The Astronomical Journal, 164:258 (27pp), 2022 December https://doi.org/10.3847/1538-3881/ac9ab3

The Detectability of Rocky Planet Surface and Atmosphere Composition with the JWST: The Case of LHS 3844b

Emily A. Whittaker[1,2], Matej Malik[2], Jegug Ih[2], Eliza M.-R. Kempton[2], Megan Mansfield[3], Jacob L. Bean[4], Edwin S. Kite[5], Daniel D. B. Koll[6], Timothy W. Cronin[6], and Renyu Hu[7,8]
[1] University of California, Los Angeles, Department of Earth, Planetary, and Space Sciences, Los Angeles, CA 90095-1567, USA; emilywhittaker@g.ucla.edu
[2] University of Maryland, Department of Astronomy, College Park, MD 20742, USA
[3] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[4] Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA
[5] Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA
[6] Department of Earth, Atmospheric, and Planetary Sciences, MIT, Cambridge, MA 02139, USA
[7] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[8] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA
Received 2022 July 12; revised 2022 September 19; accepted 2022 October 14; published 2022 November 18

Abstract

The spectroscopic characterization of terrestrial exoplanets over a wide spectral range from the near- to the mid-infrared will be made possible for the first time with the JWST. One challenge is that it is not known a priori whether such planets possess optically thick atmospheres or even any atmospheres altogether. However, this challenge also presents an opportunity, the potential to detect the surface of an extrasolar world. This study explores the feasibility of characterizing with the JWST the atmosphere and surface of LHS 3844b, the highest signal-to-noise rocky thermal emission target among planets that are cool enough to have nonmolten surfaces. We model the planetary emission, including the spectral signal of both the atmosphere and surface, and we explore all scenarios that are consistent with the existing Spitzer 4.5 μm measurement of LHS 3844b from Kreidberg et al. In summary, we find a range of plausible surfaces and atmospheres that are within 3σ of the observationless reflective metal-rich, iron-oxidized, and basaltic compositions are allowed, and atmospheres are restricted to a maximum thickness of 1 bar, if near-infrared absorbers at $\gtrsim$100 ppm are included. We further make predictions on the observability of surfaces and atmospheres and find that a small number ($\sim$3) of eclipse observations should suffice to differentiate between surface and atmospheric features. We also perform a Bayesian retrieval analysis on simulated JWST data and find that the surface signal may make it harder to precisely constrain the abundance of atmospheric species and may falsely induce a weak H_2O detection.

Unified Astronomy Thesaurus concepts: Exoplanet atmospheres (487); Exoplanet surfaces (2118); Theoretical models (2107); Infrared spectroscopy (2285)

1. Introduction

1.1. Characterizing Terrestrial Exoplanets in the Era of the JWST

During the next few years the James Webb Space Telescope (JWST) will serve as the prime observatory enabling us to characterize rocky exoplanets using spectroscopic information about the planetary thermal emission. The challenges to characterizing such planets include their smaller sizes and increased potential for atmospheric diversity, compared to more well-studied hot Jupiters. Another less-appreciated challenge is that, while it may generally be likely for terrestrial planets to have atmospheres, it is possible for an individual terrestrial exoplanet to lack an atmosphere that is sufficiently thick to be detectable. Whether or not a terrestrial exoplanet has a detectable atmosphere cannot be known a priori. This leads to a potential source of confusion in characterizing rocky worlds—it is not initially clear whether a planet's emission originates from its surface, or from its optically thick atmosphere. A third possibility is a semi-optically thin atmosphere, for which the surface is seen directly at certain wavelengths corresponding to transparent windows through the atmosphere. One such example is Titan, where parts of the surface can be seen in between CH_4 bands. (McCord et al. 2006). To optimize the scientific yield of observations of rocky exoplanets with the JWST, it is therefore crucial to understand the spectroscopic fingerprints of both their atmospheres and surfaces and how these translate to an emission signal recorded by the telescope.

