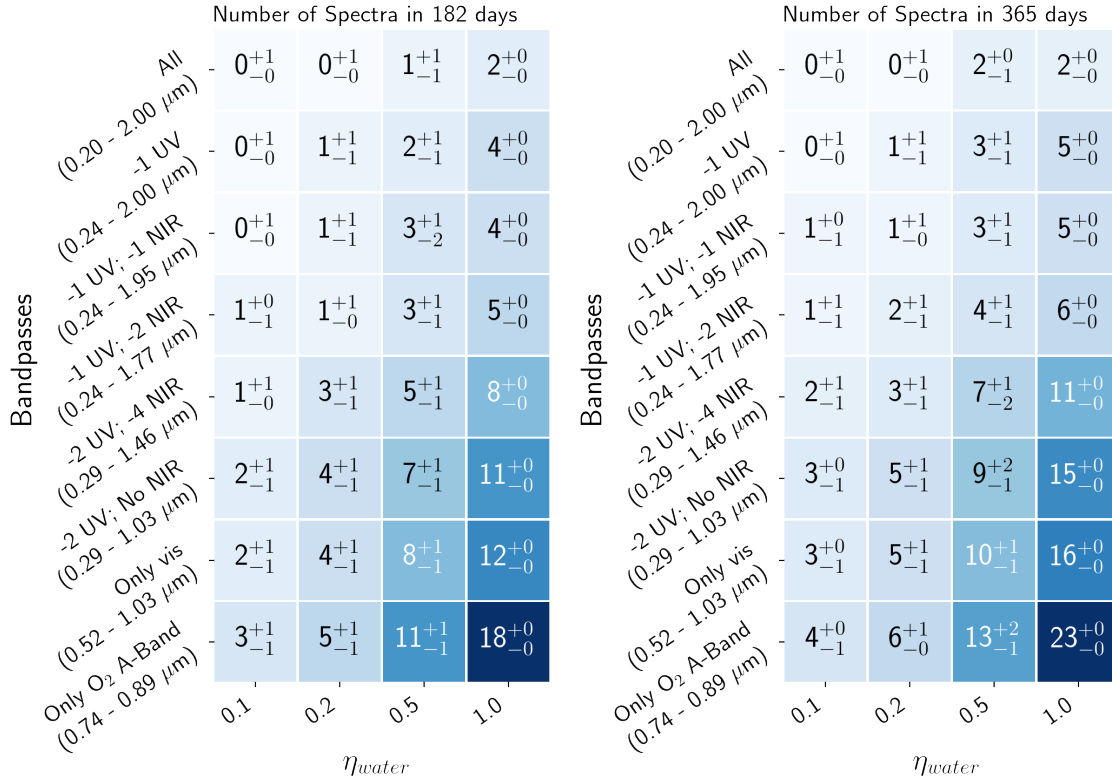


Figure 5.7 Number of spectra (SNR=8.5) observable in a 6-month or 12-month LUVOR-B habitable exoplanet characterization program for different observed wavelength ranges (spectral completeness) and assuming different η_{water} fractions. The quoted “plus-minus” values only represent the uncertainty in which specific stellar targets are randomly selected to host “interesting” planets via η_{water} .

we can increase the exo-Earth spectral yields to 24 (11) targets. We emphasize that these calculations do not describe a fully optimized observing program, but rather represent a proof-of-concept to establish feasibility and uncover avenues for further exploration.

The strategy that obtains the largest number of planetary spectra involves observing spectra of all exo-Earth candidates in order of lowest exposure time to obtain a reflectance spectrum, rather than down-selecting interesting candidates based on preliminary information. If the fraction of interesting targets (η_{water} in this case) is small, the exo-Earths selected for follow-up will on average be farther away from us and will require longer integration times to characterize due to the

small number of targets. In such a scenario piece-wise spectra that strategically focus on specific molecular signatures may be a better strategy than the “full” spectral characterization strategy laid out here. Furthermore, in a low η_{water} scenario, some time may be shifted from the 2 year nominal exo-Earth search survey to the 6 month characterization mission, as a 6 month reduction in the nominal survey is expected to decrease the exo-Earth candidate yield by only 10%, while it would double the time available for characterization. Maximizing the number of planets observed can enable powerful comparative planetology. As a result, we chose to focus on a nominal program that obtains spectra of as many planets as possible.

Figure 5.8 shows example spectral observations that LUVOIR could obtain in this mock observing program. The top panel shows the modern Earth’s spectrum with clearly visible Rayleigh scattering and gaseous absorption due to O_3 , O_2 , and H_2O , although CO_2 and CH_4 would not be detectable. The second panel shows a version of the Proterozoic Earth with low atmospheric oxygen levels with detectable Rayleigh scattering and gaseous absorption due to H_2O , CH_4 , and CO_2 , but O_3 is only detectable with UV coverage short of $0.3 \mu\text{m}$. The third panel shows a version of the Archean Earth with a hydrocarbon haze, which is detectable via a visible and UV haze absorption feature, CH_4 , H_2O , and CO_2 are also detectable. The fourth (bottom) panel shows the spectrum of Venus with a prominent absorption feature in the UV/blue due to still unknown UV absorber, along with numerous CO_2 features throughout the NIR.

It is worth noting that part of LUVOIR’s full habitable planet candidate characterization strategy includes measuring planet masses using on-board astrometry with HDI. Planet masses are essential for this stage of characterization, but are not included as part of our work here. The time required for these astrometric measurements is included in the LUVOIR Final Report as *Signature Science Case #5: The formation of planetary systems* (Appendix B.6.2 of [LUVOIR Mission Concept Study Team, 2019](#)). Although the exposure times are negligible compared to those presented here for spectra, the time costs are dominated by overheads to perform multi-epoch observations for each planet throughout their orbits.

5.4.1 Targeted versus Agnostic Observing Strategies

The distinction between $\eta_{\text{water}} = 100\%$ and $\eta_{\text{water}} < 100\%$ provides insights into the difference between a *statistical search* for habitability and biosignatures versus a *system science* target-based approach. That is, using an observable planetary characteristic—such as the presence of water vapor—to down-select targets necessarily biases the original sample of exo-Earth candidates in order to test a given hypothesis on specific targets. The unbiased sample (represented by $\eta_{\text{water}} = 100\%$ for simplicity) is larger and can be used to probe underlying population trends for planets in the HZ (e.g. are there trends in the CO_2 abundance with semi-major axis? [Checlair et al., 2019](#)). While the biased sample ($\eta_{\text{water}} < 100\%$) can be used to probe the environments of individually selected targets (e.g. are there biosignatures and can we rule out false positives?). In the system science approach, we intentionally have fewer targets because we focus exclusively on what are deemed to be the most interesting/promising targets, such that we can test a hypothesis. In the statistical approach, we benefit from as large and unbiased a sample as possible. These are two complementary approaches to exoplanet characterization; they both seek answers on the nature of exoplanets, but by asking fundamentally different questions. Perhaps ideally, we would create an observing program for LUVOIR that optimally features both approaches.

The spectra in this observing program for $\eta_{\text{water}} = 100\%$ may offer a first glimpse at both characterization approaches working together as the spectra are sufficiently high SNR and broad enough in wavelength coverage to allow for a robust constraint on biosignatures and biosignature false positives, thereby testing the hypothesis that life exists as we know it. However, by including all exo-Earth candidates we achieve a large enough sample to conduct a rigorous characterization campaign capable of placing statistical constraints on any underlying comparative planetology trends identified between the planets.

5.4.2 Future Work

The results presented here can be improved upon in the future to increase confidence and robustness for habitability and biosignature detection surveys. First, the design specifications for LUVOIR and

its instruments are guaranteed to change even if NASA decides to prioritize one of the LUVOIR-A or LUVOIR-B architectures. Not only will such changes make these results obsolete, but there may be opportunities to optimize the design to increase the yield of exo-Earth spectra—a potentially crucial back-and-forth. Additionally, the synthetic spectra “observed” in this work and shown in Figure 5.8 should be studied in detail to further quantify how the telescope and instrument designs, exposure times, and S/N directly affect the inference of habitability and biosignatures. Instead of tuning our S/N to calibrate our spectra to the retrieval results from Feng et al. (2018), future work should reproduce and improve upon the work of Feng et al. (2018), using a state-of-the-art atmospheric retrieval code to interpret the many synthetic spectra produced from our calculations. This would help to verify the inferred constraints on molecular abundances and to articulate optimal future strategies. In the next chapter we will describe a new atmospheric retrieval model that has been specifically designed to complete future projects such as this.

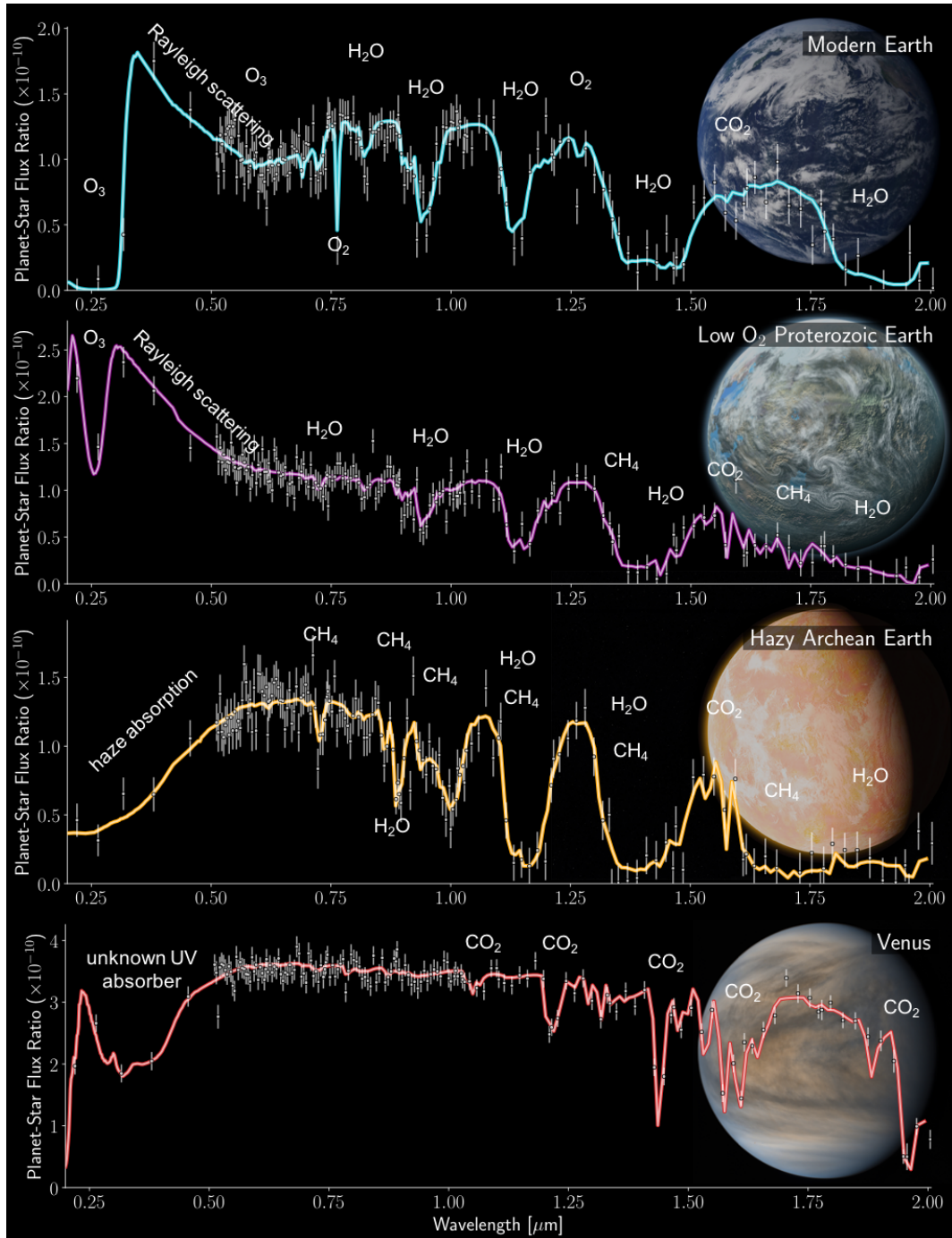


Figure 5.8 Example reflected light spectra of the modern Earth (top panel), low O_2 Proterozoic Earth (second panel), hazy Archean Earth (third panel), and Venus (bottom panel) as they would be observed, as exoplanets, in this theoretical LUVOIR observing program.

Part III

RETRIEVING THE ATMOSPHERES OF TERRESTRIAL EXOPLANETS

Chapter 6

DEVELOPMENT OF A TERRESTRIAL EXOPLANET ATMOSPHERIC RETRIEVAL MODEL

Portions of this chapter are in preparation for publication in the The American Astronomical Society Journals in collaboration with Victoria S. Meadows, Michael R. Line, David Crisp, and Tyler D. Robinson.

Summary

Exoplanet atmospheric retrieval models have become invaluable tools for the interpretation of observed exoplanet spectra, however surprisingly little work has been done to develop and test retrieval methods for *terrestrial* exoplanets despite the excitement for upcoming spectral observations of terrestrial exoplanets with *JWST*. In this chapter, we develop and describe a novel exoplanet atmospheric retrieval algorithm built specifically for terrestrial exoplanets whose atmospheres and surfaces may span a broad range of temperatures, surface pressures, and compositions. Our retrieval algorithm is built around the physically rigorous line-by-line radiative transfer code, SMART, originally built for Solar System terrestrial remote sensing. Our new SMART Exoplanet Retrieval code, or SMARTER, has a modular design featuring a suite of different forward and inverse models that enable atmospheric inference for a variety of different observing techniques—including transmission, emission, and reflectance spectroscopy—coupled to an assortment of different Bayesian retrieval methods—including optimal estimation, Markov chain Monte Carlo, nested sampling, and a novel machine learning approach. We validate SMARTER against transmission spectrum observations of Earth observed in solar occultation, demonstrating key capabilities and challenges for interpreting the spectra of terrestrial exoplanets with complex atmospheres.

6.1 Introduction

Exoplanet spectroscopy offers a window into the atmospheric state of tens to hundreds of planets only recently discovered, all but promising a scientific revolution in our understanding of atmospheric formation, evolution, composition, climate, chemistry, and dynamics (Burrows, 2014). Spectroscopic observations of rocky, Earth-size exoplanets orbiting the smallest of M dwarfs are expected to reach new heights in precision with the James Webb Space Telescope (*JWST*) when it

launches, facilitating the study of primary and secondary atmospheres and enabling the first search for habitability and biosignatures, especially in the TRAPPIST-1 system (Morley et al., 2017; Krissansen-Totton et al., 2018a; Wunderlich et al., 2019; Lustig-Yaeger et al., 2019b). Concepts are under development for a next generation space-based telescope that will expand these science cases to Earth-like exoplanets orbiting Sun-like stars using direct imaging spectroscopy, expanding our focus beyond M dwarf planets and requiring even greater levels of spectral precision (HabEx Study Team, 2019; LUVOIR Mission Concept Study Team, 2019). However, the atmospheric retrieval models required to interpret exoplanet spectra have largely focused their efforts on extra-solar gas giants due to the current data that exists for such planets (Line et al., 2013; Kreidberg et al., 2014, 2015; Line and Parmentier, 2016), leaving a substantial gap in terrestrial exoplanet retrieval capabilities on the metaphorical eve of finally acquiring terrestrial data¹.

A significant amount of work has been done to develop atmospheric retrieval models capable of interpreting the transmission and thermal emission spectra of gas giant exoplanets (Irwin et al., 2008; Madhusudhan and Seager, 2009; Line et al., 2013; Benneke and Seager, 2013; Waldmann et al., 2015; Cubillos et al., 2016; Lavie et al., 2016; Howe et al., 2017; Kitzmann et al., 2019; Zhang et al., 2019; Damiano and Hu, 2020). Some retrieval models focus on a data-driven approach that allows the atmosphere to take on any form as long as it optimally fits the spectrum (e.g. Madhusudhan and Seager, 2009), while other retrieval codes enforce thermal equilibrium chemistry so that the best-fitting spectra maintain plausible chemical self-consistency (e.g. Kreidberg et al., 2015). The assortment of different retrieval models and their respective physical and statistical approaches have proven to be an essential ingredient in numerous recent exoplanet atmospheric characterization discoveries, including the measurement of atmospheric water vapor in the atmospheres of hot Jupiters and super-Neptunes (Wakeford et al., 2013; Fraine et al., 2014; Kreidberg et al., 2015, 2018), and the seemingly ubiquitous presence of clouds and hazes in the atmospheres of exoplanets from super-Earths to hot Jupiters (Berta et al., 2012; Pont et al., 2013; Kreidberg et al., 2014; Barstow et al., 2017; Pinhas et al., 2019). Furthermore, efforts to validate

¹We note that the persistent *JWST* delays have created an oft overlooked *steady-state eve*.

the different retrieval models against one another have led to modeling improvements and to a general consensus of code agreement (Kreidberg et al., 2015; Barstow et al., 2020). Despite the well-established importance and recent successes of retrieval modeling, relatively little work has been done to develop and test retrieval methods for small, rocky, terrestrial exoplanets.

Due to the recognized difficulty of studying Earth-size exoplanet atmospheres and the lack of any spectrally-resolved observations of these bodies, only a few studies over the past decade have specifically used retrieval models to characterize the atmospheres of terrestrial exoplanets using *JWST* (Barstow et al., 2016; Barstow and Irwin, 2016; Krissansen-Totton et al., 2018a). Barstow et al. (2016) used the NEMESIS retrieval model (Irwin et al., 2008) to determine if *JWST* can distinguish an Earth-like from a Venus-like exoplanet, thus highlighting challenges associated with distinguishing intrinsically distinct atmospheres using only low resolution, low signal-to-noise transit spectra. The discovery of the TRAPPIST-1 system (Gillon et al., 2016, 2017b) opened a new door for terrestrial exoplanet atmospheric characterization, as the seven confirmed planets form a set of terrestrial targets whose atmospheres could be studied with *JWST* observations (Morley et al., 2017; Lustig-Yaeger et al., 2019b). Barstow and Irwin (2016) performed simulated *JWST* retrievals of the inner TRAPPIST-1 planets assuming that they had Earth-like atmospheres² and concluded that a dedicated campaign of 30-60 transits could detect ozone. Krissansen-Totton et al. (2018a) expanded on these studies by using the NEMESIS retrieval and determined that *JWST* may be able to detect signs of anoxic biosignatures in the atmosphere of TRAPPIST-1e, taking an important step towards biosignature retrieval for planets that are not modern-Earth-analogs. Indeed the discovery of the TRAPPIST-1 system has reoriented the field towards a perspective that terrestrial exoplanet atmospheric characterization may in fact be achievable with *JWST*, even if only an initial assay.

Most of the past terrestrial exoplanet atmospheric retrieval efforts have focused primarily on using Earth's atmospheric composition as a guide to build science cases for different future telescopes and/or observing/instrumentation techniques that may be required to characterize such

²Although the inner TRAPPIST-1 planets are not in the HZ.

rocky worlds. Some researchers have performed terrestrial retrieval studies on thermal emission spectra spanning mid-IR wavelengths to build the case for studying the thermal structure and composition of Earth-like planets using a future infrared interferometer (Von Paris et al., 2013; Léger et al., 2019) or a JWST-like successor (Quanz et al., 2019; Tremblay et al., 2020). Recently, Tremblay et al. (2020) upgraded the CHIMERA retrieval algorithm (Line et al., 2013, 2014; Line and Parmentier, 2016) to study TRAPPIST-1e with Earth-like atmospheric conditions to make the case for biosignature detection with the next-generation mid-IR telescope concept, OST (Origins Space Telescope Study Team, 2019). Other papers have undertaken terrestrial exoplanet retrievals to build the case for directly-imaging Earth-like exoplanets to search for signs of habitability and biosignatures in light reflected off planetary surfaces and atmospheres (Brandt and Spiegel, 2014; Feng et al., 2018). In particular, Feng et al. (2018) performed seminal work to develop a reflected-light retrieval model and assess their ability to retrieve gases from an Earth-like exoplanet observed with future direct imaging missions (WFIRST/HabEx/LUVOIR). This work demonstrated that HabEx and LUVOIR could indeed detect and constrain the abundance of H₂O, O₂, and O₃, and provided key spectral resolution and SNR constraints for the telescope design teams. Additionally, Wang et al. (2017a) investigated whether high-dispersion coronagraphy (high-contrast coronagraphy combined with high-resolution spectroscopy) can be used to directly image Earth-like exoplanets orbiting M dwarfs by leveraging the increased planet-to-star flux contrast at high spectral resolution—a method that could be performed with future ground-based telescopes³. These past works make strong cases for developing future telescopes that can study Earth-like planet atmospheres, but have largely neglected non-Earth-like atmospheric compositions.

The possible diversity among terrestrial exoplanet atmospheres stands as a major challenge for atmospheric retrieval modeling, despite also being an exciting opportunity to significantly advance the fields of planetary science and astrobiology through comparative planetology. In particular, the impacts of unknown/unconstrained bulk atmospheric composition have yet to be fully explored

³Although this work did not perform atmospheric retrievals, they suggested that integrating atmospheric retrieval into the cross-correlation template-matching analysis framework could help mitigate errors introduced by mismatched templates.

with retrieval models. The subsequent degeneracies among different bulk compositions are likely to confound efforts to understand the nature of terrestrial atmospheres. In the future, the ability to robustly infer biosignatures and rule out false positives will depend on our understanding of the atmospheric context in which a potential biosignature is found (Meadows, 2017; Meadows et al., 2018b; Catling et al., 2018). Meeting this goal will require a thorough census of atmospheric gases attained via spectral inference with atmospheric retrieval models. It will also require a self-consistent treatment of the physics and chemistry of terrestrial atmospheres to synthesize this information. The self-consistent thermal equilibrium chemistry retrieval approaches that have been applied to extrasolar gas giants (e.g. Kreidberg et al., 2015) are not always appropriate for terrestrial planet atmospheres, which are often sufficiently cool or thin that photochemistry or geologic outgassing drive the atmospheric composition out of thermal equilibrium. However, these methods do demonstrate the potential benefits of thermal and chemical constraints on the terrestrial planet retrieval approach.

Atmospheric retrievals will clearly play a crucial role in the near future to analyze the spectra of terrestrial exoplanets and attempt to robustly infer indicators of habitability and biosignatures. However, despite all of the work that has been done to establish atmospheric retrievals as commonplace in the exoplanet field, much work remains to develop a robust atmospheric retrieval framework for the interpretation of terrestrial exoplanet spectra with *JWST*, and beyond it. A model is needed that (1) does not treat terrestrial planets as small-scale Jovian planets nor as strictly Earth-like, but is built from the ground up to handle the physical complexity and chemical composition diversity expected for terrestrial exoplanets; (2) is prepared to address these challenges while analyzing upcoming *JWST* observations; (3) has a consistent framework to compare, contrast, and combine different observing techniques that are applicable to future telescopes; (4) can operate at high resolution to investigate promising new techniques applicable to ground-based 30-m class telescopes; and (5) can incorporate robust self-consistent physics and chemistry to lay the foundation for a Bayesian framework for future habitability and biosignature detection.

In this chapter we describe and validate such a model for the interpretation of terrestrial exoplanet atmospheres. In §6.2 we introduce our novel terrestrial retrieval model, including both the

forward (§6.2.3) and inverse models (§6.2.4). In §6.3 we detail our efforts to validate the model using spacecraft observations of Earth. We summarize and discuss our findings in §6.4.

6.2 The SMART Exoplanet Retrieval (SMARTER) Model

The purpose of a remote sensing retrieval model is to infer physical and chemical information about a planet’s atmosphere and surface given observations of the planet’s emitted, reflected, or transmitted spectrum. A typical retrieval model does this by assuming an initial atmospheric state, and then modifying that state to simulate the planetary spectrum that is consistent with the observations. Generally, this is an iterative process, in which the simulated spectrum for an assumed atmosphere and surface state is compared to the observed spectrum, and the state is updated to reduce the differences between the simulated and observed spectra. This process is repeated until the best possible fit to the planetary spectrum, and by extension, the most probable picture of the planetary properties, is obtained. Figure 6.1 shows a diagram of the standard atmospheric retrieval workflow used in this work.

In more detail, atmospheric retrieval modeling is Bayesian statistics applied to fitting planetary spectra. A retrieval begins with an initial guess for the physical and chemical state of the planet’s atmosphere and surface (e.g. atmospheric temperature structure and gas abundances, surface temperature and albedo, clouds, etc.), which is called the initial *state vector*, $\mathbf{x}_0$. A radiative transfer model is used to generate a high-resolution synthetic spectrum of the planet described by the parameters in the state vector. The spectrum is then convolved with a response function for the instrument that was used to make the observation, and mapped to the appropriate spectral resolution and wavelength range of the data. The combined radiative transfer and instrument simulator models are together referred to as the *forward model*. The model spectrum, $\mathbf{F}(\mathbf{x})$, is directly compared to the observations, $\mathbf{y}$. A hypothetical perfect model’s fit to noiseless data would give $\mathbf{F}(\hat{\mathbf{x}}) = \mathbf{y}$, where $\hat{\mathbf{x}}$ in this case would be the best possible retrieved atmospheric parameters. However, for any real application the data will have measurement errors, which must be taken into account, such that the relevant metric of comparison between the observed and the simulated spectra is a *cost* (or *log-likelihood*) function. The canonical cost function is written as the sum of two terms, the first

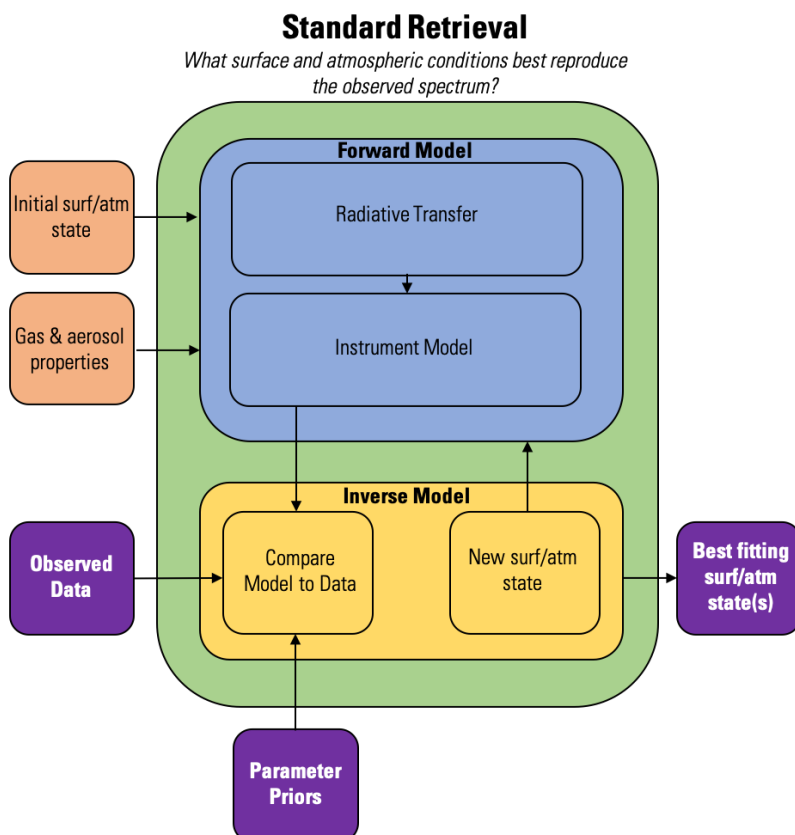


Figure 6.1 Diagram of the standard retrieval workflow used by SMARTER. An initial guess for the surface and atmospheric state is passed to the forward model, along with gas and aerosol optical properties. The forward model calculates the spectrum and uses an instrument model to degrade it to the quality of the observed data, before passing the spectrum to the inverse model for comparison with the observed spectrum via Bayes' theorem. The inverse model then determines a new state for the planet surface and atmosphere, and runs it back through the forward model. The retrieval iterates in this way, until the inverse model's specific convergence criteria are met and a best fitting set of parameters (or the full posterior distributions) are obtained.

describes the “goodness of the fit” relative to the noise in the spectrum (*chi-squared*), while the second penalizes states that deviate from prior constraints on the atmosphere given other previous measurements or physical constraints. Depending on the value of the cost function for a given state, $\mathbf{x}_i$, that state will either be accepted as the retrieved solution, $\hat{\mathbf{x}} = \mathbf{x}_i$, or modified and fed back into the forward model for another iteration, $\mathbf{x}_{i+1}$. The logic that drives the retrieval process to converge to the best fitting atmospheric state (or set of states) and iterates $\mathbf{x}$ through parameter

space is called the *inverse model*. In the remainder of this section, we will describe how SMARTER treats each of the aforementioned retrieval components.

In the next section we introduce Bayes’ Theorem, which is the underlying statistical framework used by retrieval algorithms. Then, in Section 6.2.2, we describe in detail the radiative transfer code, SMART, used within our retrieval forward model that enables terrestrial planet spectra to be simulated. In Section 6.2.3, we explain how our retrieval forward model operates for exoplanet transmission, emission, and reflectance spectroscopy. Finally, in Section 6.2.4, we detail the many different inverse modeling approaches that can be used within our retrieval framework.

6.2.1 Bayes’ Theorem

At the heart of an atmospheric retrieval algorithm is the Inverse Model—an elegant mathematical formalism based on Bayes’ Theorem—that is used to update our understanding about an observed system by reconciling *new data* from observations of the system with *prior knowledge* on the system from previous measurements or physical constraints. For instance, in the context of Earth remote sensing, one may wish to update our knowledge of Earth’s atmospheric CO₂ abundance by interpreting a recently observed spectrum from an orbiting satellite. However, we also want to leverage the fact that previous measurements of Earth’s CO₂ may have reported the abundance as 415 ± 3 ppm during the same month in the previous year⁴. Bayes’ Theorem provides the statistical framework for reconciling the previously inferred uncertainties on CO₂ with the new spectral data and its uncertainty so that we can provide an informed update to the CO₂ abundance. This is necessary because the CO₂ abundance and the uncertainty on it cannot be directly measured, but instead must be *inferred* by fitting a simulated spectrum, generated with a radiative transfer model using the correct amount of CO₂, to the spectral observations, using a retrieval model. Therefore, Bayes’ Theorem offers a method to propagate the measurement uncertainties from the spectrum through to the inferred quantities, such as the CO₂ abundance.

Formally, Bayes’ Theorem states that the *posterior probability* of our model parameters, $\mathbf{x}$ (e.g.

⁴Although this is based on values acquired online (<https://www.co2.earth/daily-co2>), this example is purely pedagogical.

which may include CO₂), given some observed data, y (e.g. our measured spectrum), is

$$\underbrace{\mathcal{P}(x|y)}_{\text{posterior}} = \frac{\overbrace{\mathcal{P}(y|x)}^{\text{likelihood}} \overbrace{\mathcal{P}(x)}^{\text{prior}}}{\underbrace{\mathcal{P}(y)}_{\text{evidence}}}. \quad (6.1)$$

We can interpret $\mathcal{P}(x|y)$ as “the probability of x given y ”, where we read “ $\mathcal{P}$ ” as “the probability” and the delimiter “|” as “given”. The posterior can be thought of as the updated information after taking the data into account. The term, $\mathcal{P}(y|x)$, is called the *likelihood* and represents the probability that our data would exist, given our choice of atmospheric model parameters. Although this is conceptually backwards from what we are actually trying to estimate (think *inverse* modeling), as we will see later in this section, the likelihood is mathematically well-defined as it simply quantifies how well our current state fits the observed spectrum; *how likely is it that our data are noisy samples of our current synthetic spectrum?* The term, $\mathcal{P}(x)$, is called the *prior probability* and represents the probability of our model parameters at any particular atmospheric state given previous measurements and/or physical constraints. The term in the denominator, $\mathcal{P}(y)$, is called the *marginal likelihood* or the *evidence* and it acts to normalize the equation by integrating the likelihood over the entire parameter space (e.g. all possible values of x). As we will see, the evidence term is often neglected for Bayesian parameter estimation because it is unnecessary to normalize Bayes’ Theorem to find estimates of each parameter (e.g. histograms need not be normalized to identify where they peak). However, the evidence term becomes important for model comparison applications, when one wishes to identify which specific forward model, among many, does the data provide the strongest evidence.

The broad applicability of Bayes’ Theorem stems from the way it attempts to mathematically balance the influx of new information (via the likelihood) with previously held beliefs (via the priors), thereby formalizing the process of changing one’s mind with the advent of compelling new data. However, despite its relatively simple form, Bayes’ Theorem is conceptually tricky because each term represents a probability distribution that can be visualized using a histogram, but we do not typically hold intuition for performing algebra using probability distributions. Consequently,

it is often necessary to evaluate Bayes' Theorem numerically by picking one choice of $\mathbf{x}$ at a time (e.g. a single choice for the CO₂ abundance), evaluate the prior and likelihood, and then calculate the posterior probability at that state. After many choices of $\mathbf{x}$, the full shape of the posterior probability distribution can be constructed.

A bit of careful mathematical manipulation helps to transform Bayes' Theorem from the abstract into a form that can be related back to our retrieval problem, and easily evaluated numerically. In practice, characterizing the full shape of the posterior distribution, as well as estimating maximum *a posteriori* solutions (best-fitting solutions), can be done without the use of the marginal likelihood to normalize the equation, such that Bayes' theorem simply becomes

$$\mathcal{P}(\mathbf{x}|\mathbf{y}) \propto \mathcal{P}(\mathbf{y}|\mathbf{x})\mathcal{P}(\mathbf{x}). \quad (6.2)$$

Taking the logarithm of both sides of the equation gives

$$\log \mathcal{P}(\mathbf{x}|\mathbf{y}) \propto \log \mathcal{P}(\mathbf{y}|\mathbf{x}) + \log \mathcal{P}(\mathbf{x}) \quad (6.3)$$

which is now in a form that is easier to evaluate numerically. Next, let's examine the log-likelihood term, which we can write as

$$\mathcal{L} = \log \mathcal{P}(\mathbf{y}|\mathbf{x}) = -\frac{1}{2}[\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_e^{-1}[\mathbf{y} - \mathbf{F}(\mathbf{x})] + \text{const}. \quad (6.4)$$

(Line et al., 2014), where $\mathbf{F}(\mathbf{x})$ is the *forward model* that generates a spectrum to be compared to the data, using the assumed properties of the planetary environment ($\mathbf{x}$), thereby mapping the state vector onto the spectral dimension of the data ($\mathbf{y}$). $\mathbf{S}_e$ is the data error matrix, which traditionally contains uncorrelated measurement errors along the diagonal, but may also have off-diagonal correlations⁵ (Line et al., 2013). For data with uncorrelated, Gaussian errors, the log-likelihood can

⁵Correlated errors can occur in a spectrum due to correlations between neighboring pixels on a CCD, wavelength dependent systematic effects, etc, but are neglected in the bulk of this work.

be simplified and written as

$$\mathcal{L} = -\frac{1}{2} \sum_{i=1}^N \left(\frac{y_i - \mathbf{F}(\mathbf{x})_i}{\sigma_i} \right)^2 \quad (6.5)$$

$$= -\frac{1}{2} \chi^2 \quad (6.6)$$

where the sum runs over all spectral elements i , σ_i is the i -th measurement error (i.e. the square-root of the diagonal elements of $\mathbf{S}_e$), and we have recognized that this expression is related to the traditional χ^2 (*chi-squared*) metric for “goodness of the fit”. Thus, we see that the likelihood term in Bayes’ Theorem can be intimately related to χ^2 , and “chi-squared minimization” problems are really the same as “maximum likelihood estimation” problems. However, one difference between such optimization problems and the application of Bayesian statistics is the use of priors.

Prior probability distributions must be selected carefully due to the explicit dependence of the posterior upon the chosen distribution. Prior constraints are critical in a remote sensing retrieval problem because there are more unknowns in the surface-atmosphere state than there are constraints in the measured spectrum. Thus, informed priors can be a strength when they can help provide additional constraints to a retrieval problem that is using an observed spectrum with large uncertainties, but priors may also be a weakness if incorrect previous information or misinformed physical notions lead the retrieval away from the correct result. In general, the log-prior probability from Equation 6.3 can be written as

$$\log \mathcal{P}(\mathbf{x}) = -\frac{1}{2} [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a] \quad (6.7)$$

(Line et al., 2014), where $\mathbf{x}_a$ is the *a priori* state vector and $\mathbf{S}_a$ is the *a priori* covariance matrix, which specifies the level of uncertainty in the prior for each state vector element with respect to other elements in the state vector. However, for many exoplanet atmospheric characterization problems, no prior measurements have been made and thus the field is currently ignorant to the environmental state of most exoplanet atmospheres. In such cases, it is traditional to impose uninformative/uniform/flat priors on each parameter, which are finite between two strict boundaries and

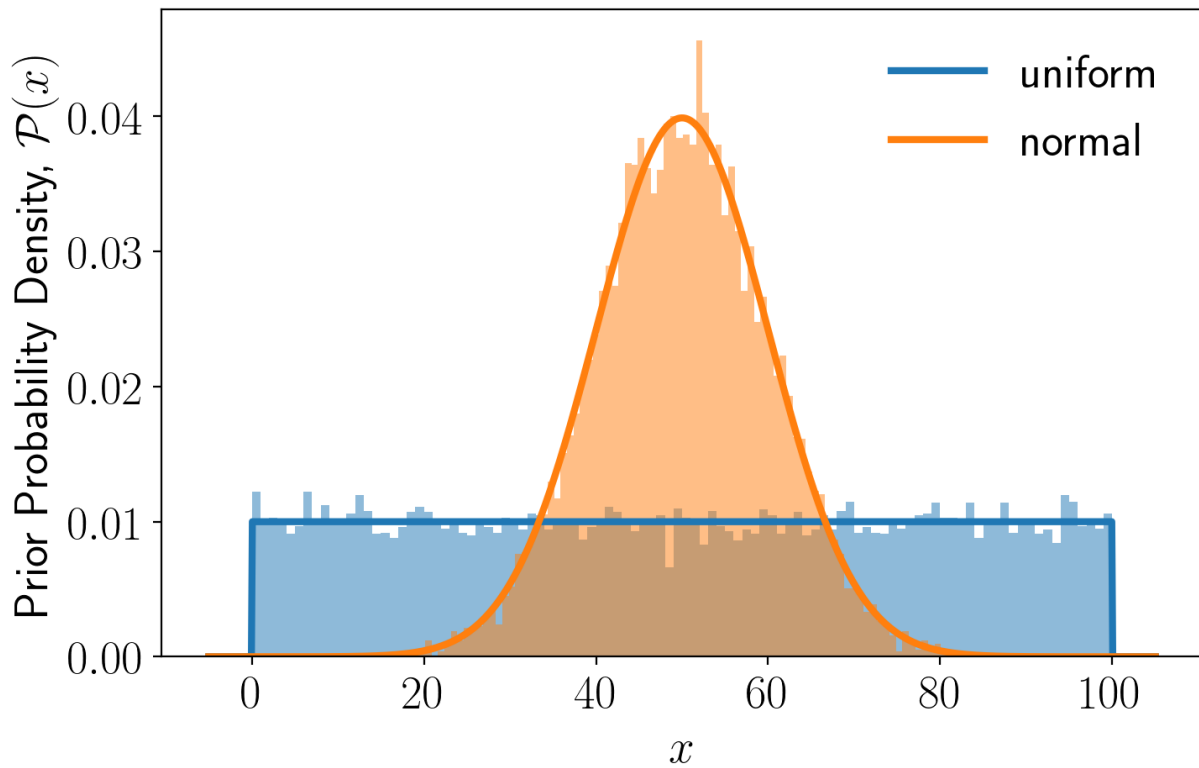


Figure 6.2 Example of uniformly (blue) and normally (orange) distributed priors. The solid lines show the analytic probability density functions for each distribution, while the histograms show 10,000 random samples from each.

zero outside of that range. For instance, the surface albedo of a planet must be between $[0,1]$: all values in that range are allowed at equal probability and values outside that range are disallowed. It may be desirable to impose physical or chemical constraints using the prior to favor or disfavor certain parameter values. However, it is also very common to use Gaussian/normal priors in situations where there have been previous measurements that have been used to constrain any of the free parameters in the state vector. In this case, the results reported from previous investigations can become the priors used in the current investigation. Figure 6.2 visualizes the difference between flat/uniform priors and Gaussian/normal priors for a random variable, x . The SMARTER model currently includes both of these types of priors, but can easily be expanded to support additional distributions.

Equations 6.3 and 6.4 demonstrate that actually calculating a numerical value for the posterior probability fundamentally relies on the ability to run the forward model, $F(\mathbf{x})$. Atmospheric retrievals use radiative transfer models to simulate interactions of light with planets given the physical, chemical, and optical properties of the atmospheric gases, aerosols, and surface to produce a spectrum to compare with observations. The quality of the information retrieved about an exoplanet atmosphere is only as good as the radiative transfer model used as the forward model. We use the Spectral Mapping Atmospheric Radiative Transfer (SMART) model in our forward model, which is a powerful tool for simulating planetary spectra.

6.2.2 *The Spectral Mapping Atmospheric Radiative Transfer (SMART) Model*

SMART is a one-dimensional, line-by-line, multi-stream, multiple-scattering radiative transfer model (developed by D. Crisp; Meadows and Crisp, 1996; Crisp, 1997). The one dimension in SMART refers to the vertical extent of the atmosphere. The “streams” refer to the zenith angles at which radiative transfer through the atmosphere is calculated. SMART’s multi-stream capabilities can extend from two-streams up to at least 64 streams (half upwelling and half downwelling). In exoplanet remote sensing applications, the upward streams are used (observer zenith angles). By default, these are calculated at Gaussian quadrature computational points that are used to solve the equation of transfer using the discrete ordinate method and integrated to get the top of atmosphere (TOA) planetary flux. SMART’s multiple-scattering calculations account for aerosol scattering phase functions, allowing for rigorous treatments of clouds and hazes. At its core, the version of SMART adopted here uses the DIScrete Ordinate Radiative Transfer (DISORT; Stamnes et al., 2017) model to solve the radiative transfer equation⁶.

SMART features a novel modification to solve the equation of radiative transfer that significantly improves computational efficiency over the traditional line-by-line approach. Before performing any radiative transfer calculations, SMART pre-processes all of the input optical proper-

⁶Although, a new version of SMART is currently under development and testing that uses the linearized discrete ordinate radiative transfer (LIDORT; Spurr et al., 2001) method and its vectorized counterpart (VLIDORT; Spurr, 2006).

ties on their native, fully resolved, high-resolution wavenumber grid to identify wavenumbers with sufficiently similar optical properties along the entire optical path (Meadows and Crisp, 1996). SMART then combines similar intervals into a broadband, quasi-monochromatic bin for which it is only required to solve the radiative transfer equation once (West et al., 1990). As a result of this binning procedure, SMART can reduce the number of calls to DISORT by a factor of ~ 10 -1000 depending on the user-specified binning tolerance. After the radiative transfer calculations are completed, SMART maps the spectral results back to the full resolution wavenumber grid, and outputs them for the user.

Relatively recently upgrades to SMART added the ability to perform transmission spectral calculations for transiting exoplanets (Misra et al., 2014; Robinson, 2017a). The transit transmission spectrum calculations may be performed using one of three methods (Robinson, 2017a): (1) a simple geometric approach with light traveling along straight paths, (2) a ray tracing method that includes the refraction of stellar rays passing through the dense lower regions of the planetary atmosphere, and (3) a Monte Carlo numerical integration to account for photon multiple-scattering by aerosols. However, the effects of multiple-scattering on transmission spectra are only significant for planets extremely close to the star (Robinson, 2017a), and the Monte Carlo scattering calculations are extremely computationally expensive, so unless otherwise noted, we typically use the ray tracing method with refraction.

SMART also features the ability to output layer-dependent flux Jacobians and top-of-atmosphere radiance Jacobians. A radiance Jacobian is a matrix whose elements contain the partial derivatives of the top of atmosphere radiance field with respect to each input atmospheric property at every vertical level in the atmosphere. Thus, the top-of-atmosphere radiance Jacobians encode information on how the observable spectrum responds to small changes in the atmospheric state, enabling rapid extrapolation of observed behavior without resorting to a full radiative transfer calculation. SMART allows Jacobians to be computed for the temperature structure, each gas abundance and aerosol optical depth profile, and the surface albedo. This information may be used to increase the speed of a retrieval and/or allow the use of more sophisticated retrieval inverse methods, such as Hamiltonian Monte Carlo (HMC). Transit depth Jacobians (derivatives of the effective absorbing

altitude) are not currently supported by SMART, but are a subject of interest for future work.

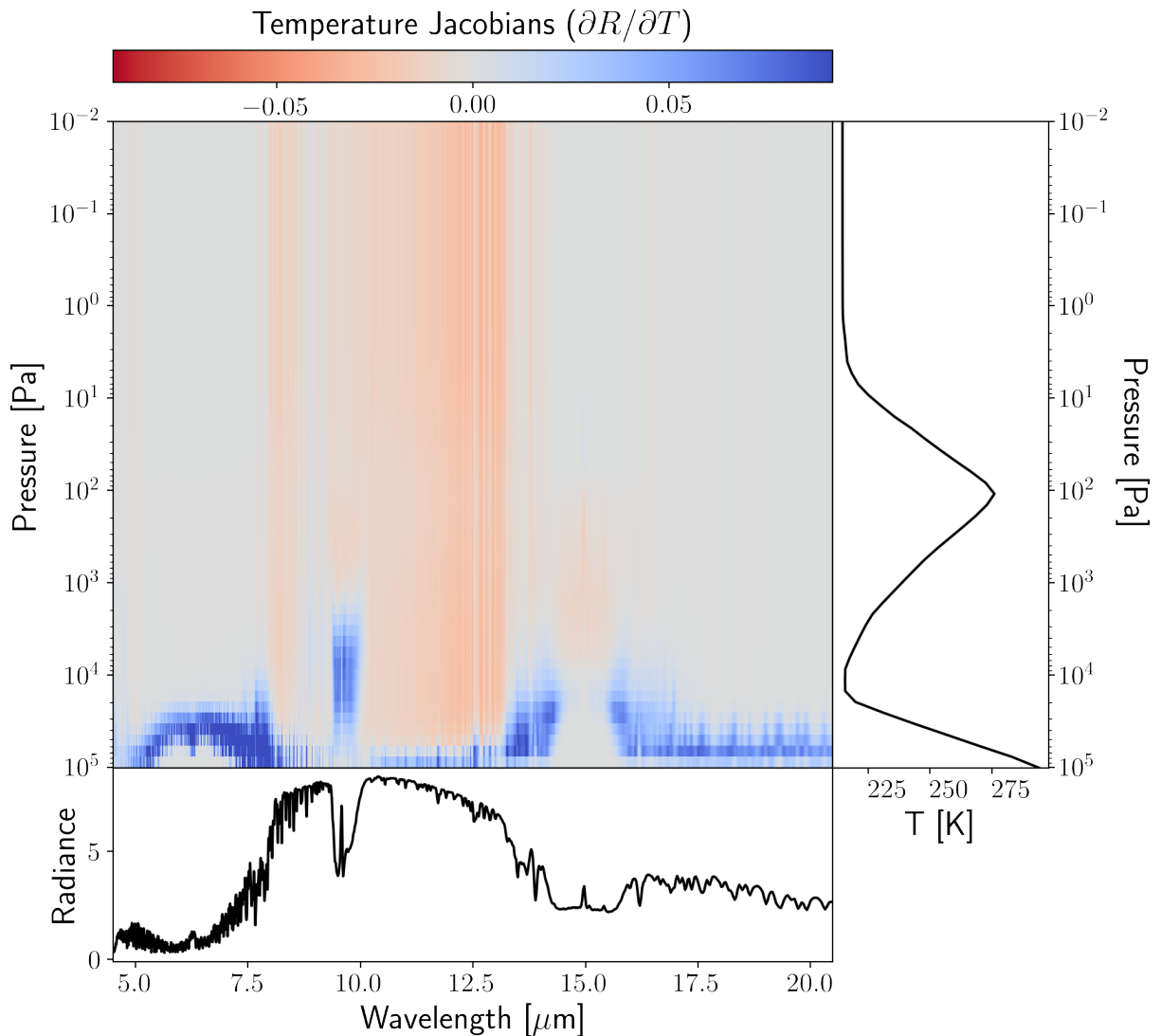


Figure 6.3 Temperature Jacobian for the Earth's thermal emission radiance spectrum. The central panel shows the full Radiance Jacobian matrix as a color contour to emphasize positive (blue) and negative (red) departures from zero. The lower panel shows the radiance spectrum in units of $\text{W/m}^2/\mu\text{m/sr}$, which shares the x-axis with the Jacobian. The right panel shows the temperature-pressure structure of Earth's atmosphere, which shares the y-axis with the Jacobian. Positive/Blue (negative/red) regions of the Jacobian indicate that the radiance will increase (decrease) as the temperature at those atmospheric pressure levels increases. This highlights the atmospheric pressures that shape the observable spectrum.

Figure 6.3 shows the thermal emission radiance Jacobian for the Earth’s atmospheric temperature as a function of both wavelength and pressure. The Jacobian can be used to predict the change in radiance at a given wavelength when the temperature of the atmosphere changes at a given pressure level. For reference, the bottom panel shows Earth’s thermal radiance and the right panel shows Earth’s vertical temperature structure. The radiance Jacobian, shown in the middle, displays how the emitted radiance spectrum will change at each wavelength due to changes in the vertical temperature profile. Visual inspection of the temperature Jacobian provides insight into where in the atmosphere (e.g. which pressure levels) the observable thermal spectrum is sensitive. For example, the blue regions highlight the positive Jacobian—where increases in the temperature structure will lead to increases in the thermal emission. This makes sense in the troposphere where the thermal continuum is emitted. However, note how increasing the temperature of the tropopause will increase the emission within the $9.6\ \mu\text{m}$ ozone band and therefore decrease the overall depth of the ozone absorption feature. These temperature Jacobians are precisely the information needed to expand a planet’s radiance spectrum to first order in temperature, thereby enabling small perturbations in the radiance to be calculated due to small changes in the TP profile, without needing to re-run the radiative transfer model.

SMART was developed for Solar System terrestrial planet atmospheres and has been rigorously validated on the Earth (Robinson et al., 2011) and Venus (Meadows and Crisp, 1996; Arney et al., 2014). SMART has been used to simulate the spectra of a diversity of exoplanet environments, including the spectra of M dwarf habitable planets (Segura et al., 2003, 2005; Meadows et al., 2018a) and uninhabitable planets (Lincowski et al., 2018, 2019), habitable haze-enshrouded exoplanets (Arney et al., 2016, 2017), and high-altitude clouds in the H_2 -dominated mini-Neptune exoplanet, GJ1214b (Charnay et al., 2015b). SMART also functions as the radiative transfer core for the VPL Climate Model, which uses the flux Jacobians to advance radiative time steps without recomputing the radiative transfer (Robinson and Crisp, 2018; Lincowski et al., 2018).

LBLABC

SMART depends on the Line-By-Line ABsorption Coefficient (LBLABC) code to calculate molecular vibrational-rotational transition absorption coefficient input files for the radiative transfer (developed by D. Crisp; [Meadows and Crisp, 1996](#)). LBLABC takes information about the input atmospheric state—the temperature and pressure profile, vertical gas mixing ratio, mean molecular weight of the atmosphere, and radius and surface gravity of the planet—and combines it with line parameter data from HITRAN, HITEMP and other sources, along with isotope information, to calculate the gas absorption coefficients as a function of pressure, temperature, and wavenumber. LBLABC uses nested spectral grids to fully resolve the narrow cores of each absorption line at every atmospheric layer over a very large range of line-center distances. Line contributions are calculated up to 1000 cm^{-1} from the line center for H_2O and CO_2 to capture the full wings of strong lines; all other gases have a 50 cm^{-1} wing cutoff to speed up calculations where contribution from the far wings is negligible. LBLABC employs different line profiles as a function of distance from the line core. LBLABC uses a Voigt line shape ([Humlíček, 1982](#)) within 40 Doppler half-widths of the line center, and a van Vleck-Weisskopf profile ([Van Vleck and Weisskopf, 1945](#)) beyond that for all gases except H_2O , CO_2 , and O_2 . Empirically determined χ correction factors are used in the far wings for H_2O ([Clough et al., 1989](#)) and CO_2 ([Fukabori et al., 1986](#); [Perrin and Hartmann, 1989](#); [Pollack et al., 1993](#)) to reproduce the effects of quantum-mechanical line mixing, a collisional effect that occurs for overlapping lines of a given molecular species ([Rao, 2012](#)). For O_2 , the super-Lorentzian line shape from [Hirono and Nakazawa \(1982\)](#) is used with an exponent of 1.958. LBLABC features the ability to simulate line absorption coefficients for fixed temperature offsets from the given TP profile, which SMART can interpolate between for added flexibility when deviating from an initial temperature structure. This can be used to generate a single set of line absorption coefficients that may be used over a broad range of atmospheric temperature structures that are explored by the retrieval, without requiring LBLABC to be rerun.

6.2.3 *Forward Model*

SMART functions as the radiative transfer workhorse within the SMARTER forward model. This allows the retrieval to use all of the capabilities that SMART has as a forward model to retrieve the properties of terrestrial exoplanets. This is achieved by “wrapping” the retrieval framework around the Python interface for SMART, thereby enabling the user to decide which SMART input parameters will be held fixed, as either quality control parameters or explicit physical assumptions, and which input parameters will be “free” and inferred by the retrieval.

However, due to the ability of SMART to simulate different types of spectra for planets with highly complex atmospheric descriptions, we initially develop multiple baseline retrieval forward models that can each be used for one specific type of observation, and which simplify the physical descriptions of the atmospheric inputs to SMART. The three baseline forward models are for transit transmission spectroscopy, secondary eclipse emission spectroscopy, and directly-imaged reflected spectroscopy. The physical simplifications used in these models are motivated by the lack of precision expected from upcoming terrestrial exoplanet data, but also allow for (1) comparisons with other current retrieval models and (2) an assessment of how/what future exoplanet data motivate adding the physical complexity back into the model. Each of the baseline forward models uses SMART, and each share a common set of procedures (via a base class in Python called the `SMARTForwardModel`) for generating new input states, running SMART, and parsing outputs. The baseline forward models are described next.

Transit Spectroscopy

Transit transmission spectroscopy is the current standard for exoplanet spectral characterization and, as we have demonstrated in Chapters 2 and 3, it will be the only observational technique capable of characterizing the atmospheres of temperate HZ terrestrial exoplanets with *JWST*.

Our baseline transmission spectroscopy forward model is highly flexible in terms of the molecular constituents that can be included in the atmosphere, but it also makes numerous physical simplifications and assumptions. This forward model can include any gas that is in the HITRAN

2016 database. The vertical gas volume mixing ratio profiles can be specified in one of two ways. Either the gases are assumed to be evenly mixed in the atmospheric column, in which case there is one free parameter associated with each gas: the evenly-mixed volume mixing ratio. Alternatively, the user may provide a vertical “template profile” and a factor with which to scale the profile. The temperature profile has the same two options: isothermal or scaled profile. The surface pressure can be set fixed or included as a free parameter in the state vector.

The user may decide to employ a “filler gas” to occupy the remaining volume of the atmosphere after all of the other free-parameter gases are specified. That is, the VMR of the filler gas is given by,

$$f_{\text{fill}} = 1 - \sum_{i=1}^N f_i \quad (6.8)$$

where f_i is the VMR of the i th gas, N is the total number of non-filler gases included in the atmosphere, and we have assumed that these constitute the entire volume of the atmosphere. This implicitly enforces a common prior that is used in atmospheric retrievals, which specifies that the VMR of all gases must not sum to greater than unity. Any molecule that can be used as a free parameter can alternatively be used as the filler gas, however, typically N_2 is used, although Ar and CO_2 are also reasonable choices that a user may select. The filler gas must be set prior to starting the retrieval and cannot be changed dynamically during the retrieval.