Even before the age of the JWST, there have been a small number of spectroscopic observations of terrestrial exoplanets. For example, the transmission spectra of planets b, c, d, e, f, and g of the TRAPPIST-1 system were observed with the Hubble Space Telescope (HST) Wide Field Camera 3 (WFC3) and Spitzer Infrared Array Camera (IRAC) Channel 2 (de Wit et al. 2016; Gillon et al. 2017; de Wit et al. 2018). These observations allowed us to constrain the atmospheric properties of these planets to a certain degree, as their relatively featureless spectra are inconsistent with cloudless, hydrogen-dominated envelopes indicating either high mean molecular weight or cloudy atmospheres. Still, even a no-atmosphere case for the rocky planets cannot be ruled out by these data. More recently, Swain et al. (2021) observed the transmission spectrum of GJ 1132b with HST/WFC3 and concluded a low mean molecular weight atmosphere that exhibits spectral signatures of HCN and CH_4 as well as aerosol scattering. However, Mugnai et al. (2021) and Libby-Roberts et al. (2022) analyzed the same GJ 1132b data and both obtained a

featureless spectrum as the best fit, giving the same implications as for the atmospheres of the TRAPPIST-1 planets. To date, there has been no definitive detection of a terrestrial exoplanet atmosphere (Mugnai et al. 2021).

Terrestrial exoplanets that are found to not possess atmospheres could present the first opportunities to characterize the surface composition of planets beyond our solar system. The analogs of airless or nearly airless rocky bodies in the solar system are Mercury, Mars, the Moon, and various asteroids. Before they could be spatially well resolved, many of these objects were studied using the spectra of reflected solar radiation or of thermal emission. The lunar mare are generally dark and absorb strongly at 1 and 2 μm, indicating that they are basaltic in composition (Pieters 1978), while the lunar highlands are bright and relatively featureless, pointing to a plagioclase feldspar composition (Pieters 1986). The spectral features of the Martian surface, together with its visually red appearance and its polarization properties, show that it is rich in ferric oxides (McCord & Westphal 1971). Spectra of Mercury's surface initially indicated that it had a similar composition to that of the lunar highlands due to its relatively flat spectrum (Blewett et al. 1997). However, recent observations of Mercury from spacecraft have pointed to a surface that is closer to ultramafic (Nittler et al. 2011)—the more recent revision demonstrates the challenges in using remote sensing data to uniquely constrain planetary surface composition. Present-day Earth possesses a granitoid surface, a tertiary crust that is a record of plate tectonics and the incorporation of water in subducted crustal materials. All of these solar system examples present plausible surface compositions for terrestrial exoplanets. Additional possibilities include a metal-rich surface, which would be indicative of a world with the mantle stripped off (Hu et al. 2012), among others, as discussed in Section 4.2 of Mansfield et al. (2019).

At the limit of airless planets, Hu et al. (2012) proposed to use reflected and emitted light spectra of terrestrial exoplanets to constrain their surface compositions. However, the surface characterization may be complicated by the fact that such planets could possess substantial atmospheres, which could obscure the surface from view. We therefore need to explore whether an exoplanet's surface composition can be constrained in the absence of an atmosphere, and whether degeneracies exist between determining atmospheric and surface properties for cases in which an overlying atmosphere is present. This is the purpose of our current work.

1.2. The Rocky Super-Earth LHS 3844b

In this study, we focus on the specific test case of the hot, rocky exoplanet LHS 3844b. We assess our ability to characterize the planet with the JWST with a focus on plausible atmospheric and surface compositions, given our current knowledge of the planet's properties. LHS 3844b, discovered in 2018 by the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015), is a presumed tidally locked super-Earth with a radius of 1.303 ± 0.022 $R_{\rm Earth}$ and orbits its red dwarf M5-type host star every 11 hr (Vanderspek et al. 2019). With an emission spectroscopy metric (ESM) of 28 (Kempton et al. 2018), LHS 3844b provides one of the strongest emission signals for any exoplanet that is believed to have a dayside made of solid rock (in contrast to even hotter magma worlds), making it a prime target for future characterization. Indeed, LHS 3844b will be observed during three secondary eclipses with the JWST Mid-InfraRed Instrument (MIRI) Low-Resolution Spectrometer (LRS) as part of Cycle 1 of the General Observer program.

Due to the very high irradiation level from the host star, which promotes atmospheric escape, any hydrogen-dominated primordial atmosphere of LHS 3844b is thought to have been lost over the course of its lifetime (Owen & Wu 2013, 2017). However, without observational constraints it is difficult to assess the presence of secondary atmospheres as these can be produced through many pathways during the planet's evolution, e.g., degassing during accretion (Elkins-Tanton & Seager 2008), vaporization of a molten mantle (Schaefer et al. 2012), reactions between nebula gas and magma (Kite & Schaefer 2021), high-energy impacts (Lupu et al. 2014), and volcanism (Gaillard & Scaillet 2014). However, there are also a number of potential conditions that would prevent the formation and retention of a thick secondary atmosphere on LHS 3844b or similar planets, including (i) a volatile-poor formation and associated inhibition of late-stage degassing (Kite & Barnett 2020), (ii) stripping of the atmosphere due to impacts (Schlichting et al. 2015), (iii) primordial atmosphere loss, as heavier molecules will be dragged away alongside H_2 when the flux is high (Hunten et al. 1987), and (iv) a high stellar UV flux preventing the revival of an atmosphere (Kite & Barnett 2020).

Empirical data appear to strengthen the picture that LHS 3844b does not possess a thick atmosphere. Namely, Spitzer phase curve observations with IRAC Channel 2 at 4.5 μm by Kreidberg et al. (2019) measured a high dayside brightness temperature of 1040 ± 40 K, an undetectably small emission from the nightside, and a phase curve offset consistent with zero. As the main driver of heat redistribution is large-scale atmospheric circulation, the finding of a negligible longitudinal heat transport points toward the absence of a thick atmosphere (Showman et al. 2013; Koll 2022). Furthermore, comparing atmospheric model predictions to the measured eclipse depth of 380 ± 40 ppm, Kreidberg et al. (2019) narrowed the possible surface pressure to 10 bars or less. Simulating atmospheric mass loss over the planet's history, Kreidberg et al. (2019) further concluded that only with an initial water inventory of at least $\sim$100 Earth oceans could atmospheric erosion be prevented. Additionally, more recent ground-based observations using 13 transits in the optical to near-infrared by Diamond-Lowe et al. (2020) reveal a flat transmission spectrum, which is also consistent with a bare-rock scenario. In particular, the transmission spectrum observations ruled out a hydrogen-dominated atmosphere with high confidence (unless there are high-altitude clouds that would also flatten the spectrum) and disfavored a water-steam atmosphere with a surface pressure $\geqslant 0.1$ bar.

The most promising strategy to constrain the surface composition of rocky exoplanets is through secondary eclipse observations, which probe the planetary reflection and thermal emission spectrum. In contrast to the transmission spectrum, the planetary surface signal is imprinted on the emission spectrum in wavelengths where the atmosphere is optically thin. In addition, as the thermal emission signal scales with the planetary temperature, secondary eclipse spectroscopy appears to be especially well suited for atmospheric detection and characterization of hot planets, such as LHS 3844b (Morley et al. 2017; Koll et al. 2019). Due to these reasons, we model the planetary emission spectrum and make predictions on the

Table 1
Numerical Values for the Planetary and Stellar Parameters used in this Study

Planetary and Stellar Parameters						
R_{pl} (R_{Earth})	g_{pl} (cm s^{-2})	a (au)	R_{star} ($R_{\odot}$)	T_{star} (K)	$\log g_{\star}$ (cm s^{-2})	[M/H]$_{\star}$
1.246[a]	1600[b]	0.00622[a]	0.178[b]	3036[a]	5.136	0[a]