The mean molecular weight of the atmosphere can be included as a fitting parameter, or it may be calculated self-consistently. If the mean molecular weight is included in the state vector, then each evaluation of the forward model must specify it explicitly. If the user opts to calculate the mean molecular weight self-consistently, then it is determined during each call to the forward model (after the filler gas is added) as the pressure weighted mean molecular weight,

$$\mu = \frac{\sum_j^{N_P} P_j (\sum_i^{N_{\text{gas}}} m_i f_{i,j})_j}{\sum_j^{N_P} P_j}, \quad (6.9)$$

where P_j is the j th of N_P total pressure layers in the atmosphere, m_i is the molecular mass of the i th of N_{gas} in the atmosphere, and $f_{i,j}$ is the volume mixing ratio of the i th gas at the j th pres-

sure. Although the mean molecular weight is physically a function of altitude/pressure, SMART currently requires that it be specified as a constant throughout the vertical column. The pressure-weighting in Equation 6.9 biases towards the higher pressure, lower atmosphere, rather than the thin upper atmosphere. However, for evenly mixed atmospheres, these pressure-weighting subtleties have no effect. The use of a filler gas enables the mean molecular weight to be calculated self-consistently, as we will explore in greater detail in Chapter 7.

In addition to the atmosphere, the transit forward model also allows for the planet radius and planet mass to be included in the state vector. Although the planet radius is measured by the transmission spectrum, the radius included in the transit forward model is the planet’s solid body/surface radius, specified in units of Earth-radii. This surface radius corresponds to the surface pressure (whether or not it is fixed or free), and represents the “zero point” above which the atmospheric absorption extends radially outward. The planet mass is not an easily measured quantity for a transiting terrestrial exoplanet (as it requires either extremely precise RV or TTVs), nor is it easy to constrain from transmission spectra. However, the planet mass does affect the scale height of the atmosphere, via the surface gravity, and is therefore degenerate with both the temperature structure and the mean molecular weight of the atmosphere. This allows mass to vertically increase or decrease the scale of all absorption features in the transmission spectrum. In the transit forward model state vector, planet mass is specified in units of Earth-mass and it is converted to surface gravity before passed to SMART for the radiative transfer. Note that we do not make a distinction between the mass contained within the solid body of the planet versus the mass of the entire planet, due to the negligible mass of the atmosphere for Earth-like⁷ planets (e.g. the atmosphere makes up less than 1 ppm of Earth’s mass) compared to realistic constraints on terrestrial exoplanet masses.

The user can select which transit geometry model from [Robinson et al. \(2016\)](#) to use, enabling different degrees of physical rigor. The fastest model uses geometric optics, and tends to be the starting point for retrievals. The ray tracing model with atmospheric refraction is also computationally efficient, and although it is still under investigation (see Chapter 7), we encourage its use.

⁷We also note that gaseous planets have almost all of their mass below the $\tau = 1$ surface.

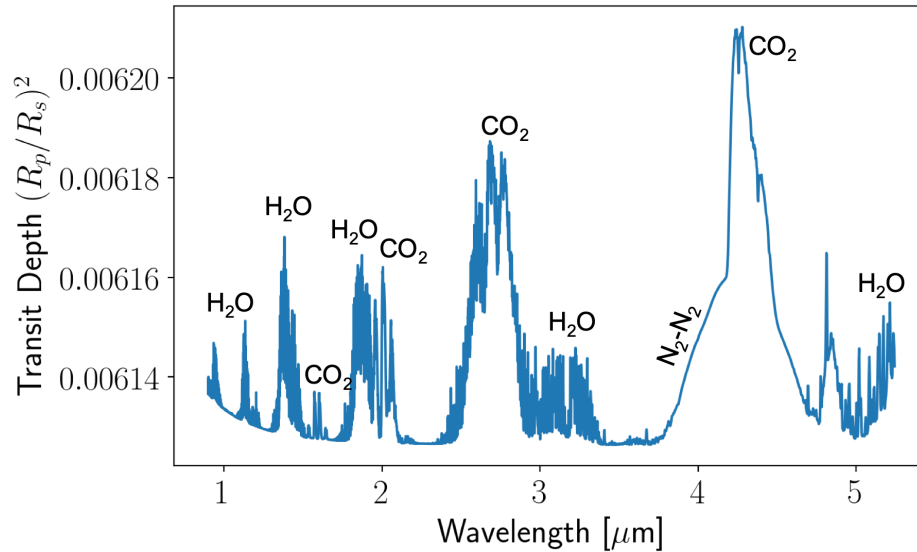


Figure 6.4 Example transmission spectrum output from the SMARTER transit forward model. This spectrum corresponds to an Earth-size planet transiting a TRAPPIST-1-size star. The planet atmosphere is modeled with a 200 K isothermal TP profile with an evenly mixed atmospheric column with 10 ppm H_2O , 100 ppm CO_2 , and the remaining bulk atmosphere composed of N_2 .

The Monte Carlo forward scattering transit model is extremely slow, but it can in practice be used to investigate how forward scattering aerosols affect the retrieval of gas and aerosol properties.

Outputs from the transit forward model are provided either in units of the fractional transit depth, $(R_p/R_s)^2$, or in kilometers for the effective transit altitude. Figure 6.4 shows an example output spectrum from the transit forward model. The atmosphere for this example is extremely simplistic, containing only trace H_2O and CO_2 with N_2 as the filler gas, but it serves as the basic introduction to the SMARTER transit code for new users. Despite the simplicity, the spectrum still shows CO_2 and H_2O absorption features not unlike those seen in the Earth’s atmosphere. The N_2 – N_2 CIA feature is also visible in the spectrum (e.g. [Schwieterman et al., 2015b](#)).

Emission Spectroscopy

Secondary eclipse thermal emission spectroscopy is expected to be a promising method to characterize hot terrestrial exoplanets with *JWST* and is an important forward model component to

SMARTER. The emission forward model operates very similarly to the transit forward model, and makes many of the same assumptions regarding the atmospheric composition. The emission forward model also allows the planet mass and radius to be fixed or retrieved. The main difference is that the emission forward model allows for significantly more flexibility in terms of the TP profile, for which the user can provide a custom profile that can remain fixed or be scaled by a retrievable factor, or a TP profile parameterization can be used.

The first option for the TP profile parameterization is to use Poisson’s adiabatic equation of state (Catling and Kasting, 2017). The adiabatic temperature profile in the convective region of the atmosphere is taken to be

$$T = T_0 \left(\frac{P}{P_0} \right)^{\alpha(\gamma-1)/\gamma}, \quad (6.10)$$

where P_0 is a reference pressure, T_0 is the temperature at the reference pressure, γ is the ratio of specific heats (c_p/c_v), and α is a control parameter to modify the adiabat due to latent heat release or non-constant specific heats (Sagan, 1962; Robinson and Catling, 2012). For a dry adiabat $\alpha = 1$, but typically varies between 0.6-0.9 and represents the average ratio of the true lapse rate to the dry adiabatic lapse rate in the planet’s convective region (Catling and Kasting, 2017). In addition to the parameters in Equation 6.10, we also define the tropopause pressure, P_1 , and set the TP profile to be isothermal for pressures (altitudes) lower than (above) this. These five parameters are used to specify an adiabatic TP.

Figure 6.5 shows an example input adiabatic TP profile and the resulting mid-IR emission spectrum for an Earth-like planet around TRAPPIST-1 observed in secondary eclipse. The continuum thermal radiation emanates from near the surface (dashed purple line). As the thermal emission leaves the atmosphere, it is absorbed by molecules—notably O_3 at $9.6 \mu\text{m}$ and CO_2 at $15 \mu\text{m}$ —in the cooler stratosphere, such that the flux at the bottom of the absorption bands corresponds to the temperature in the atmosphere where the absorption predominately occurred (dashed cyan line). Inversely, this aspect of thermal radiation makes emission spectroscopy sensitive to retrieving the TP profile. However, this TP profile lacks the ability to model a stratospheric temperature inversion, which gives Earth a distinctive emission spike at the center of the $15 \mu\text{m}$ CO_2 band, and is a

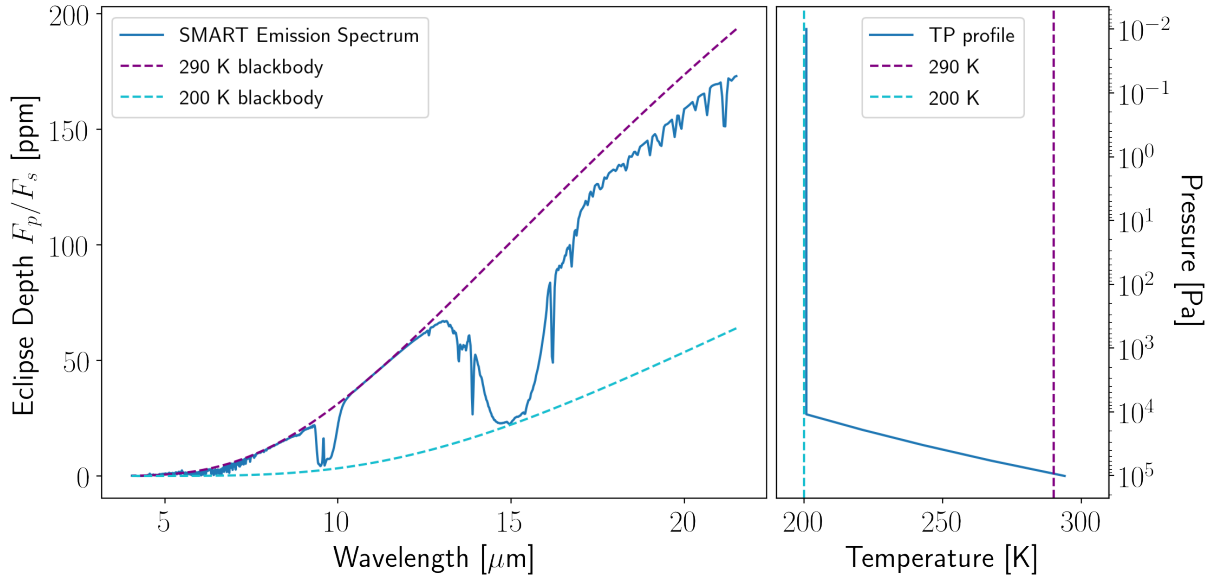


Figure 6.5 Example emission spectrum output from the SMARTER emission forward model. This spectrum corresponds to an Earth-size planet eclipsed by TRAPPIST-1. The planet atmosphere is modeled with an adiabatic tropospheric TP with an isothermal stratosphere and with an evenly mixed atmospheric column with 1 ppm O_3 , 10 ppm H_2O , 100 ppm CO_2 , and the remaining bulk atmosphere composed of N_2 . The dashed lines correspond to isothermal blackbody spectra at the surface (purple) and stratospheric (cyan) temperatures.

major caveat when using this TP parameterization.

The second TP profile parameterization is a three-channel, analytic radiative equilibrium model that is based off of the work of [Guillot \(2010\)](#), derived formally by [Parmentier and Guillot \(2014\)](#), and is used heavily in the atmospheric retrieval community ([Line et al., 2013](#); [Feng and Robinson, 2017](#); [Taylor et al., 2020](#)). The atmospheric temperature structure for a plane-parallel atmosphere in local thermodynamic equilibrium that receives incident stellar flux from above and internal heat flux from below, can be written as a function of the optical depth τ ,

$$T^4(\tau) = \frac{3T_{\text{int}}^4}{4} \left(\frac{2}{3} + \tau \right) + \frac{3T_{\text{irr}}^4}{4} (1 - \alpha) \zeta_{\gamma_1}(\tau) + \frac{3T_{\text{irr}}^4}{4} \alpha \zeta_{\gamma_2}(\tau), \quad (6.11)$$

with,

$$\zeta_{\gamma_i} = \frac{2}{3} + \frac{2}{3\gamma_i} \left[1 + \left(\frac{\gamma_i \tau}{2} - 1 \right) e^{(-\gamma_i \tau)} \right] + \frac{2\gamma_i}{3} \left(1 - \frac{\tau^2}{2} \right) E_2(\gamma_i \tau), \quad (6.12)$$

where γ_1 and γ_2 are ratios of the mean opacity in the two visible bands, respectively, to the mean IR opacity, α partitions the incident flux between the two visible streams, $E_2(\gamma_i \tau)$ is the second-order exponential integral function, T_{int} is the internal temperature, and T_{irr} is the irradiation temperature for the stellar flux at the top of atmosphere. We can express the irradiation temperature in terms of the stellar radius R_s , the stellar effective temperature T_s , and the semi-major axis of the planet a ,

$$T_{\text{irr}} = \beta \left(\frac{R_s}{2a} \right)^{1/2} T_s, \quad (6.13)$$

where β is a factor of order unity that represents the planet albedo, emissivity, and day-night heat redistribution—all the factors which may act to make the TP profile as hotter or colder. Equations 6.11, 6.12, and 6.13 can be related back to the atmospheric pressure by invoking hydrostatic equilibrium and assuming a linear relationship between pressure and the optical depth,

$$\tau = \frac{\kappa_{\text{IR}} P}{g}, \quad (6.14)$$

where κ_{IR} is the IR opacity and g is the gravity. Together, Equations 6.11, 6.12, 6.13, and 6.14 can be used to construct a TP profile using five free parameters: γ_1 , γ_2 , κ_{IR} , α , and β .

Other TP profiles, such as the parameterization used by [Madhusudhan and Seager \(2009\)](#) (see also [Barstow and Heng, 2020](#)), could also be implemented.

Reflectance Spectroscopy

Reflected-light spectroscopy stands to be one of the most promising future avenues for rocky exoplanet characterization, as discussed in Chapters 4 and 5, due to its sensitivity to the lower atmosphere and surface of transiting and non-transiting planets orbiting a diversity of stellar types, including Sun-like stars. Despite the fact that no current telescopes can directly observe light reflected off an exoplanet, the future telescopes LUVOIR ([LUVOIR Mission Concept Study Team](#),

2019) or HabEx (HabEx Study Team, 2019) would be designed to do just this for Earth-like exoplanets. We have developed a baseline reflectance spectroscopy forward model for SMARTER that can be used to assess science cases for these next-generation telescope concepts.

The SMARTER reflectance spectroscopy forward model builds off of the functionality introduced for the transmission and emission models. We adopt the same TP profile parameterizations that were previously discussed for emission and transmission spectroscopy. That is, users can choose to use a TP profile that is isothermal (e.g. Feng et al., 2018), a modified adiabatic model (e.g. Robinson and Catling, 2012), the analytic radiative-convective equilibrium model from Parmentier and Guillot (2014), or a fixed profile specified by the user.

For the surface reflectance, our baseline model uses a grey geometric surface albedo that assumes phase-independent Lambertian scattering. Users may instead specify a fixed surface albedo that may contain wavelength dependent information. In future work we will investigate optimal parameterizations for allowing a fully agnostic wavelength-dependent surface albedo to be retrieved. Following Feng et al. (2018), we fit for the $\log_{10}$ of the grey surface albedo, which implicitly specifies an *a priori* preference for low albedo surfaces when we use a uniform prior in this logged parameter space. This is generally consistent with the concept that planetary surfaces are more likely to be poor reflectors than efficient reflectors, but we caution that this approach may have consequences in the posteriors. Figure 6.6 shows the uniform prior on the log of the grey surface albedo, transformed to linear space where we can see the preference for low albedo surfaces.

We use a similar cloud description to that of Feng et al. (2018). The cloud model includes four parameters that can be retrieved: three that specify the optical depth profile and one that controls the cloud fraction. We construct a cloud optical depth profile using the cloud-top pressure (P_{cld}), the pressure width over which the cloud extends (ΔP_{cld}), and the total cloud optical depth (τ_{cld}) between P_{cld} and $P_{cld} + \Delta P_{cld}$. The cloud fraction (f_{cld}) is used to create an observed reflectance spectrum that is a linear combination between the fully cloudy spectrum and a fully clear spectrum where the cloud has been removed. When the cloud fraction parameter is used, the SMARTER reflectance forward model must call SMART twice, once with the cloud and once without. This

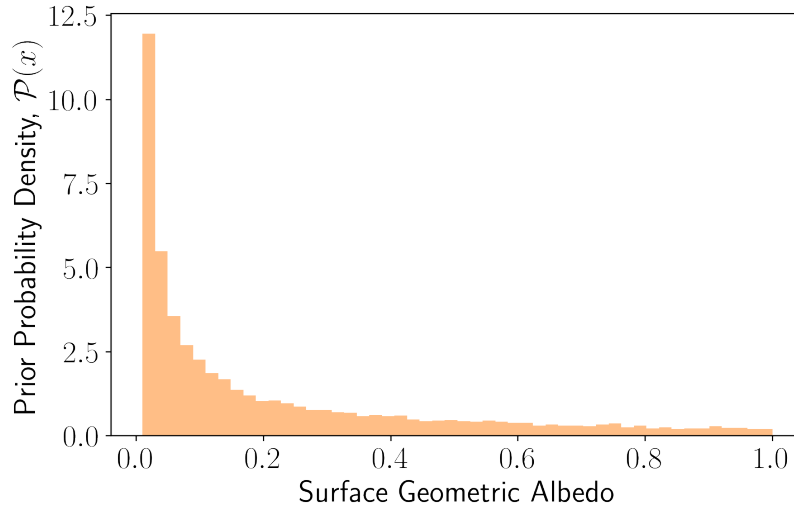


Figure 6.6 Grey surface albedo prior for the reflectance spectroscopy forward model. Using a uniform prior on the $\log_{10}$ surface albedo, exerts a bias towards lower albedo surfaces.

imposes an additional computational expense for the model.

At present, only optical properties for water clouds are included in the SMARTER reflectance forward model, but either cirrus clouds or stratocumulus clouds may be used. For cirrus (water-ice) clouds, we use the optical properties from B. Baum’s Cirrus Optical Property Library (Baum et al., 2005), with a combination of 45% solid columns, 35% plates, and 15% 3D bullet rosettes, and a mean particle diameter of $100 \mu\text{m}$. For stratocumulus (water-liquid) clouds, we use refractive indices from Hale and Querry (1973) to calculate optical properties from Mie scattering theory assuming a two-particle gamma distribution in particle size with $a = 5.33$, $b = 1.18$, and a mean radius of $4.07 \mu\text{m}$. These two water cloud types have already been validated against Earth observations using SMART as part of the VPL Earth Model (Robinson et al., 2010, 2011, 2014a). However, we note that by using optical properties for cirrus and stratocumulus clouds, our cloud model differs from that of Feng et al. (2018), which used a minimally parametric grey cloud model with a fixed single scattering albedo $\bar{\omega} = 1$ and asymmetry parameter $\bar{g} = 0.85$. Thus, our cloud model and the Feng et al. (2018) cloud model both use the same four parameters to describe water clouds, but our approach may scale more realistically into the NIR where the water clouds begin

to show non-negligible absorption.

Our baseline direct imaging reflectance retrieval model uses a simplified viewing geometry. First, we assume a single solar zenith angle of 60° to approximate the quadrature phase viewing geometry. Users may modify the fixed SZA to approximate other phase angles, but the current model does not allow the phase angle to be a free parameter. We direct readers to [Nayak et al. \(2017\)](#) for a thorough assessment of retrieval degeneracies between planet radius, albedo, and phase. Second, we nominally perform 8-stream radiative transfer, which provides four upwelling radiance streams that are integrated in quadrature to produce a TOA planet flux spectrum. Users may also opt to increase (or decrease) the number of streams. The TOA planet flux is divided by the stellar flux incident at the TOA and scaled by geometric factors to calculate the planet-to-star flux ratio,

$$\frac{F_p}{F_s} = \underbrace{\Phi(\alpha) \frac{F_{p,TOA}}{F_{s,TOA}}}_{\text{reflectance}} \left(\frac{R_p}{a} \right)^2, \quad (6.15)$$

where R_p is the planet radius, a is the planet semi-major axis, and $\Phi(\alpha)$ is the planet phase function, which we assume to be a fixed factor of $1/\pi$ at quadrature for Lambertian scattering ([Robinson et al., 2016](#); [Feng et al., 2018](#)). By default, the SMARTER reflectance forward model returns outputs in the form shown in Equation 6.15, which is most applicable to exoplanet observations. However, the model can instead return the reflectance, as highlighted in Equation 6.15, which is more applicable to Solar System remote sensing observations.

We demonstrate the SMARTER reflectance retrieval model by comparing it with the high fidelity Earth spectrum generated using the VPL's 3D Spectral Earth Model ([Robinson et al., 2011](#)). Figure 6.7 shows the spectrum of the Earth from the validated VPL Earth Model alongside our fit to it using the SMARTER reflectance model. The VPL Earth model spectrum was generated by [Robinson et al. \(2011\)](#) using spatially- and vertically-resolved satellite data for the gas mixing ratios, temperature structures, and cloud coverage and optical depths, as inputs to the SMART radiative transfer. The SMART outputs were then disk integrated to produce high resolution, phase dependent spectra of the Earth. As a result of this rigor, the VPL Earth model spectrum is taken to

Table 6.1. Parameters used in the fit to the VPL Earth Model spectrum shown in Figure 6.7

Parameter	Description	Retrieved Value	Feng et al. (fiducial)
T_{iso}	isothermal temperature (K)	255	250
$\log P_0$	surface pressure (bar)	$\log(1.075)$	$\log(1)$
$\log \text{H}_2\text{O}$	water VMR	$\log(2.85 \times 10^{-3})$	$\log(3 \times 10^{-3})$
$\log \text{O}_3$	ozone VMR	$\log(7.88 \times 10^{-7})$	$\log(7 \times 10^{-7})$
$\log \text{O}_2$	oxygen VMR	$\log(0.185)$	$\log(0.21)$
$\log \text{CO}_2$	carbon dioxide VMR	$\log(5.90 \times 10^{-3})$	—
R_p	planet radius ($R_\oplus$)	1.02	1
M_p	planet mass ($M_\oplus$)	1.09	1
$\log A_s$	surface albedo	$\log(0.0455)$	$\log(0.05)$
$\log P_{cld}$	cloud-top pressure (bar)	$\log(0.575)$	$\log(0.6)$
$\log \Delta P_{cld}$	cloud thickness (bar)	$\log(0.0952)$	$\log(0.1)$
$\log \tau_{cld}$	cloud optical depth	$\log(14.5)$	$\log(10)$
f_{cld}	cloud coverage	0.44	$\log(0.5)$

Note. — $\log$ refers to $\log_{10}$

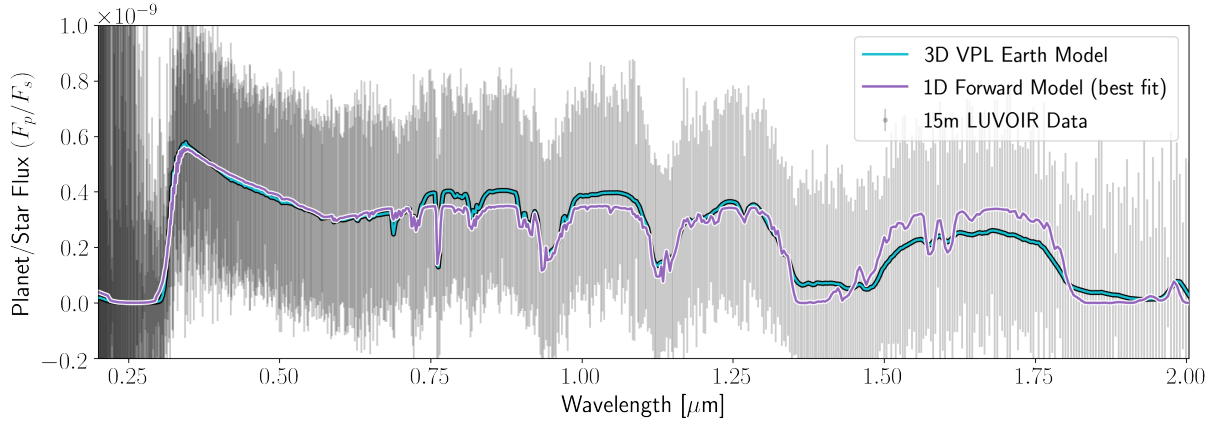


Figure 6.7 Earth’s reflectance spectrum from the 3D VPL Earth Model (cyan line) compared to the SMARTER 1D retrieval forward model (purple line). The faint error bars approximate the observational uncertainty on the VPL Earth Model spectrum that may be achieved by the 15m LUVOIR telescope concept. The 1D forward model was calculated using best fitting parameters from a retrieval to the synthetic LUVOIR data that was initialized using the fiducial parameters from [Feng et al. \(2018\)](#), see their Table 1 and Figure 4 for reference).

be representative of the “true” Earth. We calculated the data uncertainty that may be expected for 15-meter LUVOIR observations of the Earth around the Sun at a distance of 10 pc. We opted to use a higher-than-nominal spectral resolution of $\mathcal{R} = 400$ for purely aesthetic purposes. We then performed a simple non-linear, least-squares fit to the synthetic Earth data using the SMARTER reflected-light model, which we initialized at the fiducial planetary state given in Table 1 of [Feng et al. \(2018\)](#). The parameters used in our fit, and their initial and final values are provided in Table 6.1. Our best-fitting state is only a small perturbation from the [Feng et al. \(2018\)](#) state, indicating good agreement between both our model and the [Feng et al. \(2018\)](#) model, and our model and the [Robinson et al. \(2011\)](#) Earth Model. The discrepancies seen in Fig. 6.7 between the VPL Earth Model and the SMARTER fit are small relative to the approximated exoplanet-analog noise, indicating that this baseline model is sufficiently complex to study the capabilities of a LUVOIR-like mission. However, the residuals occur predominately in the spectral continuum, which highlights the shortcomings of our assumed grey surface albedo model compared to the wavelength-dependent surface albedo of the real Earth.

Instrument Models

The forward RT model generates spectra at the resolution needed to capture the wavelength dependence of the source (stellar spectra, blackbody) as well as the optical properties of the surface and atmosphere. However, most realistic instruments measure the transmitted, scattered or emitted radiation at a lower spectral resolution. The instrument model convolves the full-resolution synthetic spectrum generated by the forward RT model with the instrument spectral response function, to simulate the instrument’s spectral resolution and sampling. The instrument model can also simulate the polarimetric response of the instrument for polarized calculations⁸. This generates a convolved synthetic spectrum that can be compared directly to the observed spectrum.

Within SMARTER, a base `Instrument` class is used to implement multiple different instrument modules that can be swapped between for different retrieval use-cases. By default, SMARTER uses the simple “top hat” convolution used within the CHIMERA retrieval (Line et al., 2013). This convolves the high resolution spectrum with a rectangle, thereby applying equal weight to all high-res spectral points within each low-resolution bin interval. SMARTER will automatically determine the low resolution bin widths by assuming them to be the spacing between each wavelength point in the dataset.

We have also implemented a straightforward linear interpolation instrument model, which interpolates the high resolution model outputs onto the low resolution data grid. This approach is invalid on its own, but is appropriate to use in tandem with the convolution slit functions built into SMART. For instance, SMART typically outputs spectra after convolving the internal spectrum with a top hat or triangular slit function with a user-specified FWHM, and then writes spectral quantities to a file at a user-specified wavenumber sampling resolution. In this way, SMART can be provided the exact spectrograph FWHM, and then output spectra at, or slightly higher than, the observed sampling resolution of the data. Then, all that is needed to compare with the data is to linearly interpolate the model spectrum onto the data spectral grid. We find this instrument model particularly useful for high resolution data, as will be discussed further in Section 6.3 for Earth

⁸However, polarization is not considered in this work.

validation. Note that more sophisticated (or realistic) instrument spectral response functions can be implemented as a subclass to the `SMARTER Instrument` and applied after the fact by convolving a high resolution spectrum (with or without an internal slit function) from `SMART` with a custom instrument line shape (ILS).

6.2.4 Inverse Model

We employ a multi-pronged inverse modeling approach similar to that of the CHIMERA retrieval model (Line et al., 2013, 2014; Kreidberg et al., 2015; Feng and Robinson, 2017; Tremblay et al., 2020). Each inverse model attempts to solve Bayes’ Theorem to characterize the shape of the posterior distribution. However, each method makes different assumptions and differ in their numerical approach, which affects both the statistical robustness of the retrieved posteriors and the number of calls to the forward model required to reach convergence. As a result, four inverse models are currently available and can be used with any forward model within the `SMARTER` framework. These retrieval models are: Optimal Estimation (OE), Nested Sampling (NS), Markov Chain Monte Carlo (MCMC), and a novel machine learning approach called *approxposterior* (AP). A typical use-case may involve using OE to find a single best-fitting state (MAP solution) and then initializing MCMC or AP using this initial solution to explore the full shape of the posterior distribution starting in a useful/informed region of parameter space. We describe each of these retrieval approaches next and outline under what circumstances they are typically best used.

Optimal Estimation

Optimal Estimation has a long history of usage within the Earth and Solar System remote sensing field (Rodgers, 1976, 2000; Irwin et al., 2008). OE attempts to find the *maximum a posteriori* (MAP) solution by minimizing the cost function, or the negative log-probability function given in Equation 6.3. This is very similar to “ χ^2 minimization”, which yields *maximum likelihood estimates*, except the use of priors makes it a MAP estimate. OE uses local evaluations of the cost function’s gradient to iterate towards regions of parameter space that are higher likelihood, thereby

minimizing the cost function and finding the best fit solution. OE makes the strong assumption that the posterior for each parameter is Gaussian, which need not be physical. This assumption allows it to converge in relatively few (10s to 100s) forward models evaluations, but it can bias the shape of the posterior distribution (e.g. in cases where only an upper or lower limit on a parameter can be constrained). The OE approach of minimizing the cost function is subject to converging to local minima, making it highly sensitive to the initial conditions of the retrieval.

There are many numerical approaches that can perform OE. In SMARTER, we implement OE using the flexible `scipy.optimize.minimize()` function within the SciPy Python package (Virtanen et al., 2019). This allows users to access SciPy’s entire suite of methods that can be used to minimize a scalar function, in this case, the negative log-probability: $-\log \mathcal{P}(\mathbf{x}|\mathbf{y})$. By default, we use the “Nelder-Mead” algorithm (Nelder and Mead, 1965; Wright, 1996; Gao and Han, 2012), which is a direct search method for solving non-linear optimization problems, although we find similar results using other methods. Due to the tendency of OE approaches to find local minima, we typically run many optimizations in parallel using randomized initial states. We then take the best fit among the ensemble as the global MAP solution. This approach can quickly estimate the parameter values that best fit the spectrum, but it fails to provide robust measures of the uncertainty on each retrieved parameter, so we rarely employ only OE for science cases.

Figure 6.8 shows example results from an OE retrieval of synthetic data from the transit spectrum shown in Fig. 6.4. For this case, 28 randomly initialized Nelder-Mead optimizations were performed in parallel, with the resulting best-fit spectra and retrieved parameter states displayed in Fig. 6.8. The global MAP solution (darkest blue line/points) is visibly a very good fit to the noisy spectrum and it gets quite close to the “true” input state. However, there is generally a large spread among the best fitting spectra and the corresponding parameter states. This emphasizes why we spawn many OE retrievals in parallel from random initial starting positions. It may seem tempting to interpret the spread in best fitting spectra and retrieved parameters as relating to the uncertainty on the data points and inferred parameters, but this is not formally the case. Instead, we must use other techniques that properly account for uncertainties in parameter estimations.

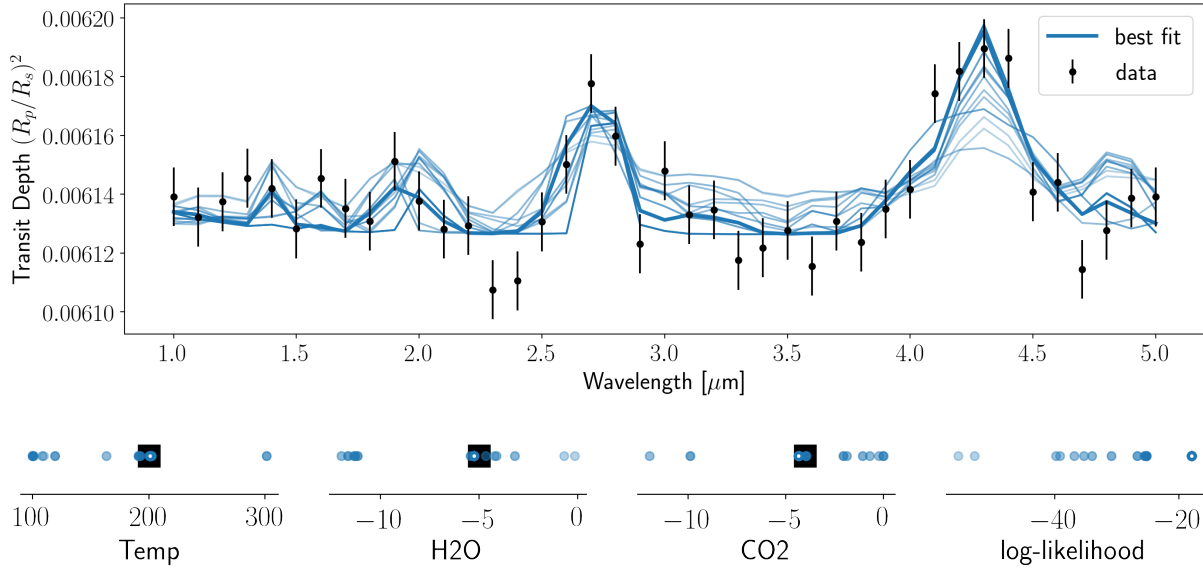


Figure 6.8 Optimal estimation retrieval results for the transit forward model example shown in Fig. 6.4. The top panel shows the best fitting spectra from 28 randomly initialized Nelder-Mead optimizations. The bottom panels show the corresponding best fit parameter states, with the “true” state given by the black square, and the global MAP solution is the darkest blue point and is indicated with the white central dot. Both panels are color-coded according to the log-likelihood, which is shown in the bottom right panel.

Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) has become the gold standard for Bayesian inference in astrophysics over the last 15 years due to its ability to directly sample high-dimensional posterior distributions, thereby facilitating the propagation of measurement uncertainties through complex physical models and enabling the assessment of parameter covariances. The SMARTER model uses the affine-invariant MCMC ensemble sampler originally proposed by [Goodman and Weare \(2010\)](#) and developed into the stable, well-supported, and easy to use Python package, EMCEE ([Foreman-Mackey et al., 2013](#)).

Briefly, the EMCEE algorithm works by using K “walkers” to execute a random walk to explore the N -dimensional parameter volume, a process that builds up representative samples from the posterior distribution over time. At each iteration, each walker has a proposed update to their position

in parameter space (state vector), which is based on the location of all other walkers relative to that walker’s current state. Proposal steps that move to regions of parameter space that are higher probability (better fit to the data) are accepted as the next position in the “chain”, however proposals that move to lower probability are either rejected or accepted based on the relative decrease in probability weighed against a randomly drawn number. In this way, the algorithm maintains a detailed balance between accepting and rejecting proposed steps such that the ultimate record of the chain is a set of samples from the unnormalized posterior distribution. We refer the reader to [Foreman-Mackey et al. \(2013\)](#) and [Goodman and Weare \(2010\)](#) for a complete description of the method.

Figure 6.9 shows an example “trace plot” from the same synthetic data shown in Fig. 6.4 and retrieved with OE in Fig. 6.8. Trace plots are used to diagnose convergence for MCMC retrievals because they show the evolution of the ensemble of walkers (color lines) as a function of iteration (or MCMC step) for each retrieval parameter. Initially (at iteration 0), the walkers are dispersed randomly across the entire parameter space, as seen on the far left of each sub-panel. As the MCMC progresses, the walkers quickly constrict into a tighter mass as they random walk into the high-probability region of parameter space. This initial constriction of MCMC walkers is referred to as the “burn-in,” beyond which the chains no longer possess information on their initial conditions. As a result, the burn-in is discarded in all subsequent analyses. Beyond the burn-in, we take each step in the MCMC chain to be a true sample from the unnormalized posterior distribution, such that the density of their random walks is proportional to the posterior probability density. The right column of subplots Fig. 6.9 shows the MCMC chain collapsed along the step-dimension into histograms of the most-likely parameter values. The histograms, shown for each parameter, are integrated over the values of all other parameters in the state vector, and are referred to as *marginalized posterior distributions*⁹. However, trace plots do not display the covariant relationships between multiple parameters, for which we will require another type of plot.

Figure 6.10 shows a “corner plot” from the same MCMC retrieval as Fig. 6.9. Like the trace

⁹This is not strictly correct because this plot example *includes* the burn-in for demonstration purposes, and is therefore not formally proportional to posterior density.

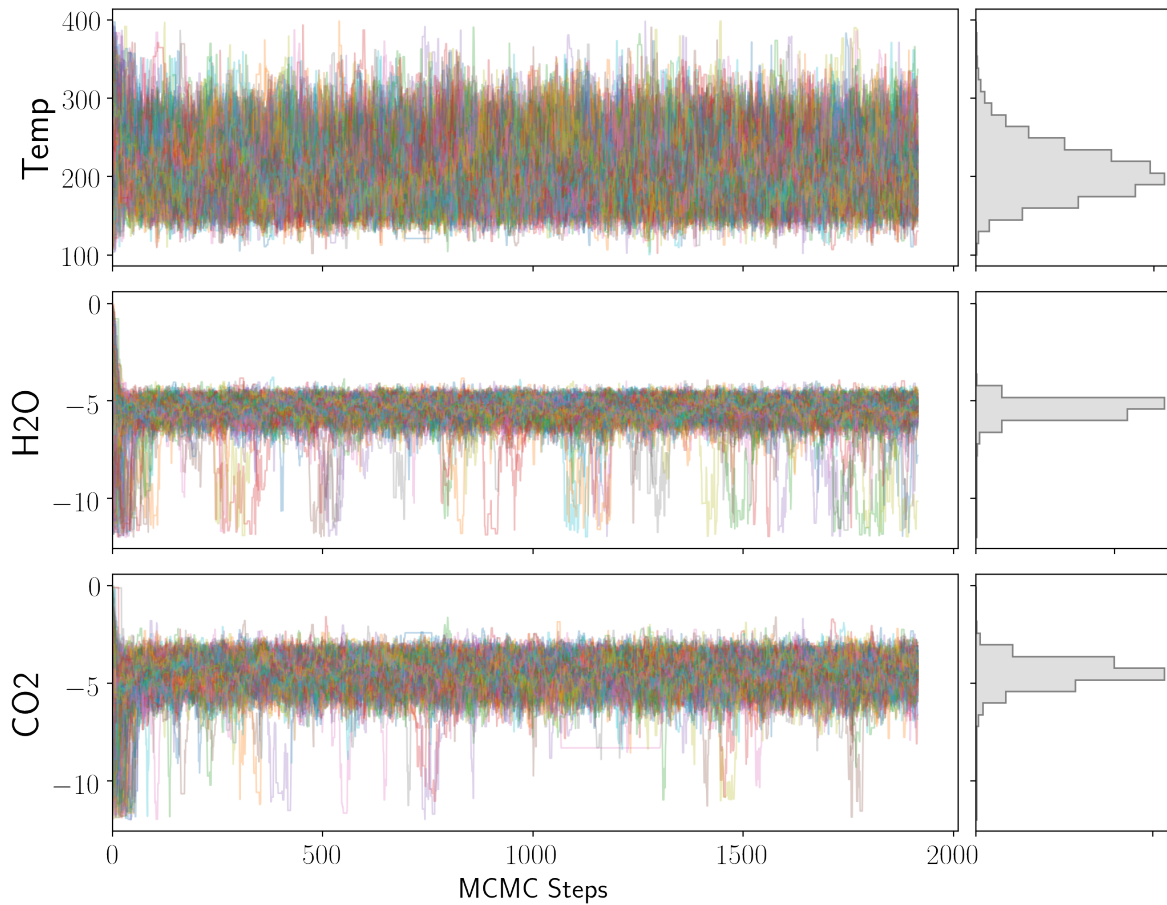


Figure 6.9 MCMC retrieval “trace plot” example showing the evolution of the ensemble of walkers (color lines) as a function of iteration for each retrieval parameter (sub-panel rows). The right column of subplots shows the walker positions, collapsed along the step-dimension into histograms of the most-likely parameter values. This type of plot is not typically provided in published papers, but is usually produced to aid in determining whether an MCMC simulation has converged.

plot, the subplots along the diagonal show the marginalized posterior distributions for each retrieved parameter. However in a corner plot, the off-diagonal subplots show the covariances between pairs of parameters. For example, the bottom left subplot shows the covariance between CO_2 and the isothermal temperature, which displays a negative correlation. If the temperature increases (decreases), then the CO_2 abundance can decrease (increase) to maintain high probability fits to the spectrum. This makes sense physically due to a degeneracy between the atmospheric scale height

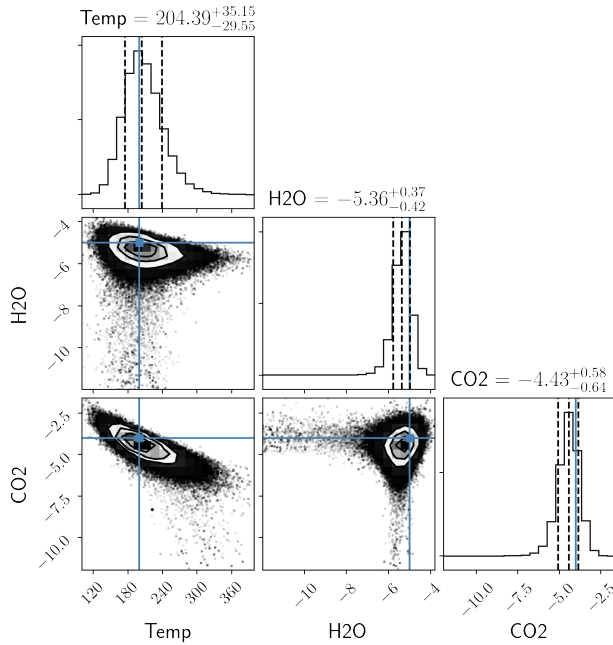


Figure 6.10 Example “corner plot” from the same MCMC retrieval as Fig. 6.9. Subplots along the diagonal show the marginalized posterior distributions for each retrieved parameters, while off-diagonal subplots show the covariances between pairs of parameters. Contours are used to indicate the most dense regions of parameter space and points are used to show posterior samples in lower probability regions.

and gas abundances: increasing the temperature will increase the scale height and thus increase the size of all absorption features in the spectrum, thereby requiring a decrease in the CO_2 abundance to bring the CO_2 absorption features back down to their best-fitting form. Thus, corner plots are an essential output from an MCMC retrieval that not only report the marginalized posteriors, but also display parameter covariances that can expose degeneracies and lead to physical insights.

Figure 6.11 shows the spread in spectral fits to the observed spectrum that are consistent with the posterior distributions derived with MCMC. The uncertainty envelopes around the median spectrum highlight the wavelengths where the model varies substantially in response to the inferred parameter uncertainties shown in Fig. 6.10.

MCMC methods can provide rigorous statistical results for inference problems, however there are a few noteworthy caveats that limit their usefulness within the SMARTER framework. First, although affine-invariant MCMC sampling can scale to inverse problems with a large number of parameters, MCMC is computationally expensive and is often only tractable to use with forward models that are quick to evaluate (e.g. less than 5 seconds). This presents a major difficulty when MCMC is used with the physically complex SMART forward model. Second, MCMC struggles

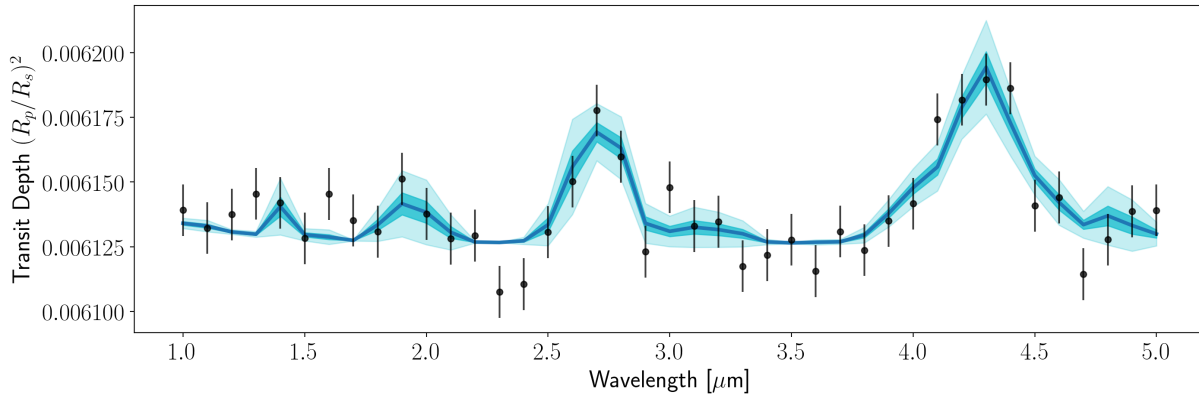


Figure 6.11 Distribution in spectra sampled from the posterior for the transit example retrieval case. The dark blue line shows the median model fit to the data and the cyan envelopes show the 1σ and 3σ confidence intervals about the median.

to converge for problems with multiple, well-separated posterior modes. In such cases, the walkers will fall into the closest mode and be unable to randomly walk to the other modes. This makes the number of samples within each mode reflective of the initial walker placement in parameter space relative to each mode, rather than the respective posterior probability of each mode, as one would desire. For multi-modal posterior spaces the Nested Sampling inverse model offers a superior approach.

Nested Sampling

Nested Sampling is an efficient and robust approach for Bayesian inference that is able to derive full posterior distributions, even for challenging multi-modal distributions, and which calculates the Bayesian evidence (denominator term in Equation 6.1) for model comparison applications (Skilling, 2004). In SMARTER, we use the MULTINEST nested sampling algorithm (Feroz and Hobson, 2008; Feroz et al., 2009, 2013) using the Python wrapper, PYMULTINEST (Buchner et al., 2014). Nested sampling has been growing in popularity within the exoplanet atmospheric retrieval community, with most of the common retrieval codes now using it as their preferred inverse model (e.g. Benneke and Seager, 2013; Line and Parmentier, 2016; Lavie et al., 2016; Rocchetto et al., 2016; MacDonald and Madhusudhan, 2017; Barstow et al., 2020). This broad adoption stems

from its efficiency—requiring $\sim 10\times$ fewer forward model evaluations compared with traditional MCMC sampling—and its applicability in model selection problems, both of which are important for atmospheric retrievals.

One of the key aspects of the nested sampling approach is the direct numerical calculation of the Bayesian evidence term. The evidence term from Equation 6.1 is the integral of the likelihood times the prior over the entire parameter space volume,

$$\mathcal{Z} = \mathcal{P}(\mathbf{y}) = \int \mathcal{L}(\mathbf{x})\mathcal{P}(\mathbf{x})d^D\mathbf{x}. \quad (6.16)$$

MULTINEST transforms this multidimensional integral into a one-dimensional integral that can be approximated numerically using the following weighted sum,

$$\mathcal{Z} \approx \sum_{i=1}^M \mathcal{L}_i w_i, \quad (6.17)$$

which is evaluated iteratively by stepping i forward through parameter space along nested contours of equal likelihood for M total iterations, where the weights w_i are given by the reduction in the prior volume between two successive iterations. The sum is performed by having N “active” or “live” points in parameter space that are initially ($i = 0$) drawn from the prior. The likelihoods for the active points are evaluated and then sorted. The point of lowest likelihood is removed from the active set (becoming “inactive”), and a new draw from the prior replaces the active point as long as it has a higher likelihood than the point it replaced. By enforcing that new active points have a higher likelihood, the prior volume is systematically reduced as a function of iteration, constraining the problem within successive contours of equal, but progressively increasing, likelihood. Thus, the MULTINEST algorithm travels through nested shells of likelihood as the prior volume is reduced. The procedure terminates once the evidence is determined to some user-specified precision (typically 0.5 in log-evidence).

Since the evidence is the normalization term for the calculation of the posterior density via Bayes’s Theorem, once found, the evidence can be used to generate posterior inferences for pa-

parameter estimation as well as to calculate *Bayes factors* for qualitatively assessing the *evidence* favoring one model’s fit to data over another. First, posterior inferences can be determined at each point for which the likelihood was evaluated by calculating the posterior weight,

$$p_j = \frac{\mathcal{L}_j w_j}{\mathcal{Z}}, \quad (6.18)$$

for each active and inactive point, j . Second, the evidence can be used to assess the relative probability between two different forward modeling approaches. If we perform retrievals with two models, M_1 and M_2 , each with their own set of parameters, then the relative probability of the two competing models can be written,

$$\frac{\mathcal{P}(M_1|\mathbf{y})}{\mathcal{P}(M_2|\mathbf{y})} = \frac{\mathcal{Z}_1 \mathcal{P}(M_1)}{\mathcal{Z}_2 \mathcal{P}(M_2)} = \mathcal{B}_{12}, \quad (6.19)$$

where we have assumed that the ratio of priors is unity in the absence of meaningful prior information (Benneke and Seager, 2013). The term $\mathcal{B}_{12}$ is called a Bayes factor and it requires the evidence to be calculated for both models. A large Bayes factor indicates more evidence in favor of M_1 over M_2 , where “weak” ($>2.1\sigma$), “moderate” ($>2.7\sigma$), and “strong” ($>3.6\sigma$) preference for M_1 requires Bayes factors of >3 , >12 , and >150 , respectively (Trotta, 2008; Benneke and Seager, 2013). We will make use of Bayes factors to compare retrieval simulations with and without certain molecules to quantify the detection significance of that gas from the spectrum.

Figure 6.12 shows a trace plot of the evolution of the MULTINEST algorithm. As the simulation progresses, the live points (color points in left panels) move from a broad prior volume with low posterior weights to a reduced volume of parameter space with high posterior weights. Unlike MCMC, the density of live points in MULTINEST is not proportional to the posterior probability density, so the weights are used to calculate posterior inferences, as shown in the marginal posteriors in the right panels.

Figure 6.13 (left corner plot) uses the same coloring as Fig. 6.12, but now shows the live points projected onto axes between pairs of parameters to visualize how the nested sampling technique

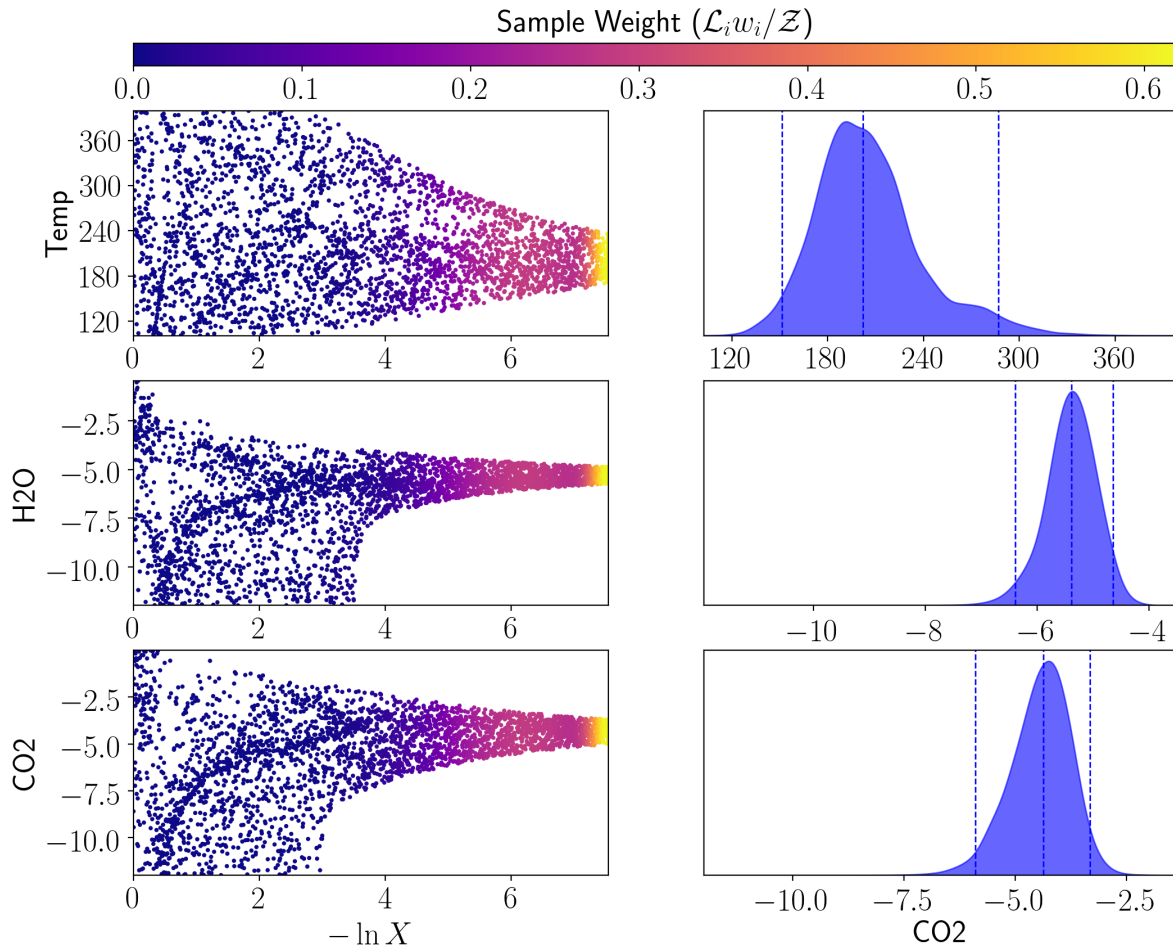


Figure 6.12 MULTINEST retrieval “trace plot” example showing the reduction in prior volume as the algorithm converges on the higher-likelihood region of parameter space. The left column of plots shows each parameter’s values as a function of the negative prior volume ($-\ln X$), which increases monotonically as the sampling volume decreases and the algorithm progresses forward. The points are colored by their relative posterior weight and represent the location and progression of the live points. The right column of plots shows the marginalized posterior distributions.

converges towards higher likelihood regions of parameter space using nested likelihood shells. Although the covariances can be seen visually by the point colors, the posterior weights must be used to transform the active and inactive points into posterior density inferences, which are shown in Fig. 6.13 in the right corner plot. Comparing the MULTINEST corner plot in Fig. 6.13 to the MCMC corner plot in Fig. 6.10 shows that the two methods are in excellent agreement.

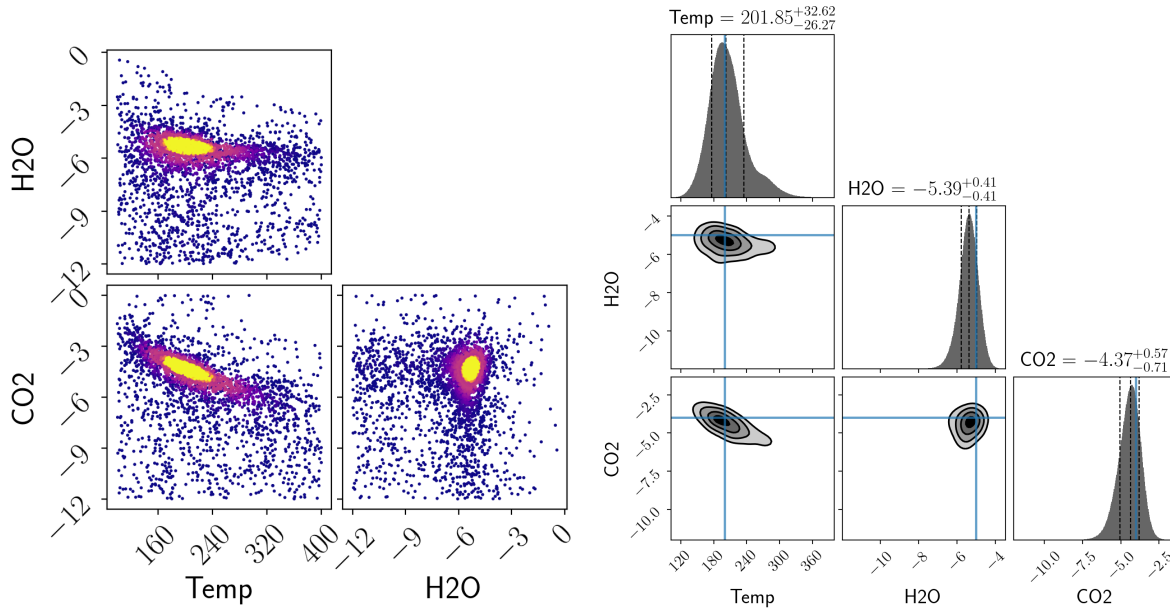


Figure 6.13 Corner plots showing the MULTINEST live points colored by their respective posterior weight (left corner plot; same coloring as Fig. 6.12) and contours of posterior probability (right corner plot; same quantities as 6.10).

Machine Learning with ApproxPosterior

The computational expense associated with repeatedly calling our forward model when running a typical retrieval algorithm motivates the development of a retrieval framework that uses machine learning to increase the speed of convergence. Indeed, machine learning approaches are beginning to show promise for increasing the speed of atmospheric retrieval inference problems (e.g. Zingales and Waldmann, 2018; Cobb et al., 2019). At present, these methods rely on generating large spectral datasets to train neural networks, which can perform atmospheric retrievals much faster than traditional physical models because the neural network replaces the radiative transfer model to compute the spectrum. However, the reliance on a single massive training set is a shortcoming considering how often new spectroscopic laboratory data (e.g. Fortney et al., 2019) or additional physical complexity must be added to the radiative transfer model, thereby requiring a new training set to be generated. We elected to mitigate the computational cost of our physically robust forward model using an alternative, but complementary approach, by performing approximate Bayesian

inference using the APPROXPOSTERIOR (AP) code (Fleming and VanderPlas, 2018; Fleming et al., 2020).

The open-source APPROXPOSTERIOR Python code employs an active learning algorithm to iteratively improve its predictive accuracy by expanding its training set while minimizing the expense of acquiring new samples. Unlike previous atmospheric retrieval machine learning approaches that learn on simulated spectra to replace the expensive radiative transfer model, AP learns the *posterior density* using a supervised learning Gaussian process (GP) regression active learning model. APPROXPOSTERIOR uses the Python code GEORGE to perform fast and efficient GP regression (Ambikasaran et al., 2015).

Given an initial training set of retrieval state vectors paired with their corresponding log posterior probabilities, AP will train a GP using maximum likelihood estimation to optimize the GP hyperparameters. AP then selects a new point to add to the training set by optimizing an objective function that attempts to identify regions of parameter space where the GP predicts high posterior probability—where model predictions are likely to match the observations—but that also have a large GP uncertainty so that augmenting the training set will optimally improve the predictive accuracy. Two approaches are implemented for the APPROXPOSTERIOR point selection utility function: the Bayesian Active Learning for Posterior Estimation (BAPE) algorithm from Kandasamy et al. (2017) and the Adaptive Gaussian process approximation for Bayesian inference with expensive likelihood functions (AGP) method from Wang and Li (2018). The true forward model is evaluated for each point that is selected and the new likelihood value is added to the training set. After m new points have been added to the training set, MCMC is used to sample the GP surrogate model’s posterior distribution. The algorithm then proceeds through the process of selecting another m new training points until either $n_{\max}$ iterations (each adding another m points to the training set) are reached or the posterior distributions obtained from MCMC have stabilized for multiple consecutive iterations. Fleming et al. (2020) suggest that APPROXPOSTERIOR can accurately produce posterior distributions in $\sim 1000\times$ fewer forward model evaluations than using MCMC. We also highlight the fact that APPROXPOSTERIOR generates a unique training set for each retrieval simulation such that it minimizes the number of calls to the forward model. This

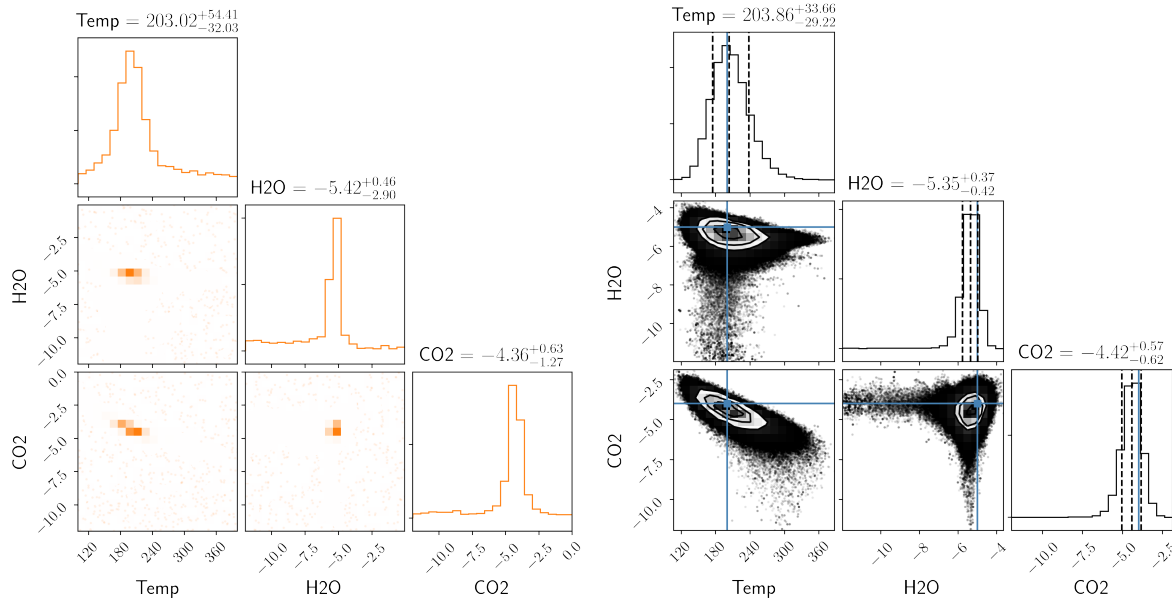


Figure 6.14 Corner plots showing the APPROXPOSTERIOR training set (left corner plot) and MCMC posterior samples from the GP surrogate model (right corner plot) for the transit example case.

feature makes APPROXPOSTERIOR robust to necessary line list and forward model upgrades because a new retrieval can be initiated after the upgrade in the exact way that one would re-run the retrieval analysis with MCMC or NS.

Figure 6.14 shows two corner plots associated with an APPROXPOSTERIOR retrieval of the transit example case discussed throughout this section. The left corner plot shows all the samples from the APPROXPOSTERIOR training set, including 50 initial training points randomly drawn from the priors and roughly 5,000 points selected by APPROXPOSTERIOR during run time. The right corner plot shows the posterior distribution from an MCMC using the APPROXPOSTERIOR GP surrogate likelihood model, which was run for 10,000 steps using 300 walkers, and is extremely well converged. Comparing these approximate posterior distributions to those generated for the same retrieval problem using MCMC (see Fig. 6.10) and MULTINEST (see Fig. 6.13), we find strikingly good agreement, particularly between MCMC and APPROXPOSTERIOR, which suggests that the numerical differences between MULTINEST and MCMC are larger in this case. From

the APPROXPOSTERIOR MCMC result, we can derive the approximate number of MCMC steps and SMARTER forward model evaluations that would be required for the full (non-approximate) forward model to converge. Our APPROXPOSTERIOR results suggest that the autocorrelation time for this MCMC chain is $\tau \approx 60$ steps. To achieve a diligent chain length of $50 \times \tau$ would require $(50 \tau / \text{walker} \times 60 \text{ steps} / \tau \times 300 \text{ walkers}) \sim 900,000$ forward model evaluations, compared to just over 5,000 using APPROXPOSTERIOR. Although this is an excellent $\sim 200 \times$ speed increase, we likely ran APPROXPOSTERIOR significantly longer than necessary in this example, implying that we could achieve even better performance in future simulations.

In summary, APPROXPOSTERIOR appears capable of increasing the efficiency of atmospheric retrieval problems by a factor of 200-1000 while producing full posterior distributions that maintain good agreement with MCMC and MULTINEST. These features make it an excellent addition to the SMARTER inverse modeling arsenal and atmospheric retrieval models in general. In particular, APPROXPOSTERIOR helps facilitate the vision of the SMART/SMARTER model being used to apply physically robust, but slow, forward modeling to terrestrial exoplanet atmospheric retrieval. However, APPROXPOSTERIOR has not been tested on moderate to high dimensionality inference problems (e.g. greater than 7 dimensions), and therefore more work is required to further validate and demonstrate the capabilities of APPROXPOSTERIOR for atmospheric retrievals, and to potentially further refine the APPROXPOSTERIOR algorithm.

Selecting an Inverse Model

Selecting the best or most efficient inverse model will depend on the specific retrieval problem at hand. As stated previously, OE is fast because it can converge in tens to hundreds of forward model evaluations, but it assumes Gaussian posterior PDFs, which can fail to capture the complicated posterior morphologies and parameter degeneracies that are common in exoplanet atmospheric retrieval modeling. Therefore, OE is only recommended for (a) quick preliminary assessments of the best-fit, (b) initializing other inverse models, such as MCMC or AP, and (c) cases where the forward model may be so physically complex that OE is the only computationally tractable inverse model. MCMC, NS, and AP relax the assumption about the shape of the derived posterior

PDFs. However, MCMC requires thousands to millions of forward model evaluations, which may be too computationally demanding for many SMARTER forward model configurations. Furthermore, MCMC has difficulty converging when the posterior contains multiple distinct modes. In such instances (e.g. see the refraction case in Chapter 7), NS is optimal because it can identify and characterize multi-modal posteriors. Additionally, NS can converge with approximately two orders of magnitude fewer forward model evaluations compared to MCMC and it also directly calculates the Bayesian evidence, making it a versatile workhorse for many SMARTER applications. Finally, AP is promising for retrievals with a relatively small number of parameters ($\lesssim 7$) due to the steep rise in training set size required for higher dimensional problems, coupled to the fact that re-training the GP becomes computationally expensive for large training sets. However, AP may be particularly useful for retrievals using a forward model that is exceptionally complex and computationally expensive. At present, this is the first atmospheric retrieval study to use AP so, unlike the other widely used inverse models, more work will be required to validate the technique and carve out optimal use-cases.

6.3 SMARTER Validation using Earth’s Transmission Spectrum

In this section we validate SMARTER on exoplanet-analog transmission spectrum observations of Earth constructed from satellite data. We use the high spectral resolution, high S/N empirical transit transmission spectrum of Earth between 2-14 μm presented in [Macdonald and Cowan \(2019\)](#), which was constructed from solar occultation observations to reproduce the limb-averaged observer perspective that is analogous to transiting exoplanets. To validate our retrieval, we bin the data to relatively low resolution to investigate the entire spectral range, drawing parallels to upcoming *JWST* observations, and we fit the native high-resolution spectrum over smaller wavelength intervals, highlighting key insights and challenges that may arise for future exoplanet observations. In addition to validating SMARTER, we perform a retrieval on the strength of $\text{O}_2\text{--N}_2$ CIA relative to $\text{O}_2\text{--O}_2$ CIA at 6 μm to validate the approach used to include $\text{O}_2\text{--N}_2$ CIA in non-Earth-like exoplanet atmospheric modeling.

We describe the transmission spectrum from [Macdonald and Cowan \(2019\)](#) in detail next. Then

we outline our retrieval experiments on high- and low-resolution versions of the spectrum, followed by a presentation of our results. Finally we present a discussion of our findings and conclude.

6.3.1 Observed Data

Macdonald and Cowan (2019) processed solar occultation measurements of the Earth’s atmosphere observed by the Canadian low-Earth orbit satellite, SCISAT, using the *Atmospheric Chemistry Experiment Fourier Transform Spectrometer* (ACE-FTS; Bernath et al., 2005; Hughes et al., 2014). ACE-FTS is a high resolution spectrograph that observes between 750–4400 wavenumbers with a full-width-at-half-maximum (FWHM) of 0.02 cm^{-1} and a sampling resolution of 0.0025 cm^{-1} . Spectra were recorded of the Sun passing through the Earth’s atmosphere as a function of time during sunrises and sunsets. Dividing out the spectrum of the Sun without the Earth’s atmosphere enables the transmittance, $\mathcal{T}$, of Earth’s atmosphere to be measured as a function of both wavelength and altitude. These data products support the main objective of the mission to study trace gases in the Earth’s atmosphere. Macdonald and Cowan (2019) calculated the transmission spectrum of the Earth as if it were an exoplanet by integrating the ACE-FTS transmittance data from the bottom to the top of the atmosphere using the approach described in Robinson et al. (2014b) for Titan and Dalba et al. (2015) for Saturn. Macdonald and Cowan (2019) selected for only cloud-free sightlines through the Earth’s atmosphere and averaged datasets from five latitudinal regions—Arctic summer and winter ($60 - 85^\circ\text{ N}$), mid-latitude summer and winter ($30 - 60^\circ\text{ N}$), and the tropics ($30^\circ\text{ S} - 30^\circ\text{ N}$)—to mimic the limb-integrated stellar light transmitted through exoplanet atmospheres during primary transit.

The Earth’s observed transmission spectrum is of exquisite quality. Figure 6.15 shows the empirical transmission spectrum of Earth from Macdonald and Cowan (2019)¹⁰. The broad wavelength coverage spanning $2\text{--}14\text{ }\mu\text{m}$ contains spectral features from over a dozen different molecules (Bernath et al., 2005). The top panel shows the spectrum at the native, high resolution and binned to a lower resolution of $\Delta\lambda = 0.02\text{ }\mu\text{m}$. The zoomed subpanels focus on the $4.3\text{ }\mu\text{m}$ CO_2 band and

¹⁰The published spectrum was acquired from the authors by request for this work. [Click here to download the data file.](#)

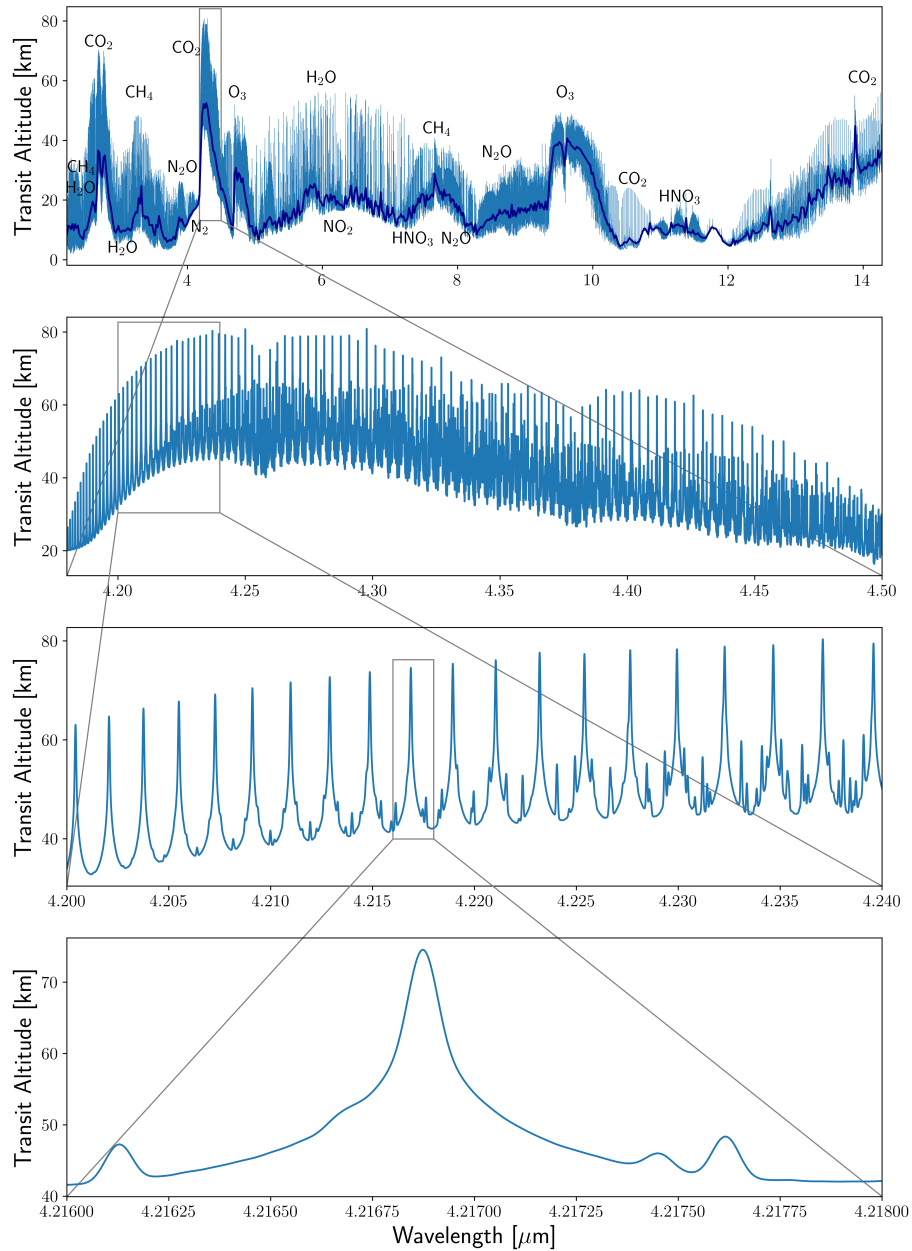


Figure 6.15 Empirical transmission spectrum of the cloudless Earth from [Macdonald and Cowan \(2019\)](#). The top panel shows the full wavelength range in high-resolution (light blue) and binned to a lower resolution ($\Delta\lambda = 0.02 \mu\text{m}$; dark blue). The lower panels zoom into the $4.2 \mu\text{m}$ CO_2 band, showing progressively smaller wavelength ranges and highlighting the fully resolved spectral lines with no apparent observational noise.

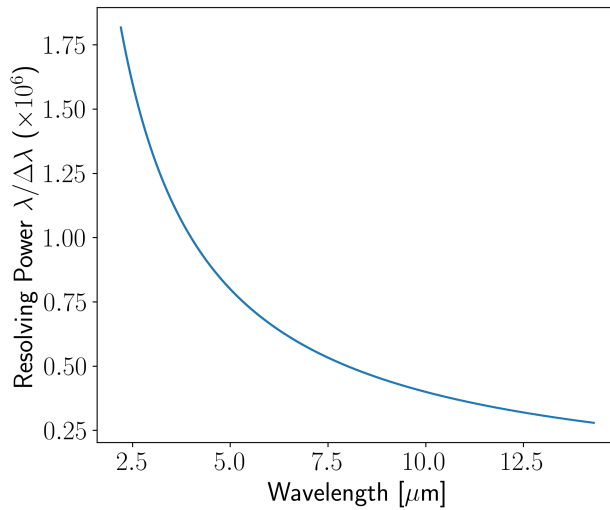


Figure 6.16 Spectral resolution of Earth’s observed transmission spectrum from [Macdonald and Cowan \(2019\)](#). The resolving power is plotted in millions, and is shown decreasing as a function of wavelength. This is a characteristic relationship that occurs because the observed spectrum has a fixed FWHM in units of wavenumbers (cm^{-1}).

showcase the extremely high resolution and high SNR of the spectrum. Each spectral line within each molecular band is fully resolved. The bottom-most panel shows that there is no visible scatter to the data that would be associated with observational noise. This spectrum is extremely low noise and high-resolution, and should present an important validation challenge for any spectral retrieval code and for transmission spectroscopy, in general.

Figure 6.16 shows the spectral resolution of Earth’s observed transmission spectrum. This is calculated from the spacing between the spectral samples and does not reflect the actual FWHM of the spectrometer. From the known spectral resolution of ACE-FTS, we find that observed spectrum oversamples the FWHM by a factor of 8. We use this information to inform the wavenumber sampling resolution for our SMART retrieval forward modeling when we fit the high resolution spectrum.

6.3.2 Retrieval Experimental Setup

We consider the validation of SMARTER using the spectrum shown in Fig. 6.15 from a few different perspectives to test and highlight different challenges and capabilities. In our first experiment, we attempted to fit the broad wavelength, low-resolution spectrum shown in the top panel of Fig. 6.15 (dark blue line) using Earth’s average 1D atmospheric temperature and gas VMR profiles as a

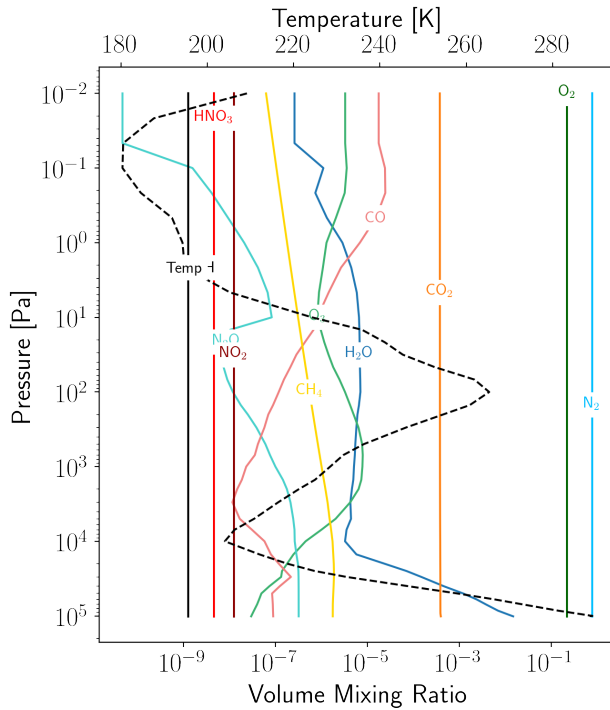


Figure 6.17 Earth's 1D atmospheric thermal structure and composition. The temperature structure (dashed black) is shown along the top x-axis and the gas mixing ratios are shown along the bottom x-axis, both as a function of pressure.

starting point, but allowed the gas mixing ratios to vary by a constant factor. In our second experiment, we turned to the high-resolution spectrum in the $6\ \mu\text{m}$ window ($5\text{--}7\ \mu\text{m}$) and attempted to fit it again by retrieving gas VRM profile multiplicative constants. For these validation experiments, we used non-linear least squares minimization to retrieve MAP solutions. Due to the lack of any known observational errors and the presence of good prior information, we found this approach to be reliable, but we discuss other approaches in Section 6.3.4.

We assume a globally-averaged 1D atmospheric structure for Earth. Figure 6.17 shows Earth's average 1D atmospheric vertical thermal structure and composition acquired from satellite data for the VPL Earth Model (Robinson et al., 2011) to reproduce the EPOXI flyby of Earth in March and April of 2008. We use these profiles as the starting point for our Earth retrieval validation, but allow the retrieval to solve for a factor by which each gas profile can be scaled to improve the fit (e.g. similar to Barstow et al., 2016). In addition to the gases included in the VPL Earth Model (H_2O , CO_2 , O_3 , O_2 , N_2 , CH_4 , CO , and N_2O), we also include HNO_3 and NO_2 because they corrected notable residuals in our initial tests. For HNO_3 and NO_2 , we use evenly mixed

atmospheric profiles as shown in Fig. 6.17, initialized using best guesses from the literature. For HNO_3 we use a measurement from Hanke et al. (2003), for NO_2 we use approximately 1% the regulatory standard provided by the EPA¹¹.

We used a nominal SMART simulation setup for our high and low resolution Earth retrievals that remained fixed throughout these experiments. We fixed the Earth’s radius and surface gravity at 6371 km and 9.81 m/s^2 , respectively. We also placed the Earth at 1 AU from the Sun, which has a stellar radius of $1 R_\odot$. We employed the ray tracing transit transmission model within SMART (Robinson, 2017a), but we did not include atmospheric refraction. We note that we did perform a forward model test with refraction enabled and with the angular size of the Sun set to an arbitrarily large value to mimic the SCISAT’s ability to access extremely refracted rays through the lower atmosphere that are below the critical refraction limit seen for exoplanets (García Muñoz et al., 2012; Bétrémieux and Kaltenegger, 2014; Misra et al., 2014), but we observed no refraction effects in the low resolution spectrum. We set the transit impact parameter to zero, and we turned off the stellar and thermal source functions to only simulate the transmission spectrum. Finally, we used an optical depth binning parameter for SMART of 0.75, which is large, but our tests have shown that the transmission spectrum incurs negligible errors even with large spectral mapping bin tolerances. For these tests we opted to have our retrieval forward model return the transmission spectrum in units of altitude in kilometers above the surface, rather than convert the data into units of transit depth, $(R_p/R_s)^2$.

6.3.3 Retrieval Validation Results

Low Resolution Spectrum

We begin by fitting the low resolution spectrum shown in dark blue in Fig. 6.15. Figure 6.18 shows our MAP solution to Earth’s full $2 - 14 \mu\text{m}$ transmission spectrum. We use semi-transparent colors to highlight the absorption features due to all molecules included in our model. Strong absorption

¹¹<https://www.epa.gov/no2-pollution/table-historical-nitrogen-dioxide-national-ambient-air-quality-standards-naaqs>

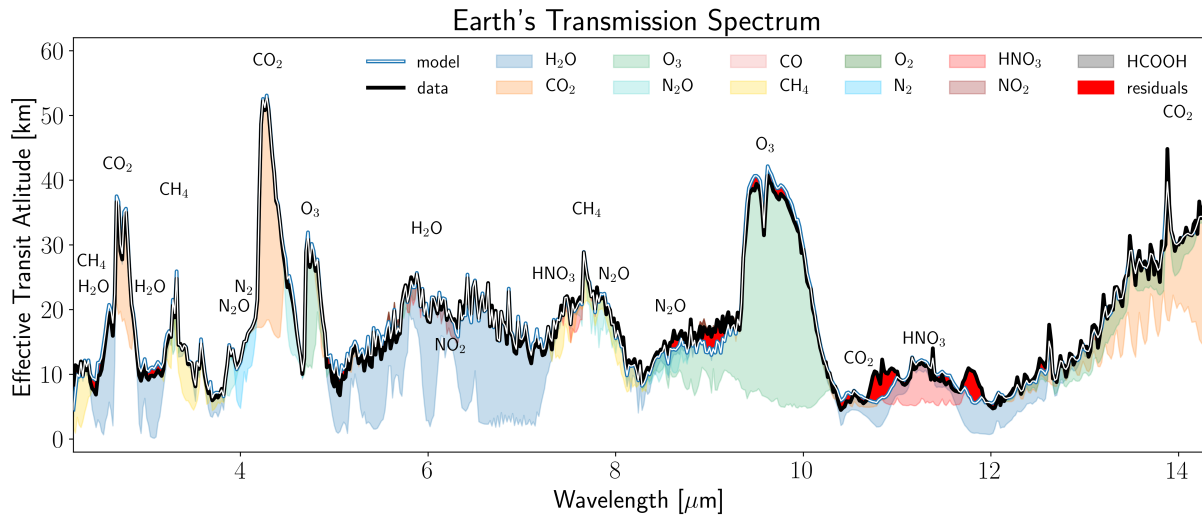


Figure 6.18 Earth’s observed cloud-free transit transmission spectrum from [Macdonald and Cowan \(2019\)](#) binned to lower resolution and fit using SMARTER. The data are shown in black, the best-fitting model spectrum is shown in white and blue, and various colors are used to indicate each molecule’s contribution to the model spectrum. Residuals between the data and model are shown in red.

features are clearly visible from H_2O , CO_2 , and O_3 , with smaller features due to CH_4 , HNO_3 , NO_2 , and N_2O . We note that the N_2 – N_2 CIA feature at $4.1 \mu\text{m}$ is also visible (see [Schwieterman et al., 2015b](#)). There is substantial agreement between our best-fitting model and the observed data. Our best-fitting spectrum has a 7.36% percent deviation on average from the empirical transmission spectrum. In addition, we derive a characteristic error of 1.33 km representing the one-sigma uncertainty on each spectral data point such that this best-fitting model would have a reduced chi-squared of unity ($\chi_{red}^2 = 1$). Figure 6.18 also highlights the less well-fit regions in bright red. The most notable residuals exist near $11 \mu\text{m}$ and $12 \mu\text{m}$ and may be due to industrially produced chemicals (see [Bernath et al., 2005](#)). There are also residuals near $9 \mu\text{m}$. Overall the residuals remain relatively small.

Table 6.2 lists the MAP mixing ratio scale factors retrieved for each molecule included in our low resolution fit. Most of the scale factors lie near unity, indicating good agreement with the known composition of Earth’s atmosphere. The molecules with large absorption features tend to

Table 6.2. Retrieved molecule VMR scale factors from MAP fit to Earth’s low-resolution (LR) and high-resolution (HR) transmission spectrum

Molecule	LR VMR Scale Factor	HR VMR Scale Factor	Earth Spatial Deviations
H ₂ O	1.205	2.268	30%
CO ₂	1.053	1.956	<1%
O ₃	1.243	1.545	25%
N ₂ O	1.157	0.747	23%
CO	0.538	—	28%
CH ₄	1.029	0.727	34%
O ₂	0.340	—	0%
N ₂	1.043	—	0%
O ₂ –N ₂	—	0.986	—

Note. — “LR” and “HR” refer to low-resolution and high-resolution, respectively.

have retrieved abundances that deviate by less than 10 – 30% from the initial assumed Earth-average profiles. However, Earth’s atmosphere naturally exhibits spatial variability for the temperature structure and many of the gas profiles, which is likely a source of systematic error in the observed and hemispherically-averaged transmission spectrum. We identify Earth’s approximate spatial variability by calculating the percent deviation among the 1D profiles from the VPL Earth Model’s atmospheric structure database (Robinson et al., 2010, 2011, 2014a; Schwieterman et al., 2015a). These values are reported in the last column of Table 6.2 for reference. For most molecules our retrieved VMR from the Macdonald and Cowan (2019) spectrum is within the spatial variance seen across the globe. A noteworthy exception to this is O₂, which exhibits effectively no spatial variability, and yet we retrieved 34% of the true abundance. This is easily reconciled by the fact that O₂ exhibits no substantial absorption features in this wavelength range, and is therefore not well constrained.

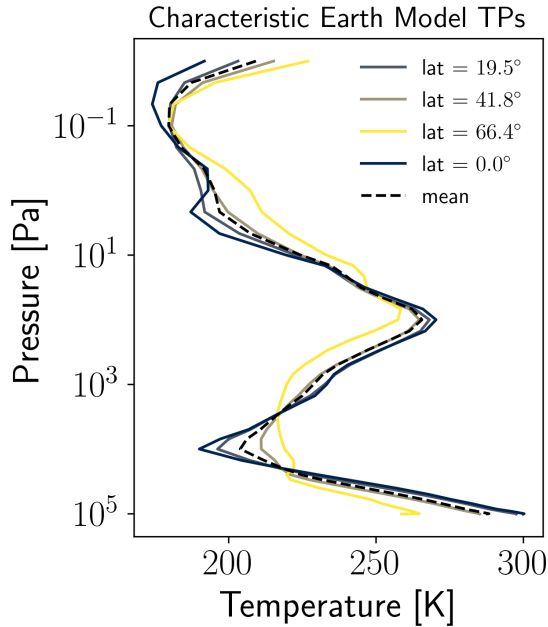


Figure 6.19 Characteristic thermal profiles of Earth’s atmosphere from the VPL Earth Model. The profiles are derived from satellite data acquired during the EPOXI fly by of Earth in 2008 (see [Robinson et al., 2011](#)). Each latitude plotted represents the globally averaged TP for all HEALPix bins with the corresponding north and south latitudes. [Click to download these TP profiles](#)

High Resolution Spectrum

We now examine the Earth transmission spectrum at the native, high spectral resolution. Here, we focus exclusively on the $5 - 7 \mu\text{m}$ region. This smaller wavelength range is more computationally tractable for fitting the extremely high resolution spectrum due to the large output file sizes from SMART and not the computational expense. In this wavelength range, our forward model can be evaluated in about 30 seconds, compared to about 3 minutes for the full wavelength range. The $5 - 7 \mu\text{m}$ range is also scientifically interesting because it contains one of the largest H_2O features in this transmission spectrum, which also overlaps with the $\text{O}_2 - \text{O}_2$ and $\text{O}_2 - \text{N}_2$ CIA bands recently highlighted in [Fauchez et al. \(2020\)](#). Thus, this wavelength range provides an opportunity to test for sensitivity to O_2 , where our low resolution spectrum was completely insensitive.

We begin this investigation from where we left off using the low resolution spectrum. We use the best fitting MAP state vector from our low resolution fit as our initial guess for our high resolution retrieval. However, in this case we make a few changes to the forward model to accommodate the high-res spectrum. First, we modify our approach for the instrument convolution by using SMART’s internal triangular slit function with the correct FWHM of the ACE-FTS instru-

ment (0.02 cm^{-1}). We output spectra on a wavenumber grid with a sampling resolution of 0.00125 cm^{-1} , which oversamples the data by a factor of about 2. Then we interpolate the model spectrum onto the observed data using the `SMARTER Interpolate()` instrument model. Interpolation is appropriate in this situation because we have already used a convolution function within SMART, so we are only aligning the model with the data. Our second modification to the forward model is to fix the O_2 and N_2 abundances at the truth and instead solve for the factor by which the $\text{O}_2\text{--N}_2$ CIA is scaled. We use the $\text{O}_2\text{--O}_2$ CIA file from HITRAN 2016 as a proxy for $\text{O}_2\text{--N}_2$ CIA, thereby enabling us to solve for the relative strength of $\text{O}_2\text{--N}_2$ CIA to $\text{O}_2\text{--O}_2$ CIA. Additionally, CO is not included because it does not show absorption features in this range. Our third modification comes as a result of trying multiple TP profiles. Figure 6.19 shows four characteristic TP profiles that we found represented a substantial portion of the variance within the VPL Earth Model (Robinson et al., 2011). Via experimentation we found that a “cold” TP profile from the 66° latitude bin (yellow line in Fig. 6.19) provides a better fit to the high resolution H_2O lines than the globally averaged TP profile that we used for the broad wavelength spectrum (black dashed line in Fig. 6.19). This demonstrates that high resolution transmission spectra have a strong sensitivity to a planet’s thermal structure, a point that we will return to later in this section.

We find remarkably good agreement between our best-fitting model spectrum and the high resolution data. Figure 6.20 shows our MAP model fit to the $6 \mu\text{m}$ region. In the top two panels it is difficult to discern the differences between the model (in blue) and the data (in black) because both lines plot on top of one another. The bottom panels of Fig. 6.20 zoom into a narrow enough region to see small deviations, and these panels highlight the contribution from various line and continuum absorbers. In the third panel, there are several well-resolved H_2O lines that make up the $6 \mu\text{m}$ absorption band and span approximately 50 km in absorbing radius from the continuum to their cores. The forth panel zooms into the region between H_2O lines where the opacity is low and contributions to the continuum are visible from $\text{O}_2\text{--N}_2$ CIA, $\text{O}_2\text{--O}_2$ CIA, and H_2O . In this region there are also subtle absorption features due to HNO_3 and CH_4 . Between $5 - 7 \mu\text{m}$ our best-fitting spectrum has a 4% percent deviation on average from the empirical transmission spectrum and we derive a characteristic error of 0.96 km representing the one-sigma uncertainty on each spectral

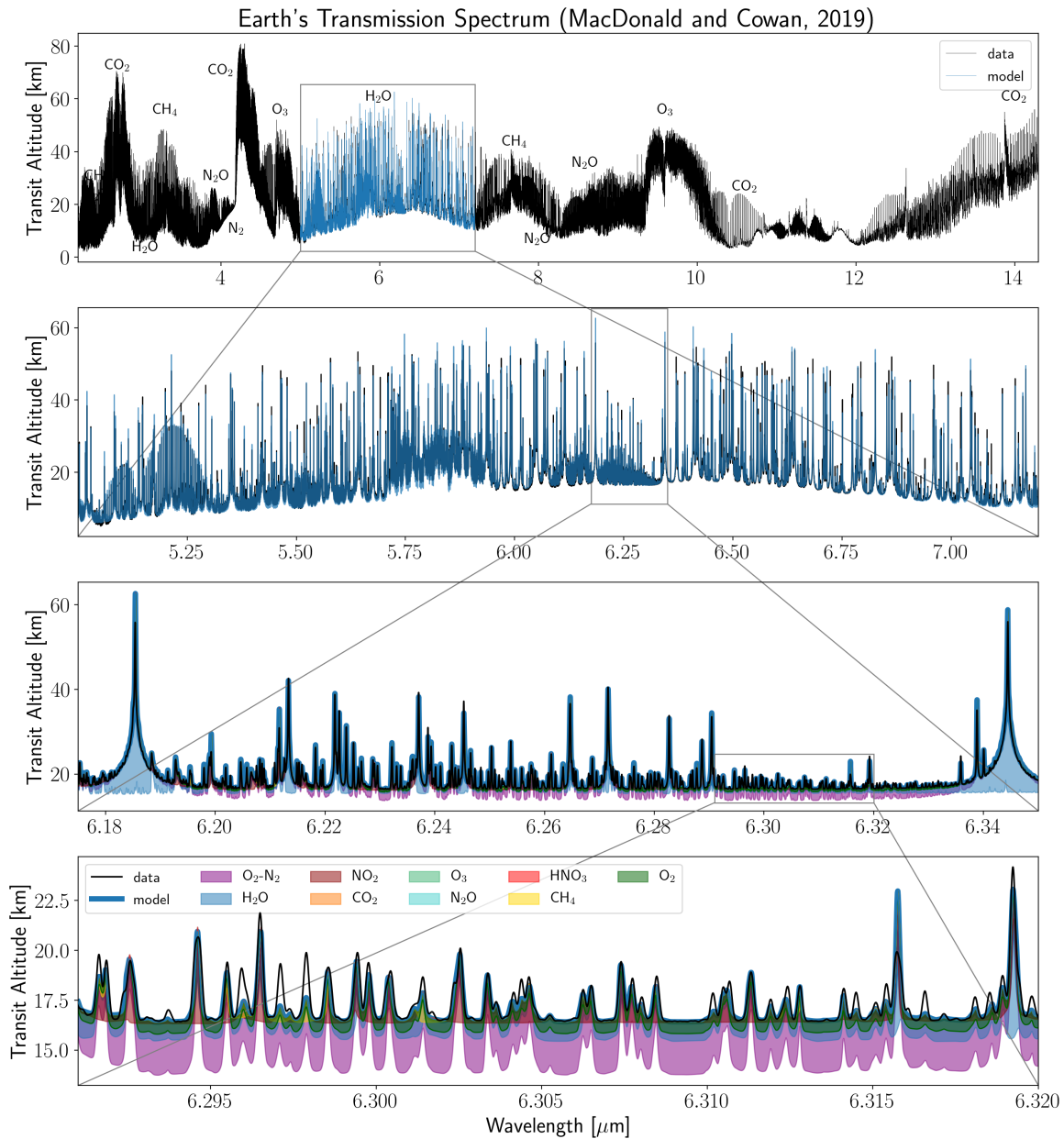


Figure 6.20 Earth's observed transit transmission spectrum from [Macdonald and Cowan \(2019\)](#) fit using SMARTER at the native spectral resolution. The data are shown in black, the best-fitting model spectrum is shown in blue, and various colors are used in the bottom panels to indicate each molecule's contribution to the model spectrum. In the bottom panel, HNO_3 , CH_4 , and H_2O show subtle absorption features above a $\text{O}_2\text{-N}_2$ CIA, $\text{O}_2\text{-O}_2$ CIA, and H_2O continuum.

data point such that this model best-fit would produce a reduced chi-squared of unity ($\chi_{red}^2 = 1$).

Table 6.2 lists the retrieved scale factors from for our high resolution spectral fit in the third column. Despite the remarkably good fit to the data, in this case we see a greater difference from the initial Earth atmospheric structure than seen in our low resolution case. In particular, we require the H₂O and CO₂ profiles to be enhanced by factors of $\sim 2.2\times$ and $\sim 2\times$, respectively. It is concerning that the retrieval is able to fit the data so well (residuals under 1 km), and yet the retrieved abundances are discrepant by more than the expected propagated measurement error. Although H₂O exhibits spatial variability, since CO₂ effectively does not, these discrepancies may point to a degeneracy between the TP structure and gas abundances that can cause retrieval biases.

We find that the O₂–N₂ CIA scale factor is near unity, indicating that it is comparable in intensity to the 6 μm O₂–O₂ CIA band. We caution about the uniqueness of these results due to the fact that the H₂O continuum, O₂–O₂ CIA, and O₂–N₂ CIA all overlap, and we already see a potential degeneracy between our TP structure and H₂O abundance. This degeneracy will need to be fully explored before confident values for the O₂–N₂ CIA strength can be reliably retrieved from this spectrum. However, we do find that this result is consistent with the O₂–N₂ CIA intensity reported in Moreau et al. (2001), which is compared with the 6 μm O₂–O₂ CIA band from HITRAN 2016 (Gordon et al., 2017) in Fig. 6.21. Thus, despite the caveats from our retrievals, Moreau et al. (2001) provide a compelling laboratory measurement and a theoretical basis for concluding that O₂–N₂ CIA compares in strength to O₂–O₂ CIA at 6 μm .

6.3.4 Discussion & Conclusion

We used SMARTER to fit the empirical transmission spectrum of Earth from Macdonald and Cowan (2019) as a validation case for our new atmospheric retrieval code. In general, we found exceptionally good model fits to the spectra both in low resolution across the full 2–14 μm wavelength range and in high resolution in the 5–7 μm wavelength range, with model deviations under 10%. Our retrieved molecular abundances from the low resolution spectrum are consistent with the Earth’s atmospheric composition within the range of spatial variability seen across the globe. Specifically, we retrieved abundances for H₂O, CO₂, O₃, N₂O, CH₄, and N₂ that are within 25% of their average

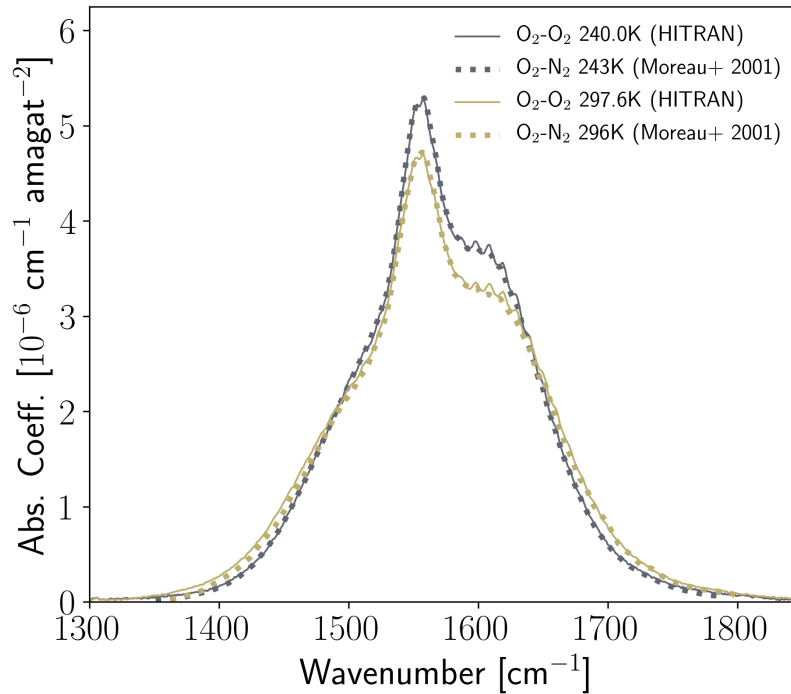


Figure 6.21 A comparison between the $6\mu\text{m}$ $\text{O}_2\text{--N}_2$ CIA band from [Moreau et al. \(2001\)](#) and the $6\mu\text{m}$ $\text{O}_2\text{--O}_2$ CIA band from HITRAN 2016 ([Gordon et al., 2017](#)). At Earth-like temperatures and pressures, the $\text{O}_2\text{--N}_2$ and $\text{O}_2\text{--O}_2$ CIA bands appear to agree in intensity.

Earth values; CO is within 50% and O_2 is within 70%. Although our fit to the spectrum is even better at high resolution than at low resolution (deviations down to 4% from 7%), our retrieved molecular abundances deviate more from the global averages. In particular, our high resolution fit required a colder TP profile representative of higher latitudes, and we retrieved a H_2O abundance that exceeds the global average by $2.2\times$, a CO_2 abundance that exceeds the global average by $2\times$, and an O_3 abundance that exceeds the global average by 50%.

Our best-fit to the high resolution spectrum (Figure 6.20) does not recover molecular abundances at the level of precision expected given (1) how good the fit is and (2) how high the S/N is. Our decision to use a different TP profile indicates that such high resolution transmission spectra are sensitive to atmospheric thermal structure, but likewise illuminates degeneracies between the TP structure and the gas abundances. These complicating factors must be investigated in future

works aimed at retrieving a characteristic altitude-dependent TP profile. However, it is possible that even such an approach may still impose biases in gas abundances because the globally-averaged exoplanet-analog transmission spectrum from [Macdonald and Cowan \(2019\)](#) may contain information from more than one characteristic TP profile, similar to the findings of [Caldas et al. \(2019\)](#) who used a GCM to produce a 3D exoplanet transmission spectrum. These biases are likely to be exacerbated at high spectral resolution.

We ran additional retrieval tests using more sophisticated inverse models, but these encountered trouble efficiently exploring parameter space. In particular, both EMCEE and MULTINEST appeared to converge to very specific results, relatively nearby their initial parameter states. We suspect that the lack of observational errors contributed to our sampling difficulties, such that the “basin of attraction” around our local MAP solutions presented an insurmountable barrier for MCMC and Nested Sampling to explore beyond. This interpretation is consistent with another caveat that we were able to find multiple MAP solutions depending on our initial starting state and our choice for TP profile. The spacing between these local minima greatly exceeded the parameter uncertainties that we derived. It may be possible for future studies to use our model calibrated errors of ~ 1 km to normalize the likelihood calculations and perhaps improve the retrieval’s ability to traverse degeneracies in this parameter space. Nonetheless, it is also likely that we have encountered modeling uncertainties that exceed the observational uncertainties as a result of using a single atmospheric structure to fit an inherently 3D atmosphere, as discussed above.

The subtle deviations between our best-fitting model and the observed data can be related to an exoplanet-analog transmission spectrum to reveal the unobtainable quality of data that would be required to warrant a more sophisticated model than the one we have used. For instance, an observed transmission spectrum for a transiting Earth-Sun analog exoplanetary system would require transit depth uncertainties of 10 ppb before our fit would be considered unacceptable. An exo-Earth transiting the much smaller star, TRAPPIST-1 ([Gillon et al., 2016](#)), would still require observational errors under 1 ppm. We emphasize that achieving such high S/N transmission spectra is not feasible in the near term with *JWST*—particularly at the spectral resolutions explored here. Nevertheless, this work demonstrates that a physically rigorous retrieval model designed to

interpret exoplanet observations can properly interpret extremely high quality Earth observations (modulo insensitivities to some species inherent in the observed spectral range) and allow an exploration of fundamental degeneracies, while also leaving substantial room for exoplanet observations to increase in quality over the next few decades.

6.4 Discussion & Conclusion

In this chapter we have detailed the methodology behind our novel terrestrial exoplanet atmospheric retrieval code, SMARTER, and validated it against an empirical transmission spectrum of Earth. SMARTER features a physically rigorous line-by-line radiative transfer code that is used within a retrieval forward model to produce exoplanet transmission, emission, and reflected spectra. SMARTER has been developed with four versatile inverse models for evaluating Bayes's Theorem for parameter estimation. Our validation using Earth observations demonstrated the successful operation of the SMARTER code as a terrestrial exoplanet atmospheric retrieval model, but it also opened up numerous avenues for future investigation, including the exploration of TP profile retrievals and degeneracies using high resolution transmission spectrum observations. Despite the work already completed, there remain numerous opportunities for further model development and validation going forward. We will briefly describe these efforts.

6.4.1 Continued Validation using Titan's Transmission Spectrum

The transmission spectrum of Titan observed using *Cassini* solar occultations and published in [Robinson et al. \(2014b\)](#) provides an excellent opportunity to further validate SMARTER using a haze covered world. This is particularly important because the Earth transmission spectrum from [Macdonald and Cowan \(2019\)](#) purposefully excluded clouds, focusing only on clear-sky paths through the atmosphere. Furthermore, aerosols are a key ingredient for understanding exoplanet transmission spectra.

We have begun an initial retrieval validation pilot study using the [Robinson et al. \(2014b\)](#) Titan transmission spectrum. Figure 6.22 shows Titan's NIR transmission spectrum and our best fitting

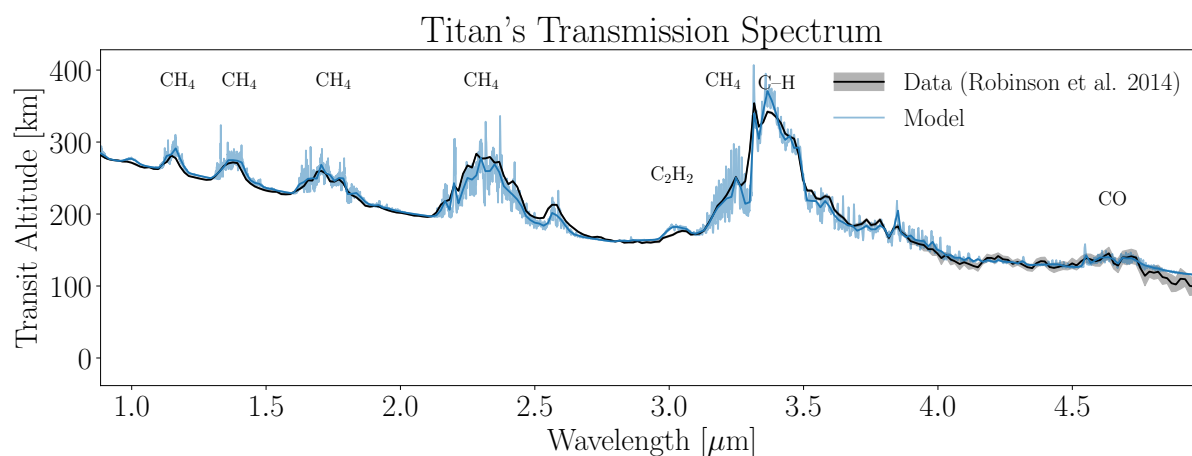


Figure 6.22 Titan’s observed transit transmission spectrum from Robinson et al. (2014b) fit using SMARTER. The model assumes a fractal aggregate hydrocarbon haze that gives the spectral continuum its characteristic linearly decreasing slope. Molecular absorption features are seen from methane (CH_4), acetylene (C_2H_2), carbon monoxide (CO), and a C–H stretch transition modeled using propane (C_3H_8) as a proxy for all aliphatic hydrocarbons.

model spectrum. The spectral continuum exhibits a linearly decreasing slope with wavelength that is characteristic of Titan’s hydrocarbon hazes. We fit for the optical depth profile of two fractal aggregate hydrocarbon haze modes (Arney et al., 2016)—an upper and a lower haze. Both fractal aggregate haze modes use optical constants for organic tholins (Khare et al., 1984) and assume 50 nm monomers. The upper haze mode is a $6\text{ }\mu\text{m}$ aggregate composed of over one million monomers with a fractal dimension of 2.4 and the lower haze mode is a 70 nm aggregate composed of ~ 3 monomers with a fractal dimension of 1.5. Figure 6.22 displays the maximum likelihood haze solution fit only to the continuum (blue), with molecular absorption bands of CH_4 , C_2H_2 , and CO added into the radiative transfer after the fit, fixed at $0.1\times$ the average mixing ratio profiles derived from *Cassini* measurements. Titan’s 1D atmospheric structure is shown in Figure 6.23 for reference. Additionally, there is a C–H stretching mode of aliphatic hydrocarbon chains near $3.4\text{ }\mu\text{m}$ (Robinson et al., 2014b; Bellucci et al., 2009; Maltagliati et al., 2015), which we are able to fit remarkably well using C_3H_8 (propane) as a proxy for all aliphatic hydrocarbons (at $2\times$ the *Cassini* VMR). Despite use of *a priori* knowledge of Titan’s atmospheric and haze compositions,

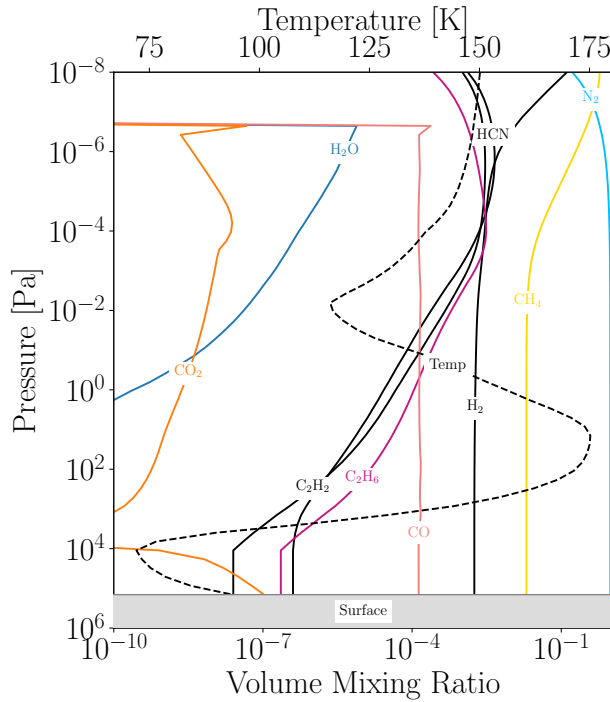


Figure 6.23 Titan's 1D atmospheric thermal structure and composition. The temperature structure (dashed black) is shown along the top x-axis and the gas mixing ratios are shown along the bottom x-axis, both as a function of pressure.

this fit is a promising initial step, and already exceeds the rigor of published literature on this exoplanet-analog spectrum.

The lack of consistency between the absorption features using the *Cassini* gas mixing ratios and our haze model fit indicates that our haze fit does not have the correct vertical optical depth profile. The fact that we are over predicting the strength of the gaseous absorption features using the *Cassini* mixing ratios, and therefore adjusting them to be lower by an order of magnitude, indicates that our haze fit should be shifted to higher altitudes to obscure more of the atmospheric column. Future retrievals will need to fit the haze optical depth profiles and the gas mixing ratios at the same time so as to fully explore the degeneracies between the haze parameters and the molecular abundances.

6.4.2 Developing a Physically Augmented Atmospheric Retrieval Code

The exoplanet atmospheric retrieval literature has already demonstrated the importance of identifying, exploring, and attempting to mitigate against degeneracies between atmospheric parameters

(e.g. [Feng et al., 2016](#); [Line and Parmentier, 2016](#); [Caldas et al., 2019](#)). These concerns grow larger for terrestrial exoplanets with a greater range of plausible bulk and trace gas atmospheric compositions and for data with larger uncertainties that allow for more atmospheric solutions to describe the same dataset. As a result, it will become increasingly important to attempt to use prior knowledge about the physical and chemical processes within planetary atmospheres that have been refined in the Solar System to attempt to discriminate between and/or rule out possible atmospheric composition scenarios. SMARTER has been designed to integrate and investigate such a physically augmented atmospheric retrieval approach, which will necessarily add more physical complexity to the retrieval modeling suite.