Notes. As the planet's mass is unknown, the surface gravity is approximated assuming an Earth-like bulk density. Note that the surface gravity affects the emission spectrum more weakly than the planetary transmission spectrum. The stellar surface gravity is calculated via Newton's law of gravity using a stellar mass of $M_{\star} = 0.158\ M_{\odot}$ (Kreidberg et al. 2019).
[a] Vanderspek et al. (2019).
[b] Kreidberg et al. (2019).

distinguishability of plausible atmospheres and surfaces and on the inference of the atmospheric composition of LHS 3844b using secondary eclipse observations in the current work.

2. Methods

2.1. Radiative Transfer Modeling

We generate model spectra of LHS 3844b with different assumed atmospheric species and surfaces with the open-source, 1D radiative transfer (RT) code HELIOS[9] (Malik et al. 2017, 2019a, 2019b), which simulates the atmosphere in radiative–convective equilibrium and the corresponding planetary emission. In contrast to most radiative transfer models used in the exoplanet field, HELIOS takes the radiative effects of the surface on both the atmosphere and the resulting planetary spectrum into account. Previously, only a gray surface albedo, constant over all wavelengths, could be modeled with HELIOS, as described in detail in Malik et al. (2019a). We have since expanded the code's functionality to include a surface albedo as a function of the wavelength, enabling the consideration of realistic surface albedos in the model. We describe updates to the HELIOS code and the new method of modeling the surface in Appendix A.

In HELIOS the radiative transfer calculation is performed twice. First, the atmospheric temperature profile and the surface temperature in radiative–convective equilibrium are obtained. For this step, we use the k-distribution method with 420 wavelength bins between 0.245 and 10^5 μm (distributed evenly in the wavenumber space) and 20 Gaussian points within each bin. The k coefficients are calculated from high-resolution gaseous opacities (see Section 2.3) with the ktable program that is included in the HELIOS software package. As in this work each atmospheric setup consists of only a single main infrared absorbing species alone, or in combination with a less absorbing background gas, the induced error by mixing k coefficients from different sources via the correlated-k approximation is marginal. Once the equilibrium state is found, HELIOS is used in post-processing mode to generate the emission spectrum using opacity sampling with a resolution of $R = 4000$.

Convection is treated in HELIOS via convective adjustment with the adiabatic coefficient $\nabla_{ad} = (\partial \log T / \partial \log P)_S$—where T is the temperature, P is the pressure, and S is the entropy—set to 2/7 (1/4) if a diatomic (triatomic) molecule acts as the atmospheric bulk gas, according to the ideal gas approximation. We model dayside-averaged conditions and use the scaling theory for heat redistribution of Koll (2022), their Equation (10), making the day-to-night heat transport dependent on the thickness of the atmosphere. Note that the scaling of the heat redistribution subtracts the total heat that is assumed to be transported to the nightside from the incoming stellar flux in order to estimate the dayside heat content, but it does not include any vertical dependency of the day-to-night heat flow. In the no-atmosphere case, we assume no global heat transport and set the redistribution parameter accordingly to $f = 2/3$ (Burrows et al. 2008; Hansen 2008). Lastly, a numerical limitation of the HELIOS code is that it is not possible to model a surface without an overlaying atmosphere, and so we approximate this situation by including an atmosphere with $P_{surf} = 2 \times 10^{-9}$ bar, which has a $\sim 10^{-10}$ Planck mean optical depth.[10]

We use a PHOENIX stellar model (Husser et al. 2013) to simulate the spectrum of the host star, downloaded from the online spectral library[11] and linearly interpolated in effective temperature and $\log_{10} g_{\star}$ with $g_{\star}$ being the stellar surface gravity. The numerical values for the stellar parameters of LHS 3844 are listed in Table 1. The high-resolution PHOENIX spectrum is further rebinned to the HELIOS wavelength grid, conserving the wavelength-integrated flux. Further input parameters include the planetary surface gravity, semimajor axis, planet radius, stellar radius, stellar temperature, stellar surface gravity, and stellar metallicity. The adopted numerical values for these parameters are displayed in Table 1.