Figure 6.24 and Figure 6.25 show two different approaches for conducting a physically augmented terrestrial exoplanet atmospheric retrieval. The first approach is shown in Fig. 6.24 and represents a chemically self-consistent terrestrial retrieval approach, similar to the equilibrium chemistry retrieval approaches implemented in [Kreidberg et al. \(2015\)](#). In this case, we envision that the surface flux boundary conditions for a 1D photochemical model will be retrieved in the state vector instead of the gas mixing ratios. These surface fluxes will be used to generate atmospheric profiles that are consistent with photochemistry, before passing them to the radiative transfer model and on through the traditional retrieval framework. This approach seeks to determine the best-fitting spectrum that is consistent with photochemistry. However, a downside of this approach is that it will *only* consider atmospheres that are consistent with the simulated physics and chemistry, meaning that unexpected atmospheric compositions or missing physics (e.g. chemical reaction pathways) in the models will not be properly interpreted. The second approach, shown in Fig. 6.25, attempts to mitigate this problem by starting with a standard data-driven atmospheric retrieval. In this case, the retrieved atmospheric state (TP profile and gas mixing ratios) posteriors are provided as inputs into a second Bayesian inference problem that attempts to use physics and chemistry to reproduce the atmospheres retrieved from the spectrum. This approach has the benefit of “letting the data speak for itself,” but it is unclear whether uncertainties and degeneracies in the initially retrieved atmospheric state can be propagated into the second retrieval without losing information. We note that the “retrieval-on-a-retrieval” concept is analogous to how a multi-wavelength exoplanet lightcurve

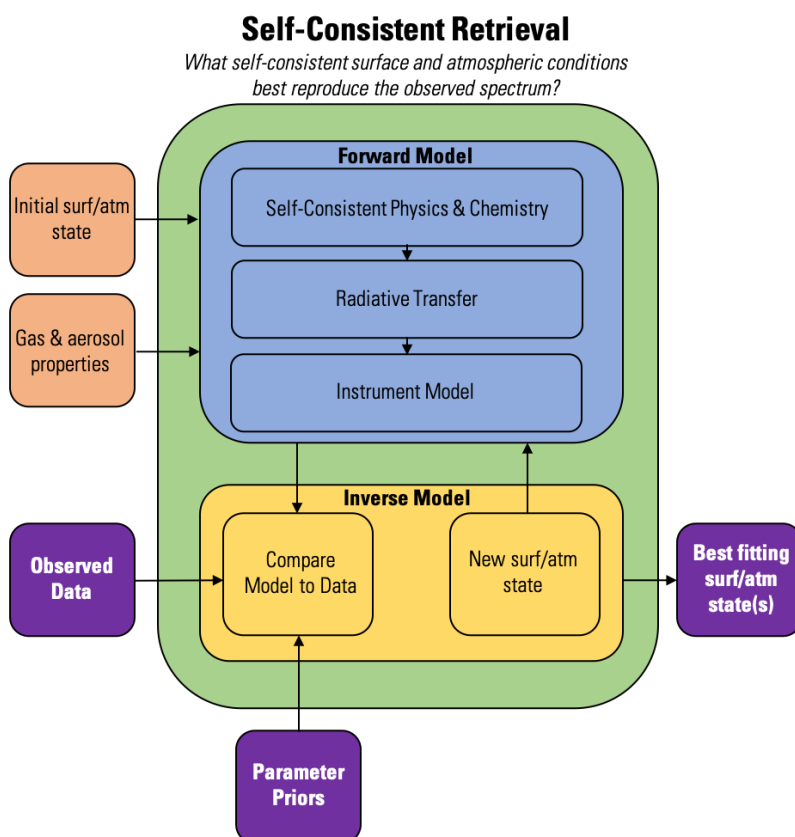


Figure 6.24 Diagram of a physically and chemically self-consistent retrieval. An initial guess for the gas fluxes from the surface is passed to the forward model, which first calculates the atmospheric composition from climate and photochemical models, and then produces a spectrum to compare with the observed data.

is initially fit using a transit model (e.g. [Mandel and Agol, 2002](#)) to retrieve transit depths at each wavelength, and then these transit depths are provided (usually via an email exchange) to an atmospheric retrieval code to fit the transmission spectrum and retrieve the atmospheric composition. Additional work is required to implement these two approaches into the SMARTER framework and assess their efficacy for retrieving accurate and unbiased results.

This physically augmented atmospheric retrieval framework may be the first step towards a viable implementation of the biosignature detection framework discussed in [Catling et al. \(2018\)](#); [Walker et al. \(2018\)](#). Ultimately, we require an astrobiological modeling tool that can synthesize

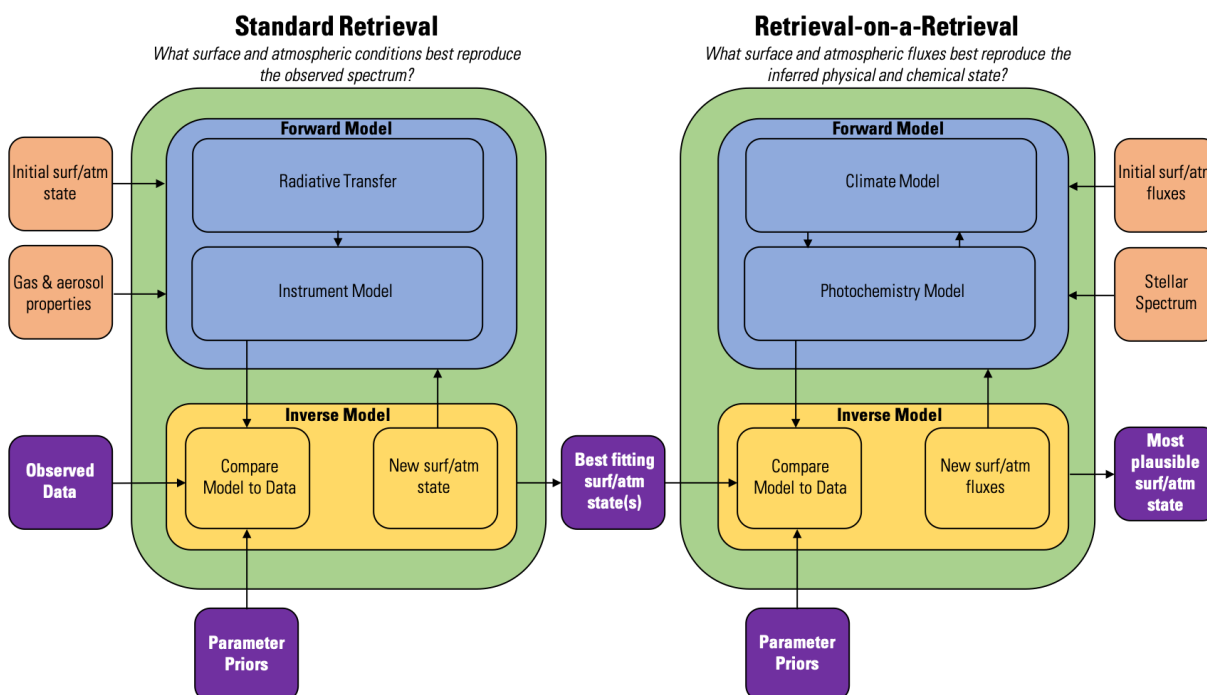


Figure 6.25 Diagram of a “retrieval-on-a-retrieval” approach to physically interpreting the retrieved atmospheric composition. An initial data-driven “standard” retrieval is used to infer constraints on the atmospheric composition. These atmospheric constraints are then passed into a second “retrieval” that attempts to fit climate and photochemistry models to the inferred atmospheric vertical profiles.

astronomical data with knowledge of planetary processes to detect surface and atmospheric biosignatures and rule out false positives. The step of ruling out false positives is so fundamental (e.g. see [Meadows et al., 2018b](#)) because it challenges us to consider all of the plausible abiotic mechanisms that could produce a given signal, before we declare a successful biosignature has been detected. To meet this lofty goal we need a thorough census of atmospheric gases that can be interpreted through the lens of photochemistry and surface outgassing. The physically augmented retrieval approaches discussed above offer two practical methods for developing a biosignature assessment framework for the future.

Chapter 7

RETRIEVING TRAPPIST-1E'S ATMOSPHERIC COMPOSITION WITH JWST

Portions of this chapter are in preparation for publication in the The American Astronomical Society Journals in collaboration with Victoria S. Meadows, Michael R. Line, and David Crisp, and Tyler D. Robinson.

Summary

The TRAPPIST-1 system will provide an excellent opportunity to study the atmospheres of Earth-size planets with the upcoming James Webb Space Telescope (*JWST*). However, numerous challenges lie ahead for correctly interpreting their atmospheres given the broad range of possible secondary atmospheres that could explain what limited quality spectra we are able to obtain. To help understand and address these challenges, we have developed a terrestrial exoplanet atmospheric retrieval model built on the physically robust radiative transfer code used extensively by the Virtual Planetary Laboratory to simulate terrestrial exoplanet spectra. We employed this retrieval model to study the atmosphere of TRAPPIST-1e as if it were observed by an intensive *JWST* transmission spectroscopy campaign, assuming several different plausible habitable and uninhabitable archetypal planetary states. We find that the CH_4 – CO_2 disequilibrium biosignature pair suggested by [Krissansen-Totton et al. \(2018a\)](#) may also be inferred if the planet possesses a modern Earth-like atmosphere due to the increased photochemical lifetime of CH_4 . However, enhanced CH_4 absorption features only worsen the already dim prospects for detecting H_2O , which could complicate the interpretation of planetary habitability. We also found that stacking 25 transits could lead to a weak detection of CO_2 in a cloudy Venus-like atmosphere, or alternatively reveal O_2 – O_2 CIA in an oxygen-dominated post-runaway atmosphere. We found that adding refraction to our transmission retrievals modifies the known degeneracy between reference radius and pressure allowing for an additional mode of interpretation wherein the spectral continuum is controlled by refraction rather than an opaque medium. We identified a strong implicit prior on the mean molecular weight caused by the common assumption of a background gas that can significantly affect the retrieved temperature via scale height degeneracies. Finally, we verified that the TRAPPIST-1e planet mass constraints from TTVs are sufficient to enable dozens of observed transmission spectra to be stacked without incurring additional posterior uncertainties on atmospheric parameters.

7.1 Introduction

The TRAPPIST-1 system (Gillon et al., 2017b,a; Luger et al., 2017c) offers one of the best opportunities to characterize temperate, terrestrial exoplanets in the era of *JWST* (Barstow and Irwin, 2016; Morley et al., 2017; Batalha et al., 2018; Lustig-Yaeger et al., 2019b). This fortuitously well-aligned system (e.g. Luger et al., 2017a) of at least seven Earth-size planets transiting their Jupiter-size M8 host star features relatively deep transit depths, which enabled the system to be detected from the ground and will soon enable their atmospheres to be studied. The seven known planets span a range in incident stellar flux that includes and brackets the habitable zone (HZ), so that the atmospheric characterization of this system may offer the second ever empirical data point on the HZ after the Solar System, and the first ever for an M dwarf system. However, in-depth atmospheric characterization for the TRAPPIST-1 planets is estimated to require a significant observing time investment with *JWST*. The interpretation of their atmospheres, should they have them, promises to be challenging, both theoretically and observationally. In this work, we focus on the retrieval of information about secondary atmospheres for the innermost HZ planet, TRAPPIST-1e, and assess the outlook for constraining the planet’s atmospheric composition.

Numerous studies have explored the detectability of different archetypal atmospheric compositions for the TRAPPIST-1 planets with *JWST*. Many of the postulated atmospheric compositions for the TRAPPIST-1 planets possess CO₂, which has strong near- and mid-IR absorption features that *JWST* may be capable of detecting (Morley et al., 2017; Batalha et al., 2018; Wunderlich et al., 2019; Lustig-Yaeger et al., 2019b; Fauchez et al., 2019). For the temperate TRAPPIST-1 planets, this may be the best method to confirm the existence of an atmosphere. However, forecasting *JWST*’s ability to characterize the atmospheric composition of these planets—beyond the presence of CO₂—depends heavily on the *a priori* unknown nature of their atmospheres. Climate and photochemical models of the TRAPPIST-1 planets are particularly useful for making physically informed predictions about the surface and atmospheric temperature structure, clouds, and trace gas composition of M dwarf planet upper atmospheres (Lincowski et al., 2018; Wunderlich et al., 2019; Meadows et al., 2020), which can inform observing strategies on wavelength range,

spectral resolution, and instrument choice. However, little work has been done to explore the challenges and degeneracies associated with performing full atmospheric retrievals for TRAPPIST-1e, or terrestrial exoplanets in general, using spectra generated across a diverse set of atmospheric compositions with photochemically self-consistent modeling.

A few key atmospheric retrieval studies have shaped the outlook for characterizing terrestrial exoplanet atmospheres with *JWST*, but these are limited to only a few specific possible atmospheric compositions and have tended to focus exclusively on biosignatures. Prior to the discovery of the outer four TRAPPIST-1 planets, [Barstow and Irwin \(2016\)](#) assessed the ability to characterize the atmospheres of TRAPPIST-1 b, c and d if they possess a modern Earth atmospheric composition. They found that at least 30 transits using both the NIRSpec Prism and MIRI LRS would be required to detect O₃. Although anticipating habitable Earth-like conditions for the inner TRAPPIST-1 planets may be unrealistic, [Barstow and Irwin \(2016\)](#) demonstrated how difficult detecting biosignatures from oxygenic photosynthesis will be with *JWST*, even for the most optimistic planetary system. [Krissansen-Totton et al. \(2018a\)](#) performed simulated retrievals of TRAPPIST-1e to investigate the potential to characterize an anoxic biosphere, like the Earth once had during the Archean geologic eon ([Krissansen-Totton et al., 2018b](#)). They found that the combination of CH₄ and CO₂—a disequilibrium biosignature pair—may be detectable in a stacked transmission spectrum from ~10 observed transits. These studies have indicated that biosignature detection in the TRAPPIST-1 system, while technically possible, will require an extensive observational campaign along with an exhaustive consideration of all atmospheric conditions that can plausibly be retrieved.

A significant amount of work has been done to develop atmospheric retrieval models capable of interpreting the observed transmission and thermal emission spectra of gaseous exoplanets ([Irwin et al., 2008](#); [Madhusudhan and Seager, 2009](#); [Line et al., 2013](#); [Benneke and Seager, 2013](#); [Waldmann et al., 2015](#); [Cubillos et al., 2016](#); [Lavie et al., 2016](#); [Howe et al., 2017](#); [Kitzmann et al., 2019](#); [Zhang et al., 2019](#); [Damiano and Hu, 2020](#)). Although many of these models can simulate high mean molecular weight atmospheres, the lack of terrestrial exoplanet spectra to date has led to relatively little work in this regime. At nearly every turn, studies have shown that physically

oversimplifying atmospheres for the purpose of spectral inference can introduce retrieval biases due to 1D, 2D, and 3D atmospheric effects. This has been reported for tidally locked planets that may have complicated day-night effects at the terminator (Caldas et al., 2019; MacDonald et al., 2020; Pluriel et al., 2020), vertical and spatial pressure-temperature profile effects (Feng et al., 2016; Rocchetto et al., 2016; Blecic et al., 2017; Taylor et al., 2020), and cloud distribution and composition effects (Line and Parmentier, 2016; Mai and Line, 2019). Although the upcoming *JWST* signal-to-noise on terrestrial exoplanet spectra is likely insufficient to warrant fully 3D atmospheric and spectral modeling for inference, these studies indicate that retrievals must move towards more physically rigorous atmospheric prescriptions to avoid biased inferences. Moreover, the expected compositional diversity amongst terrestrial environments is an area ripe for oversimplification, which deserves focus in light of upcoming *JWST* observations.

In this work, we describe a new terrestrial-focused exoplanet retrieval model and confront some of the opportunities and challenges associated with retrieving the atmospheric composition of TRAPPIST-1e with *JWST*. Our retrieval model leverages a physically sophisticated radiative transfer model that is admittedly slow to evaluate. This presents challenges for inference, but also creates opportunities to test common retrieval assumptions. We focus our retrieval efforts on simulated *JWST* transmission spectra that could result from a relatively intensive observing strategy (e.g. from the T-1 Community Initiative; Gillon et al., 2020). However, we examine multiple different intrinsic planet environmental states spanning from habitable modern and Archean Earth-like atmospheres (Meadows et al., 2020, *in prep*) to uninhabitable post-ocean-loss Venus-like or oxygen-dominated worlds (Lincowski et al., 2018). Our retrieval tests are the first of their kind to use synthetic *JWST* data generated from terrestrial atmospheric compositions that are consistent with climate and photochemical predictions, enabling an *a posteriori* assessment of different plausible atmospheric archetypes. In addition to retrieving various types of secondary atmospheres, we investigate the sensitivity to atmospheric bulk composition and mean molecular weight, refraction, vertical gas and thermal profiles, and planet mass constraints to assess retrieval bias.

We describe our simulated *JWST* observations of TRAPPIST-1e next in Section 7.2. We then detail our new terrestrial exoplanet atmospheric retrieval model in Section 7.3. In Section 7.4,

we present our results on the characterization of TRAPPIST-1e’s atmosphere and the assessment of biases for the retrieval of secondary atmospheres. We discuss our findings in Section 7.5, and conclude in Section 7.6.

7.2 Simulated JWST Data

We use the PANDEXO noise model to produce synthetic *JWST* transmission spectrum observations of TRAPPIST-1e (Batalha et al., 2017; Batalha et al., 2018). At its core, PANDEXO uses PANDEIA (Pontoppidan et al., 2016)—the STScI’s official exposure time calculator for *JWST*—and pysynphot (STScI Development Team, 2013) to simulate *JWST* exoplanet transmission spectroscopy. We refer the reader to Batalha et al. (2017) for a thorough description of PANDEXO. We build off the findings of previous works that have indicated optimal *JWST* observing modes to characterize the atmosphere of TRAPPIST-1e (Batalha et al., 2018; Lustig-Yaeger et al., 2019b). Specifically, we primarily consider transmission spectrum observations between $1 - 5 \mu\text{m}$ using the NIRSpec Prism (Bagnasco et al., 2007; Ferruit et al., 2014) with the partial saturation strategy outlined in Batalha et al. (2018). We do not consider emission spectroscopy in this work due to the low expected emission at these wavelengths and low precision expected for HZ exoplanets with *JWST* (Barstow et al., 2016; Lustig-Yaeger et al., 2019b,a). Our retrieval approach involves fitting the transmission spectra at the native spectral resolution of each *JWST* instrument, so we do not by default bin the observed spectra to lower resolution. We adopt the common strategy for simulated retrieval studies of generating the synthetic data centered at the “truth” (mean of the Gaussian uncertainties), so the inferred posteriors are not biased by a single noise instance (Feng et al., 2018; Mai and Line, 2019; MacDonald et al., 2020), but rather demonstrate any biases in the retrieval model. Furthermore, we assume that *JWST* operates nominally, with no time or wavelength dependent systematic effects, and assume no stellar variability (Morris et al., 2018a) or spectral contamination (e.g. Rackham et al., 2018; Zhang et al., 2018; Morris et al., 2018c; Wakeford et al., 2019), which are important but beyond the scope of this work.

We apply PANDEXO noise modeling to the TRAPPIST-1 transmission spectrum models from Lincowski et al. (2018) and Meadows et al. (2020, *in prep*), which possess climate and photo-

chemically self-consistent atmospheres. The atmospheric compositions under consideration are all *secondary* atmospheres with high mean molecular weights compared to hydrogen and helium dominated primary atmospheres, but vary widely in their bulk composition and habitability. We consider clear and cloudy CO₂ dominated Venus-like atmospheres and a desiccated O₂ dominated atmosphere from [Lincowski et al. \(2018\)](#), both of which are uninhabitable atmospheric composition predictions based on extreme water loss during the super-luminous pre-main-sequence phase of TRAPPIST-1 (e.g. [Luger and Barnes, 2015](#)). We also consider numerous atmospheric states for inhabited biospheres from [Meadows et al. \(2020, in prep\)](#), including a clear and cloudy modern Earth-like planet with the contemporary surface methane flux, an Archean Earth-like planet with no biogenic oxygen and enhanced biogenic methane, and an Archean sulfur organic biosphere. As a final comparison case, we consider TRAPPIST-1e with the modern Earth atmospheric composition as a baseline for calculations that neglect photochemistry. The spectral models for TRAPPIST-1e are generated assuming the mean planet mass ($M_p = 0.772 M_\oplus$) and radius ($R_p = 0.910 R_\oplus$) reported in [Grimm et al. \(2018\)](#). Note that this radius constraint is used as the solid body surface radius in the input spectral models. We discuss the implications of this in Section 7.5. We assume that the planetary environment and therefore the observed spectrum do not vary during or between periods of observation.

As a baseline for comparison, we assume that TRAPPIST-1e has 25 transits observed. Accordingly, we produce an input noised spectrum for each atmospheric composition considered, where each represents one single instance of a possible planetary environment. The ensemble of spectra with expected levels of noise allow us to examine the atmospheric similarities and differences we may expect to retrieve with *JWST*, given the spread of different bulk atmospheric states, habitability, and presence of biosignatures. Figure 7.1 shows a representative transmission spectrum of TRAPPIST-1e with a photochemically self-consistent modern Earth-like atmosphere observed with the *JWST* NIRSpec Prism for 25 transits. The spectrum is shown at the native resolution of the NIRSpec Prism and therefore displays relatively large, but numerous, error bars. Although our primary retrieval assessment maintains a uniform number of observed transits for each atmospheric case considered (25 transits), we also perform retrievals of the clear modern Earth-like case

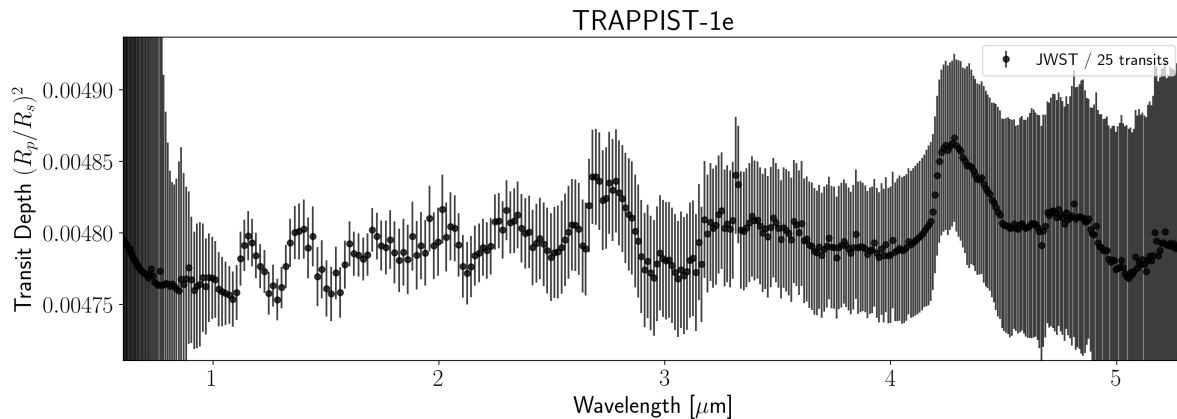


Figure 7.1 Simulated transmission spectrum observations of TRAPPIST-1e with the *JWST* NIR-Spec Prism assuming the planet possesses a photochemically self-consistent modern Earth-like atmosphere. The noise calculations assume that TRAPPIST-1e had 25 transits observed and coadded.

with a varied number of observed transits to examine the evolution of the posterior atmospheric constraints with increasingly higher precision data.

7.3 Retrieval Model

In this work, we use a novel atmospheric retrieval code developed specifically to interpret the spectra of terrestrial exoplanets. Our retrieval model differs from other state-of-the-art retrieval codes primarily in the treatment of radiative transfer, for which we use the robust line-by-line Spectral Mapping Atmospheric Radiative Transfer (SMART) model (Meadows and Crisp, 1996; Crisp, 1997). The SMART Exoplanet Retrieval (SMARTER) code faces a number of computational challenges owing to the rigor of the radiative transfer, so it has been developed to use a variety of different Bayesian retrieval approaches to solve the inverse problem. In this section, we first describe the SMARTER forward model used for *JWST* transmission spectroscopy (§ 7.3.1), and then describe the SMARTER inverse modeling suite (§ 7.3.2).

7.3.1 Forward Model

The SMARTER forward model was designed with a modular approach capable of leveraging the many radiative transfer capabilities of SMART to perform exoplanet retrievals. As a result, SMARTER can run atmospheric retrievals using transmission, thermal emission, and reflected-light spectra. However, in this work we will focus exclusively on the transmission spectroscopy capability as it is most relevant to the upcoming atmospheric characterization of rocky exoplanets with *JWST*.

Radiative Transfer with SMART

A complete description of SMART is provided in Chapter 6 Section 6.2.2. Briefly, SMART is a one-dimensional, line-by-line, multi-stream, multiple-scattering radiative transfer model (developed by D. Crisp; [Meadows and Crisp, 1996](#); [Crisp, 1997](#)). SMART's multiple-scattering calculations account for aerosol scattering phase functions, allowing for rigorous treatments of clouds and hazes. At its core, SMART uses the DIScrete Ordinate Radiative Transfer (DISORT; [Stamnes et al., 2017](#)) model to solve the radiative transfer equation.

SMART features a novel modification to solve the equation of radiative transfer that significantly improves computations efficiency over the traditional line-by-line approach. Before performing any radiative transfer, SMART pre-processes the all of the input optical properties on their native, fully resolved, high-resolution wavenumber grid to identify wavenumbers with sufficiently similar optical properties along the entire optical path ([Meadows and Crisp, 1996](#)). SMART then combines these points into a single, broadband, quasi-monochromatic bin for which it is only required to solve the radiative transfer equation once ([West et al., 1990](#)). As a result of this binning procedure, SMART can reduce the number of calls to DISORT by a factor of ~ 10 -1000 depending on the user-specified binning tolerance. After the radiative transfer calculations are completed, SMART maps the spectral results back to the full resolution wavenumber grid, and outputs them for the user.

Recent upgrades to SMART added transmission spectra capabilities for transiting exoplanets

(Misra et al., 2014; Robinson, 2017a). The transmission spectrum calculations may be performed using one of three methods (Robinson, 2017a): (1) a simple geometric approach using straight paths, (2) a ray tracing method that includes the refraction of stellar rays passing through the atmosphere, and (3) Monte Carlo numerical integration to account for forward-scattering off aerosol particles. The Monte Carlo scattering is extremely computationally expensive, so unless otherwise noted, we use the ray tracing method, and we explore the effects of refraction in Section 7.4.

SMART depends on the Line-By-Line ABsorption Coefficient (LBLABC) code to calculate molecular vibrational-rotational transition absorption coefficient input files for the radiative transfer (developed by D. Crisp; Meadows and Crisp, 1996). LBLABC takes information about the input atmospheric state—the TP profile, vertical gas mixing ratio, mean molecular weight of the atmosphere, and radius and surface gravity of the planet—and combines it with HITRAN line-parameter and isotope information to calculate the gas absorption coefficients as a function of pressure, temperature, and wavenumber. LBLABC uses nested spectral grids to fully resolve the narrow cores of each absorption line and to resolve the far wings of the lines at every atmospheric layer over a very large range of line-center distances. Line contributions are calculated up to 1000 cm^{-1} from the line center for H_2O and CO_2 to capture the full wings of strong lines; all other gases have a 50 cm^{-1} wing cutoff to speed up calculations where contribution from the far wings is negligible. LBLABC employs different line profiles as a function of distance from the line core. LBLABC uses a Voigt line shape (Humlíček, 1982) within 40 Doppler half-widths of the line center, and a van Vleck-Weisskopf profile (Van Vleck and Weisskopf, 1945) beyond that for all gases except H_2O , CO_2 , and O_2 . Empirically determined χ correction factors are used in the far wings for H_2O (Clough et al., 1989) and CO_2 (Fukabori et al., 1986; Perrin and Hartmann, 1989; Pollack et al., 1993) to reproduce the effects of quantum-mechanical line mixing, a collisional effect that occurs for overlapping lines of a given molecular species (Rao, 2012). For O_2 , the super-Lorentzian line shape from Hirono and Nakazawa (1982) is used with an exponent of 1.958. LBLABC features the ability to simulate line absorption coefficients for fixed temperature offsets from the given TP profile, which SMART can interpolate between for added flexibility when deviating from an initial temperature structure. This can be used to generate a single set

of line absorption coefficients that may be used over a broad range of atmospheric temperature structures that are explored by the retrieval, without requiring LBLABC to be rerun.

SMART was developed for Solar System terrestrial planet atmospheres and has been rigorously validated on the Earth (Robinson et al., 2011) and Venus (Meadows and Crisp, 1996; Arney et al., 2014). SMART has been used to simulate the spectra of a diversity of exoplanet environments, including the spectra of M dwarf habitable planets (Segura et al., 2003, 2005; Meadows et al., 2018a) and uninhabitable planets (Lincowski et al., 2018, 2019), habitable haze-enshrouded exoplanets (Arney et al., 2016, 2017), and high-altitude clouds in the H₂-dominated mini-Neptune exoplanet, GJ1214b (Charnay et al., 2015b). SMART also functions as the radiative transfer core for the VPL Climate Model (Robinson and Crisp, 2018; Lincowski et al., 2018).

Baseline Transmission Forward Model

We construct a baseline transit transmission spectroscopy forward model designed in preparation for terrestrial exoplanet transit observations with *JWST*. This forward model calls SMART for the radiative transfer, but it also makes numerous simplifying assumptions to limit the volume of parameter space that the retrieval must traverse. First, we assume that the 1D atmosphere is isothermal with evenly mixed gas abundance profiles. Although such an atmosphere is unphysical, and indeed mixing ratio biases have been reported in retrieval tests of hot Jupiter atmospheres (Rocchetto et al., 2016), transmission spectroscopy is only weakly sensitive to the atmospheric thermal structure (Kempton et al., 2017), and it remains unclear if *JWST* will have the precision to justify TP profile retrievals for rocky exoplanets. We will revisit this discussion in Section 7.5.

Constraining the bulk composition of terrestrial exoplanet atmospheres will be a major challenge for the foreseeable future. This is primarily due to the lack of absorption features for inert gases (e.g. H₂, He, N₂, Ar, etc.) even at large abundance, the saturation of absorption features for other gases even at small to moderate abundance (e.g. CO₂, CH₄, H₂O), and the multi-dimensional scale height degeneracy between the atmospheric mean molecular weight, temperature, and surface gravity. Current, retrieval models tend to confront the problem of bulk atmospheric composition by treating the mean molecular weight in one of two ways (Barstow and Heng, 2020). The first

and most commonly used approach derives the mean molecular weight from the gases in the atmosphere by assuming the planet has a primary atmosphere with a solar composition mixture of H₂-He as the background gas (e.g. [Benneke and Seager, 2013](#); [Greene et al., 2016](#); [Rocchetto et al., 2016](#); [Irwin et al., 2020](#); [Lavie et al., 2017](#); [Mollière et al., 2019](#)) or assuming the planet has a secondary atmosphere with N₂ as the assumed background gas (e.g. [Barstow et al., 2016](#); [Krissansen-Totton et al., 2018a](#); [Changeat et al., 2019](#); [Barstow et al., 2020](#)). An alternative approach is to directly fit for the atmospheric mean molecular weight to account for any spectrally inactive bulk gases (e.g. [Tremblay et al., 2020](#)).

Our forward model allows for the existence of a filler gas that enables the mean molecular weight to be calculated self-consistently or the mean molecular weight can be retrieved as a free parameter. Any combination of gases in the HITRAN 2016 database ([Gordon et al., 2017](#)) can be used as the filler gas or included as a free parameter in the state vector. The VMR of the filler gas is given by,

$$f_{\text{fill}} = 1 - \sum_{i=1}^N f_i \quad (7.1)$$

where f_i is the VMR of the i th gas, N is the total number of non-filler gases included in the atmosphere, and we have assumed that these constitute the entire volume of the atmosphere. This implicitly enforces the commonly used prior that requires the VMR of all gases sum to unity. The filler gas must be set prior to starting the retrieval and cannot be exchanged with another gas during the retrieval. For our baseline model in this work, we use N₂ as the filler gas and calculate the mean molecular weight self-consistently, however, we explore the implications of this assumption in Section 7.4. The bottom of the atmosphere can be specified by a fixed surface pressure or it can be fit for as a free parameter.

In addition to the atmospheric parameters, we also allow for the planet radius and planet mass to be included in the retrieval as free parameters. The radius included in the transit forward model is the planet’s solid body/surface radius. This surface radius corresponds to the surface pressure (whether or not it is fixed or free), and represents the “zero point” above which the atmospheric absorption extends radially outward. The planet mass is not an easily measured quantity for a

transiting terrestrial exoplanet (as it requires either extremely precise RV or TTVs), nor is it easy to constrain from transmission spectra. However, the planet mass does affect the scale height of the atmosphere, via the surface gravity, and is therefore degenerate with both the temperature structure and the mean molecular weight of the atmosphere. This allows mass to vertically increase or decrease the scale of all absorption features in the transmission spectrum, thereby potentially propagating uncertainty into the inferences of other parameters (Changeat et al., 2019). In our transit forward model, planet mass is converted to surface gravity before it is passed to SMART for the radiative transfer. Note that we do not make a distinction between the mass contained within the solid body of the planet versus the mass of the entire planet, due to the negligible mass of the atmosphere for Earth-like planets¹ (e.g. the atmosphere makes up less than 1 ppm of Earth’s mass) compared to realistic constraints on terrestrial exoplanet masses.

State Vector & Priors

For the purposes of this retrieval study, we use a nominal eight parameter state vector which includes the isothermal temperature T , the surface pressure P , the solid-body planet radius R_p , and the five gases: H_2O , CO_2 , CH_4 , O_2 , and O_3 . All other parameters are specified and held fixed. In Section 7.4, we also explore the effects of fitting for the planet mass M_p , mean molecular weight μ , and we examine the inclusion of a few additional gases, most notably CO. Table 7.1 shows the free retrieval parameters under consideration in this work and their respective prior probabilities. All parameters have uninformative uniform priors except for the planet mass for which we use a Gaussian prior with mean and standard deviation reported in Grimm et al. (2018).

7.3.2 Inverse Model

We employ a multi-pronged inverse modeling approach similar to that of the CHIMERA retrieval model (Line et al., 2013, 2014; Kreidberg et al., 2015; Feng and Robinson, 2017; Tremblay et al., 2020). Each inverse model attempts to solve Bayes’ Theorem to characterize the shape of the poste-

¹We also note that gaseous planets have almost all of their mass below the $\tau = 1$ surface.

Table 7.1. Retrieval parameters and their respective priors.

Parameter	Description	Prior
$\log(\text{H}_2\text{O})$	H_2O volume mixing ratio	$\mathcal{U}(-10, 0)$
$\log(\text{CO}_2)$	CO_2 volume mixing ratio	$\mathcal{U}(-10, 0)$
$\log(\text{CH}_4)$	CH_4 volume mixing ratio	$\mathcal{U}(-10, 0)$
$\log(\text{O}_2)$	O_2 volume mixing ratio	$\mathcal{U}(-10, 0)$
$\log(\text{O}_3)$	O_3 volume mixing ratio	$\mathcal{U}(-10, 0)$
T_0	isothermal temperature (K)	$\mathcal{U}(100, 400)$
$\log(P_0)$	surface pressure (Pa)	$\mathcal{U}(2, 6)$
R_0	solid-body surface radius ($R_\oplus$)	$\mathcal{U}(0.85, 0.95)$
M_p	planet mass ($M_\oplus$)	$\mathcal{N}(0.772, 0.077)^{\text{a}}$
μ	mean molecular weight (g/mol)	$\mathcal{U}(10, 45)$

Note. — $\mathcal{U}$ and $\mathcal{N}$ denote uniform and normal distributions, respectively.

^aGrimm et al. (2018)

rior distribution. However, each method makes different assumptions that differ in their numerical approach and affect both the statistical robustness of the retrieved posteriors and the number of calls to the forward model required to reach convergence. As a result, four inverse models are currently available and can be used with any forward model within the SMARTER framework. These retrieval models are: Optimal Estimation (OE) using `scipy.optimize.minimize` (Virtanen et al., 2019), Nested Sampling (NS) using MULTINEST (Feroz and Hobson, 2008; Feroz et al., 2009, 2013) via the PYMULTINEST Python wrapper (Buchner et al., 2014), Markov Chain Monte Carlo (MCMC) using EMCEE (Foreman-Mackey et al., 2013), and a novel active learning approach called APPROXPOSTERIOR (AP; Fleming and VanderPlas, 2018; Fleming et al., 2020). These methods are described in greater detail in the preceding chapter.

In this work we primarily use nested sampling to derive posterior distributions. Nested sampling has been growing in popularity within the exoplanet atmospheric retrieval community, with most of the common retrieval codes now using it as their preferred inverse model (e.g. Benneke and Seager, 2013; Line and Parmentier, 2016; Lavie et al., 2017; Rocchetto et al., 2016; MacDonald and Madhusudhan, 2017; Barstow et al., 2020). This broad adoption stems from its efficiency—requiring $\sim 10\times$ fewer forward model evaluations compared with traditional MCMC sampling—and its applicability in model selection problems, both of which are important for atmospheric retrievals. The MULTINEST algorithm iterates forward through nested shells of likelihood, progressively focusing on smaller and smaller regions of the prior volume. The procedure terminates once the marginal likelihood (evidence) is determined to some user-specified precision (we use the recommended value of 0.5 in log-evidence). We also use the evidence to perform Bayes factor and detection significance calculations for each molecule included in the retrieval following the approach outlined in Benneke and Seager (2013).

Active Learning with APPROXPOSTERIOR

The computational expense associated with repeatedly calling a radiative transfer model within a typical retrieval algorithm motivates the development of retrieval frameworks that use machine learning to increase the speed of performance. Indeed, machine learning approaches are beginning

to show promise for increasing the speed of atmospheric retrieval inference problems (e.g. [Zingales and Waldmann, 2018](#); [Cobb et al., 2019](#)). At present, these methods rely on generating large spectral datasets to train neural networks, which can perform atmospheric retrievals much faster than traditional physical models because the neural network replaces the radiative transfer model used to compute the spectrum. However, the reliance on a single massive training set is a potential shortcoming of these previous approaches considering how often new spectroscopic laboratory data (e.g. [Fortney et al., 2019](#)) or additional physical complexity must be added to the retrieval forward model, thereby requiring a new training set to be generated. We elected to mitigate the computational cost of our physically robust forward model using an alternative, but complementary approach, by performing approximate Bayesian inference using the APPROXPOSTERIOR code ([Fleming and VanderPlas, 2018](#); [Fleming et al., 2020](#)).

The open-source APPROXPOSTERIOR Python code employs an active learning algorithm to iteratively improve its predictive accuracy by expanding its training set while minimizing the expense of acquiring new samples. Unlike previous atmospheric retrieval machine learning approaches that learn on simulated spectra to replace the expensive radiative transfer model, AP learns the *posterior density* using a supervised learning Gaussian process (GP) regression active learning model. APPROXPOSTERIOR uses the Python code, GEORGE, to perform fast and efficient GP regression ([Ambikasaran et al., 2015](#)).

Given an initial training set of retrieval state vectors paired with their corresponding log posterior probabilities, AP will train a GP using maximum likelihood estimation to optimize the GP hyperparameters. AP then selects a new point to add to the training set by optimizing an objective function that attempts to identify regions of parameter space where the GP mean predicts high posterior probability, where model predictions are likely to match the observations, but that also have a large GP uncertainty so that augmenting the training set will optimally improve the predictive accuracy. Two approaches are implemented for the APPROXPOSTERIOR point selection utility function: the Bayesian Active Learning for Posterior Estimation (BAPE) algorithm from [Kandasamy et al. \(2017\)](#) and the Adaptive Gaussian process approximation for Bayesian inference with expensive likelihood functions (AGP) method from [Wang and Li \(2018\)](#). The true forward

model is evaluated for each point that is selected and the new likelihood value is added to the training set. After m new points have been added to the training set, MCMC is used to sample the GP surrogate model’s posterior distribution. The algorithm then proceeds through the process of selecting another m new training points until either $n_{\max}$ iterations (each adding another m points to the training set) are reached or the posterior distributions obtained from MCMC have stabilized.

APPROXPOSTERIOR features a number of user-controlled parameters that dictate how the algorithm builds the training set and optimizes the GP. Our approach for using APPROXPOSTERIOR in our atmospheric retrieval code is to build an initial training set in parallel composed of 100 points per dimension in the state vector. We then elect to have APPROXPOSTERIOR select $m = 50$ new points per iteration for $n_{\max} = 50$ total iterations. We optimize the objective function 5 times for each selected point to increase likelihood of finding the global best design point to add to the training set. Unless otherwise stated, we use an exponential squared kernel for the GP to enforce smoothness in our approximated posterior distributions. We optimize the GP hyperparameters after every 5th new point has been added to the training set, during which we perform 10 randomly initialized optimizations to increase the likelihood of finding the global best selection of GP hyperparameters. We present a validation of APPROXPOSTERIOR in Section 7.4.7 and discuss prospects for its future usage for exoplanet atmospheric characterization in Section 7.5.3.

7.4 Results

Our results begin with a systematic retrieval assessment of molecular detections and abundance constraints that may be achievable in a 25 transit NIRSpec Prism campaign to observe TRAPPIST-1e under various assumptions about the true state of the observed Earth-sized exoplanet (Section 7.4.1). We then compare our retrieved results to the true 1D vertical atmospheric profiles to assess the quality, accuracy, and biases of information that we may be able to learn about this planet atmospheres (Section 7.4.2). Next, we perform experiments to test the sensitivity to atmospheric refraction (Section 7.4.3), planet mass constraints from TTV measurements (Section 7.4.4), assumptions about the atmospheric bulk composition (Section 7.4.5), and random noise instances in retrieval tests (Section 7.4.6). Finally, we validate the APPROXPOSTERIOR active learning model

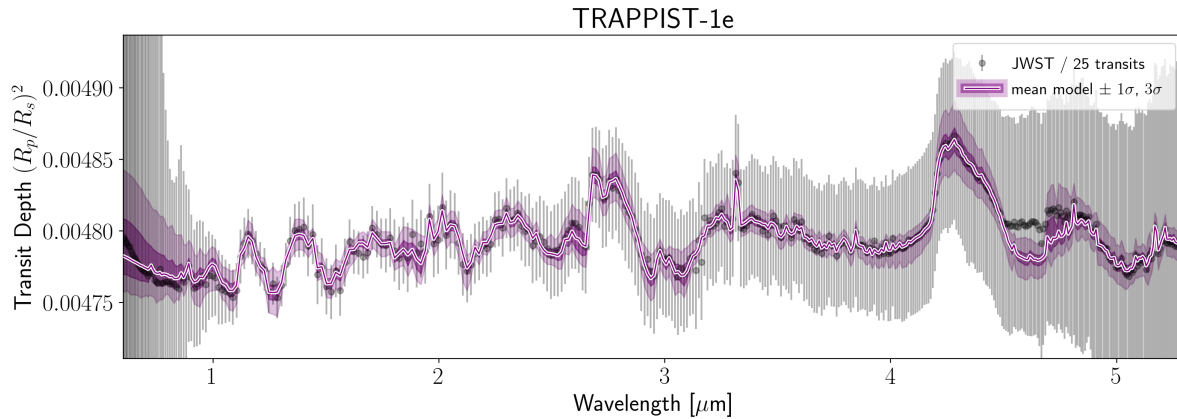


Figure 7.2 Fit to synthetic JWST/NIRSpec Prism data of TRAPPIST-1e assuming the observed planet possesses a clear Earth-like atmosphere. The purple envelopes show the 1σ and 3σ uncertainty about the median retrieved spectral fit.

for atmospheric retrievals (7.4.7).

7.4.1 Atmospheric Composition, Habitability & Biosignatures

We now assess the ability to detect molecules in the atmosphere of TRAPPIST-1e, retrieve their abundance, and distinguish between different archetypal atmospheric composition scenarios with *JWST* including signs of habitability and biosignatures. Figure 7.2 shows the breadth of spectral models that fit a synthetic transmission spectrum of TRAPPIST-1e. The fit to the spectrum is shown at the native resolution of the JWST/NIRSpec Prism instrument and is extremely good despite the simplifications made to our retrieval forward model compared to the capabilities of the full radiative transfer model. The 1σ and 3σ uncertainty envelopes around the median retrieved spectrum show a similar magnitude of model uncertainty at all wavelengths, except below $1\ \mu\text{m}$ where the data are sufficiently low S/N to propagate greater uncertainty into the inferred spectrum.

Figure 7.3 shows marginalized posterior distributions for the spectral fit shown in Fig. 7.2 for the clear modern Earth-like atmosphere. The marginalized posteriors show well-defined peaks for the H_2O , CO_2 , and CH_4 abundances, the isothermal temperature, the planet radius, and the reference pressure. O_2 and O_3 are only constrained by upper limits. The 2D posterior projections

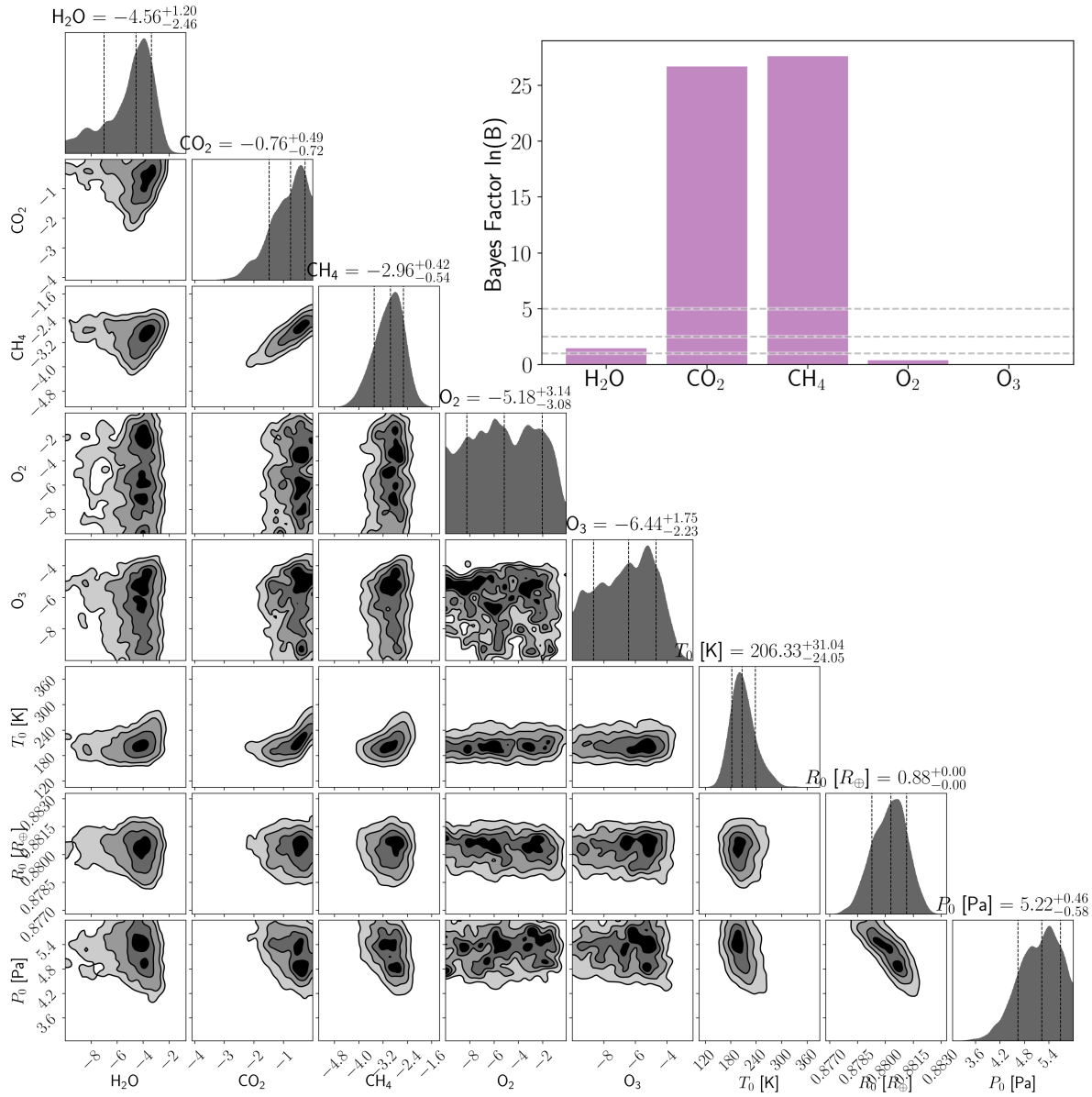


Figure 7.3 Corner plot showing 1D and 2D marginalized posterior distributions for the retrieval of the clear modern Earth-like transmission spectrum shown in Fig. 7.2. The upper right panel shows Bayes factors for models excluding the listed molecule to quantify the evidence provided by the spectrum in support of that molecule. From smaller to larger values of $\ln B$, the horizontal dashed lines show the limits for “weak” (2.1σ), “moderate” (2.7σ) and “strong” (3.6σ) detections, respectively.

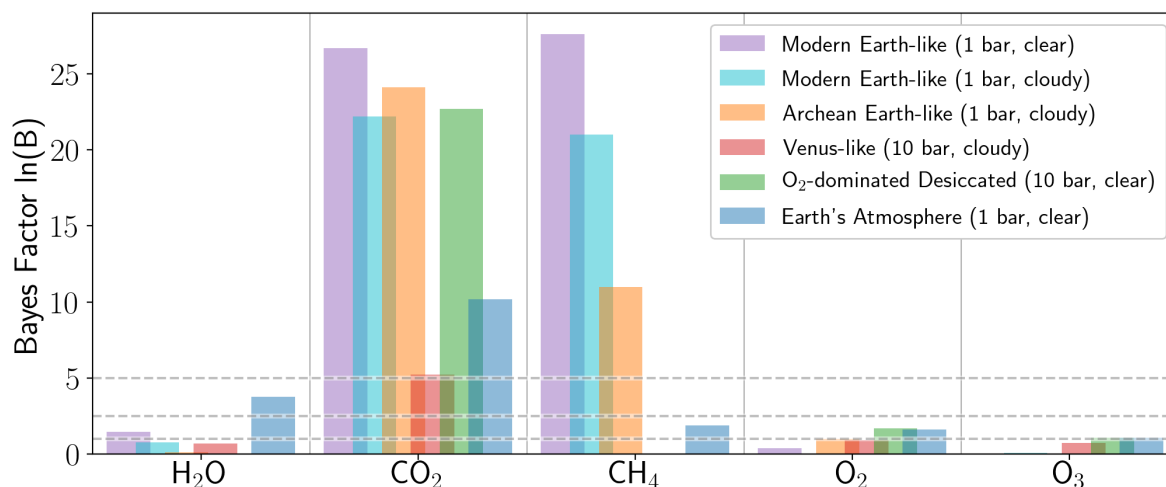


Figure 7.4 Bayes factors showing the evidence in favor of each molecule from retrievals on 25 transits of TRAPPIST-1e assuming multiple plausible atmospheric compositions.

reveal a strong positive correlation between the CO₂ and CH₄ abundance and a negative correlation between reference pressure and radius. CO₂ and CH₄ also show a slight positive correlation with temperature towards large abundance values. The upper right panel of Fig. 7.3 shows Bayes factors from model comparisons indicating the significance of molecular detections, with larger values giving stronger support for the model with the listed molecule included. We find CO₂ and CH₄ to be strongly detected ($>7\sigma$), while H₂O is only weakly detected, and O₂ and O₃ are not detected at all.

We repeated the retrieval analyses shown in Figures 7.2 and 7.3 assuming several additional atmospheric compositions for TRAPPIST-1e that would give rise to different observed transmission spectra. Figure 7.4 shows the resulting Bayes factors supporting the detection of each molecule included in the study, for each atmospheric composition considered. Each of the self-consistent secondary atmospheres considered possesses CO₂, which is strongly detected across the board. For both the modern and the Archean Earth-like atmospheres, biogenic CH₄ is also strongly detected in 25 transits. The actual modern Earth atmosphere, which is included for comparison but is photochemically inconsistent with the TRAPPIST-1 UV environment, has enhanced O₂ and sup-

pressed CH_4 relative to the self-consistent modern Earth-like atmosphere. This unrealistic scenario leads to an under prediction in the strength and detectability of CH_4 and an over prediction in the detectability of H_2O , which would otherwise be obscured by the photochemically enhanced CH_4 . In the Venus-like atmosphere with H_2SO_4 clouds, CO_2 is still strongly detected above the cloud deck, but no other gases warrant inclusion in the model. The post-ocean-loss O_2 -dominated atmosphere has strongly detected CO_2 features and weakly detected O_2 from the 1.06 and 1.27 μm O_2 – O_2 CIA features (see [Schwieterman et al., 2016b](#); [Lincowski et al., 2018](#); [Lustig-Yaeger et al., 2019b](#); [Fauchez et al., 2020](#)).

7.4.2 Comparison to Heterogeneous Vertical Profiles

Figure 7.5 compares our retrieved atmospheric parameters to the underlying true 1D vertical atmospheric profiles from [Lincowski et al. \(2018\)](#) and [Meadows et al. \(2020, in prep\)](#) that were used to generate the synthetic *JWST* data. From the top row of subplots to the bottom, we show cases for the clear Earth-like, cloudy Archean, clear 10 bar CO_2 , and the 10 bar desiccated O_2 atmospheres. Each subplot compares the vertically homogeneous retrieved constraints for one parameter in the state vector to the vertically resolved self-consistent profile shown as a function of pressure.

In general, we find good agreement between our vertically homogeneous atmospheres and the 1D true profiles. The uncertainty in our retrieved isothermal temperature profiles and evenly mixed gas volume mixing ratio profiles tend to bracket the variance seen in the true 1D profiles, with a particularly strong weighting towards low atmospheric pressures between $10 - 10^4$ Pa ($10^{-4} - 0.1$ bars). In many cases, CO_2 and CH_4 are (nearly) evenly mixed in the original 1D atmosphere and their median retrieved abundances agree well with these values. Other gases—such as H_2O , O_3 in the modern Earth-like case, and O_2 in the Archean-like case—exhibit substantially varying vertical profiles, for which the retrieval recovers an upper limit for that gas that is consistent with the maximum along the vertical profile. The isothermal temperature tends towards temperatures in the stratosphere, reflecting the region most sensitive to in transmission spectroscopy. However, the two 10 bar cases show anomalously low retrieved temperatures compared to their respective 1D TP profiles.

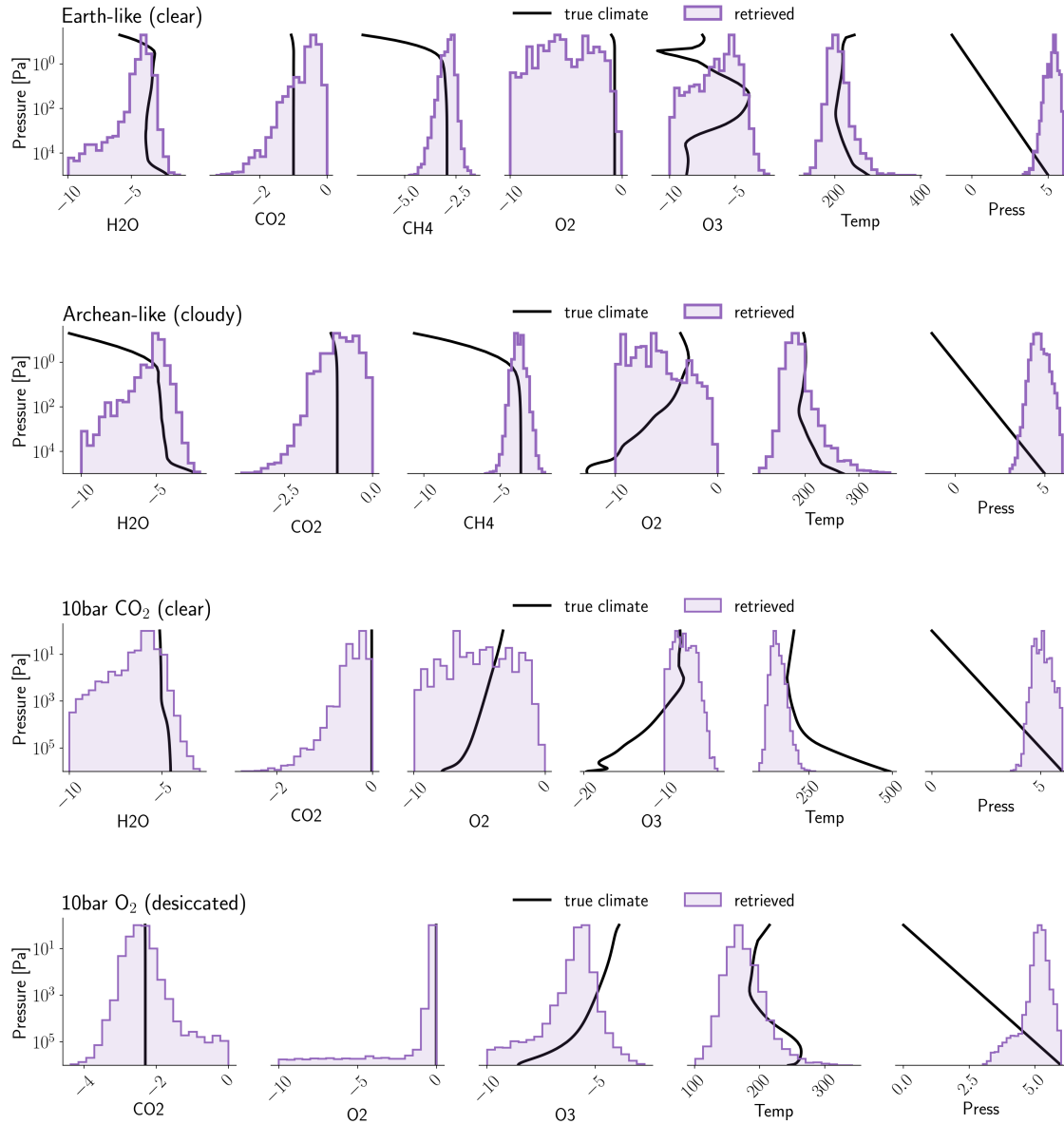


Figure 7.5 Retrieved marginalized posterior distributions compared to the 1D vertical atmospheric profiles used to generate the synthetic *JWST* spectra. Each horizontal row of plots shows a different atmospheric case. Each subplot shows a normalized histogram of the 1D marginalized posterior, compared to that parameters true vertical profile shown as a function of pressure along the vertical y-axis.