2.2. Atmospheric and Surface Compositions Considered

For the atmospheric models of LHS 3844b we choose N_2, O_2, and CO_2 as potential dominant atmospheric species. From an empirical perspective, there are two solar system rocky bodies with thick atmospheres that have N_2 as the main constituent: Earth, which is of similar size to LHS 3844b, and Titan. There are a number of hypothetical pathways that may lead to substantial amounts of atmospheric N_2 (see discussion in Lammer et al. 2019). For instance, N_2 outgassing can potentially counteract nitrogen fixation by lightning, meteoritic impacts, and high-energy UV or flares, leading to atmospheric N_2 buildup (Chameides & Walker 1981; Mikhail & Sverjensky 2014; Airapetian et al. 2016). Another theoretical prediction is an O_2-dominated atmosphere, which is a potential product of an initially water-rich atmosphere that lost most of its hydrogen due to high irradiation from the host star (Wordsworth & Pierrehumbert 2014; Luger & Barnes 2015; Wordsworth et al. 2018). The high stellar irradiation places LHS 3844b close to the "cosmic shoreline" (Zahnle & Catling 2017), assuming an

[9] https://github.com/exoclime/helios

[10] According to our tests, an atmosphere with $P_{surf} < 10^{-6}$ bar does not noticeably affect the surface temperature or planetary emission found in the model.
[11] ftp://phoenix.astro.physik.uni-goettingen.de/HiResFITS/

Earth-like bulk density. The idea here is that if this planet started with a thick envelope and a sufficient water abundance, it is expected to have undergone a runaway greenhouse effect at one point during its history, possibly ending up as a Venus analog (Kane et al. 2014).

Additional absorbers in our models include CO_2, CO, SO_2, CH_4, and H_2O. They are all found in the atmospheres of Earth, Mars, and Venus, and they are predicted to be ubiquitous and major constituents in rocky exoplanet atmospheres as well, be it from mantle degassing (Schaefer & Bruce 2011; Schaefer et al. 2012; Liggins et al. 2022), meteoritic accretion (Lupu et al. 2014; Zahnle et al. 2020), photochemistry (Gao et al. 2015), or late-stage volcanism (Gaillard & Scaillet 2014). Additionally, all of these species exhibit strong absorption bands in the infrared, providing an opportunity to be detected with the JWST.

Our atmospheric grid consists of the following gas combinations. On one side we explore oxygen-rich and oxidizing atmospheres: bulk O_2 with H_2O, bulk O_2 with CO_2, bulk O_2 with SO_2, and a pure CO_2 scenario; and on the other side we explore nitrogen-rich and reducing atmospheres: bulk N_2 with CO, bulk N_2 with CH_4, bulk N_2 with CO_2, and a pure CO scenario.

For the surface contribution, six realistic surfaces compositions are tested: metal-rich (pyrite), feldspathic (97% Fe-plagioclase, 3% augite), ultramafic (60% olivine, 40% enstatite), basaltic (76% plagioclase, 8% augite, 6% enstatite, 5% glass, 1% olivine), granitoid (40% K-feldspar, 35% quartz, 20% plagioclase, 5% biotite), and iron-oxidized (50% nanophase hematite, 50% basalt), which are ubiquitous in the solar system and common products of geological processes (see Section 1 and further discussion in Hu et al. 2012). We use the geometric albedo spectra from Hu et al. (2012) for these surfaces, pre-tabulated between 0.3 and 25 μm, which were obtained from experimental data taken on rock powders and extended by radiative transfer modeling. Although these albedo spectra are mineralogically realistic, spectra of the surface of Mars and the Moon differ from that of mineral mixtures in the laboratory due to various effects (e.g., Carli et al. 2015; Bishop et al. 2019). We extrapolate the tabulated albedo data to the whole wavelength range of the RT calculation by using the albedo value at 0.3 μm for all smaller wavelengths and the value at 25 μm for all larger wavelengths. Lastly, we include a uniform gray surface with a geometric albedo of 0.1 as control to test the impact of the nongray albedo variation of realistic surfaces. The albedo spectra used in this work are shown in Figure 1.