7.4.3 Refraction

We compared the effects of including refraction in the transmission spectrum forward model using many different atmospheric cases for TRAPPIST-1e. For this test, we kept the same 8 parameter state vector as in our previous tests, and fit the same fiducial 25 transit NIRSpec Prism transmission spectrum. The only change made was the inclusion of the refraction of stellar rays passing through the transiting planet’s atmosphere using the ray tracing model described in [Robinson \(2017a\)](#). We reiterate here that the effects of refraction are included in the simulated *JWST* data from the more sophisticated spectral models generated from self-consistent climate and photochemistry atmospheric calculations.

Figure 7.6 shows marginalized posterior distributions for each model parameter to compare retrieved results when refraction is included. Each row of plots corresponds to a different assumed atmospheric composition for TRAPPIST-1e and each subpanel corresponds to one retrieved parameter. In general, we find very similar results to our cases without refraction. However, with refraction included, the molecular abundances and the isothermal temperature are constrained more tightly, particularly evident for CH_4 and O_3 in the Earth-like case. The solid body radius and the reference/surface pressure have slightly looser constraints, which is more evident for the higher pressure atmospheres considered. With refraction included, the planet radius posterior also shows a shoulder for the 1 bar Earth-like case, which transforms into a secondary mode for the three other higher pressure cases shown. In all cases, the median retrieved temperature shifts to higher temperatures when refraction is included. Moreover, the deeper into the atmosphere that the transmission spectrum intrinsically probes, the greater the shift to hotter temperatures (note that the 10 bar cloudy CO_2 atmosphere has a median inferred reference pressure at <0.1 bar).

Refraction also appears to modify the degenerate relationship between the reference pressure and other key planet and atmospheric parameters. Figure 7.7 shows two-dimensional posterior projections for the reference pressure versus the radius, temperature, and the O_3 and O_2 abundances for the 10 bar O_2 atmospheric case. The covariance between pressure and radius exhibits a significant change in morphology when refraction is included in the model, which helps elucidate

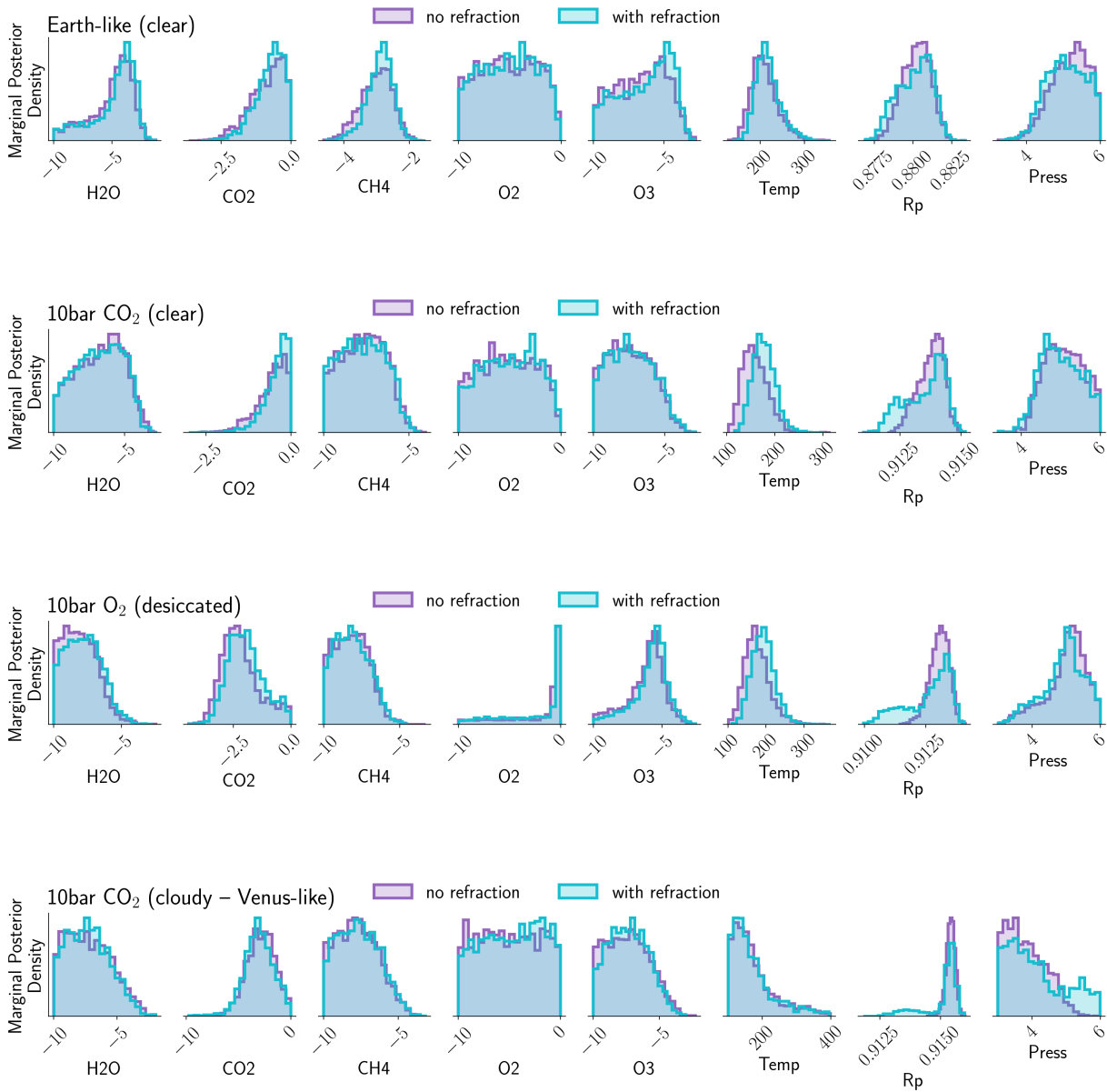


Figure 7.6 Marginalized posterior distributions for retrievals with (cyan) and without (purple) refraction in the transmission spectrum forward model. Simulations were performed on simulated data assuming 25 transits were observed with JWST/NIRSpec Prism and the planet possesses a clear Earth-like atmosphere (first panel), a 10 bar clear CO₂ atmosphere (second panel), a 10 bar desiccated O₂ atmosphere (third panel), or a 10 bar cloudy CO₂ atmosphere (fourth panel).

the secondary radius mode seen in Fig. 7.6. This change in the form of the degeneracy occurs at pressures higher than 10^5 Pa (1 bar), but negligibly affects the posteriors at lower pressures. Furthermore, this effect is not substantial in the 1D marginalized pressure posteriors shown Fig. 7.6, despite causing a substantial change to the pressure-radius covariance, which explains the change to the 1D marginalized radius posteriors. The covariance between pressure and the other parameters shown in Fig. 7.7 is more subtle, but generally show how the tighter constraints on the inferred parameters noted in Fig. 7.6 occur at higher pressures.

The tendency for refraction to affect changes in atmospheric inferences, preferentially at higher pressures, is consistent with the underlying physical picture of refraction in transiting exoplanet atmospheres (Bétrémieux and Kaltenegger, 2013; Misra et al., 2014). For thin atmospheres where the planet surface, or likewise an optically thick cloud/haze, is controlling the transmission spectrum continuum, the reference pressure and the surface radius are nearly independent and can be (loosely) constrained by the spectrum continuum. However, for thicker atmospheres without aerosols, the spectrum continuum is controlled by the critical refraction pressure (Bétrémieux and Kaltenegger, 2013; Misra et al., 2014), which obscures the atmosphere below. In such a case, if refraction is neglected from the retrieval, the radius and reference pressure can still fit the spectrum as if the continuum is an optically thick medium. However, with refraction included the surface radius and reference pressure are degenerate, unseen, and can be exchanged—increasing the surface pressure and decreasing the radius—to nearly indefinite depths into the atmosphere, while the critical refraction limit maintains the continuum at a fixed pressure.

We also compare marginal likelihoods to examine the evidence in favor of the refraction model. Although we are able to include refraction in our model without increasing the dimensionality of the inference, we find that Bayes factors do not favor the atmospheric refraction model over the model without refraction. This occurs because we find equally good fits to the spectra when we do not include refraction, and therefore the marginal likelihoods for both models remain consistent. This result holds when increasing the number of observed transits to 50. However, achieving adequate fits to data that yield biased results is a major concern for thick, cloud-free atmospheres.

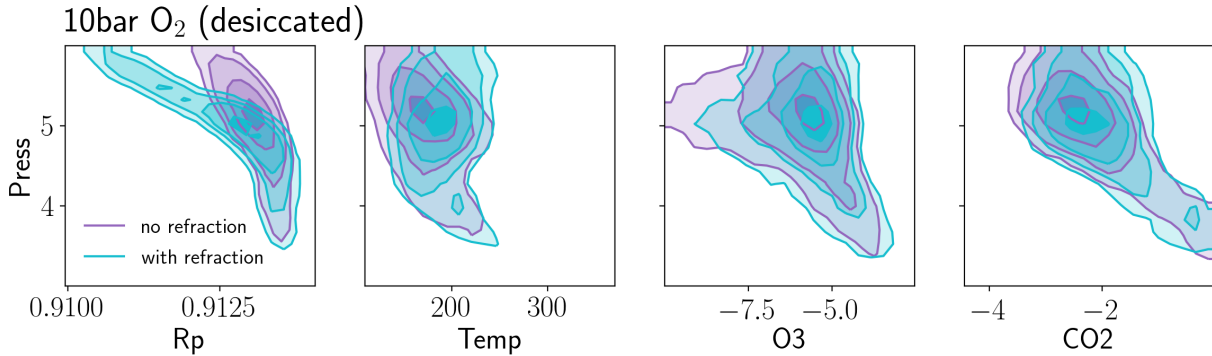


Figure 7.7 Two-dimensional marginalized posterior projections for the fit to the 25 transit data of the 10 bar dry O_2 atmospheric composition using a model with (cyan) and without (purple) refraction included in the forward model. Refraction modifies the covariance between the reference pressure and other key parameters, particularly the solid-body planet radius, at higher pressures.

7.4.4 Planet Mass & TTV Constraints

We test the effect of including planet mass in the retrieval similarly to our tests for refraction. We consider the clear Earth-like case observed with 25 and 50 transits. We impose a Gaussian prior on the planet mass using the TTV mass constraints from [Grimm et al. \(2018\)](#). Note that not including mass in these retrieval tests is identical to assuming a delta-function prior on mass because it is optimistically fixed to the known value. We also examine the impact of expanding the TTV mass prior by a factor of two (in standard deviation) to relate the results to systems for which the inferred mass is less well constrained.

We find that the exceptional TTV constraints on the mass of TRAPPIST-1e are sufficient to effectively reproduce the same posteriors as those without mass included. Figure 7.8 shows results from two retrieval tests: 25 and 50 observed transits (top and bottom panels, respectively). The marginals show very close agreement between the retrievals with and without mass included. However, we do find that marginalizing over the mass constraints propagates uncertainty into the retrieved temperature. The 1σ uncertainty on the isothermal temperature increased by 18% for the 25 transit dataset, and it increased by 44% for the 50 transit data. Therefore, retrievals that do not marginalize over the planet mass uncertainties may retrieve artificially tight temperature con-

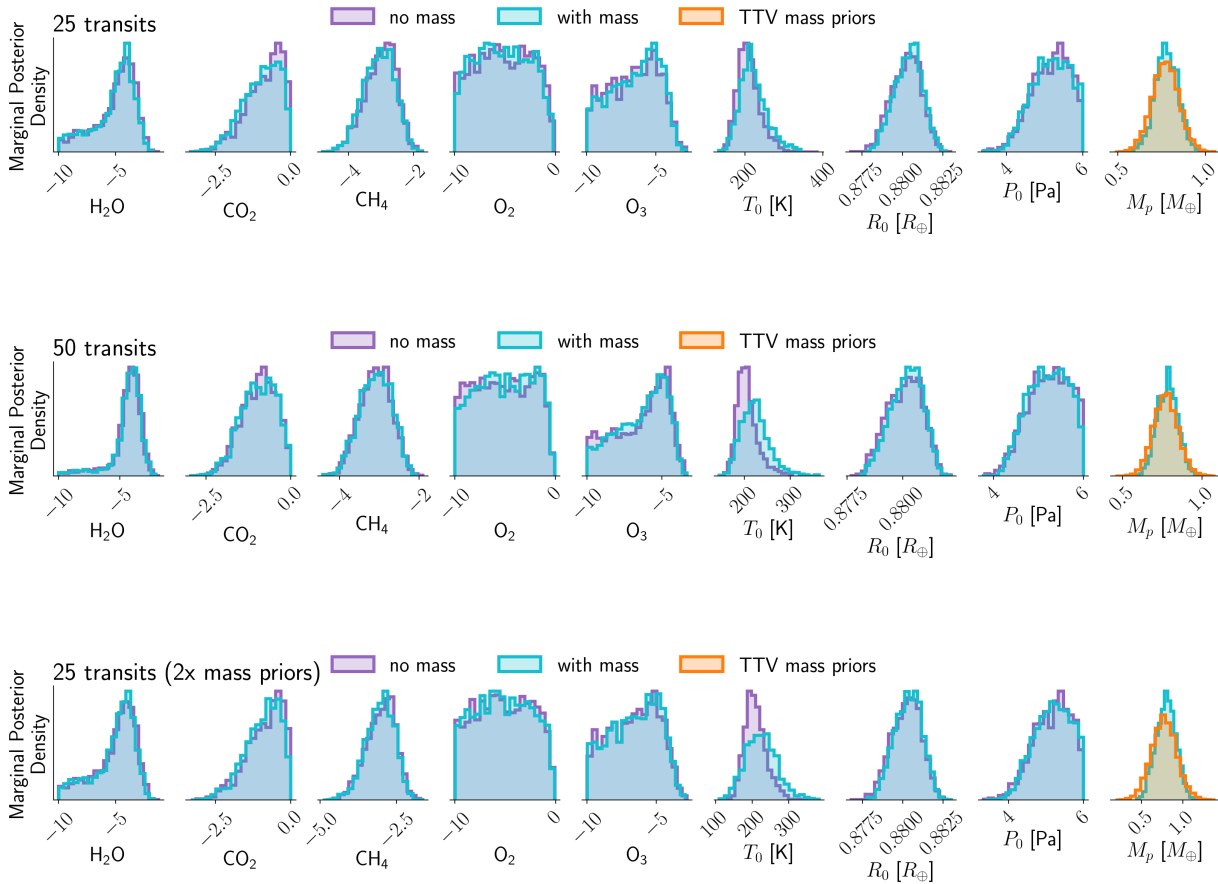


Figure 7.8 Marginalized posterior distributions for retrievals with (cyan) and without (purple) planet mass included in the state vector. Retrievals were performed on simulated data of TRAPPIST-1e with an Earth-like (clear-sky) atmosphere assuming 25 transits (top and bottom panels) and 50 transits (middle panel) were observed with JWST/NIRSpec Prism. Gaussian priors obtained from TTV measurements are used for the planet mass (shown in orange); the bottom panel uses a Gaussian prior with twice the standard deviation.

straints, particularly for high S/N spectra. Additionally, the 50 transit dataset appears to constrain the planet mass slightly beyond the constraints acquired from TTV analyses.

Decreasing the precision of the mass prior also leads to less precise temperature constraints. The bottom panel of Fig. 7.8 shows the results after doubling the standard deviation on the planet mass Gaussian prior. In this case, the temperature posterior is more poorly constrained, showing a

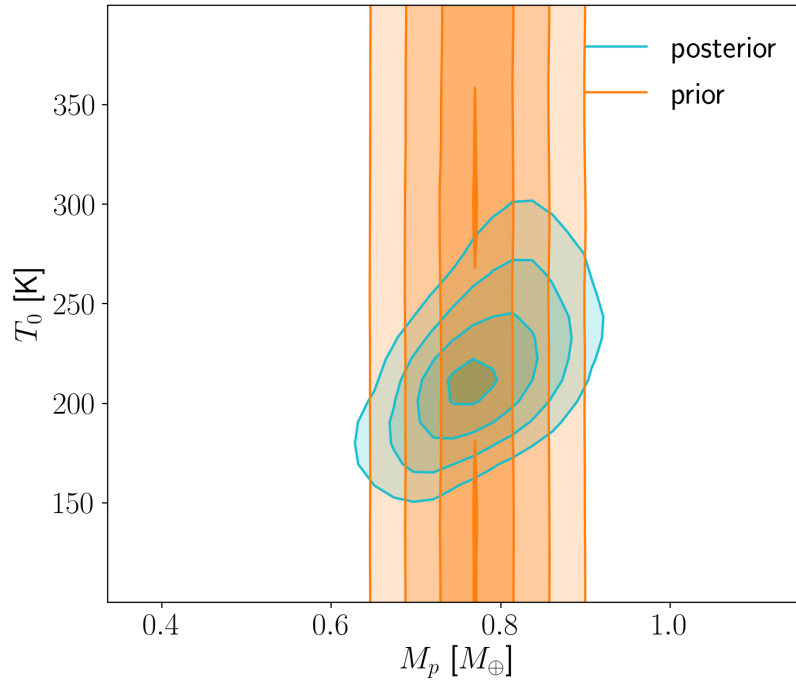


Figure 7.9 Two-dimensional marginalized prior (orange) and posterior (cyan) constraints on the planet mass and temperature from the 25 transit data of the Earth-like atmospheric composition. A positive correlation is evident between the mass and temperature.

52% increase in the 1σ temperature uncertainty, which resembles our findings in the middle panel when we doubled the number of observed transits. That is, not including mass can lead to biased constraints on the inferred temperature if the mass uncertainty is large relative to the precision on the spectrum.

These findings are the result of a degeneracy between planet mass and temperature, via the scale height of the atmosphere. Figure 7.9 shows prior and posterior probability distributions as a function of temperature and mass for the nominal 25 transit retrieval case with TTV priors on the mass. The scale height degeneracy between mass and temperature is clearly evident in the posterior as a positive correlation. Thus, as the prior constraints on planet mass expand or contract—due to TTV measurements—or the precision of the spectrum changes—due to more or less transit observations—the posterior constraints on the mass and temperature will change along the contours of this degeneracy.

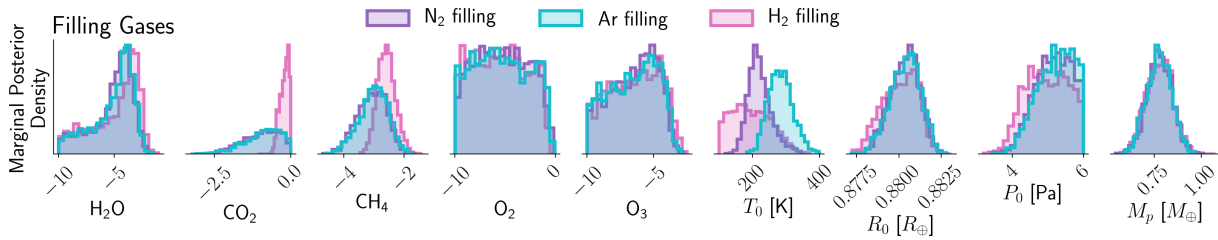


Figure 7.10 Marginalized posterior distributions for retrievals with three different assumed VMR filling gases: N_2 (purple), Ar (cyan), and H_2 (pink). Retrievals were performed on simulated data of TRAPPIST-1e with an Earth-like (clear-sky) atmosphere assuming 25 transits were observed with JWST/NIRSpec Prism.

7.4.5 Mean Molecular Weight & The Bulk Atmospheric Composition

We now revisit our assumption to use N_2 as a background gas to fill in the volume of the atmosphere not expressed by the gases in the state vector. This approach is used extensively within the atmospheric retrieval community with mixtures of H_2 -He to fill the bulk composition of gas giants with primary atmospheres (e.g. [Benneke and Seager, 2013](#); [Greene et al., 2016](#); [Irwin et al., 2020](#); [Lavie et al., 2017](#); [Mollière et al., 2019](#)), and has also recently been applied to high mean molecular weight atmospheres using N_2 ([Barstow et al., 2020](#)). Using our fiducial 25 transit clear Earth-like NIRSpec transmission spectrum, we perform two additional retrievals using Ar and H_2 as the assumed filling gas to test for biases with higher (Ar) and lower (H_2) molecular weight gases compared to N_2 . We include planet mass in these experiments.

We find that the choice of filling gas can bias the retrieved temperature, but physically-motivated choices do not appear to significantly affect the inference of gas abundances. Figure 7.10 shows one-dimensional marginalized posterior results from retrievals of the clear Earth-like TRAPPIST-1e spectrum, assuming N_2 (28.014 g/mol), Ar (38.948 g/mol), and H_2 (2.106 g/mol) VMR filling. The marginalized posteriors for Ar and N_2 are consistent with one another for all parameters except for the temperature, which shows distinctly inconsistent retrieved distributions. Using H_2 as a filling gas shows inconsistent retrieved CO_2 and CH_4 abundances, in addition to an inconsistent

temperature. These findings manifest from the dependent relationship between the filling gas and the mean molecular weight and the scale height degeneracy between the molecular weight and temperature. Being heavier than N_2 , Ar will contract the scale height unless the temperature increases to compensate. Similarly, H_2 is much lighter than N_2 , forcing the temperature to decrease substantially to compensate. However, this pushes the temperature up against the lower prior bound set at 100 K, and as a result, we see increases to the CO_2 and CH_4 abundances in an attempt to increase the molecular weight and still fit the features of these spectrally active gases in the spectrum. We can weakly rule out the H_2 model with a Bayes factor of $\ln B = 1.8$ when compared to the N_2 and Ar filled cases, but these heavier gases cannot be distinguished in our Bayesian hypothesis test.

We find that using filling gases in a retrieval places a strong implicit prior on the mean molecular weight of the atmosphere. Figure 7.11 shows prior and posterior samples from our three filling gas tests projected onto axes for the mean molecular weight and VMR of the filling gas. Since these two parameters are not fitting parameters, the displayed samples are not explicitly drawn from the state vector, but are simply calculated from it for each prior and posterior sample. The central panel of Fig. 7.11 indicates that there is an *a priori* preference for the mean molecular weight to take on values corresponding to the available end-member bulk atmospheric compositions. That is, as the VMR of the filling gas increase towards unity and it becomes the bulk atmospheric constituent, the mean molecular weight of the atmosphere necessarily converges to the molecular weight of the filling gas. On the other hand, as the filling gas decreases in abundance the atmospheric mean molecular weight follows trajectories towards a linear combination of the available gases in the state vector. Furthermore, the (log) uniform priors used for the free gases have a bias towards small abundances², which transforms via Equation 7.1 into a bias towards large abundances for the filling gas, as seen in the top panel of Fig. 7.11. Ultimately, this implicit prior on the filling gas abundance causes strongly peaked priors on the mean molecular weight that are centered on the weight of the filling gas with weak tails extending the breadth of the available gases in the state vector.

²Our choice of priors give a 50-50 chance of each gas falling below 10^{-5} VMR.

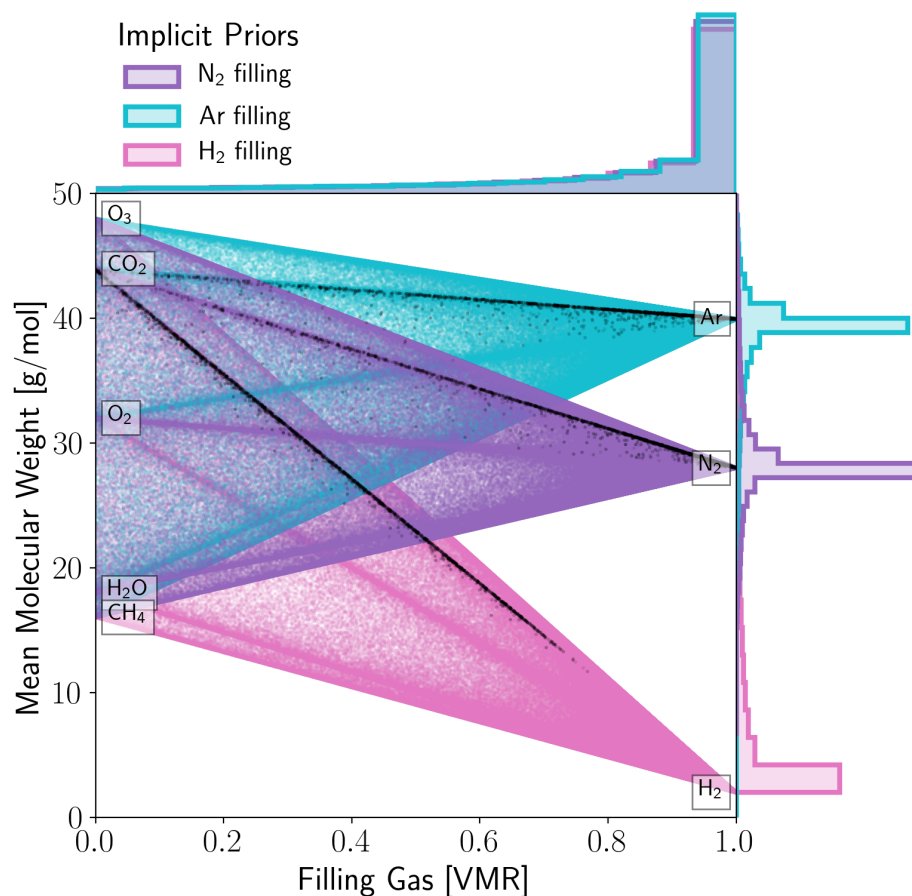


Figure 7.11 Two-dimensional marginalized prior (color points and histograms) and posterior (black points) samples/distributions for the VMR of the filling gas and the atmospheric mean molecular weight, assuming N₂ (purple), Ar (cyan), and H₂ (pink) as the filling gas. The central panel shows samples from the prior as color points and equally weighted samples from the posterior as black points, indicating an *a priori* preference for the mean molecular weight to take on values corresponding to the available end-member bulk atmospheric compositions. The top and right histograms show one-dimensional prior projections for the filling gas VMR and the mean molecular weight, respectively. The choice of filling gas imposes an implicit prior that propagates a substantial preference for the molecular weight of the atmosphere into posterior inferences.

Despite the implicit priors on mean molecular weight and their propagation into the inferred temperature, high S/N data may be able to overcome this effect. The posterior samples shown in Fig. 7.11 (black dots) generally fall along a line connecting the molecular weight of CO₂ with that of the filling gas. When Ar and N₂ are used the mean molecular weight posterior peaks near the weight of each gas as it is found to compose the bulk of the atmosphere. When H₂ is used the opposite is seen and the mean molecular weight tends towards CO₂ as the H₂-dominated atmosphere is rejected. This indicates that retrievals can overcome these strong implicit priors and identify bulk compositions that vary from the filling gas if (a) the true bulk gas is included in the state vector and (b) the spectral data are high enough S/N.

7.4.6 *Random Noise Instances*

We now briefly revisit the validity of our technique to sample the spectral data at the mean of the Gaussian errors, without adding random Gaussian noise. As previously stated, this approach has been used in other retrieval studies to efficiently acquire posteriors that are consistent with having performed many retrievals each with a different noise instance (e.g. [Feng et al., 2018](#); [Mai and Line, 2019](#); [MacDonald et al., 2020](#)). However, this method can only be applied to retrieval studies on synthetic data, so it warrants continued investigation to validate the results against random noise instances.

We apply random Gaussian noise to the 25 transit NIRSpec Prism spectrum of TRAPPIST-1e with a modern Earth-like atmosphere and assess the posterior and the model comparison agreement with the retrieval results found using the mean spectrum. Figure 7.12 shows the retrieval model fit to the noisy spectrum and the resulting posteriors compared to the case with no Gaussian noise added. The figure also shows the detection significance of each molecule in the spectrum from an evidence-based model comparison test. Although the posterior distributions are clearly not identical, they largely agree within the inferred uncertainties and they generally yield the same broad interpretation. In some cases the mean retrieval approach shows tighter constraints, such as for the H₂O abundance and the reference pressure, while for other parameters the random noise instance displays artificially enhanced constraints, such as for O₃. However, the parameters that

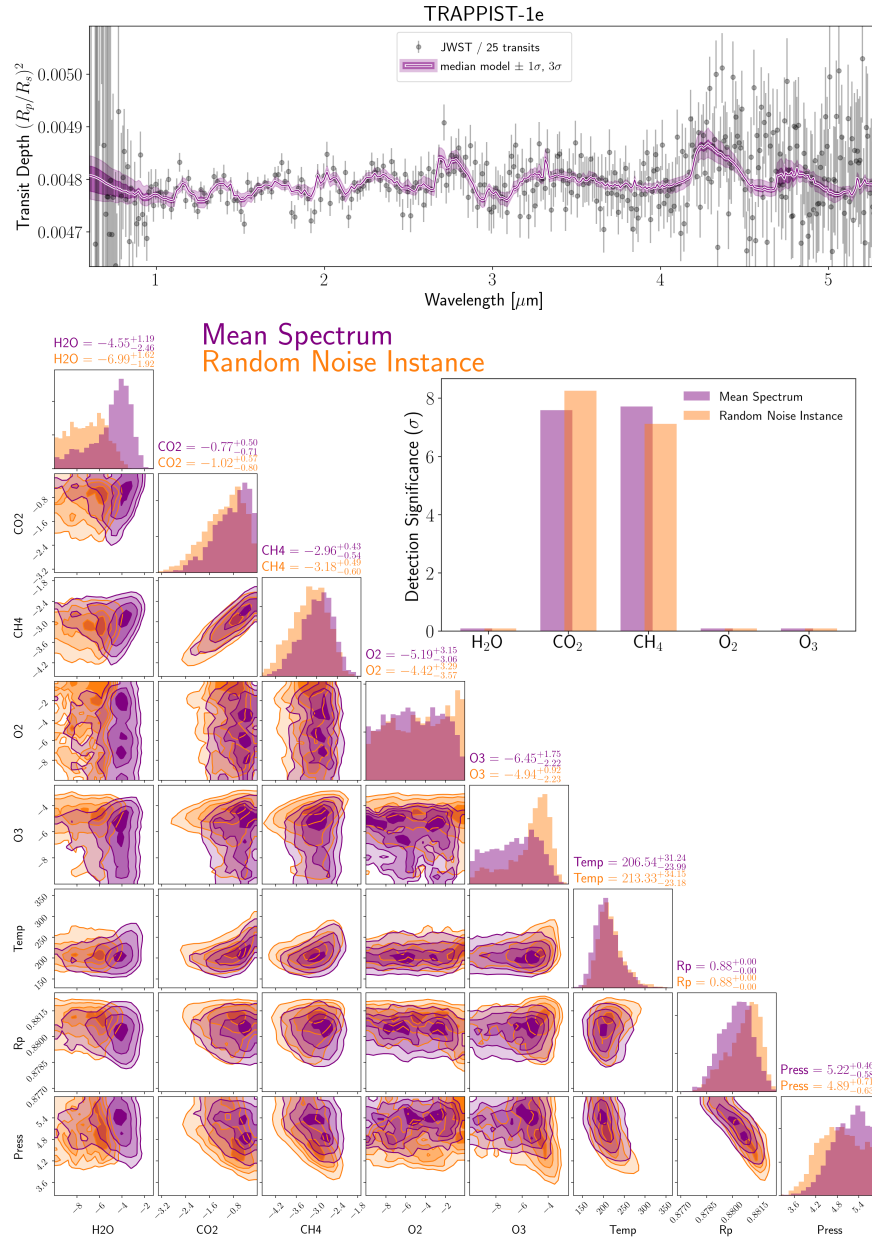


Figure 7.12 Comparison between retrievals with and without Gaussian noise added to the spectrum. The top panel shows the synthetic *JWST* spectrum and the median retrieved spectrum with 1σ and 3σ uncertainty envelopes. The corner plot shows the 1D and 2D marginalized posteriors for the mean spectrum retrieval (purple) and the random noise instance spectrum retrieval (orange). The middle-right inset axis shows the detection significance of each molecule from an Bayesian evidence-based model comparison for each noise approach.

are well constrained—the CO_2 and CH_4 abundances, the temperature, the radius, and the reference pressure—appear to have much less bias in their posteriors. Moreover, the large disagreement in H_2O VMR and the false constraints on O_2 and O_3 are easily flagged as non-detections by the Bayesian hypothesis tests for each molecule, which show remarkably good agreement between the two noise approaches. Thus, we find the mean spectrum retrieval approach to be valid for predictive assessments, particularly when paired with evidence-based model comparisons.

7.4.7 APPROXPOSTERIOR Validation

We now explore retrieval results from the clear Earth-like TRAPPIST-1e case using the APPROXPOSTERIOR machine learning method. In particular, we examine both the agreement between posteriors obtained with APPROXPOSTERIOR versus those obtained with MCMC and NS and the efficiency of each method in terms of forward model evaluations.

Our first test compares MCMC and AP using a simplified state vector with only three parameters. In this case, we fit for the isothermal temperature, the H_2O abundance, and the CO_2 abundance. All other parameters are held fixed at their true simulated values. Figure 7.13 shows 1D and 2D posterior projections for retrieval results obtained using MCMC (purple) and APPROXPOSTERIOR (orange). We successfully derive statistically identical posteriors for our three atmospheric model parameters in just 4000 forward model evaluations using APPROXPOSTERIOR compared to 3,000,000 evaluations needed with MCMC, an $750\times$ speed improvement.

Our second test uses the 8 parameter state vector used throughout this work to fit the fiducial TRAPPIST-1e clear sky Earth-like transmission spectrum. In this case, we started APPROXPOSTERIOR with an initial training set of 800 points randomly drawn from the priors and allowed APPROXPOSTERIOR to select another 1600 points using its active learning algorithm. For this test, we compare APPROXPOSTERIOR against MULTINEST. Figure 7.14 shows the resulting 1D and 2D marginalized posterior results for comparison. We can see significant discrepancies between APPROXPOSTERIOR and MULTINEST in both the 1D marginal posteriors and the 2D covariances. For example, the most well-constrained gas— CO_2 and CH_4 —show Gaussian posteriors that are offset from the MULTINEST results by greater than 1σ , although the CO_2 – CH_4 covariance is still

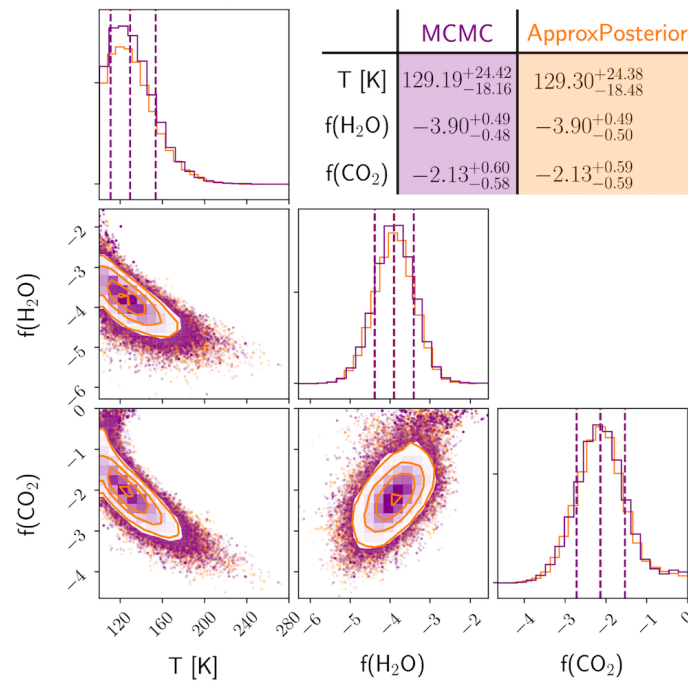


Figure 7.13 Corner plot showing the 1D and 2D marginalized posteriors for a simplified retrieval scenario simulated using both MCMC (purple) and AP (orange). For this low-dimensional case, AP derives statistically identical posteriors as MCMC using $\sim 1000 \times$ fewer forward model evaluations.

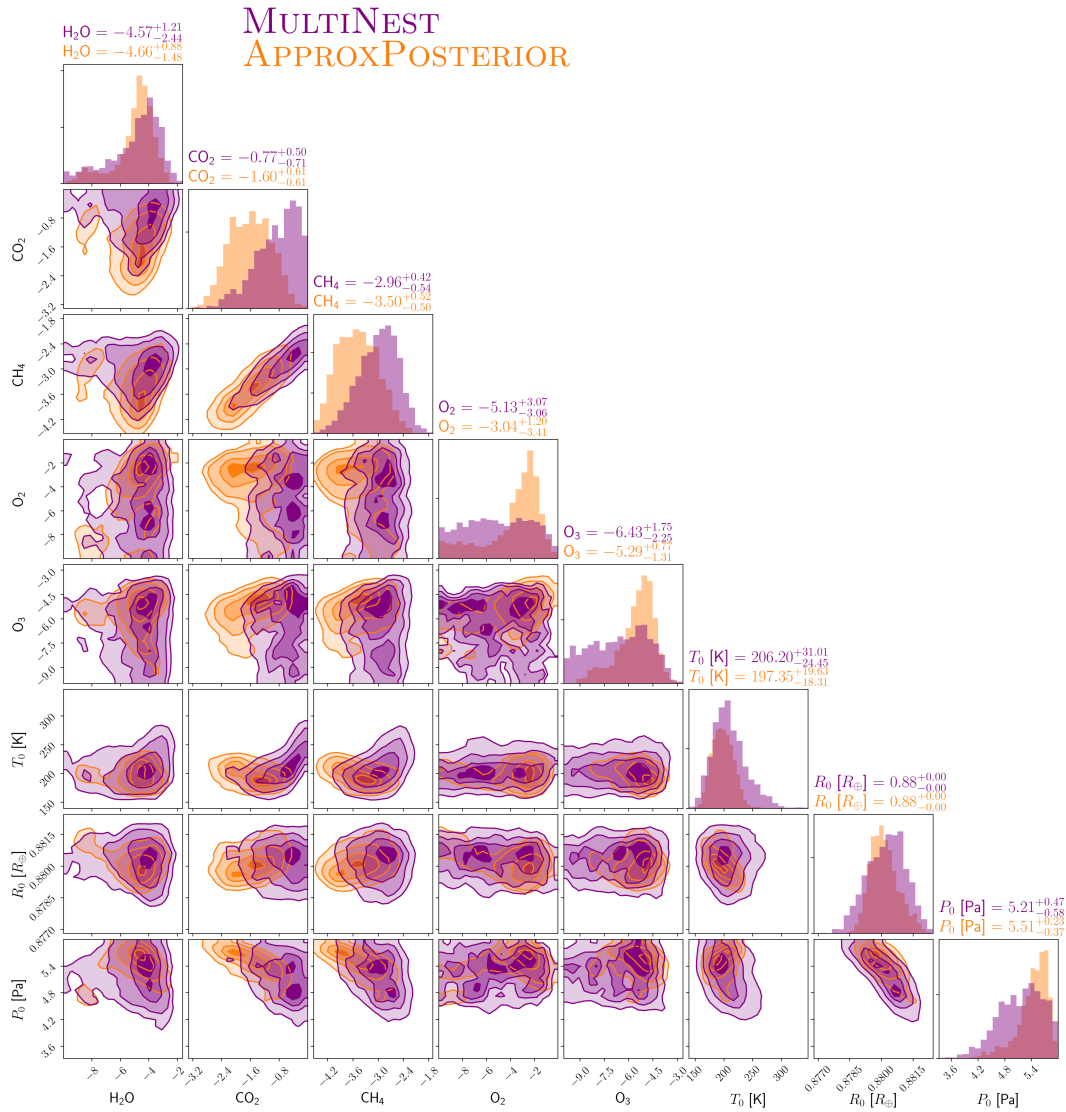


Figure 7.14 Posterior comparison between retrievals using MULTINEST (purple) and APPROXPOSTERIOR (orange).

captured. Parameters for which MULTINEST gives an upper limit (e.g. O_2 and O_3) are artificially over-constrained by APPROXPOSTERIOR, and show a Gaussian peak that declines correctly along the upper limit, but fails to capture the extended tail of the distribution at low abundances. As a result, APPROXPOSTERIOR derives posteriors that are incorrectly over-constrained.

7.5 Discussion

7.5.1 Characterizing the Atmosphere of TRAPPIST-1e

Our results further illuminate an emerging picture that *JWST* transmission spectroscopy can be used to detect the presence of secondary atmospheres for the temperate TRAPPIST-1 planets, in particular, by detecting the strong $4.3\ \mu\text{m}$ CO_2 band (Morley et al., 2017; Wunderlich et al., 2019; Lustig-Yaeger et al., 2019b; Fauchez et al., 2019; Pidhorodetska et al., 2020). We underscore the robustness of such a CO_2 detection even in the presence of high altitude global clouds by demonstrating that CO_2 can still be confidently detected ($>3.6\sigma$) above the H_2SO_4 aerosols in a self-consistent Venus-like atmosphere for TRAPPIST-1e in 25 observed transits with NIRSpec, although no evidence of the aerosols are detected to distinguish them from a solid surface, as cautioned by Lustig-Yaeger et al. (2019a). Beyond CO_2 , the characterization of TRAPPIST-1e becomes a much more difficult endeavor.

We find strong agreement with the growing consensus that transmission spectroscopy with *JWST* appears ill-suited to assess the habitability of TRAPPIST-1e and, by extension, other HZ planets transiting M dwarfs. Previous works have shown that H_2O features may be difficult to detect in transit due to H_2O 's scarcity in the upper atmosphere of habitable planets where transmission spectra are most sensitive (Robinson, 2018; Lustig-Yaeger et al., 2019b). The detectability of H_2O is further reduced by the obscuring effects of H_2O clouds and photochemical hazes (Fauchez et al., 2019; Komacek et al., 2020; Pidhorodetska et al., 2020; Suissa et al., 2020). In addition to these complicating effects for habitability assessments, we find that photochemically enhanced CH_4 (see Segura et al., 2003, 2005; Lincowski et al., 2018; Wunderlich et al., 2019; Meadows et al., 2020) produces large MIR spectral features that overlap with crucial H_2O bands, decreasing the sensitivity to H_2O for remote sensing retrievals.

However, even if habitability eludes our spectroscopic searches, it may still be possible to observationally confirm that a planet is uninhabitable. Methods have been proposed to distinguish the atmospheres of planets that do not have surface oceans by detecting $\text{O}_2 - \text{O}_2$ CIA as evidence of pre-main-sequence water loss (Schwieterman et al., 2016b; Lincowski et al., 2018; Lustig-Yaeger

et al., 2019b; Fauchez et al., 2020), detecting isotopologue bands from Venus-level isotopic fractionation as evidence of ocean loss (Lincowski et al., 2019), or detecting an $\text{H}_2\text{SO}_4 - \text{H}_2\text{O}$ haze layer or SO_2 , which would be incompatible with surface liquid water (Loftus et al., 2019). We confirm and build upon the results of Lustig-Yaeger et al. (2019b), finding that 25 transits of TRAPPIST-1e with NIRSpec would be sufficient to weakly detect ($>2.1\sigma$) the presence of O_2 via the NIR $\text{O}_2 - \text{O}_2$ CIA bands and infer O_2 as the bulk composition of the atmosphere.

The ability for *JWST* to detect biosignature gases appears quite limited and may ultimately depend on our ability to interpret the simultaneous presence of CH_4 and CO_2 . Krissansen-Totton et al. (2018a) report that 10 transits of TRAPPIST-1e may be sufficient to detect CO_2 and constrain the CH_4 abundance enough to rule out known nonbiological CH_4 production, potentially indicating a biosignatures for an anoxic Archean Earth-like planet. We infer similar results to Krissansen-Totton et al. (2018a) regarding the detection of CH_4 and CO_2 and the retrieval of their order-of-magnitude abundances for the self-consistent Archean Earth-like environment, both with and without significant sulfur organics, studied by Meadows et al. (2020, *in prep*). However, our findings suggest that H_2O may be significantly more difficult to detect than reported by Krissansen-Totton et al. (2018a). Like previous works, our retrieval results are largely insensitive to O_2 and O_3 for a modern Earth-like TRAPPIST-1e planet (Barstow and Irwin, 2016; Krissansen-Totton et al., 2018a). However, as in Meadows et al. (2020, *in prep*), we find that CH_4 and CO_2 may also be easily constrained for TRAPPIST-1e if it possesses a photochemically self-consistent modern Earth-like atmosphere. Interestingly, because signs of oxygenic photosynthesis are effectively undetectable, retrievals may appear quite similar if the planet has a modern Earth-like atmosphere or a haze-free Archean Earth-like atmosphere (for Archean hazes e.g., see Arney et al., 2016, 2017, 2018). Thus, the simultaneous presence of abundant CO_2 and CH_4 appears to be an intriguing disequilibrium indicator that Earth-like biospheres may readily express in their spectra *throughout much of their evolution* if such life is present on planets around late M dwarfs like TRAPPIST-1.

7.5.2 Terrestrial Exoplanet Atmospheric Retrieval Subtleties

We find that the common approach of using an isothermal TP structure and evenly mixed vertical gas abundances is sufficient to provide order-of-magnitude constraints that are generally consistent with the 1D vertical profiles from self-consistent climate and photochemistry models used to generate our synthetic *JWST* transmission spectra. For instance, despite the strong vertical dependence of O_3 in a modern Earth-like atmosphere, we are able to retrieve an upper limit on the ozone abundance that matches the peak stratospheric O_3 volume mixing ratio. Other gases do not exhibit as substantial of vertical gradients and our retrieval succeeds in retrieving their approximate abundances across the relatively narrow pressure region probed by transmission spectroscopy. However, based on our retrieval of the stratospheric H_2O abundance rather than the troposphere or surface H_2O vapor content, we caution against interpreting the atmosphere retrieved from a terrestrial exoplanet transmission spectrum as representative of the entire atmosphere or the surface conditions.

We find that retrievals which do not include atmospheric refraction using a ray tracing radiative transfer model, may introduce appreciable biases for some inferred atmospheric parameters. All of our synthetic *JWST* transmission spectra observations used the refraction model, but we ran retrieval models with and without refraction. Unsurprisingly, we find the greatest effects from refraction for thick (10 bar) aerosol-free atmospheres. Nonetheless, all of our simulations with refraction lead to slightly tighter abundance constraints with a slightly higher inferred isothermal temperature.

Moreover, we find that refraction modifies the degeneracy between the solid-body surface radius and the corresponding reference pressure (e.g. [Welbanks and Madhusudhan, 2019](#)) to create multi-modal radius posteriors. That is, the spectral continuum can be interpreted in one of two ways: an optically thick medium (planet surface or cloud deck) or the critical refraction limit (see [B  tr  mieux and Kaltenegger, 2014](#); [Misra et al., 2014](#)). The former corresponds to the solution that is obtained without refraction, while the latter is a direct consequence of refraction that enables a degenerate anti-correlation between the reference radius and pressure to maintain the critical refraction limit as the fixed spectral continuum while allowing the unseen surface pressure to be

arbitrarily high. These refraction effects are subtle, but may grow in importance as transmission spectra are obtained for a broader range of exoplanets, particularly cooler planets, that could have relatively clear atmospheres.

We find that the TRAPPIST-1e mass constraints from TTV measurements are precise enough at present to enable a substantial spectral characterization effort using *JWST* without propagating uncertainty into the atmospheric inference. We explored how the marginalizing over the planet mass constraints compares to retrievals artificially fixed at the true mass. In a 25 transit survey of TRAPPIST-1e, the mass from [Grimm et al. \(2018\)](#) provide the same posteriors as those with the fixed mass, but for a 50 transit survey, the current mass constraints do begin to decrease the precision of the retrieved temperature. However, decreasing the precision of the [Grimm et al. \(2018\)](#) mass by a factor of two led to a $\sim 50\%$ increase in the temperature uncertainty. Thus, because the TRAPPIST-1 system is amenable to remarkably precise planet mass characterization via TTVs, we do not expect mass uncertainties to hinder the atmospheric characterization of these planets. This assessment is in agreement with the more detailed findings presented in [Batalha et al. \(2019\)](#). However, in general, including planet mass uncertainties in terrestrial exoplanet atmospheric retrievals is important to faithfully explore scale height degeneracies, particularly for systems with poor mass constraints.

Uncovering the bulk composition of secondary atmospheres, and remaining unbiased when the bulk composition cannot be constrained, is poised to be one of the greatest challenges of terrestrial exoplanet atmospheric characterization. Throughout much of this work, we employed the common assumption that N_2 fills the remaining content of the atmosphere after all other gases in the state vector are accounted for. However, we demonstrated that this approach places a strong implicit prior on the mean molecular weight of the atmosphere that propagates artificial constraints through to temperature inferences by way of the scale height degeneracy between mean molecular weight, temperature, and surface gravity. These results echo similar findings reported and discussed by [Benneke and Seager \(2012\)](#). We tested for sensitivity to the choice of filling gas and found that trace gas abundances are unlikely to be biased by this assumption, as long as the trace gases are not hijacked by the retrieval to overcome a poor choice of filling gas (e.g. see H_2 filling in Fig. 7.10).

These effects decrease as more gases are added to the state vector. The primary concern for this bulk composition bias is the mean molecular weight and the atmospheric TP structure.

We advocate a nuanced approach for specifying the bulk atmospheric composition and mean molecular weight going forward. For instance, explicitly fitting for the mean molecular weight, as done by Tremblay et al. (2020), is robust against this implicit prior bias. However, this approach fails to self-consistently leverage information about the bulk composition of the atmosphere that may actually be constrained by the spectrum, such as how a large posterior CO₂ abundance should naturally sway the mean molecular weight towards the weight of CO₂, or how the presence of N₂–N₂ CIA should be correlated with bulk atmospheric N₂ (Schwieterman et al., 2015b). One approach could be to report retrieval results for multiple choices in filling gas. Another promising approach may be to combine the two strategies of explicitly fitting for the mean molecular weight and using a VMR filling gas. This allows the prior on mean molecular weight to be flat, uninformed, and unbiased, but the prior covariance matrix may still possess off diagonal terms to account for the dependence of the mean molecular weight on the retrieved gas abundances. In practice, this can be achieved by fitting for the molecular weight of the filler gas subject to a uniform prior. A third option could be to reparameterize the atmospheric composition using a centered-log-ratio transformation, as discussed in Benneke and Seager (2012) to mitigate the bias caused by the asymmetric filling gas parameterization. However, regardless of the approach, it is likely that a family of solutions to the bulk composition will be retrieved.

Many of the challenges associated with terrestrial atmospheric retrievals circle back to the scale height degeneracy between temperature, mean molecular weight (bulk atmospheric composition), and surface gravity (planet mass). We highlighted a few of these challenges, but neglected others that will further complicate the retrieval of secondary atmospheres. Iyer and Line (2020) found that stellar spectral contamination (see Rackham et al., 2018; Zhang et al., 2018; Morris et al., 2018b,c; Wakeford et al., 2019) can bias parameters associated with the scale height, and accounting for it in a retrieval adds another degenerate parameter to the problem. Additionally, Line and Parmentier (2016) found that non-uniform cloud cover is also degenerate with the mean molecular weight, and it remains unclear how much this will complicate the characterization of secondary

atmospheres. Thus, further exploration of these entangled degeneracies is essential for correctly interpreting the atmospheres of small planets with secondary atmospheres.

7.5.3 *Active Learning for Atmospheric Retrievals*

We find that active learning can provide a pathway towards efficient approximate Bayesian inference in circumstances with slow to evaluate forward models. For low dimensional retrieval problems, we find speed improvements that approach three orders of magnitude without sacrificing posterior accuracy. However, our 8 parameter validation test demonstrates the limitations of this approach when the dimensionality of the problem grows larger, which is a common shortcoming of machine learning atmospheric retrieval approaches (Nixon and Madhusudhan, 2020).

We advocate for continued investigation into active learning for atmospheric retrievals. This approach may enable exoplanet atmospheric retrievals using physical models that would typically not be considered computationally feasible, such as with 1D coupled climate-chemistry models, multiple/many 1D atmospheric profiles, and potentially even 3D GCMs. In cases where the addition of more physical and chemical modeling can decrease the number of fitting parameters, active learning may excel. However, more work is needed to scale the impressive results seen for low dimensional problems up to the higher dimensional problems that are commonplace in the field.

7.6 *Conclusion*

We have developed a new terrestrial exoplanet atmospheric retrieval model and applied it to test *JWST*'s capacity to characterize the atmosphere of TRAPPIST-1e in a 25 transit observational campaign. Our retrieval model uses a physically robust, line-by-line radiative transfer model designed to fit medium to high resolution spectra, including *JWST* spectra at the native instrumental resolution. In agreement with previous studies on *JWST* transmission spectra of TRAPPIST-1e, we found that CO₂ may be detected relatively easily in just a few observed transits (e.g. Morley et al., 2017; Lustig-Yaeger et al., 2019b), H₂O may be quite difficult to detect (>25 transits) (e.g. Fauchez et al., 2019; Komacek et al., 2020; Pidhorodetska et al., 2020; Suissa et al., 2020), and

the CH_4 – CO_2 disequilibrium pair seen for the Archean Earth (Krissansen-Totton et al., 2018b) may be the best chance for *JWST* to detect any indication of a biosignature (Krissansen-Totton et al., 2018a). Moreover, we demonstrated that the CH_4 – CO_2 disequilibrium biosignature is also detectable for a modern Earth-like environment for TRAPPIST-1e because of the role photochemistry plays to increase the CH_4 abundance above that seen in the Earth’s atmosphere (Segura et al., 2005; Meadows et al., 2018a). Furthermore, since evidence of oxygenic photosynthesis cannot be constrained, the Archean Earth and the modern Earth may look quite similar in the eyes of *JWST* (neglecting Archean hazes e.g., see Arney et al., 2016, 2017, 2018). If instead TRAPPIST-1e possesses a post-ocean-loss oxygen-dominated atmosphere, we are able to identify O_2 as the bulk atmospheric constituent and retrieve its approximate abundance using the NIR O_2 – O_2 CIA features (Schwartz et al., 2016; Lincowski et al., 2018; Lustig-Yaeger et al., 2019b; Fauchez et al., 2020).

We also tested a few common atmospheric retrieval assumptions to gauge *JWST*’s sensitivity for TRAPPIST-1e. We formally identified a strong implicit prior on the atmospheric mean molecular weight that results from assuming any background/filler gas in a retrieval model. Although this did not affect abundance posteriors in most of our tests, it can yield substantially biased temperature constraints, and should be a chief concern as we transition into characterizing exoplanets with high mean molecular weight secondary atmospheres. In addition, we identified a modest retrieval bias for thick aerosol-free atmospheres due to the omission of refraction in transmission spectrum forward modeling that affects the radius-pressure degeneracy and can affect retrieved temperature constraints. We found reasonably good agreement between our retrieved evenly-mixed gas abundance profiles and the vertically-resolved self-consistent profiles used to generate our synthetic *JWST* spectra; vertical resolution does not appear necessary for terrestrials in the *JWST* era. Finally, we found that the TRAPPIST-1e planet mass from TTV measurements (Grimm et al., 2018) is sufficient to propagate no additional posterior uncertainty into a 25 transit campaign of TRAPPIST-1e with *JWST*. This work has illuminated opportunities and challenges ahead for the atmospheric retrieval of TRAPPIST-1e and other Earth-sized planets in the era of *JWST* and beyond.