2.3. Opacities & Scattering Cross Sections

Molecular opacities are calculated with `HELIOS`-K (Grimm & Heng 2015; Grimm et al. 2021). Included are O_2 (Gordon et al. 2017), H_2O (Polyansky et al. 2018), CO (Li et al. 2015), CO_2 (Rothman et al. 2010), CH_4 (Yurchenko et al. 2017), and SO_2 (Underwood et al. 2016). The opacities are calculated on a wavenumber grid with a resolution of 0.01 cm^{-1}, assuming a Voigt profile with a wing cutoff at 100 cm^{-1}. For H_2O, CH_4, and SO_2 we include the default pressure broadening coefficients provided by the Exomol database[12] (Tennyson et al. 2020). For O_2, CO, and CO_2 we use the High-resolution Transmission broadening formalism with self-broadening only. Further included is collision-induced absorption (CIA) by O_2–O_2, O_2–CO_2, CO_2–CO_2, N_2-N_2, and N_2–CH_4 (Richard et al. 2012) and Rayleigh scattering of H_2O, O_2, N_2, CO_2, and CO (Cox 2000; Sneep & Ubachs 2005; Wagner & Kretzschmar 2008; Thalman et al. 2014).

2.4. Simulating JWST Observations

To simulate observations with the JWST, the Python package `Pandexo`[13] (Batalha et al. 2017) is used. We focus on two JWST instruments that will be suitable and are planned to be utilized for rocky exoplanet characterization, MIRI LRS and NIRSpec G395H. Being on the long-wave end of the JWST capabilities, the bandpass of MIRI LRS (5.02–13.86 μm) maximizes the recorded secondary eclipse depth (planet-to-star contrast), as the emission of the hotter star falls off faster with the wavelength than the planetary emission. Although the relative eclipse depth is smaller in the NIRSpec G395H range (2.87–5.18 μm), a larger stellar flux at these wavelengths leads to a smaller overall photon noise. Additionally, many atmospheric species possess strong spectral signatures in the near-infrared range, making NIRSpec G395H one of the preferred choices for exoplanet observations.

When simulating observations with `Pandexo`, we insert the model values from the `HELIOS` spectrum (consistent with noise-free data), set a given instrument, resolution, and number of secondary eclipse measurements, and add the simulated statistical noise as error bars. For all tests done over the entire instrument wavelength range, points are rebinned using `Pandexo`'s integrated functionality, which defines new wavelength bins of a constant resolution and uses the included "uniform_tophat_mean" function to calculate the new bin-averaged function values together with the their uncertainty. In some cases, in addition to using the whole instrument range, we also focus on the detectability of individual spectral features by sampling over the width of these features, manually applying `Pandexo`'s rebinning procedure. As for the systematic noise floor, we assume a value of 30 ppm based on tentative estimates (Matsuo et al. 2019; Schlawin et al. 2020, 2021). To test this assumption, the same analysis as shown in Figure C7 has been conducted with the noise floor set to 0 ppm, which has led to indistinguishable results compared to our nominal setup. Therefore, we assume that the choice of noise floor has negligible impact on our results.

2.5. Bayesian Retrieval Modeling

We retrieve the simulated JWST spectrum using a modified version of the Bayesian retrieval code `PLATON` (Zhang et al. 2018, 2020). Our version, first utilized in Ih & Kempton (2021), differs from the main branch of `PLATON` in that it allows for measuring abundances of multiple species during retrieval (the so-called "free" retrieval). `PLATON` natively supports custom abundance profiles during forward models, but is configured to only perform equilibrium chemistry retrievals, in which the mixing ratios of all species are prescribed by the metallicity and C-to-O ratio defined globally, and temperature and pressure per layer. We relax this restriction by allowing the abundances of each species to be included as a retrieved parameters. All other details regarding radiative transfer remain the same as the original implementation.