Chapter 8
DISCUSSION

8.1 *On the Era of Terrestrial Exoplanet Science*

As we transition into the era of terrestrial exoplanet science, it is important that we proceed with well-posed hypotheses that can be realistically tested with upcoming astronomical instruments. In Chapter 1, we introduced four major milestones for terrestrial exoplanet atmospheric characterization as key scientific questions that we seek to answer for rocky exoplanets. These questions illuminate a path from simply detecting the presence of an atmosphere, to characterizing its structure and composition, through to identifying planetary habitability, and confirming the existence of biosignatures. Ultimately, the large astronomical sample of exoplanets will enable these terrestrial milestones to be explored statistically—from a comparative planetology perspective—to unveil the prevalence of terrestrial exoplanet atmospheres, atmospheric archetypes, habitable planets, and inhabited worlds in our corner of the galaxy. We will now revisit these major themes to provide an outlook for the future using the novel from this thesis and other recent work.

8.1.1 *Detecting Terrestrial Exoplanet Atmospheres*

Being able to determine whether a terrestrial planet has an atmosphere is likely the single most important initial assay required to understand its environment, potential for habitability, and to plan for the most effective future observations. In this dissertation we have explored several different techniques that could be used to search for and confirm the presence of an atmosphere, including: spectroscopic detection of the presence of molecular features, and the use of thermal emission data to infer the presence of reduced day-night temperature contrasts or planetary temperatures enhanced by an atmospheric green house effect. In particular, our work highlighted how the upcoming *JWST* mission will confront the challenge of detecting terrestrial exoplanet atmospheres.

In Chapter 2 (Lustig-Yaeger et al., 2019b), we showed that *JWST* can be used to detect terrestrial exoplanet atmospheres around the smallest of stars, and the TRAPPIST-1 system is our current best target. We input simulations of plausible post-ocean-loss and habitable TRAPPIST-1 planetary environments into *JWST* noise models, and determined optimal strategies for transmission and emission spectroscopic studies of the TRAPPIST-1 planets. Our results indicate that highly

evolved (post-ocean or -atmosphere loss) TRAPPIST-1 terrestrial atmospheres should be detectable with *JWST*, usually in 10 or fewer transits, using the NIRSpec Prism instrument. However, if the TRAPPIST-1 planets are aerosol enshrouded Venus-like worlds then 20-30 stacked transits may be required to detect the tenuous atmosphere above the clouds. We also reaffirmed these findings in Chapter 7.

We found that these atmospheric detections largely follow from the detection of CO₂ absorption features in the transmission spectrum—particularly the very strong 4.3 μm band and the slightly weaker 2.7 μm band. The evolved terrestrial atmospheres modeled for the TRAPPIST-1 planets by [Lincowski et al. \(2018\)](#) are oxidized and share in common the presence of CO₂, as do the atmospheres of Venus, Earth, and Mars. Furthermore, the strong 4.3 μm and 15 μm CO₂ bands saturate their cores at relatively low CO₂ abundances (e.g. at 280 ppm for the pre-industrial Earth), making them effective tracers of an atmosphere whether CO₂ is a trace gas or a bulk gas. As a result of its potential prevalence and strong absorption bands, we posit that CO₂ is a robust gas for detecting the oxidized atmospheres of terrestrial exoplanets. Terrestrial planets with more reducing atmospheres, such as the Archean Earth ([Arney et al., 2016, 2017](#)) or Titan, may alternatively store some of their carbon budget in CH₄, which also possesses numerous strong and detectable NIR absorption features. In Chapter 7, we showed that CH₄ features in Archean Earth-like planets may be just as detectable as CO₂ in the NIR ([Meadows et al., 2020](#)). Therefore, we expect that *JWST*'s NIRSpec Prism should be an excellent instrument for detecting terrestrial exoplanet atmospheres using transmission spectroscopy, whether the observed terrestrial exoplanets in the TRAPPIST-1 system are oxidized and possess strong CO₂ features, or are reducing and possess strong CH₄ features (with a negligible H₂ fraction).

In [Meadows et al. \(2018a\)](#), we found an analogous result that CO₂ may also be used to detect the atmosphere of the non-transiting planet Proxima Centauri b. In this case, we must rely on the thermal phase curve of the planet (assumed synchronously rotating) as it orbits the star and reveals different sub-stellar longitudes. This generates a periodic signal on the orbital period of the planet with an amplitude related to the difference in thermally emitted flux from the day- and night-sides of the planet. The amplitude of a thermal phase curve varies as a function of wavelength much like

the thermal spectrum of the planet and allows for absorption features from strong absorbers in the atmosphere to be detected (Selsis et al., 2011). To detect CO₂, we found that 10-100 hours of *JWST* integration time would be needed at both full and new phase and in each of the F1130W, F1500W, and F1800W MIRI filters. These potential future observations of Proxima Centauri b are extremely expensive compared to the transit transmission spectroscopy exposure time estimates required to detect the HZ TRAPPIST-1 planet atmospheres (Lustig-Yaeger et al., 2019b). This underscores the utility of the transiting geometry for atmospheric detection and characterization, and the difficulty that we will face characterizing non-transiting exoplanet atmospheres, particularly for HZ and cooler planets that do not have a strong thermal emission signal.

Similar to thermal phase curves, day-side thermal emission spectroscopy from secondary eclipse observations offers an appealing alternative path to detecting transiting terrestrial atmospheres for highly irradiated exoplanets with strong thermal signals. In Chapter 2 (Lustig-Yaeger et al., 2019b), we demonstrated that detecting absorption features from eclipse spectra of the TRAPPIST-1 planets is possible for the innermost planet, TRAPPIST-1b; is plausible for TRAPPIST-1c; but quickly becomes infeasible by planet d; and is hopeless for the cooler planets beyond that. For reference, transmission spectroscopy is feasible for all seven TRAPPIST-1 planets because it is less sensitive to the planets' atmospheric temperature, but as previously detailed, it is quite sensitive to atmospheric CO₂ absorption features. Our secondary eclipse approach relied on detecting the deviations from blackbody emission caused by molecular absorption features in the spectrum, which we demonstrated is quite difficult for terrestrial exoplanets. Recently, Mansfield et al. (2019) and Koll et al. (2019) presented a simple, but potentially powerful, new approach to efficiently detect terrestrial exoplanet atmospheres using secondary eclipse photometry. By assuming that a planet with a thick atmosphere will efficiently recirculate heat to its nightside and conversely an airless planet will not, they posit that rocky planets with significant secondary eclipse depths, and therefore have very hot day-sides, must be (nearly) airless. This logic extends to Venus-like planets with clouds that reflect a significant portion of incident stellar flux because—whether by heat redistribution to the night-side or by high albedo clouds—the presence of an atmosphere acts to lower the broadband thermal emission signal from the day-side relative to an airless planet. Although

this is a promising new technique, it relies heavily on *a priori* planetary assumptions to convert a non-detection of thermal emission into the detection of an atmosphere. Furthermore, this technique appears to neglect potential atmospheric windows through which strong thermal emission may emerge from the greenhouse-heated lower atmosphere (e.g. [Lincowski et al., 2018](#)). Such windows confound the broadband assumptions put forth in [Mansfield et al. \(2019\)](#) and [Koll et al. \(2019\)](#), which does not preclude the usefulness of their techniques, but provide a necessary caveat. As a result, we recommend that *JWST* observations be used to validate these techniques against more traditional atmospheric detection methods using wavelength-dependent transmission and/or emission spectroscopy.

8.1.2 Characterizing Terrestrial Exoplanet Atmospheres

Characterizing the nature and composition of terrestrial exoplanet atmospheric and surface environments will offer a glimpse into the physical, chemical, and dynamical processes that sculpt them over time. Ultimately, the observable differences between exoplanetary atmospheres may be used to discriminate different evolutionary states and the plausible histories that each planet might have experienced. These observable differences may be macro in scale, such as the bulk atmospheric composition being CO₂-, N₂-, or O₂-dominated, while other differences may be more subtle, such as the precise abundance of CO₂ or the presence of geologically outgassed SO₂. In this dissertation, we identified numerous possible opportunities to distinguish between different terrestrial exoplanet atmospheres with upcoming observations, but also cautioned about degenerate observables that are likely to prevent the unique interpretation of many terrestrial environments in the near future.

In Chapter 2 ([Lustig-Yaeger et al., 2019b](#)), we found that *JWST* will be capable of broad brush atmospheric characterization in the TRAPPIST-1 system (and other similar, but yet to be discovered systems), but may face significant challenges distinguishing key atmospheric characteristics. In general, one consequence of CO₂ being an effective tool to detect the presence of an atmosphere is that CO₂ is a weak discriminant between different atmospheric compositions. For example, CO₂ may be detectable for all the TRAPPIST-1 planets in under 10 transits if they do not possess clouds, and this result held true across a massive range in CO₂ abundances from 400 ppm to 96%. How-

ever, we showed that detecting the presence of O_2 – O_2 CIA could be used to identify O_2 -dominated atmospheres, ruling out a CO_2 -dominated atmospheres, and implicating a post-ocean-loss atmosphere. We showed in Chapter 7 that detecting O_2 (at 3σ) as the bulk gas for TRAPPIST-1e would require ~ 25 transits, but no other gases (bulk or trace) beyond CO_2 can be detected in fewer than 10 transits, except CH_4 in an Archean or a modern Earth-like biosphere—a point that we will return to in Section 8.1.4. As a result, in the *JWST* era, it may be possible to use key spectral diagnostics to distinguish between different atmospheric archetypes, but it will remain quite difficult to draw definitive distinctions between subtly different terrestrial exoplanet atmospheres. This may limit terrestrial comparative planetology in the near term to only focusing on substantial differences between secondary atmospheres.

Uncovering the nature and composition of terrestrial exoplanet atmospheres is likely to also be challenging due to the presence of clouds. The obscuring effects of clouds has already garnered appreciation in the exoplanet community in their flattening of hot Jupiter and super Earth transmission spectra (Berta et al., 2012; Ehrenreich et al., 2014; Knutson et al., 2014; Kreidberg et al., 2014; Nikolov et al., 2015; Sing et al., 2016). It should come as no surprise that terrestrial exoplanet atmospheric characterization will also contend with clouds, especially considering Earth’s haze and cloud enshrouded sister, Venus.

In particular, in Chapter 3 (Lustig-Yaeger et al., 2019a) we highlighted challenges that temperate Venus-like exoplanets may pose to transmission spectroscopy characterization efforts. Indeed, in Chapter 2 (Lustig-Yaeger et al., 2019b) we found that even if the TRAPPIST-1 planets have high altitude Venus-like clouds, *JWST* should have the sensitivity to detect CO_2 features from the tenuous atmosphere above the clouds. However, Venus-like sulfuric acid clouds appear quite featureless in a transmission spectrum despite the presence of a broad but weak H_2SO_4 absorption feature near $3.4\ \mu\text{m}$ (Meadows et al., 2018a; Arney et al., 2017), which provides little observational leverage to confirm their existence. As a result, the flat spectral continuum upon which the CO_2 features absorb may be reasonably interpreted as either a cloud deck or a solid surface. This cloud-surface degeneracy may further complicate the interpretation of archetypal planets, such that cloudy Venus-like atmospheres may be indistinguishable from thin Mars-like atmospheres.

We further posit that for an ensemble of exoplanets this cloud-surface ambiguity may propagate into a degenerate statistical trends between cloud formation and condensation processes (if the atmospheres are interpreted as thick and cloudy) or atmospheric escape physics (if the atmospheres are interpreted as thin and eroded).

Refraction is also capable of controlling and flattening the transmission spectrum continuum, muting the absorption features from molecules, and potentially further confusing atmospheric interpretations. In Chapter 7, we added refraction to a transmission spectrum retrieval for the first time and found that it modifies the degeneracy between the solid body planet radius and the reference/surface pressure. This adds another interpretation into the cloud-surface degeneracy wherein the continuum is not caused by a cloud deck or a solid surface, but rather, is caused by the critical refraction limit (Misra et al., 2014; B  tr  mieux and Kaltenegger, 2014). This limit exists in transmission spectra because refraction causes light from the star to bend as it passes through the planetary atmosphere. At a critical density in the lower atmosphere, the angular deflection caused by refraction is sufficiently large that it would require the source of light seen by the observer to have emerged from beyond the angular extent subtended by the finite-sized star. As a result, the spectrum exhibits a flat continuum from the refraction limit and contains no information from the atmosphere below. However, because refraction depends on the geometry of the planet-star system and the bulk nature of the planetary atmosphere (Misra et al., 2014), it should be possible to correctly diagnose when it is the dominant effect controlling the continuum. Nevertheless, if we are fortunate enough to observe terrestrial exoplanets with thick atmospheres that do not possess high altitude aerosols, then our interpretation of the lower atmosphere via the continuum will contend with a cloud-surface-refraction degeneracy.

8.1.3 *Identifying Habitable Exoplanets*

The search for habitable exoplanets forms a cornerstone of modern exoplanetary astronomy as we pursue Earth-like environments with stable liquid water on their surfaces. Due to the habitability being a surface phenomenon (as defined in this work), an unambiguous identification of a habitable planet faces numerous observational challenges. In the near term, *JWST* will be used to attempt

to indirectly assess the habitability of exoplanets, like TRAPPIST-1 e, f and g, by observing the transmission spectrum of their upper atmospheres and extrapolating the inferred composition down to the surface, or where we think the surface is. Later on, a direct-imaging mission like LUVOIR will be capable of performing direct habitability assessments by characterizing the composition of the surface and lower atmosphere, and searching for ocean glint. The work in this dissertation has provided crucial insights into both future paths to identifying habitable exoplanets.

The aforementioned challenges for precise terrestrial exoplanet atmospheric characterization in the *JWST* era are likely to also limit habitability assessments in the near future, where numerous factors are likely to complicate the remote inference of surface liquid water. [Robinson \(2017b\)](#) describe multiple approaches to detect surface liquid water on rocky exoplanets, including both direct and indirect methods. Due to its insensitivity to ocean glint for nearby HZ M dwarf planets (e.g. Proxima Centauri b; [Meadows et al., 2018a](#)), *JWST* will be limited to indirect methods to assess planetary habitability, such as surface pressure and temperature constraints that are consistent with a liquid water surface or a water vapor mixing ratio that increases towards the surface ([Robinson, 2017b](#)), which would still remain inconclusive. However, we have already seen that emission spectroscopy of HZ terrestrials, like TRAPPIST-1e, will likely be infeasible with *JWST*, ruling out the potential to use surface inferences from thermal emission. As a result, transmission spectroscopy appears to be the only viable path to study rocky exoplanets in the HZ with *JWST*, a result which is supported by our findings in Chapter 2 ([Lustig-Yaeger et al., 2019b](#)).

Transmission spectroscopy with *JWST* is likely to be insensitive to, or at least ambiguous about, the surface conditions of terrestrial exoplanets, and therefore provide little to no evidence to establish planetary habitability. In Chapter 2 ([Lustig-Yaeger et al., 2019b](#)) and Chapter 7 we demonstrated that the presence of water clouds and/or CH₄ absorption features have the potential to completely obscure the NIR H₂O features, especially at the relatively high altitudes in the atmosphere that transmission spectroscopy probes. Numerous other recent studies have also shown that clouds may significantly obscure H₂O features ([Fauchez et al., 2019](#); [Komacek et al., 2020](#); [Pidhorodetska et al., 2020](#); [Suisa et al., 2020](#)). In addition to difficulties detecting atmospheric water vapor from transmission spectra, the aforementioned cloud-surface-refraction degeneracy is

likely to make ambiguous any potential inference of surface temperature and pressure—e.g. *are we even seeing the surface?* This issue is further compounded by the scale height degeneracies associated with transmission spectroscopy retrievals explored in Chapter 7. Specifically, uncertainties in the surface gravity (planetary mass) and atmospheric mean molecular weight (atmospheric bulk composition) will propagate directly into the scale temperature retrieved from a transmission spectrum. These intrinsic factors lay bare the likely insurmountable difficulties that *JWST* will face for exoplanet habitability assessments, and ultimately provide a warning about relying on future transmission spectroscopy missions to succeed in the footsteps of *JWST*.

The HabEx (HabEx Study Team, 2019) and LUVOIR (LUVOIR Mission Concept Study Team, 2019) direct imaging mission concepts provide a compelling pathway towards robust habitability assessments in the future. Not only does direct imaging offer greater sensitivity to the planetary surface environment than transmission spectroscopy, but high albedo clouds can increase the spectral contrast and S/N on atmospheric absorption features (e.g. Rugheimer et al., 2015), thereby turning the tables on the effect of clouds and increasing the likelihood of success for atmospheric characterization. Furthermore, by observing light reflected off rocky exoplanets, these possible future direct imaging missions will be capable of directly inferring habitability by detecting the non-Lambertian scattering signature of ocean glint at crescent phase (Williams and Gaidos, 2008; Robinson et al., 2010).

In Chapter 4 (Lustig-Yaeger et al., 2018), we developed a novel observing strategy that could yield a robust ocean detection in the future using a direct-imaging mission. This approach combines two previously published ocean detection techniques—ocean glint (Sagan et al., 1993; Williams and Gaidos, 2008; Robinson et al., 2010) and rotational mapping (Cowan et al., 2009; Fujii et al., 2011; Cowan and Strait, 2013; Fujii et al., 2017)—into one rigorous method to identify the unique observables of oceans on the surface of an exoplanet. Our method leverages the time-dependent glint signal, which may be modulated in time by Lambertian scattering land masses, to longitudinally map the oceans as they rotate into and out of the illuminated planet crescent (Lustig-Yaeger et al., 2018). We demonstrated that reflected observations of the planet at both quadrature and crescent phases can be used to identify and decouple the Lambertian and specular (glint) scatter-

ing surfaces, thereby enhancing the glint signal by $\sim 2.5\times$ that reported by [Robinson et al. \(2010\)](#) in the disk average, and enabling surfaces with unique scattering phase functions to be spatially mapped. With this new technique and its enhancement in sensitivity, we estimate that a 15 m space telescope like LUVOIR would be able to search for oceans on roughly 10 exoplanets orbiting in the HZ of the nearest G, K, and M dwarfs. Thus, it is clear that a future mission capable of observing the reflected light of rocky exoplanets would enable a survey of exoplanet habitability.

8.1.4 *Confirming Exoplanet Biosignatures*

Detecting life outside the Solar System will require a rigorous search for biosignatures on the surfaces and in the atmospheres of rocky exoplanets. Conducting such a search will necessarily require a thorough characterization of the planet's atmospheric composition both to detect biosignature gases that may implicate biology and to rule out false positive gases that may indicate that any biosignature gas is produced abiotically by the environment. Ascertaining the lines of evidence that point to exobiology is the pinnacle of exoplanet atmospheric characterization, not just because the discovery would be so grand, but because it will fundamentally require each characterization milestone answered about the exoplanet environment, from detecting the atmosphere, to characterizing its nature and composition, through to assessing its surface habitability. In this dissertation, we have examined key pieces of exoplanet biosignature detection in the near and far term that help to shape the outlook for this monumental endeavor going forward.

Our work has shown that confirming an inhabited world via biosignature detection will be challenging, but longshot scenarios in the *JWST* era and careful planning for the next generation of space telescopes is likely to place this goal within grasp. The general issues discussed for terrestrial exoplanet characterization and habitability assessment in the *JWST* era, such as surface insensitivity, also apply to biosignature detection. In Chapter 7, we elucidated additional challenges associated with biosignature detection using *JWST*, and also outlined intriguing plausible outcomes. In agreement with previous studies ([Barstow et al., 2016](#); [Krissansen-Totton et al., 2018a](#)), we found that *JWST* will not be capable of robustly detecting the biosignatures O_2 and O_3 associated with oxygenic photosynthesis, even for the most optimistic current target, TRAPPIST-1e. The photo-

chemically self-consistent atmospheric models of [Meadows et al. \(2020\)](#) only strengthen this result because of the suppression of O_3 formation.

However, we also reached a similar conclusion to [Krissansen-Totton et al. \(2018a\)](#) that the CH_4 – CO_2 disequilibrium biosignature pair may be detectable for TRAPPIST-1e if it possesses an Archean Earth-like anoxic biosphere. We built upon this finding by showing that it is also possible to infer the CH_4 – CO_2 biosignature from transmission spectrum observations of TRAPPIST-1e if it possesses a modern Earth-like biosphere. This result follows from the photochemically-enhanced CH_4 abundance expected for Earth-like planets orbiting M dwarfs ([Segura et al., 2005](#); [Meadows et al., 2018a](#)). This raises the question of whether we would be able to distinguish Earth’s biosphere *through time* around an M dwarf. That is, can we count on the simultaneous presence of abundant CH_4 and CO_2 as a general biosignature, or are there false positive scenarios—such as geologic outgassing of CH_4 in an oxidized atmosphere or significant H_2O photolysis and O_2 release in a reducing atmosphere—that could mimic this biosignature? Furthermore, [Krissansen-Totton et al. \(2018b\)](#) discussed the importance of both N_2 and surface liquid water coexisting with CH_4 and CO_2 to establish the atmospheric disequilibrium of the Archean biosphere, combined with a non-detection of CO to rule out a geological origin of the CH_4 and CO_2 . This additional context may be particularly challenging to obtain with *JWST* considering both the caveats explored in Chapter 7 related to inferring the atmospheric bulk composition (e.g. N_2) and the numerous challenges to habitability assessment previously discussed. Although more work will be needed to confront these issues, the CH_4 – CO_2 disequilibrium pair offers an intriguing and potentially detectable biosignature to investigate further with *JWST*.

However, future direct imaging missions, like the 15 m LUVOIR mission concept, are being designed to have the wavelength range and sensitivity needed for more robust habitability and biosignature detection for a larger number of targets. In Chapter 5 we sought to find the number of Earth-like exoplanets for which the future LUVOIR mission could search for biosignatures and rule out false positive biosignatures. We found that the LUVOIR architecture as defined by the [LUVOIR Mission Concept Study Team \(2019\)](#) would be capable of observing reflected spectra between the near-UV through near-IR for nearly two dozen Earth-like exoplanets in a nominal one

year dedicated survey. We also found that these yield estimates are highly sensitive to the methodology used to prioritize target selection for intensive spectroscopic follow-up after each system's initial detection. In particular, when we simply prioritized the targets in order of increasing exposure time we minimized the bias among the sample of rocky HZ exoplanets and maximized the yield. However, when we down-selected for targets that met a specific criteria in the initial detection reconnaissance, such as having detectable H₂O features in their spectrum, we significantly reduced the spectral yield and imposed a bias among the planets that we would go on to rigorously characterize. This bias results from the attempt to test the specific hypothesis that planets with water features are likely to be habitable and habitable planets are more likely to have life. This highlights the difference between an unbiased survey that may better optimize the yield to provide larger statistical sample, and a necessarily biased survey that may optimally test a specific hypothesis for a smaller number of targets. Careful consideration will be required to conduct optimized biosignature searches in the future, and significantly more work must be done to weigh and balance these competing approaches¹.

8.2 Ongoing & Future Work

The investigations described in this thesis have generated a suite of multi-purpose tools that enhance observation planning and enable the retrieval of key planetary properties from observations. Most significantly, our new SMARTER atmospheric retrieval model will continue to be developed to include additional physical models and parameterizations, to be validated on additional Solar System planets, and to be applied to test more retrieval approaches and assumptions prior to the acquisition of *JWST* data. After *JWST* launches, SMARTER will be used to interpret the spectra of the TRAPPIST-1, and other Earth-sized, exoplanets around late M dwarfs to detect and characterize their atmospheres. It will also be essential to use SMARTER to reproduce and expand upon the direct-imaging retrieval work of [Feng et al. \(2018\)](#) to explore the potential for LUVOIR or HabEx to characterize additional possible atmospheric compositions beyond modern Earth-like,

¹For recent proposed statistical characterization surveys of Earth-like exoplanets, see [Bean et al. \(2017\)](#); [Turbet et al. \(2019\)](#); [Bixel and Apai \(2020\)](#).

combining the approaches taken in Chapter 5 and Chapter 7.

Our work demonstrated the importance of considering the effects of clouds and hazes on terrestrial exoplanet atmospheric characterization, however our retrieval work with SMARTER in this arena remains unfinished. In Chapter 6, we validated SMARTER on clear-sky observations of Earth and we detailed ongoing work to validate SMARTER on the haze-afflicted transmission spectrum of Saturn’s moon, Titan (§6.4.1). Like the Earth validation, completing the Titan validation is likely to lend insights into the future characterization of hazy exoplanets. To accomplish this task and prepare for future aerosol retrievals, we plan to incorporate each of the cloud models investigated by Mai and Line (2019) into the SMARTER forward models. This will be crucial for assessing degeneracies and biases when retrieving hazes and molecules. Furthermore, these cloud models will facilitate the expansion of the synthetic TRAPPIST-1e retrieval work in Chapter 7 to studying the hazy Archean Earth-like spectrum presented in Meadows et al. (2020, *in prep*). Undoubtedly, *JWST* observations of terrestrial exoplanets will contend with cloudy atmospheres.

This work has highlighted the extent to which clouds may appear indistinguishable from solid surfaces for transit transmission observations, but significant work is still needed to understand the impact of clouds on direct imaging retrievals. In Chapter 3, we cautioned about how a cloud-top–solid-surface degeneracy could make a class of Venus-analog exoplanets with thick, cloudy atmospheres appear indistinguishable from a class of Mars-analog exoplanets with thin atmospheres using *JWST* transmission spectroscopy. A similar, albeit much more complex, degeneracy is likely to exist for directly-imaged rocky exoplanets using reflectance spectroscopy. In particular, in limited S/N spectra the wavelength-dependence of cloud extinction, surface reflectance, and the gas continuum absorption are all liable to get tangled in their interpretation. This will only be exacerbated as we develop agnostic approaches to fit the surface albedo of exoplanets that are capable of taking on any residual spectral shape. Fortunately, phase-dependent or high spectral resolution observations may hold the key to breaking this degeneracy. If LUVOIR is prioritized for future NASA funding it will be essential to investigate the extent of this degeneracy and develop approaches to identify and mitigate against it.

As we progress into the era of terrestrial exoplanet characterization it will become increasingly

important that we synthesize all available information about each observed system within the context of known physical and chemical relationships gleaned from decades of planetary science in the Solar System. As discussed in Chapter 6, SMARTER has been designed to incorporate additional physical and chemical models to expand upon the traditional atmospheric retrieval framework. Chief among these planned future developments is to incorporate the 1D photochemistry model ATMOS (Kasting et al., 1979; Segura et al., 2005; Zahnle et al., 2006; Arney et al., 2016; Schwieterman et al., 2016b; Meadows et al., 2018a; Lincowski et al., 2018), so that the retrieved atmospheric states can be related to (1) the host star’s UV spectrum that is incident on the top of atmosphere and which drives disequilibrium causing photochemical reactions, and (2) the gas fluxes from the surface of the planet that may be produced by geological and/or biological processes that must be disentangled to confidently claim that a biosignature has been detected. Furthermore, refining the APPROXPOSTERIOR machine learning inverse model within SMARTER will help enhance these physically augmented atmospheric retrieval use-cases by significantly improving the efficiency of obtaining reliable retrieval results, thereby making these complex modeling endeavors computationally tractable.

Chapter 9

CONCLUSIONS

In this thesis we have explored fundamental approaches for studying terrestrial exoplanet atmospheres at present and into the future. We dissected the broader topic of terrestrial exoplanet atmospheric characterization into four key objectives to consider for every planetary system found that hosts one or more small rocky planets. These objectives are to (1) detect the atmosphere, (2) characterize the nature and composition of the atmosphere, (3) assess habitability, and (4) identify and confirm any biosignatures. Adopting these as major reoccurring themes, we examined the potential for each key objective to be met in the near term with the *James Webb Space Telescope*, and we identified crucial capabilities that a farther future telescope will need to satisfy to improve our chances of success in the future. We used these findings as motivation and a guide to develop and validate a first-of-its-kind, physically robust, terrestrial exoplanet atmospheric retrieval model for interpreting the atmospheric composition of rocky exoplanets using spectroscopic observations with *JWST* and future telescopes when they come online. To demonstrate the capabilities of our new retrieval model, we analyzed synthetic *JWST* observations of the potentially habitable planet TRAPPIST-1e and revealed new insights and challenges for practical terrestrial exoplanet atmospheric characterization.

Our first major finding is that *JWST* will have an opportunity to detect and begin to characterize the atmospheres of Earth-size planets that transit nearby late M dwarfs, such as, in particular, the TRAPPIST-1 system (Part I). We found that if the TRAPPIST-1 planets have thick and evolved secondary atmospheres, then *JWST* will be capable of placing strong constraints on the detection of atmospheres for all seven TRAPPIST-1 planets in fewer than 10 observed transits (Chapter 2; Lustig-Yaeger et al., 2019b). In particular, our results suggest that efficient use of *JWST*'s NIRSpec Prism instrument between $1 - 5 \mu\text{m}$ is an optimal approach for detecting terrestrial atmospheres due to the presence of strong CO_2 bands at 2.0, 2.7, and $4.3 \mu\text{m}$ in transmission spectra that may be expected for any oxidized atmosphere. Near-IR transit observations with *JWST* may also provide an opportunity to detect $\text{O}_2 - \text{O}_2$ CIA bands at 1.06 and $1.27 \mu\text{m}$ to test the hypothesis of pre-main-sequence atmospheric oxygen build-up. We identified multiple ways that the presence of clouds may impede and even confound such *JWST* transmission spectroscopy efforts (Chapter 3; Lustig-Yaeger et al., 2019a). First, cloud-covered planets have flattened transmission spectra

with reduced molecular absorption feature amplitudes that may require $2 - 12\times$ more transits with *JWST* to detect their atmospheres. Second, sulfuric acid clouds in Venus-like planetary atmospheres appear sufficiently flat and featureless in transit spectra that they may be mistaken as a solid surface and therefore make it near impossible to distinguish Venus-like planets with thick-cloudy atmospheres from Mars-like planets with thin arid atmospheres. Third, water clouds in habitable atmospheres can prevent the detection of water vapor in their spectrum, limiting *JWST*'s ability to probe planetary habitability. Numerous factors—including clouds, transit geometry, and *JWST*'s sensitivity—indicate that *JWST* will be insufficient to rigorously characterize terrestrial exoplanets for signs of habitability and biosignatures.

Our second major finding is that a future space-based direct-imaging mission that is capable of observing terrestrial exoplanets in reflected light may provide an excellent opportunity to perform robust habitability and biosignature assessments (Part II). At present, the LUVOIR telescope concept stands as the optimal mission for achieving these goals. To this end, we introduced a novel combination of multi-wavelength, multi-phase, time-series direct-imaging observations and accompanying analyses that may improve the robustness of exoplanet ocean detection by spatially mapping the ocean glint signal (Chapter 4; Lustig-Yaeger et al., 2018). As an ocean-bearing planet rotates, the glint spot may appear to “blink” as Lambertian scattering continents interrupt the specular reflection from the ocean. This causes a strong source of periodic variability in crescent-phase disk-integrated reflected light curves. These light curves can be inverted to retrieve the longitudinal slice maps and apparent albedo of multiple surfaces at both quadrature and crescent phase. At crescent phase, we found that the retrieved apparent albedo of ocean-bearing longitudinal slices was increased by a factor of 5, compared to the albedo at quadrature phase, due to the contribution from glint, while the land-bearing slices exhibited no significant change in apparent albedo with phase. We posit that such multi-phase longitudinal mapping has the potential to provide a definitive ocean glint detection for between $1 - 10$ HZ exoplanets orbiting the nearest G, K, and M dwarfs using a space-based, high-contrast, direct-imaging telescope with a diameter between 6 and 15 m. Furthermore, we used the coronagraph design specifications for the two LUVOIR mission architectures to calculate the expected yield of high S/N reflectance spectra of Earth-analog exo-

planets (Chapter 5). We found that the 15 meter LUVOIR-A design would be capable of observing the spectra of approximately 24 Earth-analog exoplanets with sufficient S/N to detect Earth-like biosignatures in a dedicated one-year characterization survey. The 6 meter LUVOIR-B design would be capable of observing the spectra of 11 exo-Earths. These results only scratch the surface of what a large-aperture, space-based, coronagraph-equipped telescope could do to study terrestrial exoplanet atmospheres, habitability, and biosignatures.

Our third major accomplishment is the development, validation, and application of a novel terrestrial exoplanet atmospheric retrieval code designed to interpret the spectra of Earth-size exoplanets observed with *JWST* in the next few years, upcoming ground-based extremely large telescopes in the late 2020s, and future large-aperture space missions in the late 2030s (Part III). Our atmospheric retrieval model incorporates a line-by-line radiative transfer code that has a strong legacy of modeling the terrestrial planets in the Solar System, and provides a level of physical rigor that complements other more computationally efficient retrieval models in the literature and offers a platform for continued growth in this area (Chapter 6). We included four different flexible Bayesian inverse models that help to offset the computational expense of our physical forward model by allowing a multi-pronged approach to atmospheric parameter estimation. We validated our new retrieval model on high S/N, high spectral resolution transmission spectrum observations of Earth, which showcased the robustness of our model, but also illuminated a potential bias that can manifest from using a single 1D atmospheric structure to fit high resolution spectra of a true, 3D planet. Finally, we applied our retrieval model to simulated *JWST* transmission spectra to determine what we may be able to infer about the atmosphere of TRAPPIST-1e given a relatively substantial observing program of 25 transits (Chapter 7). Our results largely echoed our initial TRAPPIST-1 results that evidence of an atmosphere may be relatively easily accessed via CO₂ absorption features. However, we also demonstrated that the abundances of CO₂ and CH₄ can be retrieved within about an order of magnitude if the planet possesses an Archean or a modern Earth-like atmosphere, further raising the question of whether the CO₂–CH₄ disequilibrium biosignature pair can be generally relied upon without substantial evidence of other trace gases for added atmospheric context. This work also shed light on additional biases that may result from the

common retrieval assumption of a background gas, such as N_2 , and to a lesser extent, excluding the physics of atmospheric refraction when fitting the spectra of planets with thick, aerosol-free atmospheres. We underscore the utility of a focused terrestrial exoplanet retrieval model for testing various observing and modeling approaches in advance of actually acquiring such time-intensive datasets.

In conclusion, we anticipate a bright future for terrestrial exoplanetary science. In this work, we have presented a reasoned outlook and a natural pathway towards answering our burning questions about rocky planets orbiting distant stars. Rather than being The End, this is truly The Beginning of a long journey. A journey with the potential to place our human experience here on Earth in the greater context of life in the universe. Whether we find that life in the cosmos is ubiquitous and we are but one of many living havens, or we find that life in the cosmos is ever fleeting and we are alone as far as our star-gazing eyes can see, this knowledge has the potential to help us stop and reflect on our humble beginnings, to reorient our priorities, and to shape the trajectory of our civilization.

“It will not be we who reach Alpha Centauri and the other nearby stars. It will be a species very like us, but with more of our strengths and fewer of our weaknesses, more confident, farseeing, capable, and prudent. For all our failings, despite our limitations and fallibilities, we humans are capable of greatness.”

– Sagan (1997), “Pale Blue Dot”

BIBLIOGRAPHY

- Ackerman, A. S. and Marley, M. S. (2001). Precipitating Condensation Clouds in Substellar Atmospheres. *ApJ*, 556(2):872–884.
- Airapetian, V. S., Gloer, A., Khazanov, G. V., Loyd, R. O. P., France, K., Sojka, J., Danchi, W. C., and Liemohn, M. W. (2017). How Hospitable Are Space Weather Affected Habitable Zones? The Role of Ion Escape. *ApJ*, 836:L3.
- Allen, D. A. (1987). The dark side of Venus. *Icarus*, 69(2):221–229.
- Allen, D. A. and Crawford, J. W. (1984). Cloud structure on the dark side of Venus. *Nature*, 307(5948):222–224.
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., and O’Neil, M. (2015). Fast Direct Methods for Gaussian Processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38:252.
- Anglada-Escudé, G., Amado, P. J., Barnes, J., Berdiñas, Z. M., Butler, R. P., Coleman, G. A. L., de La Cueva, I., Dreizler, S., Endl, M., Giesers, B., Jeffers, S. V., Jenkins, J. S., Jones, H. R. A., Kiraga, M., Kürster, M., López-González, M. J., Marvin, C. J., Morales, N., Morin, J., Nelson, R. P., Ortiz, J. L., Ofir, A., Paardekooper, S.-J., Reiners, A., Rodríguez, E., Rodríguez-López, C., Sarmiento, L. F., Strachan, J. P., Tsapras, Y., Tuomi, M., and Zechmeister, M. (2016). A terrestrial planet candidate in a temperate orbit around Proxima Centauri. *Nature*, 536:437–440.
- Arking, A. and Potter, J. (1968). The phase curve of Venus and the nature of its clouds. *Journal of Atmospheric Sciences*, 25:617–628.
- Arney, G., Domagal-Goldman, S. D., and Meadows, V. S. (2018). Organic Haze as a Biosignature in Anoxic Earth-like Atmospheres. *Astrobiology*, 18:311–329.

- Arney, G., Domagal-Goldman, S. D., Meadows, V. S., Wolf, E. T., Schwieterman, E., Charnay, B., Claire, M., Hébrard, E., and Trainer, M. G. (2016). The Pale Orange Dot: The Spectrum and Habitability of Hazy Archean Earth. *Astrobiology*, 16(11):873–899.
- Arney, G. and Kane, S. (2018). Venus as an Analog for Hot Earths. *arXiv e-prints*, page arXiv:1804.05889.
- Arney, G., Meadows, V., Crisp, D., Schmidt, S. J., Bailey, J., and Robinson, T. (2014). Spatially resolved measurements of H₂O, HCl, CO, OCS, SO₂, cloud opacity, and acid concentration in the Venus near-infrared spectral windows. *Journal of Geophysical Research (Planets)*, 119:1860–1891.
- Arney, G. N. (2019). The K Dwarf Advantage for Biosignatures on Directly Imaged Exoplanets. *ApJ*, 873(1):L7.
- Arney, G. N., Meadows, V. S., Domagal-Goldman, S. D., Deming, D., Robinson, T. D., Tovar, G., Wolf, E. T., and Schwieterman, E. (2017). Pale Orange Dots: The Impact of Organic Haze on the Habitability and Detectability of Earthlike Exoplanets. *ApJ*, 836:49.
- Bagnasco, G., Kolm, M., Ferruit, P., Honnen, K., Koehler, J., Lemke, R., Maschmann, M., Melf, M., Noyer, G., Rumler, P., Salvignol, J.-C., Strada, P., and Te Plate, M. (2007). Overview of the near-infrared spectrograph (NIRSpec) instrument on-board the James Webb Space Telescope (JWST). In *Cryogenic Optical Systems and Instruments XII*, volume 6692 of Proc. SPIE, page 66920M.
- Baraffe, I., Chabrier, G., Allard, F., and Hauschildt, P. H. (1998). Evolutionary models for solar metallicity low-mass stars: mass-magnitude relationships and color-magnitude diagrams. *A&A*, 337:403–412.
- Baraffe, I., Homeier, D., Allard, F., and Chabrier, G. (2015). New evolutionary models for pre-main sequence and main sequence low-mass stars down to the hydrogen-burning limit. *A&A*, 577:A42.

- Barath, F. T., Barrett, A. H., Copeland, J., Jones, D. E., and Lilley, A. E. (1963). Microwave Radiometers. *Science*, 139(3558):908–909.
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., and Fletcher, L. N. (2016). Telling twins apart: exo-Earths and Venuses with transit spectroscopy. *Monthly Notices of the Royal Astronomical Society*, 458:2657–2666.
- Barstow, J. K., Aigrain, S., Irwin, P. G. J., and Sing, D. K. (2017). A Consistent Retrieval Analysis of 10 Hot Jupiters Observed in Transmission. *ApJ*, 834(1):50.
- Barstow, J. K., Changeat, Q., Garland, R., Line, M. R., Rocchetto, M., and Waldmann, I. P. (2020). A comparison of exoplanet spectroscopic retrieval tools. *MNRAS*, 493(4):4884–4909.
- Barstow, J. K. and Heng, K. (2020). Outstanding Challenges of Exoplanet Atmospheric Retrievals. *arXiv e-prints*, page arXiv:2003.14311.
- Barstow, J. K. and Irwin, P. G. J. (2016). Habitable worlds with JWST: transit spectroscopy of the TRAPPIST-1 system? *Monthly Notices of the Royal Astronomical Society*, 461:L92–L96.
- Barstow, J. K., Tsang, C. C. C., Wilson, C. F., Irwin, P. G. J., Taylor, F. W., McGouldrick, K., Drossart, P., Piccioni, G., and Tellmann, S. (2012). Models of the global cloud structure on Venus derived from Venus Express observations. *Icarus*, 217(2):542–560.
- Batalha, N., Stevenson, K., Hill, M., Fraine, J., eas342, and Cubillos, P. (2018). natasha-batalha/pandexo: Starting pandexo releases.
- Batalha, N. E., Lewis, N. K., Line, M. R., Valenti, J., and Stevenson, K. (2018). Strategies for Constraining the Atmospheres of Temperate Terrestrial Planets with JWST. *ApJ*, 856:L34.
- Batalha, N. E., Lewis, T., Fortney, J. J., Batalha, N. M., Kempton, E., Lewis, N. K., and Line, M. R. (2019). The Precision of Mass Measurements Required for Robust Atmospheric Characterization of Transiting Exoplanets. *ApJ*, 885(1):L25.

- Batalha, N. E., Mandell, A., Pontoppidan, K., Stevenson, K. B., Lewis, N. K., Kalirai, J., Earl, N., Greene, T., Albert, L., and Nielsen, L. D. (2017). PandExo: A Community Tool for Transiting Exoplanet Science with JWST HST. *Publications of the Astronomical Society of the Pacific*, 129(6):064501.
- Batalha, N. M. (2014). Exploring exoplanet populations with nasa's kepler mission. *Proceedings of the National Academy of Sciences*, 111(35):12647–12654.
- Battersby, C., Armus, L., Bergin, E., Kataria, T., Meixner, M., Pope, A., Stevenson, K. B., Cooray, A., Leisawitz, D., Scott, D., et al. (2018). The origins space telescope. *Nature Astronomy*, 2(8):596–599.
- Batygin, K., Adams, F. C., Brown, M. E., and Becker, J. C. (2019). The planet nine hypothesis. *Physics Reports*.
- Baum, B. A., Yang, P., Heymsfield, A. J., Platnick, S., King, M. D., Hu, Y. X., and Bedka, S. T. (2005). Bulk Scattering Properties for the Remote Sensing of Ice Clouds. Part II: Narrowband Models. *Journal of Applied Meteorology*, 44(12):1896–1911.
- Bean, J. L., Abbot, D. S., and Kempton, E. M. R. (2017). A Statistical Comparative Planetology Approach to the Hunt for Habitable Exoplanets and Life Beyond the Solar System. *ApJ*, 841(2):L24.
- Bellucci, A., Sicardy, B., Drossart, P., Rannou, P., Nicholson, P. D., Hedman, M., Baines, K. H., and Burrati, B. (2009). Titan solar occultation observed by Cassini/VIMS: Gas absorption and constraints on aerosol composition. *Icarus*, 201:198–216.
- Belu, A. R., Selsis, F., Morales, J.-C., Ribas, I., Cossou, C., and Rauer, H. (2011). Primary and secondary eclipse spectroscopy with JWST: exploring the exoplanet parameter space. *Astronomy & Astrophysics*, 525:A83.
- Benneke, B. (2015). Strict Upper Limits on the Carbon-to-Oxygen Ratios of Eight Hot Jupiters from Self-Consistent Atmospheric Retrieval. *arXiv e-prints*, page arXiv:1504.07655.

- Benneke, B. and Seager, S. (2012). Atmospheric Retrieval for Super-Earths: Uniquely Constraining the Atmospheric Composition with Transmission Spectroscopy. *ApJ*, 753:100.
- Benneke, B. and Seager, S. (2013). How to Distinguish between Cloudy Mini-Neptunes and Water/Volatile-dominated Super-Earths. *ApJ*, 778(2):153.
- Berdyugina, S. V. and Kuhn, J. R. (2017). Surface Imaging of Proxima b and Other Exoplanets: Topography, Biosignatures, and Artificial Mega-Structures. *ArXiv e-prints*.
- Bernath, P. F., McElroy, C. T., Abrams, M. C., Boone, C. D., Butler, M., Camy-Peyret, C., Carleer, M., Clerbaux, C., Coheur, P. F., Colin, R., DeCola, P., DeMazière, M., Drummond, J. R., Dufour, D., Evans, W. F. J., Fast, H., Fussen, D., Gilbert, K., Jennings, D. E., Llewellyn, E. J., Lowe, R. P., Mahieu, E., McConnell, J. C., McHugh, M., McLeod, S. D., Michaud, R., Midwinter, C., Nassar, R., Nichitiu, F., Nowlan, C., Rinsland, C. P., Rochon, Y. J., Rowlands, N., Semeniuk, K., Simon, P., Skelton, R., Sloan, J. J., Soucy, M. A., Strong, K., Tremblay, P., Turnbull, D., Walker, K. A., Walkty, I., Wardle, D. A., Wehrle, V., Zander, R., and Zou, J. (2005). Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.*, 32(15):L15S01.
- Berta, Z. K., Charbonneau, D., Désert, J.-M., Miller-Ricci Kempton, E., McCullough, P. R., Burke, C. J., Fortney, J. J., Irwin, J., Nutzman, P., and Homeier, D. (2012). The Flat Transmission Spectrum of the Super-Earth GJ1214b from Wide Field Camera 3 on the Hubble Space Telescope. *The Astrophysical Journal*, 747:35.
- Berta-Thompson, Z. K., Irwin, J., Charbonneau, D., Newton, E. R., Dittmann, J. A., Astudillo-Defru, N., Bonfils, X., Gillon, M., Jehin, E., Stark, A. A., Stalder, B., Bouchy, F., Delfosse, X., Forveille, T., Lovis, C., Mayor, M., Neves, V., Pepe, F., Santos, N. C., Udry, S., and Wünsche, A. (2015). A rocky planet transiting a nearby low-mass star. *Nature*, 527:204–207.
- Bétrémieux, Y. and Kaltenegger, L. (2013). Transmission Spectrum of Earth as a Transiting Exoplanet from the Ultraviolet to the Near-infrared. *ApJ*, 772:L31.

- Bétrémieux, Y. and Kaltenegger, L. (2014). Impact of Atmospheric Refraction: How Deeply can We Probe Exo-Earth's Atmospheres during Primary Eclipse Observations? *The Astrophysical Journal*, 791:7.
- Bétrémieux, Y. and Swain, M. R. (2018). The hidden depths of planetary atmospheres. *ArXiv e-prints*.
- Beuzit, J.-L., Feldt, M., Dohlen, K., Mouillet, D., Puget, P., Wildi, F., Abe, L., Antichi, J., Baruffolo, A., Baudoz, P., et al. (2008). Sphere: a planet finder instrument for the vlt. In *Ground-based and airborne instrumentation for astronomy II*, volume 7014, page 701418. International Society for Optics and Photonics.
- Bixel, A. and Apai, D. (2020). Testing Earth-like atmospheric evolution on exo-Earths through oxygen absorption: required sample sizes and the advantage of age-based target selection. *arXiv e-prints*, page arXiv:2005.01587.
- Blecic, J., Dobbs-Dixon, I., and Greene, T. (2017). The Implications of 3D Thermal Structure on 1D Atmospheric Retrieval. *ApJ*, 848(2):127.
- Bolcar, M. R., Aloezos, S., Bly, V. T., Collins, C., Crooke, J., Dressing, C. D., Fantano, L., Feinberg, L. D., France, K., Gochar, G., Gong, Q., Hylan, J. E., Jones, A., Linares, I., Postman, M., Pueyo, L., Roberge, A., Sacks, L., Tompkins, S., and West, G. (2017). The Large UV/Optical/Infrared Surveyor (LUVOIR): Decadal Mission concept design update. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 10398 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1039809.
- Bolmont, E., Selsis, F., Owen, J. E., Ribas, I., Raymond, S. N., Leconte, J., and Gillon, M. (2017). Water loss from terrestrial planets orbiting ultracool dwarfs: implications for the planets of TRAPPIST-1. *MNRAS*, 464:3728–3741.
- Bond, J. C., Lauretta, D. S., and O'Brien, D. P. (2010). Making the Earth: Combining dynamics and chemistry in the Solar System. *Icarus*, 205(2):321–337.

- Borucki, W. J., Koch, D. G., Lissauer, J. J., Basri, G. B., Caldwell, J. F., Cochran, W. D., Dunham, E. W., Geary, J. C., Latham, D. W., Gilliland, R. L., et al. (2003). The kepler mission: a wide-field-of-view photometer designed to determine the frequency of earth-size planets around solar-like stars. In *Future EUV/UV and Visible Space Astrophysics Missions and Instrumentation*, volume 4854, pages 129–140. International Society for Optics and Photonics.
- Bouchet, P., García-Marín, M., Lagage, P.-O., Amiaux, J., Auguères, J.-L., Bauwens, E., Blommaert, J. A. D. L., Chen, C. H., Detre, Ö. H., Dicken, D., Dubreuil, D., Galdemard, P., Gastaud, R., Glasse, A., Gordon, K. D., Gognaud, F., Guillard, P., Justtanont, K., Krause, O., Leboeuf, D., Longval, Y., Martin, L., Mazy, E., Moreau, V., Olofsson, G., Ray, T. P., Rees, J.-M., Renotte, E., Ressler, M. E., Ronayette, S., Salasca, S., Scheithauer, S., Sykes, J., Thelen, M. P., Wells, M., Wright, D., and Wright, G. S. (2015). The Mid-Infrared Instrument for the James Webb Space Telescope, III: MIRIM, The MIRI Imager. *PASP*, 127:612.
- Bourrier, V., de Wit, J., Bolmont, E., Stamenković, V., Wheatley, P. J., Burgasser, A. J., Delrez, L., Demory, B. O., Ehrenreich, D., Gillon, M., Jehin, E., Leconte, J., Lederer, S. M., Lewis, N., Triaud, A. H. M. J., and Van Grootel, V. (2017). Temporal Evolution of the High-energy Irradiation and Water Content of TRAPPIST-1 Exoplanets. *AJ*, 154(3):121.
- Brandt, T. D. and Spiegel, D. S. (2014). Prospects for detecting oxygen, water, and chlorophyll on an exo-Earth. *Proceedings of the National Academy of Science*, 111(37):13278–13283.
- Brogi, M., de Kok, R. J., Albrecht, S., Snellen, I. A. G., Birkby, J. L., and Schwarz, H. (2016). ROTATION AND WINDS OF EXOPLANET HD 189733 b MEASURED WITH HIGH-DISPERSION TRANSMISSION SPECTROSCOPY. *The Astrophysical Journal*, 817(2):106.
- Brogi, M., De Kok, R. J., Birkby, J. L., Schwarz, H., and Snellen, I. A. (2014). Carbon monoxide and water vapor in the atmosphere of the non-transiting exoplanet HD 179949 b. *Astronomy and Astrophysics*, 565:1–10.
- Brogi, M., Snellen, I. A., De Kok, R. J., Albrecht, S., Birkby, J., and De Mooij, E. J. (2012). The

- signature of orbital motion from the dayside of the planet τ Boötis b. *Nature*, 486(7404):502–504.
- Brown, R. A. (2005). Single-Visit Photometric and Obscurational Completeness. *ApJ*, 624:1010–1024.
- Buchner, J., Georgakakis, A., Nandra, K., Hsu, L., Rangel, C., Brightman, M., Merloni, A., Salvato, M., Donley, J., and Kocevski, D. (2014). X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *A&A*, 564:A125.
- Burke, C. J., Christiansen, J. L., Mullally, F., Seader, S., Huber, D., Rowe, J. F., Coughlin, J. L., Thompson, S. E., Catanzarite, J., Clarke, B. D., Morton, T. D., Caldwell, D. A., Bryson, S. T., Haas, M. R., Batalha, N. M., Jenkins, J. M., Tenenbaum, P., Twicken, J. D., Li, J., Quintana, E., Barclay, T., Henze, C. E., Borucki, W. J., Howell, S. B., and Still, M. (2015). Terrestrial Planet Occurrence Rates for the Kepler GK Dwarf Sample. *ApJ*, 809:8.
- Burrows, A. S. (2014). Spectra as windows into exoplanet atmospheres. *Proceedings of the National Academy of Sciences*, 111(35):12601–12609.
- Caldas, A., Leconte, J., Selsis, F., Waldmann, I. P., Bordé, P., Rocchetto, M., and Charnay, B. (2019). Effects of a fully 3D atmospheric structure on exoplanet transmission spectra: retrieval biases due to day-night temperature gradients. *A&A*, 623:A161.
- Cameron, A. C. (2001). Spot mapping in cool stars. In *Astrotomography*, pages 183–206. Springer.
- Carlson, R. W., Baines, K. H., Encrenaz, T., Taylor, F. W., Drossart, P., Kamp, L. W., Pollack, J. B., Lellouch, E., Collard, A. D., Calcutt, S. B., Grinspoon, D., Weissman, P. R., Smythe, W. D., Ocampo, A. C., Danielson, G. E., Fanale, F. P., Johnson, T. V., Kieffer, H. H., Matson, D. L., McCord, T. B., and Soderblom, L. A. (1991). Galileo Infrared Imaging Spectroscopy Measurements at Venus. *Science*, 253(5027):1541–1548.
- Catling, D. C. and Kasting, J. F. (2017). *Atmospheric evolution on inhabited and lifeless worlds*. Cambridge University Press.

- Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., Crisp, D., Robinson, T. D., DasSarma, S., Rushby, A. J., Del Genio, A., Bains, W., and Domagal-Goldman, S. (2018). Exoplanet Biosignatures: A Framework for Their Assessment. *Astrobiology*, 18(6):709–738.
- Cenadelli, D. and Bernagozzi, A. (2018). *The Discovery of the First Exoplanets*, page 134.
- Changeat, Q., Keyte, L., Waldmann, I. P., and Tinetti, G. (2019). Impact of planetary mass uncertainties on exoplanet atmospheric retrievals. *arXiv e-prints*, page arXiv:1908.06305.
- Charbonneau, D., Brown, T. M., Noyes, R. W., and Gilliland, R. L. (2002). Detection of an Extrasolar Planet Atmosphere. *The Astrophysical Journal*, 568:377–384.
- Charbonneau, D., Knutson, H. A., Barman, T., Allen, L. E., Mayor, M., Megeath, S. T., Queloz, D., and Udry, S. (2008). The Broadband Infrared Emission Spectrum of the Exoplanet HD 189733b. *The Astrophysical Journal*, 686:1341–1348.
- Charnay, B., Meadows, V., and Leconte, J. (2015a). 3D Modeling of GJ1214b’s Atmosphere: Vertical Mixing Driven by an Anti-Hadley Circulation. *The Astrophysical Journal*, 813:15.
- Charnay, B., Meadows, V., Misra, A., Leconte, J., and Arney, G. (2015b). 3D Modeling of GJ1214bs Atmosphere: Formation of Inhomogeneous High Clouds and Observational Implications. *The Astrophysical Journal Letters*, 813:L1.
- Checlair, J., Abbot, D. S., Webber, R. J., Feng, Y. K., Bean, J. L., Schwieterman, E. W., Stark, C. C., Robinson, T. D., Kempton, E., Alcabes, O. D. N., Apai, D., Arney, G., Cowan, N., Domagal-Goldman, S., Dong, C., Fleming, D. P., Fujii, Y., Graham, R. J., Guzewich, S. D., Hasegawa, Y., Hayworth, B. P. C., Kane, S. R., Kite, E. S., Komacek, T. D., Kopparapu, R. K., Mansfield, M., Marounina, N., Montet, B. T., Olson, S. L., Paradise, A., Popovic, P., Rackham, B. V., Ramirez, R. M., Rau, G., Reinhard, C., Renaud, J., Rogers, L., Walkowicz, L. M., Warren, A., and Wolf, E. T. (2019). A Statistical Comparative Planetology Approach to Maximize the Scientific Return of Future Exoplanet Characterization Efforts. *BAAS*, 51(3):404.

- Chen, H., Wolf, E. T., Kopparapu, R., Domagal-Goldman, S., and Horton, D. E. (2018). Biosignature anisotropy modeled on temperate tidally locked m-dwarf planets. *The Astrophysical Journal Letters*, 868(1):L6.
- Chen, H., Wolf, E. T., Zhan, Z., and Horton, D. E. (2019). Habitability and spectroscopic observability of warm m-dwarf exoplanets evaluated with a 3d chemistry-climate model. *The Astrophysical Journal*, 886(1):16.
- Christiansen, J. L., Ballard, S., Charbonneau, D., Madhusudhan, N., Seager, S., Holman, M. J., Wellnitz, D. D., Deming, D., A'Hearn, M. F., and EPOXI Team (2010). Studying the Atmosphere of the Exoplanet HAT-P-7b Via Secondary Eclipse Measurements with EPOXI, Spitzer, and Kepler. *The Astrophysical Journal*, 710:97–104.
- Clough, S., Kneizys, F., and Davies, R. (1989). Line shape and the water vapor continuum. *Atmospheric research*, 23(3-4):229–241.
- Cobb, A. D., Himes, M. D., Soboczenski, F., Zorzan, S., O'Beirne, M. D., Baydin, A. G., Gal, Y., Domagal-Goldman, S. D., Arney, G. N., Angerhausen, D., et al. (2019). An ensemble of bayesian neural networks for exoplanetary atmospheric retrieval. *The Astronomical Journal*, 158(1):33.
- Cooray, A. R. and Origins Space Telescope Study Team (2017). Origins Space Telescope. In *American Astronomical Society Meeting Abstracts #229*, volume 229 of *American Astronomical Society Meeting Abstracts*, page 405.01.
- Cosentino, R., Lovis, C., Pepe, F., Collier Cameron, A., Latham, D. W., Molinari, E., Udry, S., Bezawada, N., Black, M., Born, A., Buchschacher, N., Charbonneau, D., Figueira, P., Fleury, M., Galli, A., Gallie, A., Gao, X., Ghedina, A., Gonzalez, C., Gonzalez, M., Guerra, J., Henry, D., Horne, K., Hughes, I., Kelly, D., Lodi, M., Lunney, D., Maire, C., Mayor, M., Micela, G., Ordway, M. P., Peacock, J., Phillips, D., Piotto, G., Pollacco, D., Queloz, D., Rice, K., Riverol, C., Riverol, L., San Juan, J., Sasselov, D., Segransan, D., Sozzetti, A., Sosnowska, D., Stobie,

- B., Szentgyorgyi, A., Vick, A., and Weber, L. (2012). *Harps-N: the new planet hunter at TNG*, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 84461V.
- Cowan, N. B., Abbot, D. S., and Voigt, A. (2012). A False Positive for Ocean Glint on Exoplanets: The Latitude-Albedo Effect. *ApJ*, 752:L3.
- Cowan, N. B. and Agol, E. (2008). Inverting Phase Functions to Map Exoplanets. *ApJ*, 678(2):L129.
- Cowan, N. B., Agol, E., and Charbonneau, D. (2007). Hot nights on extrasolar planets: mid-infrared phase variations of hot Jupiters. *MNRAS*, 379(2):641–646.
- Cowan, N. B., Agol, E., Meadows, V. S., Robinson, T., Livengood, T. A., Deming, D., Lisse, C. M., A’Hearn, M. F., Wellnitz, D. D., Seager, S., Charbonneau, D., and EPOXI Team (2009). Alien Maps of an Ocean-bearing World. *ApJ*, 700:915–923.
- Cowan, N. B., Fuentes, P. A., and Haggard, H. M. (2013). Light curves of stars and exoplanets: estimating inclination, obliquity and albedo. *MNRAS*, 434:2465–2479.
- Cowan, N. B., Greene, T., Angerhausen, D., Batalha, N. E., Clampin, M., Colón, K., Crossfield, I. J. M., Fortney, J. J., Gaudi, B. S., Harrington, J., Iro, N., Lillie, C. F., Linsky, J. L., Lopez-Morales, M., Mandell, A. M., and Stevenson, K. B. (2015). Characterizing Transiting Planet Atmospheres through 2025. *PASP*, 127:311.
- Cowan, N. B. and Strait, T. E. (2013). Determining Reflectance Spectra of Surfaces and Clouds on Exoplanets. *ApJ*, 765:L17.
- Cox, C. and Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the sun’s glitter. *JOSA*, 44(11):838–850.
- Crisp, D. (1986). Radiative forcing of the Venus mesosphere I. Solar fluxes and heating rates. *Icarus*, 67(3):484–514.

- Crisp, D. (1997). Absorption of sunlight by water vapor in cloudy conditions: A partial explanation for the cloud absorption anomaly. *Geophys. Res. Lett.*, 24(5):571–574.
- Crisp, D., Allen, D. A., Grinspoon, D. H., and Pollack, J. B. (1991). The Dark Side of Venus: Near-Infrared Images and Spectra from the Anglo- Australian Observatory. *Science*, 253(5025):1263–1266.
- Csizmadia, S., Pasternacki, T., Dreyer, C., Cabrera, J., Erikson, A., and Rauer, H. (2013). The effect of stellar limb darkening values on the accuracy of the planet radii derived from photometric transit observations. *A&A*, 549:A9.
- Cubillos, P., Blečić, J., Harrington, J., Rojo, P., Lust, N., Bowman, O., Stemm, M., Foster, A., Lored, T. J., Fortney, J., and Madhusudhan, N. (2016). BART: Bayesian Atmospheric Radiative Transfer fitting code.
- Dalba, P. A., Muirhead, P. S., Fortney, J. J., Hedman, M. M., Nicholson, P. D., and Veyette, M. J. (2015). The Transit Transmission Spectrum of a Cold Gas Giant Planet. *ApJ*, 814(2):154.
- Dalcanton, J., Seager, S., Aigrain, S., Battel, S., Brandt, N., Conroy, C., Feinberg, L., Gezari, S., Guyon, O., Harris, W., Hirata, C., Mather, J., Postman, M., Redding, D., Schiminovich, D., Stahl, H. P., and Tumlinson, J. (2015). From Cosmic Birth to Living Earths: The Future of UVOIR Space Astronomy. *ArXiv e-prints*.
- Damiano, M. and Hu, R. (2020). ExoReL R -> : A Bayesian Inverse Retrieval Framework for Exoplanetary Reflected Light Spectra. *AJ*, 159(4):175.
- Dang, L., Cowan, N. B., Schwartz, J. C., Rauscher, E., Zhang, M., Knutson, H. A., Line, M., Dobbs-Dixon, I., Deming, D., Sundararajan, S., et al. (2018). Detection of a westward hotspot offset in the atmosphere of hot gas giant corot-2b. *Nature Astronomy*, 2(3):220–227.
- de Bergh, C., Bezard, B., Crisp, D., Maillard, J. P., Owen, T., Pollack, J., and Grinspoon, D. (1995). Water in the deep atmosphere of Venus from high-resolution spectra of the night side. *Advances in Space Research*, 15(4):79–88.

- de Bergh, C., Moroz, V. I., Taylor, F. W., Crisp, D., Bézard, B., and Zasova, L. V. (2006). The composition of the atmosphere of Venus below 100 km altitude: An overview. *Planet. Space Sci.*, 54(13-14):1389–1397.
- de Wit, J., Wakeford, H. R., Gillon, M., Lewis, N. K., Valenti, J. A., Demory, B.-O., Burgasser, A. J., Burdanov, A., Delrez, L., Jehin, E., Lederer, S. M., Queloz, D., Triaud, A. H. M. J., and Van Grootel, V. (2016). A combined transmission spectrum of the Earth-sized exoplanets TRAPPIST-1 b and c. *Nature*, 537:69–72.
- de Wit, J., Wakeford, H. R., Lewis, N. K., Delrez, L., Gillon, M., Selsis, F., Leconte, J., Demory, B.-O., Bolmont, E., Bourrier, V., Burgasser, A. J., Grimm, S., Jehin, E., Lederer, S. M., Owen, J. E., Stamenković, V., and Triaud, A. H. M. J. (2018). Atmospheric reconnaissance of the habitable-zone Earth-sized planets orbiting TRAPPIST-1. *Nature Astronomy*, 2:214–219.
- Delacroix, C., Savransky, D., Garrett, D., Lowrance, P., and Morgan, R. (2016). *Science yield modeling with the Exoplanet Open-Source Imaging Mission Simulator (EXOSIMS)*, volume 9911 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 991119.
- Delrez, L., Gillon, M., Triaud, A. H. M. J., Demory, B.-O., de Wit, J., Ingalls, J. G., Agol, E., Bolmont, E., Burdanov, A., Burgasser, A. J., Carey, S. J., Jehin, E., Leconte, J., Lederer, S., Queloz, D., Selsis, F., and Van Grootel, V. (2018). Early 2017 observations of TRAPPIST-1 with Spitzer. *MNRAS*, 475:3577–3597.
- Deming, D., Harrington, J., Laughlin, G., Seager, S., Navarro, S. B., Bowman, W. C., and Horning, K. (2007). Spitzer Transit and Secondary Eclipse Photometry of GJ 436b. *The Astrophysical Journal Letters*, 667:L199–L202.
- Deming, D., Wilkins, A., McCullough, P., Burrows, A., Fortney, J. J., Agol, E., Dobbs-Dixon, I., Madhusudhan, N., Crouzet, N., Desert, J.-M., Gilliland, R. L., Haynes, K., Knutson, H. A., Line, M., Magic, Z., Mandell, A. M., Ranjan, S., Charbonneau, D., Clampin, M., Seager, S., and Showman, A. P. (2013). Infrared Transmission Spectroscopy of the Exoplanets HD 209458b

- and XO-1b Using the Wide Field Camera-3 on the Hubble Space Telescope. *The Astrophysical Journal*, 774:95.
- Demory, B.-O., Gillon, M., de Wit, J., Madhusudhan, N., Bolmont, E., Heng, K., Kataria, T., Lewis, N., Hu, R., Krick, J., Stamenković, V., Benneke, B., Kane, S., and Queloz, D. (2016). A map of the large day-night temperature gradient of a super-Earth exoplanet. *Nature*, 532(7598):207–209.
- Des Marais, D. J., Nuth III, J. A., Allamandola, L. J., Boss, A. P., Farmer, J. D., Hoehler, T. M., Jakosky, B. M., Meadows, V. S., Pohorille, A., Runnegar, B., et al. (2008). The nasa astrobiology roadmap. *Astrobiology*, 8(4):715–730.
- Dittmann, J. A., Irwin, J. M., Charbonneau, D., Bonfils, X., Astudillo-Defru, N., Haywood, R. D., Berta-Thompson, Z. K., Newton, E. R., Rodriguez, J. E., Winters, J. G., Tan, T.-G., Almenara, J.-M., Bouchy, F., Delfosse, X., Forveille, T., Lovis, C., Murgas, F., Pepe, F., Santos, N. C., Udry, S., Wünsche, A., Esquerdo, G. A., Latham, D. W., and Dressing, C. D. (2017). A temperate rocky super-Earth transiting a nearby cool star. *Nature*, 544:333–336.
- Domagal-Goldman, S. D., Meadows, V. S., Claire, M. W., and Kasting, J. F. (2011). Using biogenic sulfur gases as remotely detectable biosignatures on anoxic planets. *Astrobiology*, 11(5):419–441.
- Dong, C., Jin, M., Lingam, M., Airapetian, V. S., Ma, Y., and van der Holst, B. (2018). Atmospheric escape from the TRAPPIST-1 planets and implications for habitability. *Proceedings of the National Academy of Science*, 115(2):260–265.
- Doyon, R., Hutchings, J. B., Beaulieu, M., Albert, L., Lafrenière, D., Willott, C., Touahri, D., Rowlands, N., Maszkiewicz, M., Fullerton, A. W., Volk, K., Martel, A. R., Chayer, P., Sivaramakrishnan, A., Abraham, R., Ferrarese, L., Jayawardhana, R., Johnstone, D., Meyer, M., Pipher, J. L., and Sawicki, M. (2012). The JWST Fine Guidance Sensor (FGS) and Near-Infrared Imager

- and Slitless Spectrograph (NIRISS). In *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, volume 8442 of Proc. SPIE, page 84422R.
- Dressing, C. D. and Charbonneau, D. (2013). The Occurrence Rate of Small Planets around Small Stars. *ApJ*, 767:95.
- Dressing, C. D. and Charbonneau, D. (2015). The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity. *ApJ*, 807(1):45.
- Drossart, P., Bézard, B., Encrenaz, T., Lellouch, E., Roos, M., Taylor, F. W., Collard, A. D., Calcutt, S. B., Pollack, J., Grinspoon, D. H., Carlson, R. W., Baines, K. H., and Kamp, L. W. (1993). Search for spatial variations of the H₂O abundance in the lower atmosphere of Venus from NIMS-Galileo. *Planet. Space Sci.*, 41(7):495–504.
- Ehrenreich, D., Bonfils, X., Lovis, C., Delfosse, X., Forveille, T., Mayor, M., Neves, V., Santos, N. C., Udry, S., and Ségransan, D. (2014). Near-infrared transmission spectrum of the warm-Uranus GJ 3470b with the Wide Field Camera-3 on the Hubble Space Telescope. *Astronomy & Astrophysics*, 570:A89.
- Farr, B., Farr, W. M., Cowan, N. B., Haggard, H. M., and Robinson, T. (2018). exocartographer: A Bayesian Framework for Mapping Exoplanets in Reflected Light. *ArXiv e-prints*.
- Faucher, T. J., Turbet, M., Villanueva, G. L., Wolf, E. T., Arney, G., Kopparapu, R. K., Lincowski, A., Mandell, A., de Wit, J., Pidhorodetska, D., et al. (2019). Impact of clouds and hazes on the simulated jwst transmission spectra of habitable zone planets in the trappist-1 system. *The Astrophysical Journal*, 887(2):194.
- Faucher, T. J., Villanueva, G. L., Schwieterman, E. W., Turbet, M., Arney, G., Pidhorodetska, D., Kopparapu, R. K., Mandell, A., and Domagal-Goldman, S. D. (2020). Sensitive probing of exoplanetary oxygen via mid-infrared collisional absorption. *Nature Astronomy*, 4:372–376.

- Feng, Y. K., Line, M. R., Fortney, J. J., Stevenson, K. B., Bean, J., Kreidberg, L., and Parmentier, V. (2016). The Impact of Non-uniform Thermal Structure on the Interpretation of Exoplanet Emission Spectra. *ApJ*, 829(1):52.
- Feng, Y. K. and Robinson, T. D. (2017). Constraining habitability and atmospheric properties through reflected light observations of terrestrial exoplanets.
- Feng, Y. K., Robinson, T. D., Fortney, J. J., Lupu, R. E., Marley, M. S., Lewis, N. K., Macintosh, B., and Line, M. R. (2018). Characterizing Earth Analogs in Reflected Light: Atmospheric Retrieval Studies for Future Space Telescopes. *AJ*, 155:200.
- Feroz, F., Hobson, M., and Bridges, M. (2009). Multinest: an efficient and robust bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398(4):1601–1614.
- Feroz, F., Hobson, M., Cameron, E., and Pettitt, A. (2013). Importance nested sampling and the multinest algorithm. *arXiv preprint arXiv:1306.2144*.
- Feroz, F. and Hobson, M. P. (2008). Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463.
- Ferruit, P., Birkmann, S., Böker, T., Sirianni, M., Giardino, G., de Marchi, G., Alves de Oliveira, C., and Dorner, B. (2014). Observing transiting exoplanets with JWST/NIRSpec. In *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*, volume 9143 of Proc. SPIE, page 91430A.
- Fleming, D. P., Barnes, R., Luger, R., and VanderPlas, J. T. (2020). On the XUV Luminosity Evolution of TRAPPIST-1. *ApJ*, 891(2):155.
- Fleming, D. P. and VanderPlas, J. (2018). approxposterior: Approximate Posterior Distributions in Python. *The Journal of Open Source Software*, 3(29):781.

- Ford, E. B., Seager, S., and Turner, E. L. (2001). Characterization of extrasolar terrestrial planets from diurnal photometric variability. *Nature*, 412:885–887.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). emcee: The mcmc hammer. *PASP*, 125:306–312.
- Foreman-Mackey, D., Hogg, D. W., and Morton, T. D. (2014). Exoplanet Population Inference and the Abundance of Earth Analogs from Noisy, Incomplete Catalogs. *ApJ*, 795:64.
- Fortney, J. J. (2005). The effect of condensates on the characterization of transiting planet atmospheres with transmission spectroscopy. *MNRAS*, 364:649–653.
- Fortney, J. J., Robinson, T. D., Domagal-Goldman, S., Del Genio, A. D., Gordon, I. E., Gharib-Nezhad, E., Lewis, N., Sousa-Silva, C., Airapetian, V., Drouin, B., et al. (2019). The need for laboratory measurements and ab initio studies to aid understanding of exoplanetary atmospheres. *arXiv preprint arXiv:1905.07064*.
- Fraine, J., Deming, D., Benneke, B., Knutson, H., Jordán, A., Espinoza, N., Madhusudhan, N., Wilkins, A., and Todorov, K. (2014). Water vapour absorption in the clear atmosphere of a Neptune-sized exoplanet. *Nature*, 513(7519):526–529.
- Fressin, F., Knutson, H. A., Charbonneau, D., O’Donovan, F. T., Burrows, A., Deming, D., Mandushev, G., and Spiegel, D. (2010). The Broadband Infrared Emission Spectrum of the Exoplanet TrES-3. *The Astrophysical Journal*, 711:374–379.
- Fu, Q. and Liou, K. (1992). On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *Journal of the Atmospheric Sciences*, 49(22):2139–2156.
- Fujii, Y., Angerhausen, D., Deitrick, R., Domagal-Goldman, S., Grenfell, J. L., Hori, Y., Kane, S. R., Pallé, E., Rauer, H., Siegler, N., Stapelfeldt, K., and Stevenson, K. B. (2018). Exoplanet Biosignatures: Observational Prospects. *Astrobiology*, 18(6):739–778.

- Fujii, Y. and Kawahara, H. (2012). Mapping Earth Analogs from Photometric Variability: Spin-Orbit Tomography for Planets in Inclined Orbits. *ApJ*, 755:101.
- Fujii, Y., Kawahara, H., Suto, Y., Fukuda, S., Nakajima, T., Livengood, T. A., and Turner, E. L. (2011). Colors of a Second Earth. II. Effects of Clouds on Photometric Characterization of Earth-like Exoplanets. *ApJ*, 738:184.
- Fujii, Y., Kawahara, H., Suto, Y., Taruya, A., Fukuda, S., Nakajima, T., and Turner, E. L. (2010). Colors of a Second Earth: Estimating the Fractional Areas of Ocean, Land, and Vegetation of Earth-like Exoplanets. *ApJ*, 715:866–880.
- Fujii, Y., Kimura, J., Dohm, J., and Ohtake, M. (2014). Geology and Photometric Variation of Solar System Bodies with Minor Atmospheres: Implications for Solid Exoplanets. *Astrobiology*, 14:753–768.
- Fujii, Y., Lustig-Yaeger, J., and Cowan, N. B. (2017). Rotational Spectral Unmixing of Exoplanets: Degeneracies between Surface Colors and Geography. *AJ*, 154:189.
- Fukabori, M., Nakazawa, T., and Tanaka, M. (1986). Absorption properties of infrared active gases at high pressures—i. CO_2 . *Journal of Quantitative Spectroscopy and Radiative Transfer*, 36(3):265–270.
- Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., Cargile, P. A., Hebb, L., Weiss, L. M., Johnson, J. A., Morton, T. D., Sinukoff, E., Crossfield, I. J. M., and Hirsch, L. A. (2017). The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. *AJ*, 154(3):109.
- Galicher, R., Rameau, J., Bonnefoy, M., Baudino, J.-L., Currie, T., Boccaletti, A., Chauvin, G., Lagrange, A.-M., and Marois, C. (2014). Near-infrared detection and characterization of the exoplanet hd 95086 b with the gemini planet imager. *Astronomy & Astrophysics*, 565:L4.
- Gao, F. and Han, L. (2012). Implementing the nelder-mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, 51(1):259–277.

- García Muñoz, A., Zapatero Osorio, M. R., Barrena, R., Montañés-Rodríguez, P., Martín, E. L., and Pallé, E. (2012). Glancing Views of the Earth: From a Lunar Eclipse to an Exoplanetary Transit. *ApJ*, 755:103.
- Garcia-Sage, K., Gloer, A., Drake, J. J., Gronoff, G., and Cohen, O. (2017). On the Magnetic Protection of the Atmosphere of Proxima Centauri b. *ApJ*, 844:L13.
- Gaskin, J. A., Swartz, D., Vikhlinin, A. A., Özel, F., Gelmis, K. E., Arenberg, J. W., Bandler, S. R., Bautz, M. W., Civitani, M. M., Dominguez, A., et al. (2019). Lynx x-ray observatory: an overview. *Journal of Astronomical Telescopes, Instruments, and Systems*, 5(2):021001.
- Gillon, M., Demory, B.-O., Van Grootel, V., Motalebi, F., Lovis, C., Cameron, A. C., Charbonneau, D., Latham, D., Molinari, E., Pepe, F. A., Ségransan, D., Sasselov, D., Udry, S., Mayor, M., Micela, G., Piotto, G., and Sozzetti, A. (2017a). Two massive rocky planets transiting a K-dwarf 6.5parsecs away. *Nature Astronomy*, 1:0056.
- Gillon, M., Jehin, E., Lederer, S. M., Delrez, L., de Wit, J., Burdanov, A., Van Grootel, V., Burgasser, A. J., Triaud, A. H. M. J., Opitom, C., Demory, B.-O., Sahu, D. K., Bardalez Gagliuffi, D., Magain, P., and Queloz, D. (2016). Temperate Earth-sized planets transiting a nearby ultracool dwarf star. *Nature*, 533:221–224.
- Gillon, M., Meadows, V., Agol, E., Burgasser, A. J., Deming, D., Doyon, R., Fortney, J., Kreidberg, L., Owen, J., Selsis, F., de Wit, J., Lustig-Yaeger, J., and Rackham, B. V. (2020). The TRAPPIST-1 JWST Community Initiative. *arXiv e-prints*, page arXiv:2002.04798.
- Gillon, M., Triaud, A. H. M. J., Demory, B.-O., Jehin, E., Agol, E., Deck, K. M., Lederer, S. M., de Wit, J., Burdanov, A., Ingalls, J. G., Bolmont, E., Leconte, J., Raymond, S. N., Selsis, F., Turbet, M., Barkaoui, K., Burgasser, A., Burleigh, M. R., Carey, S. J., Chaushev, A., Copperwheat, C. M., Delrez, L., Fernandes, C. S., Holdsworth, D. L., Kotze, E. J., Van Grootel, V., Almléay, Y., Benkhaldoun, Z., Magain, P., and Queloz, D. (2017b). Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. *Nature*, 542:456–460.

- Gilmozzi, R. and Spyromilio, J. (2007). The european extremely large telescope (e-elt). *The Messenger*, 127(11):3.
- Glasse, A., Rieke, G. H., Bauwens, E., García-Marín, M., Ressler, M. E., Rost, S., Tikkanen, T. V., Vandenbussche, B., and Wright, G. S. (2015). The Mid-Infrared Instrument for the James Webb Space Telescope, IX: Predicted Sensitivity. *PASP*, 127:686.
- Goodman, J. and Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in applied mathematics and computational science*, 5(1):65–80.
- Goody, R., West, R., Chen, L., and Crisp, D. (1989). The correlated-k method for radiation calculations in nonhomogeneous atmospheres. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 42(6):539–550.
- Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Drouin, B. J., Flaud, J. M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Perevalov, V. I., Perrin, A., Shine, K. P., Smith, M. A. H., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G., Barbe, A., Császár, A. G., Devi, V. M., Furtenbacher, T., Harrison, J. J., Hartmann, J. M., Jolly, A., Johnson, T. J., Karman, T., Kleiner, I., Kyuberis, A. A., Loos, J., Lyulin, O. M., Massie, S. T., Mikhailenko, S. N., Moazzen-Ahmadi, N., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Polyansky, O. L., Rey, M., Rotger, M., Sharpe, S. W., Sung, K., Starikova, E., Tashkun, S. A., Auwera, J. V. e., Wagner, G., Wilzewski, J., Wcisło, P., Yu, S., and Zak, E. J. (2017). The HITRAN2016 molecular spectroscopic database. *J. Quant. Spec. Radiat. Transf.*, 203:3–69.
- Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. (2005). HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622:759–771.
- Gould, A. and Loeb, A. (1992). Discovering Planetary Systems through Gravitational Microlenses. *ApJ*, 396:104.

- Gravity Collaboration, Lacour, S., Nowak, M., Wang, J., Pfuhl, O., Eisenhauer, F., Abuter, R., Amorim, A., Anugu, N., Benisty, M., Berger, J. P., Beust, H., Blind, N., Bonnefoy, M., Bonnet, H., Bourget, P., Brandner, W., Buron, A., Collin, C., Charnay, B., Chapron, F., Clénet, Y., Coudé Du Foresto, V., de Zeeuw, P. T., Deen, C., Dembet, R., Dexter, J., Duvert, G., Eckart, A., Förster Schreiber, N. M., Fédou, P., Garcia, P., Garcia Lopez, R., Gao, F., Gendron, E., Genzel, R., Gillessen, S., Gordo, P., Greenbaum, A., Habibi, M., Haubois, X., Haußmann, F., Henning, T., Hippler, S., Horrobin, M., Hubert, Z., Jimenez Rosales, A., Jocou, L., Kendrew, S., Kervella, P., Kolb, J., Lagrange, A. M., Lapeyrère, V., Le Bouquin, J. B., Léna, P., Lippa, M., Lenzen, R., Maire, A. L., Mollière, P., Ott, T., Paumard, T., Perraut, K., Perrin, G., Pueyo, L., Rabien, S., Ramírez, A., Rau, C., Rodríguez-Coira, G., Rousset, G., Sanchez-Bermudez, J., Scheithauer, S., Schuhler, N., Straub, O., Straubmeier, C., Sturm, E., Tacconi, L. J., Vincent, F., van Dishoeck, E. F., von Fellenberg, S., Wank, I., Waisberg, I., Widmann, F., Wieprecht, E., Wiest, M., Wiezorrek, E., Woillez, J., Yazici, S., Ziegler, D., and Zins, G. (2019). First direct detection of an exoplanet by optical interferometry. *Astrometry and K-band spectroscopy of HR 8799 e*. *A&A*, 623:L11.
- Greene, T., Beichman, C., Eisenstein, D., Horner, S., Kelly, D., Mao, Y., Meyer, M., Rieke, M., and Shi, F. (2007). Observing exoplanets with the JWST NIRCам grisms. In *Techniques and Instrumentation for Detection of Exoplanets III*, volume 6693 of Proc. SPIE, page 66930G.
- Greene, T. P., Kelly, D. M., Stansberry, J., Leisenring, J., Egami, E., Schlawin, E., Chu, L., Hordapp, K. W., and Rieke, M. (2017). $\lambda = 2.4$ to $5 \mu\text{m}$ spectroscopy with the James Webb Space Telescope NIRCам instrument. *Journal of Astronomical Telescopes, Instruments, and Systems*, 3(3):035001.
- Greene, T. P., Line, M. R., Montero, C., Fortney, J. J., Lustig-Yaeger, J., and Luther, K. (2016). Characterizing Transiting Exoplanet Atmospheres with JWST. *The Astrophysical Journal*, 817:17.
- Grenfell, J. L., Griesmeier, J.-M., Patzer, B., Rauer, H., Segura, A., Stadelmann, A., Stracke, B.,

- Titz, R., and Von Paris, P. (2007). Biomarker Response to Galactic Cosmic Ray-Induced NO_x And The Methane Greenhouse Effect in The Atmosphere of An Earth-Like Planet Orbiting An M Dwarf Star. *Astrobiology*, 7:208–221.
- Grillmair, C. J., Burrows, A., Charbonneau, D., Armus, L., Stauffer, J., Meadows, V., van Cleve, J., von Braun, K., and Levine, D. (2008). Strong water absorption in the dayside emission spectrum of the planet HD189733b. *Nature*, 456(7223):767–769.
- Grimm, S. L., Demory, B.-O., Gillon, M., Dorn, C., Agol, E., Burdanov, A., Delrez, L., Sestovic, M., Triaud, A. H. M. J., Turbet, M., Bolmont, É., Caldas, A., Wit, J. d., Jehin, E., Leconte, J., Raymond, S. N., Grootel, V. V., Burgasser, A. J., Carey, S., Fabrycky, D., Heng, K., Hernandez, D. M., Ingalls, J. G., Lederer, S., Selsis, F., and Queloz, D. (2018). The nature of the TRAPPIST-1 exoplanets. *A&A*, 613:A68.
- Guillot, T. (2010). On the radiative equilibrium of irradiated planetary atmospheres. *A&A*, 520:A27.
- Guyon, O. (2005). Limits of Adaptive Optics for High-Contrast Imaging. *ApJ*, 629:592–614.
- HabEx Study Team (2019). HabEx Observatory Final Report. Technical report, NASA.
- Hale, G. M. and Querry, M. R. (1973). Optical constants of water in the 200-nm to 200-micrometer wavelength region. *Appl. Opt.*, 12(3):555.
- Hanke, M., Umann, B., Uecker, J., Arnold, F., and Bunz, H. (2003). Atmospheric measurements of gas-phase HNO_3 and SO_2 using chemical ionization mass spectrometry during the MINATROC field campaign 2000 on Monte Cimone. *Atmospheric Chemistry & Physics*, 3(2):417–436.
- Hansen, J. E. and Arking, A. (1971). Clouds of Venus: Evidence for Their Nature. *Science*, 171(3972):669–672.
- Hansen, J. E. and Hovenier, J. W. (1971). The doubling method applied to multiple scattering of polarized light. *J. Quant. Spec. Radiat. Transf.*, 11(6):809–812.

- Harrington, J., Hansen, B. M., Luszcz, S. H., Seager, S., Deming, D., Menou, K., Cho, J. Y. K., and Richardson, L. J. (2006). The Phase-Dependent Infrared Brightness of the Extrasolar Planet *v* Andromedae b. *Science*, 314(5799):623–626.
- Hedelt, P., von Paris, P., Godolt, M., Gebauer, S., Grenfell, J. L., Rauer, H., Schreier, F., Selsis, F., and Trautmann, T. (2013). Spectral features of Earth-like planets and their detectability at different orbital distances around F, G, and K-type stars. *Astronomy & Astrophysics*, 553:A9.
- Hirono, M. and Nakazawa, T. (1982). The shape of spectral lines with combined impact and statistical broadenings. *Journal of the Physical Society of Japan*, 51(1):265–268.
- Hitchcock, D. R. and Lovelock, J. E. (1967). Life detection by atmospheric analysis. *Icarus*, 7(1-3):149–159.
- Howard, A. W., Marcy, G. W., Johnson, J. A., Fischer, D. A., Wright, J. T., Isaacson, H., Valenti, J. A., Anderson, J., Lin, D. N. C., and Ida, S. (2010). The Occurrence and Mass Distribution of Close-in Super-Earths, Neptunes, and Jupiters. *Science*, 330(6004):653.
- Howe, A. R., Burrows, A., and Deming, D. (2017). An Information-theoretic Approach to Optimize JWST Observations and Retrievals of Transiting Exoplanet Atmospheres. *ApJ*, 835:96.
- Hu, R., Ehlmann, B. L., and Seager, S. (2012). Theoretical Spectra of Terrestrial Exoplanet Surfaces. *ApJ*, 752:7.
- Hughes, R., Bernath, P., and Boone, C. (2014). ACE infrared spectral atlases of the Earth’s atmosphere. *J. Quant. Spec. Radiat. Transf.*, 148:18–21.
- Humlíček, J. (1982). Optimized computation of the voigt and complex probability functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 27(4):437–444.
- Hunten, D. M., Colin, L., Donahue, T. M., and Moroz, V. I. (1983). *Venus*.

- Husser, T. O., Wende-von Berg, S., Dreizler, S., Homeier, D., Reiners, A., Barman, T., and Hauschildt, P. H. (2013). A new extensive library of PHOENIX stellar atmospheres and synthetic spectra. *A&A*, 553:A6.
- Irwin, P. G. J., Parmentier, V., Taylor, J., Barstow, J., Aigrain, S., Lee, G. K. H., and Garland, R. (2020). 2.5D retrieval of atmospheric properties from exoplanet phase curves: application to WASP-43b observations. *MNRAS*, 493(1):106–125.
- Irwin, P. G. J., Sromovsky, L. A., Strong, E. K., Sihra, K., Teanby, N. A., Bowles, N., Calcutt, S. B., and Remedios, J. J. (2006). Improved near-infrared methane band models and k-distribution parameters from 2000 to 9500 cm^{-1} and implications for interpretation of outer planet spectra. *Icarus*, 181(1):309–319.
- Irwin, P. G. J., Teanby, N. A., de Kok, R., Fletcher, L. N., Howett, C. J. A., Tsang, C. C. C., Wilson, C. F., Calcutt, S. B., Nixon, C. A., and Parrish, P. D. (2008). The NEMESIS planetary atmosphere radiative transfer and retrieval tool. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 109:1136–1150.
- Iyer, A. R. and Line, M. R. (2020). The Influence of Stellar Contamination on the Interpretation of Near-infrared Transmission Spectra of Sub-Neptune Worlds around M-dwarfs. *ApJ*, 889(2):78.
- Johns, M. (2006). The giant magellan telescope (gmt). In *Ground-based and Airborne Telescopes*, volume 6267, page 626729. International Society for Optics and Photonics.
- Johns, M., McCarthy, P., Raybould, K., Bouchez, A., Farahani, A., Filgueira, J., Jacoby, G., Sheiman, S., and Sheehan, M. (2012). Giant magellan telescope: overview. In *Ground-based and Airborne Telescopes IV*, volume 8444, page 84441H. International Society for Optics and Photonics.
- Kalirai, J. (2018). Scientific discovery with the James Webb Space Telescope. *Contemporary Physics*, 59:251–290.

- Kandasamy, K., Schneider, J., and Póczos, B. (2017). Query efficient posterior estimation in scientific experiments via bayesian active learning. *Artificial Intelligence*, 243:45–56.
- Kane, S., Arney, G., Crisp, D., Domagal-Goldman, S., Goldblatt, C., Grinspoon, D., Head, J. W., Lenardic, A., Meadows, V., Unterborn, C., and Way, M. J. (2019). Venus as a Nearby Exoplanetary Laboratory. In BAAS, volume 51, page 353.
- Kane, S. R., Arney, G., Crisp, D., Domagal-Goldman, S., Glaze, L. S., Goldblatt, C., Lenardic, A., Unterborn, C., and Way, M. J. (2018). Venus: The Making of an Uninhabitable World. *arXiv e-prints*, page arXiv:1801.03146.
- Kane, S. R., Kopparapu, R. K., and Domagal-Goldman, S. D. (2014). On the Frequency of Potential Venus Analogs from Kepler Data. *ApJ*, 794(1):L5.
- Kasting, J. F., Liu, S. C., and Donahue, T. M. (1979). Oxygen levels in the prebiological atmosphere. *J. Geophys. Res.*, 84(C6):3097–3107.
- Kasting, J. F., Whitmire, D. P., and Reynolds, R. T. (1993). Habitable Zones around Main Sequence Stars. *Icarus*, 101(1):108–128.
- Kawahara, H. and Fujii, Y. (2010). Global Mapping of Earth-like Exoplanets From Scattered Light Curves. *ApJ*, 720:1333–1350.
- Kawahara, H. and Fujii, Y. (2011). Mapping Clouds and Terrain of Earth-like Planets from Photometric Variability: Demonstration with Planets in Face-on Orbits. *ApJ*, 739:L62.
- Kempton, E. M.-R., Lupu, R., Owusu-Asare, A., Slough, P., and Cale, B. (2017). Exo-transmit: An open-source code for calculating transmission spectra for exoplanet atmospheres of varied composition. *Publications of the Astronomical Society of the Pacific*, 129(974):044402.
- Kendrew, S., Scheithauer, S., Bouchet, P., Amiaux, J., Azzollini, R., Bouwman, J., Chen, C. H., Dubreuil, D., Fischer, S., Glasse, A., Greene, T. P., Lagage, P.-O., Lahuis, F., Ronayette, S.,

- Wright, D., and Wright, G. S. (2015). The Mid-Infrared Instrument for the James Webb Space Telescope, IV: The Low-Resolution Spectrometer. *PASP*, 127:623.
- Khare, B. N., Sagan, C., Arakawa, E. T., Suits, F., Callcott, T. A., and Williams, M. W. (1984). Optical constants of organic tholins produced in a simulated Titanian atmosphere - From soft X-ray to microwave frequencies. *Icarus*, 60:127–137.
- Kirkland, L. E., Herr, K. C., and Adams, P. M. (2003). Infrared stealthy surfaces: Why TES and THEMIS may miss some substantial mineral deposits on Mars and implications for remote sensing of planetary surfaces. *Journal of Geophysical Research (Planets)*, 108:5137.
- Kitzmann, D., Heng, K., Oreshenko, M., Grimm, S. L., Apai, D., Bowler, B. P., Burgasser, A. J., and Marley, M. S. (2019). Helios-r.2 – A new Bayesian, open-source retrieval model for brown dwarfs and exoplanet atmospheres.
- Knutson, H. A., Benneke, B., Deming, D., and Homeier, D. (2014). A featureless transmission spectrum for the Neptune-mass exoplanet GJ436b. *Nature*, 505:66–68.
- Knutson, H. A., Charbonneau, D., Allen, L. E., Fortney, J. J., Agol, E., Cowan, N. B., Showman, A. P., Cooper, C. S., and Megeath, S. T. (2007). A map of the day-night contrast of the extrasolar planet HD 189733b. *Nature*, 447(7141):183–186.
- Koll, D. D. B. and Abbot, D. S. (2016). Temperature Structure and Atmospheric Circulation of Dry Tidally Locked Rocky Exoplanets. *ApJ*, 825(2):99.
- Koll, D. D. B., Malik, M., Mansfield, M., Kempton, E. M. R., Kite, E., Abbot, D., and Bean, J. L. (2019). Identifying Candidate Atmospheres on Rocky M dwarf Planets via Eclipse Photometry. *arXiv e-prints*, page arXiv:1907.13138.
- Kolokolova, L., Borovoi, A., Mishchenko, M., and A’Hearn, M. (2010). Glints Observed by the EPOXI Mission and Their Value for the Remote Sensing of Exoplanet Surfaces and Atmospheres. In *AAS/Division for Planetary Sciences Meeting Abstracts #42*, volume 42 of *Bulletin of the American Astronomical Society*, page 1061.

- Komacek, T. D., Fauchez, T. J., Wolf, E. T., and Abbot, D. S. (2020). Clouds will Likely Prevent the Detection of Water Vapor in JWST Transmission Spectra of Terrestrial Exoplanets. *ApJ*, 888(2):L20.
- Kopparapu, R. K. (2013). A Revised Estimate of the Occurrence Rate of Terrestrial Planets in the Habitable Zones around Kepler M-dwarfs. *ApJ*, 767:L8.
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., Eymet, V., Robinson, T. D., Mahadevan, S., Terrien, R. C., Domagal-Goldman, S., Meadows, V., and Deshpande, R. (2013). Habitable Zones around Main-sequence Stars: New Estimates. *ApJ*, 765:131.
- Kreidberg, L., Bean, J. L., Désert, J.-M., Benneke, B., Deming, D., Stevenson, K. B., Seager, S., Berta-Thompson, Z., Seifahrt, A., and Homeier, D. (2014). Clouds in the atmosphere of the super-Earth exoplanet GJ1214b. *Nature*, 505:69–72.
- Kreidberg, L., Koll, D. D. B., Morley, C., Hu, R., Schaefer, L., Deming, D., Stevenson, K. B., Dittmann, J., Vanderburg, A., Berardo, D., Guo, X., Stassun, K., Crossfield, I., Charbonneau, D., Latham, D. W., Loeb, A., Ricker, G., Seager, S., and Vanderpek, R. (2019). Absence of a thick atmosphere on the terrestrial exoplanet LHS 3844b. *arXiv e-prints*, page arXiv:1908.06834.
- Kreidberg, L., Line, M. R., Bean, J. L., Stevenson, K. B., Désert, J.-M., Madhusudhan, N., Fortney, J. J., Barstow, J. K., Henry, G. W., Williamson, M. H., and Showman, A. P. (2015). A Detection of Water in the Transmission Spectrum of the Hot Jupiter WASP-12b and Implications for Its Atmospheric Composition. *The Astrophysical Journal*, 814:66.
- Kreidberg, L., Line, M. R., Thorngren, D., Morley, C. V., and Stevenson, K. B. (2018). Water, High-altitude Condensates, and Possible Methane Depletion in the Atmosphere of the Warm Super-Neptune WASP-107b. *ApJ*, 858(1):L6.
- Kreidberg, L. and Loeb, A. (2016). Prospects for Characterizing the Atmosphere of Proxima Centauri b. *ApJ*, 832:L12.

- Krissansen-Totton, J., Bergsman, D. S., and Catling, D. C. (2016). On Detecting Biospheres from Chemical Thermodynamic Disequilibrium in Planetary Atmospheres. *Astrobiology*, 16(1):39–67.
- Krissansen-Totton, J., Garland, R., Irwin, P., and Catling, D. C. (2018a). Detectability of Biosignatures in Anoxic Atmospheres with the James Webb Space Telescope: A TRAPPIST-1e Case Study. *AJ*, 156:114.
- Krissansen-Totton, J., Olson, S., and Catling, D. C. (2018b). Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life. *Science Advances*, 4(1):eaao5747.
- Krissansen-Totton, J., Schwieterman, E. W., Charnay, B., Arney, G., Robinson, T. D., Meadows, V., and Catling, D. C. (2016). Is the Pale Blue Dot Unique? Optimized Photometric Bands for Identifying Earth-like Exoplanets. *ApJ*, 817:31.
- Lacis, A. A. and Oinas, V. (1991). A description of the correlated-k distribution method for modelling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. *J. Geophys. Res.*, 96:9027–9064.
- Lavie, B., Mendonça, J. M., Mordasini, C., Malik, M., Bonnefoy, M., Demory, B.-O., Oreshenko, M., Grimm, S. L., Ehrenreich, D., and Heng, K. (2016). HELIOS-Retrieval: An Open-source, Nested Sampling Atmospheric Retrieval Code, Application to the HR 8799 Exoplanets and Inferred Constraints for Planet Formation. *ArXiv e-prints*.
- Lavie, B., Mendonça, J. M., Mordasini, C., Malik, M., Bonnefoy, M., Demory, B.-O., Oreshenko, M., Grimm, S. L., Ehrenreich, D., and Heng, K. (2017). HELIOS-RETRIEVAL: An Open-source, Nested Sampling Atmospheric Retrieval Code; Application to the HR 8799 Exoplanets and Inferred Constraints for Planet Formation. *AJ*, 154(3):91.
- Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A., and Sing, D. (2008). Rayleigh scattering in the transit spectrum of HD 189733b. *Astronomy & Astrophysics*, 481:L83–L86.

- Léger, A., Defrère, D., Muñoz, A. G., Godolt, M., Grenfell, J., Rauer, H., and Tian, F. (2019). Searching for atmospheric bioindicators in planets around the two nearby stars, proxima centauri and epsilon eridani—test cases for retrieval of atmospheric gases with infrared spectroscopy. *Astrobiology*, 19(6):797–810.
- Lehmer, O. R. and Catling, D. C. (2017). Rocky Worlds Limited to ~ 1.8 Earth Radii by Atmospheric Escape during a Star’s Extreme UV Saturation. *ApJ*, 845(2):130.
- Liebert, J. and Gizis, J. E. (2006). RI Photometry of 2MASS-selected Late M and L Dwarfs. *PASP*, 118:659–670.
- Lincowski, A. P., Lustig-Yaeger, J., and Meadows, V. S. (2019). Observing Isotopologue Bands in Terrestrial Exoplanet Atmospheres with the James Webb Space Telescope: Implications for Identifying Past Atmospheric and Ocean Loss. *The Astronomical Journal*, 158(1):26.
- Lincowski, A. P., Meadows, V. S., Crisp, D., Robinson, T. D., Luger, R., Lustig-Yaeger, J., and Arney, G. N. (2018). Evolved Climates and Observational Discriminants for the TRAPPIST-1 Planetary System. *ApJ*, 867:76.
- Line, M. R., Knutson, H., Wolf, A. S., and Yung, Y. L. (2014). A Systematic Retrieval Analysis of Secondary Eclipse Spectra. II. A Uniform Analysis of Nine Planets and their C to O Ratios. *The Astrophysical Journal*, 783:70.
- Line, M. R. and Parmentier, V. (2016). The Influence of Nonuniform Cloud Cover on Transit Transmission Spectra. *ApJ*, 820(1):78.
- Line, M. R., Wolf, A. S., Zhang, X., Knutson, H., Kammer, J. A., Ellison, E., Deroo, P., Crisp, D., and Yung, Y. L. (2013). A Systematic Retrieval Analysis of Secondary Eclipse Spectra. I. A Comparison of Atmospheric Retrieval Techniques. *The Astrophysical Journal*, 775:137.
- Lissauer, J. J., Dawson, R. I., and Tremaine, S. (2014). Advances in exoplanet science from kepler. *Nature*, 513(7518):336–344.

- Loftus, K., Wordsworth, R. D., and Morley, C. V. (2019). Sulfate Aerosol Hazes and SO₂ Gas as Constraints on Rocky Exoplanets' Surface Liquid Water. *arXiv e-prints*, page arXiv:1908.02769.
- Luger, R. and Barnes, R. (2015). Extreme Water Loss and Abiotic O₂ Buildup on Planets Throughout the Habitable Zones of M Dwarfs. *Astrobiology*, 15:119–143.
- Luger, R., Barnes, R., Lopez, E., Fortney, J., Jackson, B., and Meadows, V. (2015). Habitable Evaporated Cores: Transforming Mini-Neptunes into Super-Earths in the Habitable Zones of M Dwarfs. *Astrobiology*, 15:57–88.
- Luger, R., Bedell, M., Vanderspek, R., and Burke, C. J. (2019). TESS Photometric Mapping of a Terrestrial Planet in the Habitable Zone: Detection of Clouds, Oceans, and Continents. *arXiv e-prints*, page arXiv:1903.12182.
- Luger, R., Lustig-Yaeger, J., and Agol, E. (2017a). Planet-Planet Occultations in TRAPPIST-1 and Other Exoplanet Systems. *ApJ*, 851:94.
- Luger, R., Lustig-Yaeger, J., Fleming, D. P., Tilley, M. A., Agol, E., Meadows, V. S., Deitrick, R., and Barnes, R. (2017b). The Pale Green Dot: A Method to Characterize Proxima Centauri b Using Exo-Aurorae. *ApJ*, 837:63.
- Luger, R., Sestovic, M., Kruse, E., Grimm, S. L., Demory, B.-O., Agol, E., Bolmont, E., Fabrycky, D., Fernandes, C. S., Van Grootel, V., Burgasser, A., Gillon, M., Ingalls, J. G., Jehin, E., Raymond, S. N., Selsis, F., Triaud, A. H. M. J., Barclay, T., Barentsen, G., Howell, S. B., Delrez, L., de Wit, J., Foreman-Mackey, D., Holdsworth, D. L., Leconte, J., Lederer, S., Turbet, M., Almléay, Y., Benkhaldoun, Z., Magain, P., Morris, B. M., Heng, K., and Queloz, D. (2017c). A seven-planet resonant chain in TRAPPIST-1. *Nature Astronomy*, 1:0129.
- Lustig-Yaeger, J., Meadows, V. S., and Lincowski, A. P. (2019a). A Mirage of the Cosmic Shoreline: Venus-like Clouds as a Statistical False Positive for Exoplanet Atmospheric Erosion. *ApJ*, 887(1):L11.

- Lustig-Yaeger, J., Meadows, V. S., and Lincowski, A. P. (2019b). The Detectability and Characterization of the TRAPPIST-1 Exoplanet Atmospheres with JWST. *AJ*, 158(1):27.
- Lustig-Yaeger, J., Meadows, V. S., Tovar Mendoza, G., Schwieterman, E. W., Fujii, Y., Luger, R., and Robinson, T. D. (2018). Detecting Ocean Glint on Exoplanets Using Multiphase Mapping. *AJ*, 156:301.
- Lustig-Yaeger, J., Robinson, T., and Arney, G. (2019). “coronagraph”: Telescope noise modeling for exoplanets in python. *Journal of Open Source Software*, 4(40):1387.
- LUVOIR Mission Concept Study Team (2019). The LUVOIR Final Report. Technical report, NASA.
- Macdonald, E. J. R. and Cowan, N. B. (2019). An empirical infrared transit spectrum of Earth: opacity windows and biosignatures. *MNRAS*, 489(1):196–204.
- MacDonald, R. J., Goyal, J. M., and Lewis, N. K. (2020). Why is it so Cold in Here?: Explaining the Cold Temperatures Retrieved from Transmission Spectra of Exoplanet Atmospheres. *arXiv e-prints*, page arXiv:2003.11548.
- MacDonald, R. J. and Madhusudhan, N. (2017). HD 209458b in new light: evidence of nitrogen chemistry, patchy clouds and sub-solar water. *MNRAS*, 469(2):1979–1996.
- Macintosh, B., Graham, J., Barman, T., De Rosa, R., Konopacky, Q., Marley, M., Marois, C., Nielsen, E., Pueyo, L., Rajan, A., et al. (2015). Discovery and spectroscopy of the young jovian planet 51 eri b with the gemini planet imager. *Science*, 350(6256):64–67.
- Macintosh, B., Graham, J. R., Ingraham, P., Konopacky, Q., Marois, C., Perrin, M., Poyneer, L., Bauman, B., Barman, T., Burrows, A. S., et al. (2014). First light of the gemini planet imager. *proceedings of the National Academy of Sciences*, 111(35):12661–12666.
- Macintosh, B. A., Graham, J. R., Palmer, D. W., Doyon, R., Dunn, J., Gavel, D. T., Larkin, J., Oppenheimer, B., Saddlemyer, L., Sivaramakrishnan, A., et al. (2008). The gemini planet imager:

- from science to design to construction. In *Adaptive Optics Systems*, volume 7015, page 701518. International Society for Optics and Photonics.
- Madhusudhan, N. and Seager, S. (2009). A temperature and abundance retrieval method for exoplanet atmospheres. *The Astrophysical Journal*, 707(1):24.
- Mai, C. and Line, M. R. (2019). Exploring Exoplanet Cloud Assumptions in JWST Transmission Spectra. *ApJ*, 883(2):144.
- Males, J. R. and Guyon, O. (2018). Ground-based adaptive optics coronagraphic performance under closed-loop predictive control. *Journal of Astronomical Telescopes, Instruments, and Systems*, 4(1):019001.
- Mallama, A., Wang, D., and Howard, R. A. (2006). Venus phase function and forward scattering from H_2SO_4 . *Icarus*, 182:10–22.
- Maltagliati, L., Bézard, B., Vinatier, S., Hedman, M. M., Lellouch, E., Nicholson, P. D., Sotin, C., de Kok, R. J., and Sicardy, B. (2015). Titan’s atmosphere as observed by Cassini/VIMS solar occultations: CH_4 , CO and evidence for C_2H_6 absorption. *Icarus*, 248:1–24.
- Mandel, K. and Agol, E. (2002). Analytic Light Curves for Planetary Transit Searches. *ApJ*, 580(2):L171–L175.
- Mandell, A. M., Haynes, K., Sinukoff, E., Madhusudhan, N., Burrows, A., and Deming, D. (2013). Exoplanet Transit Spectroscopy Using WFC3: WASP-12 b, WASP-17 b, and WASP-19 b. *The Astrophysical Journal*, 779:128.
- Mansfield, M., Kite, E. S., Hu, R., Koll, D. D. B., Malik, M., Bean, J. L., and Kempton, E. M. R. (2019). Identifying Atmospheres on Rocky Exoplanets Through Inferred High Albedo. *arXiv e-prints*, page arXiv:1907.13150.
- Marais, D. D. and Walter, M. (1999). Astrobiology: exploring the origins, evolution, and distribution of life in the universe. *Annual review of ecology and systematics*, 30(1):397–420.

- Marov, M. Y., Avduevsky, V. S., Borodin, N. F., Ekonomov, A. P., Kerzhanovich, V. V., Lysov, V. P., Moshkin, B. Y., Rozhdestvensky, M. K., and Ryabov, O. L. (1973). Preliminary Results on the Venus Atmosphere from the Venera 8 Descent Module. *Icarus*, 20(4):407–421.
- Marshak, A., Várnai, T., and Kostinski, A. (2017). Terrestrial glint seen from deep space: Oriented ice crystals detected from the Lagrangian point. *Geophys. Res. Lett.*, 44:5197–5202.
- Maurin, A. S., Selsis, F., Hersant, F., and Belu, A. (2012). Thermal phase curves of nontransiting terrestrial exoplanets. II. Characterizing airless planets. *A&A*, 538:A95.
- Mayer, C. H., McCullough, T. P., and Sloanaker, R. M. (1958). Observations of Venus at 3.15-CM Wave Length. *ApJ*, 127:1.
- Mayor, M., Pepe, F., Queloz, D., Bouchy, F., Rupprecht, G., Lo Curto, G., Avila, G., Benz, W., Bertaux, J. L., Bonfils, X., Dall, T., Dekker, H., Delabre, B., Eckert, W., Fleury, M., Gilliotte, A., Gojak, D., Guzman, J. C., Kohler, D., Lizon, J. L., Longinotti, A., Lovis, C., Megevand, D., Pasquini, L., Reyes, J., Sivan, J. P., Sosnowska, D., Soto, R., Udry, S., van Kesteren, A., Weber, L., and Weilenmann, U. (2003). Setting New Standards with HARPS. *The Messenger*, 114:20–24.
- Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature*, 378(6555):355–359.
- McKee, R. (2016). *Story: Die Prinzipien des Drehbuchschreibens*. Alexander Verlag Berlin.
- Meadows, V., Lincowski, A., and Lustig-Yaeger, J. (2020). Prospects for Biosignature Detection with JWST. In *American Astronomical Society Meeting Abstracts*, American Astronomical Society Meeting Abstracts, page 126.03.
- Meadows, V. S. (2017). Reflections on o₂ as a biosignature in exoplanetary atmospheres. *Astrobiology*, 17(10):1022–1052.

- Meadows, V. S., Arney, G. N., Schwieterman, E. W., Lustig-Yaeger, J., Lincowski, A. P., Robinson, T., Domagal-Goldman, S. D., Deitrick, R., Barnes, R. K., Fleming, D. P., Luger, R., Driscoll, P. E., Quinn, T. R., and Crisp, D. (2018a). The Habitability of Proxima Centauri b: Environmental States and Observational Discriminants. *Astrobiology*, 18:133–189.
- Meadows, V. S. and Barnes, R. K. (2018). *Factors Affecting Exoplanet Habitability*, page 57.
- Meadows, V. S. and Crisp, D. (1996). Ground-based near-infrared observations of the Venus night-side: The thermal structure and water abundance near the surface. *J. Geophys. Res.*, 101:4595–4622.
- Meadows, V. S., Reinhard, C. T., Arney, G. N., Parenteau, M. N., Schwieterman, E. W., Domagal-Goldman, S. D., Lincowski, A. P., Stapelfeldt, K. R., Rauer, H., DasSarma, S., Hegde, S., Narita, N., Deitrick, R., Lustig-Yaeger, J., Lyons, T. W., Siegler, N., and Grenfell, J. L. (2018b). Exoplanet Biosignatures: Understanding Oxygen as a Biosignature in the Context of Its Environment. *Astrobiology*, 18(6):630–662.
- Meeker, S. R., Mazin, B. A., Walter, A. B., Strader, P., Fruitwala, N., Bockstiegel, C., Szypryt, P., Ulbricht, G., Coiffard, G., Bumble, B., Collura, G., Dodkins, R., Lipartito, I., Zobrist, N., Bottom, M., Shelton, J. C., Mawet, D., van Eyken, J. C., Vasisht, G., and Serabyn, E. (2018). DARKNESS: A Microwave Kinetic Inductance Detector Integral Field Spectrograph for High-Contrast Astronomy. *ArXiv e-prints*.
- Mennesson, B., Gaudi, S., Seager, S., Cahoy, K., Domagal-Goldman, S., Feinberg, L., Guyon, O., Kasdin, J., Marois, C., Mawet, D., Tamura, M., Mouillet, D., Prusti, T., Quirrenbach, A., Robinson, T., Rogers, L., Scowen, P., Somerville, R., Stapelfeldt, K., Stern, D., Still, M., Turnbull, M., Booth, J., Kiessling, A., Kuan, G., and Warfield, K. (2016). The Habitable Exoplanet (HabEx) Imaging Mission: preliminary science drivers and technical requirements. In *Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*, volume 9904 of Proc. SPIE, page 99040L.

- Miller-Ricci, E., Seager, S., and Sasselov, D. (2009). The Atmospheric Signatures of Super-Earths: How to Distinguish Between Hydrogen-Rich and Hydrogen-Poor Atmospheres. *The Astrophysical Journal*, 690:1056–1067.
- Misra, A., Meadows, V., and Crisp, D. (2014). The Effects of Refraction on Transit Transmission Spectroscopy: Application to Earth-like Exoplanets. *The Astrophysical Journal*, 792:61.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the long-wave. *Journal of Geophysical Research: Atmospheres*, 102(D14):16663–16682.
- Mollière, P., Wardenier, J. P., van Boekel, R., Henning, T., Molaverdikhani, K., and Snellen, I. A. G. (2019). petitRADTRANS. A Python radiative transfer package for exoplanet characterization and retrieval. *A&A*, 627:A67.
- Moran, S. E., Hörst, S. M., Batalha, N. E., Lewis, N. K., and Wakeford, H. R. (2018). Limits on Clouds and Hazes for the TRAPPIST-1 Planets. *ArXiv e-prints*.
- Moreau, G., Boisssoles, J., Le Doucen, R., Boulet, C., Tipping, R. H., and Ma, Q. (2001). Experimental and theoretical study of the collision-induced fundamental absorption spectra of N₂-O₂ and O₂-N₂ pairs. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 69(2):245–256.
- Morley, C. V., Fortney, J. J., Kempton, E. M.-R., Marley, M. S., Visscher, C., and Zahnle, K. (2013). Quantitatively Assessing the Role of Clouds in the Transmission Spectrum of GJ 1214b. *The Astrophysical Journal*, 775:33.
- Morley, C. V., Fortney, J. J., Marley, M. S., Visscher, C., Saumon, D., and Leggett, S. K. (2012). Neglected Clouds in T and Y Dwarf Atmospheres. *ApJ*, 756(2):172.
- Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., and Fortney, J. J. (2017). Observing the Atmospheres of Known Temperate Earth-sized Planets with JWST. *ApJ*, 850:121.

- Morris, B. M., Agol, E., Davenport, J. R. A., and Hawley, S. L. (2018a). Possible Bright Starspots on TRAPPIST-1. *ApJ*, 857:39.
- Morris, B. M., Agol, E., Hebb, L., and Hawley, S. L. (2018b). Robust Transiting Exoplanet Radii in the Presence of Starspots from Ingress and Egress Durations. *AJ*, 156(3):91.
- Morris, B. M., Agol, E., Hebb, L., Hawley, S. L., Gillon, M., Ducrot, E., Delrez, L., Ingalls, J., and Demory, B.-O. (2018c). Non-detection of Contamination by Stellar Activity in the Spitzer Transit Light Curves of TRAPPIST-1. *ApJ*, 863(2):L32.
- Mulders, G. D., Pascucci, I., and Apai, D. (2015). A Stellar-mass-dependent Drop in Planet Occurrence Rates. *ApJ*, 798:112.
- Nayak, M., Lupu, R., Marley, M. S., Fortney, J. J., Robinson, T., and Lewis, N. (2017). Atmospheric Retrieval for Direct Imaging Spectroscopy of Gas Giants in Reflected Light. II. Orbital Phase and Planetary Radius. *PASP*, 129(3):034401.
- Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The computer journal*, 7(4):308–313.
- Nelson, B., Ford, E. B., and Payne, M. J. (2014). RUN DMC: An Efficient, Parallel Code for Analyzing Radial Velocity Observations Using N-body Integrations and Differential Evolution Markov Chain Monte Carlo. *ApJS*, 210:11.
- Nelson, J. and Sanders, G. H. (2008). The status of the thirty meter telescope project. In *Ground-based and Airborne Telescopes II*, volume 7012, page 70121A. International Society for Optics and Photonics.
- Nicodemus, F. E. (1965). Directional reflectance and emissivity of an opaque surface. *Appl. Opt.*, 4:767.
- Nikolov, N., Sing, D. K., Burrows, A. S., Fortney, J. J., Henry, G. W., Pont, F., Ballester, G. E., Aigrain, S., Wilson, P. A., Huitson, C. M., Gibson, N. P., Désert, J.-M., Lecavelier Des Etangs,

- A., Showman, A. P., Vidal-Madjar, A., Wakeford, H. R., and Zahnle, K. (2015). HST hot-Jupiter transmission spectral survey: haze in the atmosphere of WASP-6b. *Monthly Notices of the Royal Astronomical Society*, 447:463–478.
- Nixon, M. C. and Madhusudhan, N. (2020). Assessment of Supervised Machine Learning for Atmospheric Retrieval of Exoplanets. *arXiv e-prints*, page arXiv:2004.10755.
- Noecker, M. C., Zhao, F., Demers, R., Trauger, J., Guyon, O., and Kasdin, N. J. (2016). Coronagraph instrument for wfirst-afta. *Journal of Astronomical Telescopes, Instruments, and Systems*, 2(1):011001.
- Oakley, P. H. H. and Cash, W. (2009). Construction of an Earth Model: Analysis of Exoplanet Light Curves and Mapping the Next Earth with the New Worlds Observer. *ApJ*, 700:1428–1439.
- Origins Space Telescope Study Team (2019). OST Mission Concept Study Report. Technical Report August, NASA.
- Ostberg, C. and Kane, S. R. (2019). Predicting the Yield of Potential Venus Analogs from TESS and their Potential for Atmospheric Characterization. *arXiv e-prints*, page arXiv:1909.07456.
- Pallé, E., Ford, E. B., Seager, S., Montañés-Rodríguez, P., and Vazquez, M. (2008). Identifying the Rotation Rate and the Presence of Dynamic Weather on Extrasolar Earth-like Planets from Photometric Observations. *ApJ*, 676:1319–1329.
- Palmer, K. F. and Williams, D. (1975). Optical constants of sulfuric acid; application to the clouds of Venus? *Appl. Opt.*, 14(1):208–219.
- Parmentier, V. and Guillot, T. (2014). A non-grey analytical model for irradiated atmospheres. I. Derivation. *A&A*, 562:A133.
- Peacock, S., Barman, T., Shkolnik, E. L., Hauschildt, P. H., and Baron, E. (2019). Predicting

- the Extreme Ultraviolet Radiation Environment of Exoplanets around Low-mass Stars: The TRAPPIST-1 System. *ApJ*, 871:235.
- Penny, M. T., Gaudi, B. S., Kerins, E., Rattenbury, N. J., Mao, S., Robin, A. C., and Novati, S. C. (2019). Predictions of the wfirst microlensing survey. i. bound planet detection rates. *The Astrophysical Journal Supplement Series*, 241(1):3.
- Pepe, F., Mayor, M., Delabre, B., Kohler, D., Lacroix, D., Queloz, D., Udry, S., Benz, W., Bertaux, J.-L., and Sivan, J.-P. (2000). *HARPS: a new high-resolution spectrograph for the search of extrasolar planets*, volume 4008 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 582–592.
- Perrin, M. and Hartmann, J. (1989). Temperature-dependent measurements and modeling of absorption by co₂-n₂ mixtures in the far line-wings of the 4.3 μ m co₂ band. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 42(4):311–317.
- Perryman, M. (2011). *The Exoplanet Handbook*.
- Perryman, M., Hartman, J., Bakos, G. Á., and Lindegren, L. (2014). Astrometric Exoplanet Detection with Gaia. *ApJ*, 797(1):14.
- Pidhorodetska, D., Fauchez, T., Villanueva, G., Domagal-Goldman, S., and Kopparapu, R. K. (2020). Detectability of Molecular Signatures on TRAPPIST-1e through Transmission Spectroscopy Simulated for Future Space-Based Observatories. *arXiv e-prints*, page arXiv:2001.01338.
- Pinhas, A., Madhusudhan, N., Gandhi, S., and MacDonald, R. (2019). H₂O abundances and cloud properties in ten hot giant exoplanets. *MNRAS*, 482(2):1485–1498.
- Planavsky, N. J., Reinhard, C. T., Wang, X., Thomson, D., McGoldrick, P., Rainbird, R. H., Johnson, T., Fischer, W. W., and Lyons, T. W. (2014). Low Mid-Proterozoic atmospheric oxygen levels and the delayed rise of animals. *Science*, 346(6209):635–638.

- Pluriel, W., Zingales, T., Leconte, J., and Parmentier, V. (2020). Strong biases in retrieved atmospheric composition caused by day-night chemical heterogeneities. *A&A*, 636:A66.
- Pollack, J. B., Dalton, J. B., Grinspoon, D., Wattson, R. B., Freedman, R., Crisp, D., Allen, D. A., Bezard, B., DeBergh, C., Giver, L. P., et al. (1993). Near-infrared light from venus' nightside: A spectroscopic analysis. *Icarus*, 103(1):1–42.
- Pollack, J. B., Erickson, E. F., Witteborn, F. C., Chackerian, C., J., Summers, A. L., van Camp, W., Baldwin, B. J., Augason, G. C., and Caroff, L. J. (1974). Aircraft Observations of Venus' Near-Infrared Reflection Spectrum: Implications for Cloud Composition. *Icarus*, 23(1):8–26.
- Pont, F., Sing, D. K., Gibson, N. P., Aigrain, S., Henry, G., and Husnoo, N. (2013). The prevalence of dust on the exoplanet HD 189733b from Hubble and Spitzer observations. *MNRAS*, 432(4):2917–2944.
- Pontoppidan, K. M., Pickering, T. E., Laidler, V. G., Gilbert, K., Sontag, C. D., Slocum, C., Sienkiewicz, M. J., Hanley, C., Earl, N. M., Pueyo, L., Ravindranath, S., Karakla, D. M., Roberto, M., Noriega-Crespo, A., and Barker, E. A. (2016). Pandeia: a multi-mission exposure time calculator for JWST and WFIRST. In *Observatory Operations: Strategies, Processes, and Systems VI*, volume 9910 of Proc. SPIE, page 991016.
- Prusti, T., De Bruijne, J., Brown, A. G., Vallenari, A., Babusiaux, C., Bailer-Jones, C., Bastian, U., Biermann, M., Evans, D., Eyer, L., et al. (2016). The gaia mission. *Astronomy & Astrophysics*, 595:A1.
- Pueyo, L. and Norman, C. (2013). High-contrast Imaging with an Arbitrary Aperture: Active Compensation of Aperture Discontinuities. *ApJ*, 769:102.
- Quanz, S. P., Absil, O., Angerhausen, D., Benz, W., Bonfils, X., Berger, J.-P., Brogi, M., Cabrera, J., Danchi, W. C., Defrère, D., van Dishoeck, E., Ehrenreich, D., Ertel, S., Fortney, J., Gaudi, S., Girard, J., Glauser, A., Grenfell, J. L., Ireland, M., Janson, M., Kammerer, J., Kitzmann, D., Kraus, S., Krause, O., Labadie, L., Lacour, S., Lichtenberg, T., Line, M., Linz, H., Loicq,

- J., Mennesson, B., Meyer, M. R., Miguel, Y., Monnier, J., N'Diaye, M., Pallé, E., Queloz, D., Rauer, H., Ribas, I., Rugheimer, S., Selsis, F., Serabyn, G., Snellen, I., Sozzetti, A., Stapelfeldt, K. R., Triaud, A., Udry, S., and Wyatt, M. (2019). Atmospheric characterization of terrestrial exoplanets in the mid-infrared: biosignatures, habitability & diversity. *arXiv e-prints*, page arXiv:1908.01316.
- Rackham, B., Espinoza, N., Apai, D., López-Morales, M., Jordán, A., Osip, D. J., Lewis, N. K., Rodler, F., Fraine, J. D., Morley, C. V., et al. (2016). Access i: An optical transmission spectrum of gj 1214b reveals a heterogeneous stellar photosphere. *arXiv preprint arXiv:1612.00228*.
- Rackham, B. V., Apai, D., and Giampapa, M. S. (2018). The Transit Light Source Effect: False Spectral Features and Incorrect Densities for M-dwarf Transiting Planets. *ApJ*, 853:122.
- Rao, K. N. (2012). *Spectroscopy of the Earth's atmosphere and interstellar medium*. Academic Press.
- Raymond, S. N., Quinn, T., and Lunine, J. I. (2004). Making other earths: dynamical simulations of terrestrial planet formation and water delivery. *Icarus*, 168(1):1–17.
- Raymond, S. N., Quinn, T., and Lunine, J. I. (2007). High-resolution simulations of the final assembly of earth-like planets. 2. water delivery and planetary habitability. *Astrobiology*, 7(1):66–84.
- Redfield, S., Endl, M., Cochran, W. D., and Koesterke, L. (2008). Sodium Absorption from the Exoplanetary Atmosphere of HD 189733b Detected in the Optical Transmission Spectrum. *The Astrophysical Journal Letters*, 673:L87.
- Reinhard, C. T., Olson, S. L., Schwieterman, E. W., and Lyons, T. W. (2017). False Negatives for Remote Life Detection on Ocean-Bearing Planets: Lessons from the Early Earth. *Astrobiology*, 17(4):287–297.
- Richardson, L. J., Deming, D., and Seager, S. (2003). Infrared Observations during the Secondary Eclipse of HD 209458b. II. Strong Limits on the Infrared Spectrum Near 2.2 μm . *ApJ*, 597(1):581–589.

- Ricker, G. R., Latham, D. W., Vanderspek, R. K., Ennico, K. A., Bakos, G., Brown, T. M., Burgasser, A. J., Charbonneau, D., Clampin, M., Deming, L. D., Doty, J. P., Dunham, E. W., Elliot, J. L., Holman, M. J., Ida, S., Jenkins, J. M., Jernigan, J. G., Kawai, N., Laughlin, G. P., Lissauer, J. J., Martel, F., Sasselov, D. D., Schingler, R. H., Seager, S., Torres, G., Udry, S., Villaseñor, J. N., Winn, J. N., and Worden, S. P. (2010). Transiting Exoplanet Survey Satellite (TESS). In *American Astronomical Society Meeting Abstracts #215*, volume 215 of *American Astronomical Society Meeting Abstracts*, page 450.06.
- Roberge, A., Bolcar, M. R., and France, K. C. (2019). Telling the story of life in the cosmos: the LUVOIR telescope concepts. In Proc. SPIE, volume 11115 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 111150O.
- Robichon, N. (2002). Detection of Transits of Extrasolar Planets with GAIA. In Bienayme, O. and Turon, C., editors, *EAS Publications Series*, volume 2 of *EAS Publications Series*, pages 215–221.
- Robinson, T. D. (2017a). A Theory of Exoplanet Transits with Light Scattering. *ApJ*, 836:236.
- Robinson, T. D. (2017b). Characterizing Exoplanets for Habitability. *ArXiv e-prints*.
- Robinson, T. D. (2018). *Characterizing Exoplanet Habitability*, page 67.
- Robinson, T. D. and Catling, D. C. (2012). An analytic radiative-convective model for planetary atmospheres. *The Astrophysical Journal*, 757(1):104.
- Robinson, T. D. and Crisp, D. (2018). Linearized Flux Evolution (LiFE): A technique for rapidly adapting fluxes from full-physics radiative transfer models. *J. Quant. Spec. Radiat. Transf.*, 211:78–95.
- Robinson, T. D., Ennico, K., Meadows, V. S., Sparks, W., Bussey, D. B. J., Schwieterman, E. W., and Breiner, J. (2014a). Detection of Ocean Glint and Ozone Absorption Using LCROSS Earth Observations. *ApJ*, 787:171.

- Robinson, T. D., Maltagliati, L., Marley, M. S., and Fortney, J. J. (2014b). Titan solar occultation observations reveal transit spectra of a hazy world. *Proceedings of the National Academy of Science*, 111(25):9042–9047.
- Robinson, T. D., Meadows, V. S., and Crisp, D. (2010). Detecting Oceans on Extrasolar Planets Using the Glint Effect. *ApJ*, 721:L67–L71.
- Robinson, T. D., Meadows, V. S., Crisp, D., Deming, D., A’Hearn, M. F., Charbonneau, D., Liven-
good, T. A., Seager, S., Barry, R. K., Hearty, T., Hewagama, T., Lisse, C. M., McFadden, L. A.,
and Wellnitz, D. D. (2011). Earth as an Extrasolar Planet: Earth Model Validation Using EPOXI
Earth Observations. *Astrobiology*, 11:393–408.
- Robinson, T. D., Stapelfeldt, K. R., and Marley, M. S. (2016). Characterizing Rocky and Gaseous
Exoplanets with 2 m Class Space-based Coronagraphs. *PASP*, 128(2):025003.
- Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P. O., and Tinetti, G. (2016). Exploring
Biases of Atmospheric Retrievals in Simulated JWST Transmission Spectra of Hot Jupiters.
ApJ, 833(1):120.
- Rodgers, C. D. (1976). Retrieval of Atmospheric Temperature and Composition From Remote
Measurements of Thermal Radiation. *Reviews of Geophysics and Space Physics*, 14:609.
- Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice*.
- Roettenbacher, R. M. and Kane, S. R. (2017). The Stellar Activity of TRAPPIST-1 and Conse-
quences for the Planetary Atmospheres. *ApJ*, 851:77.
- Rogers, L. A. (2015). Most 1.6 Earth-radius Planets are Not Rocky. *ApJ*, 801(1):41.
- Rothman, L. S., Gordon, I. E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P. F., Birk, M.,
Bizzocchi, L., Boudon, V., Brown, L. R., Campargue, A., Chance, K., Cohen, E. A., Coud-
ert, L. H., Devi, V. M., Drouin, B. J., Fayt, A., Flaud, J. M., Gamache, R. R., Harrison, J. J.,
Hartmann, J. M., Hill, C., Hodges, J. T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy,

- R. J., Li, G., Long, D. A., Lyulin, O. M., Mackie, C. J., Massie, S. T., Mikhailenko, S., Müller, H. S. P., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E. R., Richard, C., Smith, M. A. H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G. C., Tyuterev, V. G., and Wagner, G. (2013). The HITRAN2012 molecular spectroscopic database. *J. Quant. Spec. Radiat. Transf.*, 130:4–50.
- Rothman, L. S., Gordon, I. E., Barber, R. J., Dothe, H., Gamache, R. R., Goldman, A., Perevalov, V. I., Tashkun, S. A., and Tennyson, J. (2010). HITEMP, the high-temperature molecular spectroscopic database. *J. Quant. Spec. Radiat. Transf.*, 111:2139–2150.
- Rugheimer, S., Kaltenegger, L., Segura, A., Linsky, J., and Mohanty, S. (2015). Effect of UV Radiation on the Spectral Fingerprints of Earth-like Planets Orbiting M Stars. *ApJ*, 809:57.
- Rugheimer, S., Kaltenegger, L., Zsom, A., Segura, A., and Sasselov, D. (2013). Spectral fingerprints of earth-like planets around fgk stars. *Astrobiology*, 13(3):251–269.
- Sagan, C. (1962). Structure of the lower atmosphere of Venus. *Icarus*, 1(1-6):151–169.
- Sagan, C. (1997). *Pale blue dot: A vision of the human future in space*. Random House Digital, Inc.
- Sagan, C., Thompson, W. R., Carlson, R., Gurnett, D., and Hord, C. (1993). A search for life on Earth from the Galileo spacecraft. *Nature*, 365:715–721.
- Schaefer, L. and Fegley Jr, B. (2010). Chemistry of atmospheres formed during accretion of the earth and other terrestrial planets. *Icarus*, 208(1):438–448.
- Schaefer, L., Wordsworth, R. D., Berta-Thompson, Z., and Sasselov, D. (2016). Predictions of the Atmospheric Composition of GJ 1132b. *ApJ*, 829:63.
- Schlawin, E., Rieke, M., Leisenring, J., Walker, L. M., Fraine, J., Kelly, D., Misselt, K., Greene, T., Line, M., Lewis, N., and Stansberry, J. (2017). Two NIRCcam Channels are Better than

One: How JWST Can Do More Science with NIRCams Short-wavelength Dispersed Hartmann Sensor. *PASP*, 129(1):015001.

Schwartz, J. C., Sekowski, C., Haggard, H. M., Pallé, E., and Cowan, N. B. (2016). Inferring planetary obliquity using rotational and orbital photometry. *MNRAS*, 457:926–938.

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Schwieterman, E., Meadows, V., Robinson, T. D., Lustig-Yaeger, J., Sparks, W. B., and Cracraft, M. (2016a). A synthetic high fidelity, high cadence spectral Earth database. In *AAS/Division for Planetary Sciences Meeting Abstracts*, volume 48 of *AAS/Division for Planetary Sciences Meeting Abstracts*, page 122.15.

Schwieterman, E. W., Cockell, C. S., and Meadows, V. S. (2015a). Nonphotosynthetic Pigments as Potential Biosignatures. *Astrobiology*, 15:341–361.

Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., Harman, C. E., DasSarma, S., Fisher, T. M., Arney, G. N., Hartnett, H. E., Reinhard, C. T., Olson, S. L., Meadows, V. S., Cockell, C. S., Walker, S. I., Grenfell, J. L., Hegde, S., Rugheimer, S., Hu, R., and Lyons, T. W. (2018). Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life. *Astrobiology*, 18:663–708.

Schwieterman, E. W., Meadows, V. S., Domagal-Goldman, S. D., Deming, D., Arney, G. N., Luger, R., Harman, C. E., Misra, A., and Barnes, R. (2016b). Identifying Planetary Biosignature Impostors: Spectral Features of CO and O₄ Resulting from Abiotic O₂/O₃ Production. *ApJ*, 819(1):L13.

Schwieterman, E. W., Robinson, T. D., Meadows, V. S., Misra, A., and Domagal-Goldman, S. (2015b). Detecting and Constraining N₂ Abundances in Planetary Atmospheres Using Collisional Pairs. *ApJ*, 810(1):57.

- Seager, S. and Deming, D. (2009). On the Method to Infer an Atmosphere on a Tidally Locked Super Earth Exoplanet and Upper Limits to GJ 876d. *ApJ*, 703:1884–1889.
- Seager, S. and Sasselov, D. D. (2000). Theoretical Transmission Spectra during Extrasolar Giant Planet Transits. *The Astrophysical Journal*, 537:916–921.
- Seager, S., Schrenk, M., and Bains, W. (2012). An astrophysical view of earth-based metabolic biosignature gases. *Astrobiology*, 12(1):61–82.
- Seager, S., Turnbull, M., Sparks, W., Thomson, M., Shaklan, S. B., Roberge, A., Kuchner, M., Kasdin, N. J., Domagal-Goldman, S., Cash, W., Warfield, K., Lisman, D., Scharf, D., Webb, D., Trabert, R., Martin, S., Cady, E., and Heneghan, C. (2015). The Exo-S probe class starshade mission. In *Techniques and Instrumentation for Detection of Exoplanets VII*, volume 9605 of Proc. SPIE, page 96050W.
- Segura, A., Kasting, J. F., Meadows, V., Cohen, M., Scalo, J., Crisp, D., Butler, R. A. H., and Tinetti, G. (2005). Biosignatures from Earth-Like Planets Around M Dwarfs. *Astrobiology*, 5:706–725.
- Segura, A., Krelove, K., Kasting, J. F., Sommerlatt, D., Meadows, V., Crisp, D., Cohen, M., and Mlawer, E. (2003). Ozone Concentrations and Ultraviolet Fluxes on Earth-Like Planets Around Other Stars. *Astrobiology*, 3:689–708.
- Seinfeld, J. and Pandis, S. (2006). Atmospheric chemistry and physics.
- Selsis, F., Wordsworth, R. D., and Forget, F. (2011). Thermal phase curves of nontransiting terrestrial exoplanets. I. Characterizing atmospheres. *A&A*, 532:A1.
- Shao, M., Marcy, G., Catanzarite, J. H., Edberg, S. J., Léger, A., Malbet, F., Queloz, D., Muterspaugh, M. W., Beichman, C., Fischer, D., Ford, E., Olling, R., Kulkarni, S., Unwin, S. C., and Traub, W. (2009). Astrometric Detection of Earthlike Planets. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010, page 271.

- Shields, A. L., Meadows, V. S., Bitz, C. M., Pierrehumbert, R. T., Joshi, M. M., and Robinson, T. D. (2013). The effect of host star spectral energy distribution and ice-albedo feedback on the climate of extrasolar planets. *Astrobiology*, 13(8):715–739.
- Sing, D. K., Fortney, J. J., Nikolov, N., Wakeford, H. R., Kataria, T., Evans, T. M., Aigrain, S., Ballester, G. E., Burrows, A. S., Deming, D., Désert, J.-M., Gibson, N. P., Henry, G. W., Huitson, C. M., Knutson, H. A., Lecavelier Des Etangs, A., Pont, F., Showman, A. P., Vidal-Madjar, A., Williamson, M. H., and Wilson, P. A. (2016). A continuum from clear to cloudy hot-Jupiter exoplanets without primordial water depletion. *Nature*, 529:59–62.
- Skilling, J. (2004). Nested sampling. In *AIP Conference Proceedings*, volume 735, pages 395–405. American Institute of Physics.
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., and Albrecht, S. (2010). The orbital motion, absolute mass and high-altitude winds of exoplanet HD209458b. *Nature*, 465(7301):1049–1051.
- Sole, R. V. and Munteanu, A. (2004). The large-scale organization of chemical reaction networks in astrophysics. *EPL (Europhysics Letters)*, 68(2):170.
- Soto, G., Sinha, A., Savransky, D., Delacroix, C., and Garrett, D. (2017). Starshade orbital maneuver study for wfirst. In *Techniques and Instrumentation for Detection of Exoplanets VIII*, volume 10400, page 104001U. International Society for Optics and Photonics.
- Spergel, D., Gehrels, N., Baltay, C., Bennett, D., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B., Greene, T., Guyon, O., et al. (2015). Wide-field infrared survey telescope-astrophysics focused telescope assets wfirst-afta 2015 report. *arXiv preprint arXiv:1503.03757*.
- Spurr, R. J. D. (2006). VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *J. Quant. Spec. Radiat. Transf.*, 102(2):316–342.

- Spurr, R. J. D., Kurosu, T. P., and Chance, K. V. (2001). A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *J. Quant. Spec. Radiat. Transf.*, 68:689–735.
- Stamnes, K., Tsay, S. C., Jayaweera, K., Wiscombe, W., Stamnes, S., Jin, Z., and Lin, Z. (2017). DISORT: DIScrete Ordinate Radiative Transfer.
- Stark, C. C. (2019). Exoplanet Yield Estimates for the LUVOIR Mission Concept. In *American Astronomical Society Meeting Abstracts #234*, volume 234 of *American Astronomical Society Meeting Abstracts*, page 301.15.
- Stark, C. C., Roberge, A., Mandell, A., and Robinson, T. D. (2014). Maximizing the ExoEarth Candidate Yield from a Future Direct Imaging Mission. *ApJ*, 795(2):122.
- Stephan, K., Jaumann, R., Brown, R. H., Soderblom, J. M., Soderblom, L. A., Barnes, J. W., Sotin, C., Griffith, C. A., Kirk, R. L., Baines, K. H., Buratti, B. J., Clark, R. N., Lytle, D. M., Nelson, R. M., and Nicholson, P. D. (2010). Specular reflection on Titan: Liquids in Kraken Mare. *Geophys. Res. Lett.*, 37:L07104.
- Stevenson, K. B., Désert, J.-M., Line, M. R., Bean, J. L., Fortney, J. J., Showman, A. P., Kataria, T., Kreidberg, L., McCullough, P. R., Henry, G. W., Charbonneau, D., Burrows, A., Seager, S., Madhusudhan, N., Williamson, M. H., and Homeier, D. (2014). Thermal structure of an exoplanet atmosphere from phase-resolved emission spectroscopy. *Science*, 346(6211):838–841.
- Strassmeier, K. (1994). Rotational-modulation mapping of the active atmosphere of the rs canum venaticorum binary hd 106225. *Astronomy and Astrophysics*, 281:395–420.
- STScI Development Team (2013). pysynphot: Synthetic photometry software package. Astrophysics Source Code Library.

- Suissa, G., Mandell, A. M., Wolf, E. T., Villanueva, G. L., Fauchez, T., and Kopparapu, R. k. (2020). Dim Prospects for Transmission Spectra of Ocean Earths around M Stars. *ApJ*, 891(1):58.
- Taylor, J., Parmentier, V., Irwin, P. G. J., Aigrain, S., Lee, G. K. H., and Krissansen-Totton, J. (2020). Understanding and mitigating biases when studying inhomogeneous emission spectra with JWST. *MNRAS*, 493(3):4342–4354.
- ter Braak, C. J. and Vrugt, J. A. (2008). Differential evolution markov chain with snooker updater and fewer chains. *Statistics and Computing*, 18(4):435–446.
- Thorsett, S. E., Arzoumanian, Z., and Taylor, J. H. (1993). PSR B1620-26: A Binary Radio Pulsar with a Planetary Companion? *ApJ*, 412:L33.
- Tian, F. (2015). Atmospheric escape from solar system terrestrial planets and exoplanets. *Annual Review of Earth and Planetary Sciences*, 43:459–476.
- Tremblay, L., Line, M. R., Stevenson, K., Kataria, T., Zellem, R. T., Fortney, J. J., and Morley, C. (2020). The Detectability and Constraints of Biosignature Gases in the Near- and Mid-infrared from Transit Transmission Spectroscopy. *AJ*, 159(3):117.
- Trotta, R. (2008). Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, 49(2):71–104.
- Turbet, M., Ehrenreich, D., Lovis, C., Bolmont, E., and Fauchez, T. (2019). The runaway greenhouse radius inflation effect. An observational diagnostic to probe water on Earth-sized planets and test the habitable zone concept. *A&A*, 628:A12.
- Turbet, M., Leconte, J., Selsis, F., Bolmont, E., Forget, F., Ribas, I., Raymond, S. N., and Anglada-Escudé, G. (2016). The habitability of proxima centauri b-ii. possible climates and observability. *Astronomy & Astrophysics*, 596:A112.

- Turnbull, M. C. (2015). ExoCat-1: The Nearby Stellar Systems Catalog for Exoplanet Imaging Missions. *ArXiv e-prints*.
- Underwood, D. R., Jones, B. W., and Sleep, P. N. (2003). The evolution of habitable zones during stellar lifetimes and its implications on the search for extraterrestrial life. *International Journal of Astrobiology*, 2(4):289–299.
- Van Grootel, V., Fernandes, C. S., Gillon, M., Jehin, E., Manfroid, J., Scudlaire, R., Burgasser, A. J., Barkaoui, K., Benkhaldoun, Z., Burdanov, A., Delrez, L., Demory, B.-O., de Wit, J., Queloz, D., and Triaud, A. H. M. J. (2018). Stellar Parameters for Trappist-1. *ApJ*, 853:30.
- Van Vleck, J. H. and Weisskopf, V. F. (1945). On the shape of collision-broadened lines. *Reviews of Modern Physics*, 17(2-3):227.
- Vanderspek, R., Huang, C. X., Vanderburg, A., Ricker, G. R., Latham, D. W., Seager, S., Winn, J. N., Jenkins, J. M., Burt, J., Dittmann, J., et al. (2019). Tess discovery of an ultra-short-period planet around the nearby m dwarf lhs 3844. *The Astrophysical Journal Letters*, 871(2):L24.
- Vida, K., Kővári, Z., Pál, A., Oláh, K., and Kriskovics, L. (2017). Frequent Flaring in the TRAPPIST-1 System Unsuitable for Life? *ApJ*, 841:124.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . (2019). SciPy 1.0—Fundamental Algorithms for Scientific Computing in Python. *arXiv e-prints*, page arXiv:1907.10121.
- Vogt, S. S., Allen, S. L., Bigelow, B. C., Bresee, L., Brown, B., Cantrall, T., Conrad, A., Couture, M., Delaney, C., Epps, H. W., Hilyard, D., Hilyard, D. F., Horn, E., Jern, N., Kanto, D., Keane, M. J., Kibrick, R. I., Lewis, J. W., Osborne, J., Pardeilhan, G. H., Pfister, T., Ricketts, T.,

- Robinson, L. B., Stover, R. J., Tucker, D., Ward, J., and Wei, M. Z. (1994). *HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope*, volume 2198 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 362.
- Von Paris, P., Hedelt, P., Selsis, F., Schreier, F., and Trautmann, T. (2013). Characterization of potentially habitable planets: Retrieval of atmospheric and planetary properties from emission spectra. *Astronomy & Astrophysics*, 551:A120.
- Wakeford, H. R., Lewis, N. K., Fowler, J., Bruno, G., Wilson, T. J., Moran, S. E., Valenti, J., Batalha, N. E., Filippazzo, J., Bourrier, V., Hörst, S. M., Lederer, S. M., and de Wit, J. (2019). Disentangling the Planet from the Star in Late-Type M Dwarfs: A Case Study of TRAPPIST-1g. *AJ*, 157(1):11.
- Wakeford, H. R., Sing, D. K., Deming, D., Gibson, N. P., Fortney, J. J., Burrows, A. S., Ballester, G., Nikolov, N., Aigrain, S., Henry, G., Knutson, H., Lecavelier des Etangs, A., Pont, F., Showman, A. P., Vidal-Madjar, A., and Zahnle, K. (2013). HST hot Jupiter transmission spectral survey: detection of water in HAT-P-1b from WFC3 near-IR spatial scan observations. *MNRAS*, 435(4):3481–3493.
- Waldmann, I. P., Tinetti, G., Rocchetto, M., Barton, E. J., Yurchenko, S. N., and Tennyson, J. (2015). Tau-REx I: A Next Generation Retrieval Code for Exoplanetary Atmospheres. *The Astrophysical Journal*, 802:107.
- Walker, S. I., Bains, W., Cronin, L., DasSarma, S., Danielache, S., Domagal-Goldman, S., Kacar, B., Kiang, N. Y., Lenardic, A., Reinhard, C. T., Moore, W., Schwieterman, E. W., Shkolnik, E. L., and Smith, H. B. (2018). Exoplanet Biosignatures: Future Directions. *Astrobiology*, 18(6):779–824.
- Wang, H. and Li, J. (2018). Adaptive gaussian process approximation for bayesian inference with expensive likelihood functions. *Neural computation*, 30(11):3072–3094.

- Wang, H. S., Liu, F., Ireland, T. R., Brasser, R., Yong, D., and Lineweaver, C. H. (2018). Enhanced constraints on the interior composition and structure of terrestrial exoplanets. *Monthly Notices of the Royal Astronomical Society*, 482(2):2222–2233.
- Wang, J., Mawet, D., Ruane, G., Hu, R., and Benneke, B. (2017a). Observing Exoplanets with High Dispersion Coronagraphy. I. The Scientific Potential of Current and Next-generation Large Ground and Space Telescopes. *AJ*, 153(4):183.
- Wang, S., Wu, D.-H., Barclay, T., and Laughlin, G. P. (2017b). Updated Masses for the TRAPPIST-1 Planets. *ArXiv e-prints*.
- Way, M. J., Aleinov, I., Amundsen, D. S., Chandler, M., Clune, T., Del Genio, A. D., Fujii, Y., Kelley, M., Kiang, N. Y., Sohl, L., et al. (2017). Resolving orbital and climate keys of earth and extraterrestrial environments with dynamics (rocke-3d) 1.0: a general circulation model for simulating the climates of rocky planets. *The Astrophysical Journal Supplement Series*, 231(1):12.
- Weiss, L. M. and Marcy, G. W. (2014). The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii. *ApJ*, 783(1):L6.
- Welbanks, L. and Madhusudhan, N. (2019). On Degeneracies in Retrievals of Exoplanetary Transmission Spectra. *AJ*, 157(5):206.
- West, R., Crisp, D., and Chen, L. (1990). Mapping transformations for broadband atmospheric radiation calculations. *J. Quant. Spec. Radiat. Transf.*, 43:191–199.
- Williams, D. M. and Gaidos, E. (2008). Detecting the glint of starlight on the oceans of distant planets. *Icarus*, 195:927–937.
- Wilson, C. F. and Widemann, T. (2019). Venus: Key to understanding the evolution of terrestrial planets. *arXiv e-prints*, page arXiv:1908.04269.

- Winn, J. N. (2010). *Exoplanet Transits and Occultations*, pages 55–77. University of Arizona Press.
- Wolf, E. T., Shields, A. L., Kopparapu, R. K., Haqq-Misra, J., and Toon, O. B. (2017). Constraints on climate and habitability for earth-like exoplanets determined from a general circulation model. *The Astrophysical Journal*, 837(2):107.
- Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355(6356):145–147.
- Wordsworth, R. D., Forget, F., Selsis, F., Millour, E., Charnay, B., and Madeleine, J.-B. (2011). Gliese 581d is the first discovered terrestrial-mass exoplanet in the habitable zone. *The Astrophysical Journal Letters*, 733(2):L48.
- Wordsworth, R. D., Schaefer, L. K., and Fischer, R. A. (2018). Redox Evolution via Gravitational Differentiation on Low-mass Planets: Implications for Abiotic Oxygen, Water Loss, and Habitability. *AJ*, 155:195.
- Wright, M. H. (1996). Direct search methods: Once scorned, now respectable. *Pitman Research Notes in Mathematics Series*, pages 191–208.
- Wunderlich, F., Godolt, M., Grenfell, J. L., Städt, S., Smith, A. M., Gebauer, S., Schreier, F., Hedelt, P., and Rauer, H. (2019). Detectability of atmospheric features of earth-like planets in the habitable zone around m dwarfs.
- Yang, H., Apai, D., Marley, M. S., Saumon, D., Morley, C. V., Buenzli, E., Artigau, É., Radigan, J., Metchev, S., Burgasser, A. J., et al. (2014). Hst rotational spectral mapping of two l-type brown dwarfs: variability in and out of water bands indicates high-altitude haze layers. *The Astrophysical Journal Letters*, 798(1):L13.
- Zahnle, K., Claire, M., and Catling, D. (2006). The loss of mass-independent fractionation in sulfur due to a palaeoproterozoic collapse of atmospheric methane. *Geobiology*, 4(4):271–283.

- Zahnle, K. J. and Catling, D. C. (2013). The Cosmic Shoreline. In *Lunar and Planetary Science Conference*, page 2787.
- Zahnle, K. J. and Catling, D. C. (2017). The Cosmic Shoreline: The Evidence that Escape Determines which Planets Have Atmospheres, and what this May Mean for Proxima Centauri B. *ApJ*, 843(2):122.
- Zhang, M., Chachan, Y., Kempton, E. M. R., and Knutson, H. A. (2019). Forward Modeling and Retrievals with PLATON, a Fast Open-source Tool. *PASP*, 131(997):034501.
- Zhang, Z., Zhou, Y., Rackham, B. V., and Apai, D. (2018). The Near-infrared Transmission Spectra of TRAPPIST-1 Planets b, c, d, e, f, and g and Stellar Contamination in Multi-epoch Transit Spectra. *AJ*, 156:178.
- Zhou, Y., Apai, D., Metchev, S., Lew, B. W., Schneider, G., Marley, M. S., Karalidi, T., Manjavacas, E., Bedin, L. R., Cowan, N. B., et al. (2018). Cloud atlas: Rotational modulations in the l/t transition brown dwarf companion hn peg b. *The Astronomical Journal*, 155(3):132.
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)*, 23(4):550–560.
- Zingales, T. and Waldmann, I. P. (2018). Exogan: Retrieving exoplanetary atmospheres using deep convolutional generative adversarial networks. *The Astronomical Journal*, 156(6):268.
- Zugger, M. E., Kasting, J. F., Williams, D. M., Kane, T. J., and Philbrick, C. R. (2010). Light Scattering from Exoplanet Oceans and Atmospheres. *ApJ*, 723:1168–1179.
- Zurlo, A., Vigan, A., Galicher, R., Maire, A.-L., Mesa, D., Gratton, R., Chauvin, G., Kasper, M., Moutou, C., Bonnefoy, M., et al. (2016). First light of the vlt planet finder sphere-iii. new spectrophotometry and astrometry of the hr 8799 exoplanetary system. *Astronomy & Astrophysics*, 587:A57.

Appendix A

VITA

JACOB LUSTIG-YAEGER

Curriculum Vitae

Email: jlustigy@uw.edu

GitHub: [jlustigy](#)

Web: <http://staff.washington.edu/jlustigy/>

Office Address

Department of Astronomy, University of Washington

Physics-Astronomy Bldg, B319

Box 351580

Seattle, WA 98195-1580

Education

Ph.D. Candidate in Astronomy & Astrobiology (dual-title PhD program) (2014 – Jun. 2020)

University of Washington, Seattle, WA

Advisor: Victoria Meadows

M. Sci. in Astronomy (2014 – 2016)

University of Washington, Seattle, WA

B.S. in Physics (Honors), Minor in Mathematics (2009 – 2013)

University of California, Santa Cruz, CA

Advisor: Jonathan Fortney

Research Interests

— Characterizing terrestrial exoplanets for habitability and signs of life

- Retrieving terrestrial planet atmospheres using Bayesian statistics
- Modeling telescope sensitivity to motivate future exoplanet science cases
- Developing novel methods to detect exoplanet habitability and biosignatures

Research Experience

Graduate Research Assistant: Virtual Planetary Laboratory (Sept 2014 – Present)

Terrestrial exoplanets, their atmospheres, habitability & biosignatures

- Developed, tested, and applied a physically rigorous terrestrial exoplanet retrieval model
- Simulated and analyzed radiative transfer, photochemical, climate, telescope, surface mapping, and machine learning models

Undergraduate & Postbaccalaureate Researcher: University of California, Santa Cruz (2012 – 2014)

Hot Jupiter & brown dwarf atmospheres, opacity sources, and atmospheric retrieval

- Retrieved exoplanet atmospheric compositions using emission spectra
- Applied Bayesian methods for parameter estimation
- Calculated and published weighted mean opacities for substellar and planetary atmospheres

Teaching Experience

Research Mentor: Department of Astronomy, University of Washington (Sept 2016 – Present)

Advising undergraduate students in exoplanet astronomy and astrobiology research

Teaching Assistant: Department of Astronomy, University of Washington (Sept 2014 – June

2015)

Led two biweekly sections for undergraduate students

- ASTR 101 (Spring 2015; Autumn 2014)
- ASTR 150 (Winter 2015)

Math & Writing Tutor: Learning Support Services, UCSC (Sept 2010 – June 2012)

Instructed students in college level mathematics and writing as a group and drop-in tutor

Honors & Awards

- NASA Group Achievement Award: LUVOIR Mission Concept Study Team (2019)
- Honors undergraduate thesis in physics (2013)
- University Honor, *cum laude* at University of California, Santa Cruz (2013)

Technical Skills

- Programming in Python, Julia, IDL, & Fortran
- Packaging Python code for open-source distribution
- Bayesian parameter inference and model selection
- Version control and collaboration with git & GitHub
- Data visualization with Matplotlib, Tableau, & Bokeh
- Writing and typesetting with L^AT_EX
- Extensive experience with MS Office Suite

Publications

First Authored

4. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). “[A Mirage of the Cosmic Shoreline: Venus-like Clouds as a Statistical False Positive for Exoplanet Atmospheric Erosion](#)”. *The Astrophysical Journal Letters*, 887(1), L11.
3. **Lustig-Yaeger, J.**, Robinson, T. D., & Arney, G. (2019). “[coronagraph: Telescope Noise Modeling for Exoplanets in Python](#)”. *Journal of Open Source Software*, 4(40), 1387.
2. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). “[The Detectability and Characterization of the TRAPPIST-1 Exoplanet Atmospheres with JWST](#)”. *The Astronomical Journal*, 158(1), 27.
1. **Lustig-Yaeger, J.**, Meadows, V. S., Tovar Mendoza, G., Schwieterman, E. W., Fujii, Y., Luger, R., & Robinson, T. D. (2018). “[Detecting Ocean Glint on Exoplanets Using Multi-phase Mapping](#)”. *The Astronomical Journal*, 156(6), 301.

Co-Authored

16. *Meadows, V. S., Lincowski, A. P., & **Lustig-Yaeger, J.** (2020, **in prep*). “The Feasibility of Detecting Biosignatures in the TRAPPIST-1 Planetary System with JWST”.
15. *Mansfield, M., Schlawin, E., **Lustig-Yaeger, J.**, et al. (2020, **in revision*). “Eigenspectra: A Framework for Identifying Spectra from 3D Eclipse Mapping”. *Monthly Notices of the Royal Astronomical Society*.
14. Leung, M., Meadows, V. S., & **Lustig-Yaeger, J.** (2020). “[High-Resolution Spectral Discriminants of Ocean Loss for M Dwarf Terrestrial Exoplanets](#)”. *The Astronomical Journal*, 160(1), 11.

13. Chouqar, J., Zouhair, B., Jabiri, A., **Lustig-Yaeger, J.**, Soubkiou, A., & Szentgyorgyi, A. (2020). “[Properties of Sub-Neptune Atmospheres: TOI-270 System](#)”. *Monthly Notices of the Royal Astronomical Society*, 495(1), 962-970.
12. Guzewich, S. D., **Lustig-Yaeger, J.**, Davis, C. E., Kopparapu, R. K., Way, M. J., & Meadows, V. S. (2020). “[The Impact of Planetary Rotation Rate on the Reflectance and Thermal Emission Spectrum of Terrestrial Exoplanets Around Sun-like Stars](#)”. *The Astrophysical Journal*, 893, 140.
11. Lincowski, A. P., **Lustig-Yaeger, J.**, & Meadows, V. S. (2019). “[Observing Isotopologue Bands in Terrestrial Exoplanet Atmospheres with the James Webb Space Telescope—Implications for Identifying Past Atmospheric and Ocean Loss](#)”. *The Astronomical Journal*, 158(1), 26.
10. Luger, R., Agol, E., Foreman-Mackey, D., Fleming, D. P., **Lustig-Yaeger, J.**, & Deitrick, R. (2019). “[STARRY: Analytic Occultation Light Curves](#)”. *The Astronomical Journal*, 157(2), 64.
9. Lincowski, A. P., Meadows, V. S., Crisp, D., Robinson, T. D., Luger, R., **Lustig-Yaeger, J.**, & Arney, G.N. (2018). “[Evolved Climates and Observational Discriminants for the TRAPPIST-1 Planetary System](#)”. *The Astrophysical Journal*, 867(1), 76.
8. Meadows, V. S., Reinhard, C. T., Arney, G. N., Parenteau, M. N., Schwieterman, E. W., Domagal-Goldman, S. D., Lincowski, A. P., Stapelfeldt, K. R., Rauer, H., DasSarma, S., Hegde, S., Narita, N., Deitrick, R., **Lustig-Yaeger, J.**, Lyons, T. W., & Siegler, N. (2018). “[Exoplanet Biosignatures: Understanding Oxygen as a Biosignature in the Context of Its Environment](#)”. *Astrobiology*, 18, 630-662.
7. Meadows, V. S., Arney, G. N., Schwieterman, E. W., **Lustig-Yaeger, J.**, Lincowski, A. P., Robinson, T., Domagal-Goldman, S. D., Barnes, R. K., Fleming, D. P., Deitrick, R., Luger,

- R., Driscoll, P. E., Quinn, T. R., Crisp, D. (2018). “The Habitability of Proxima Centauri b: Environmental States and Observational Discriminants”. *Astrobiology*, 18, 133-189.
6. Luger, R., **Lustig-Yaeger, J.**, & Agol, E. (2017). “Planet-Planet Occultations in TRAPPIST-1 and Other Exoplanet Systems”. *The Astrophysical Journal*, 851(2), 94.
 5. Fujii, Y., **Lustig-Yaeger, J.**, & Cowan, N. B. (2017). “Rotational Spectral Unmixing of Exoplanets: Degeneracies between Surface Colors and Geography”. *The Astronomical Journal*, 154(5), 189.
 4. Luger, R., **Lustig-Yaeger, J.**, Fleming, D. P., Tilley, M. A., Agol, E., Meadows, V. S., Deitrick, R., & Barnes, R. (2017). “The Pale Green Dot: A Method to Characterize Proxima Centauri b using Exo-Aurorae”. *The Astrophysical Journal*, 837, 63.
 3. Barnes, R., Deitrick, R., Luger, R., Driscoll, P. E., Quinn, T. R., Fleming, D. P., Arney, G., Crisp, D., Domagal-Goldman, S. D., Lincowski, A. P., **Lustig-Yaeger, J.**, & Schwieterman, E. (2017). “The Habitability of Proxima Centauri b I: Evolutionary Scenarios”. *arXiv preprint arXiv:1608.06919*.
 2. Greene, T. P., Line, M. R., Montero, C., Fortney, J. J., **Lustig-Yaeger, J.**, & Luther, K. (2016). “Characterizing transiting exoplanet atmospheres with JWST”. *The Astrophysical Journal*, 817(1), 17.
 1. Freedman, R. S., **Lustig-Yaeger, J.**, Fortney, J. J., Lupu, R. E., Marley, M. S., & Lodders, K. (2014). “Gaseous Mean Opacities for Giant Planet and Ultracool Dwarf Atmospheres over a Range of Metallicities and Temperatures”. *The Astrophysical Journal Supplement Series*, 214(2), 25.

Conference Presentations

Invited Talks

4. **Lustig-Yaeger, J.** & Meadows, V. S. (2019). [“The Era of Terrestrial Exoplanet Characterization”](#). AGU, Joint AGU-AAS Session on Frontiers in Exoplanets, San Francisco, CA, 10 Dec.
3. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). [“Prospects for Biosignatures for Rocky Planets in the Era of JWST”](#). SEEC Symposium 2019, “Rocky Exoplanets in the Era of JWST: Theory and Observation”, NASA Goddard, Greenbelt, MD, 5 Nov.
2. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). [“The Detectability and Characterization of the TRAPPIST-1 Exoplanet Atmospheres with JWST”](#). ExoPAG 20, Bellevue, WA, 23 Jun.
1. **Lustig-Yaeger, J.**, Meadows, V., Tovar, G., Schwieterman, E., & Fujii, Y. “Prospects for Mapping Terrestrial Exoplanets with LUVOIR”. Goddard Space Flight Center, October 25, 2017

Contributed Talks

7. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). [“TRAPPIST-1 and Beyond: Strategies for Characterizing Terrestrial Exoplanets and their Habitability”](#). Abstract [202-7] presented at 2019 Astrobiology Science Conference, Bellevue, WA, 24-28 June.
6. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). “Detecting and Characterizing Terrestrial Atmospheres in the TRAPPIST-1 System with JWST”. TRAPPIST-1 Conference, Liège, Belgium, 11-14 Jun.
5. **Lustig-Yaeger, J.**, Meadows, V. S., & Lincowski, A. P. (2019). [“Simulating the Detectability and Characterization of the TRAPPIST-1 Exoplanet Atmospheres with JWST”](#). American Astronomical Society Meeting Abstracts, 233, #103.03

4. **Lustig-Yaeger, J.**, Meadows, V. S., Tovar, G., Schwieterman, E., W., Fujii, Y., Luger, R., & Robinson, T. D. (2018). “[Detecting Oceans on Exoplanets with Next-Generation Telescopes](#)”. 2018 AGU Fall Meeting, Washington, D.C., 10-14 Dec. P51B-09
3. **Lustig-Yaeger, J.**, Luger, R., & Agol, E. (2017). “[Probing the Orbital Dynamics and Atmospheric Properties of the TRAPPIST-1 Planets with JWST](#)”. Habitable Worlds 2017, #4100
2. **Lustig-Yaeger, J.**, Tovar, G., Fujii, Y., Schwieterman, E., & Meadows, V. (2017). “[Mapping Surfaces and Clouds on Terrestrial Exoplanets Observed with Next-Generation Coronagraph-Equipped Telescopes](#)”. Astrobiology Science Conference, #3558
1. **Lustig-Yaeger, J.**, Line, M. R., & Fortney, J. J. (2015). “[On the Confidence of Molecular Detections in the Atmospheres of Exoplanets from Secondary Eclipse Spectra](#)”. American Astronomical Society Meeting Abstracts, 225, #124.03

Posters

8. **Lustig-Yaeger, J.**, Lincowski, A. P., & Meadows, V. S. (2017). “[Extending Atmospheric Characterization to Earth-Sized Exoplanets with JWST: Transits, Eclipses, and the TRAPPIST-1 System](#)”. Habitable Worlds 2017, Laramie, WY #4098
7. **Lustig-Yaeger, J.**, Tovar, G., Schwieterman, E. W., Fujii, Y., & Meadows, V. S. (2017). “[Detecting Oceans on Exoplanets Using Phase-Dependent Mapping with Next-Generation Coronagraph-Equipped Telescopes](#)”. Habitable Worlds 2017, Laramie, WY #4110
6. **Lustig-Yaeger, J.**, Schwieterman, E., Meadows, V., & Fujii, Y. (2016). “[Modeling Earth’s Disk-Integrated, Time-Dependent Spectrum: Applications to Directly Imaged Habitable Planets](#)”. AAS/Division for Planetary Sciences Meeting Abstracts, 48, #122.34

5. **Lustig-Yaeger, J.**, Meadows, V., Schwieterman, E. W., & Robinson, T. (2016). “[Modeling Earth’s Disk-Integrated Spectrum through a Lunar Month: Applications to Directly Imaged Habitable Exoplanets](#)”. Exoplanets I
4. **Lustig-Yaeger, J.**, Meadows, V., Line, M., & Crisp, D. (2015). “[A Novel Approach to Atmospheric Retrieval for Small Exoplanets](#)”. AAS/Division for Planetary Sciences Meeting Abstracts, 47, #416.10
3. **Lustig-Yaeger, J.**, Line, M., Fortney, J. J., & Meadows, V. (2015). “[Detecting Molecules in Exoplanet Atmospheres: Lessons Learned from Hot Jupiters](#)”. Astrobiology Science Conference, #7558
2. **Lustig-Yaeger, J.**, Line, M. R., & Fortney, J. J. (2014). “[On the Detection Significance of Molecules in Exoplanets from Secondary Eclipse Observations](#)”. Cool Stars, 18, #267
1. **Lustig-Yaeger, J.**, Fortney, J. J., Freedman, R., Marley, M. S., & Lupu, R. E. (2014). “[Gaseous Mean Opacities for Giant Planet and Brown Dwarf Atmospheres](#)”. American Astronomical Society Meeting Abstracts #223, #347.04

Public Talks

- “Proxima Centauri b: A World of Possibilities” and panel discussion with Guillem Anglada-Escude, Rory Barnes, & Olivier Guyon, UW Astrobiology & the NASA Astrobiology Institute Lecture Series, Seattle, WA. May 3, 2017.
- “BREAKING: Terrestrial Exoplanet Discovered in the Habitable Zone of Proxima Centauri” Astronomy on Tap, Peddler Brewing Company, Seattle, WA. August 24, 2016.

Press

- University of Washington, Seattle. (2019, August 13). “James Webb Space Telescope could begin learning about TRAPPIST-1 atmospheres in a single year, study indicates” [Press release].