[12] https://exomol.com/data/molecules/

[13] https://exoctk.stsci.edu/pandexo/

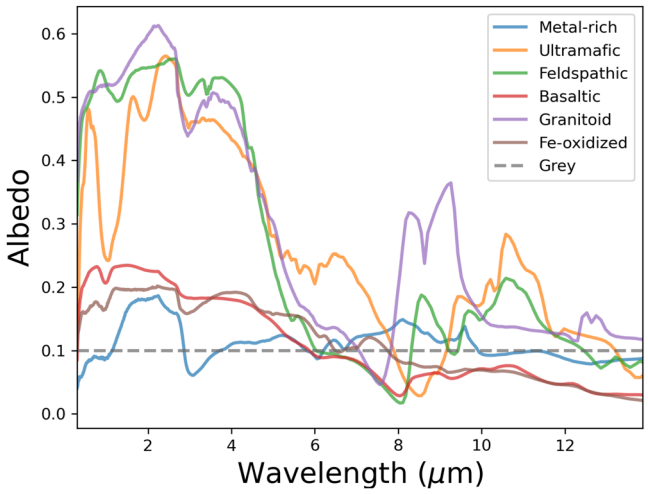

Figure 1. Nongray albedo vs. wavelength for each of the six tested surfaces as well as the gray control case with a constant 0.1 albedo (dashed). The nongray albedo extends over the wavelength range from 0.3 to 25 μm and thus captures the physically correct amount of reflected stellar light as well as thermal emissivity. The albedo data are from Hu et al. (2012).

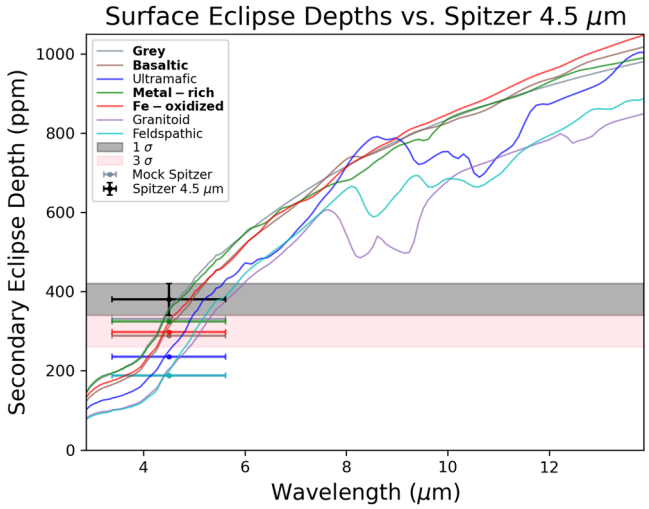

Figure 2. Secondary eclipse spectra for each of the tested surfaces and the gray albedo control case, shown at $R = 100$. The actual Spitzer 4.5 μm measurement is shown in black, where the y-axis error bar represents the uncertainty and the x-axis error bar represents the width of the bandpass. Modeled Spitzer points for each surface are shown in their respective color. We find the gray, metal-rich, iron-oxidized, and basaltic surfaces to be consistent (bold), i.e., $<3\sigma$ from the observation, and the ultramafic, granitoid, and feldspathic surfaces to be inconsistent with the Spitzer measurement. The granitoid and feldspathic mock Spitzer points are almost directly overlapping.

In our free retrieval, the atmosphere is parameterized as follows. For the composition of the atmosphere, we assume two possible background gases that do not have a spectral signature (O_2 and N_2). We then assign a parameter corresponding to the vertically fixed log-abundance of each species other than the background gases and a (linear) abundance for one of the background gases. The abundance of the other background gas is then derived as the remainder from unity. The thermal structure of the atmosphere is described by using the parametric T–P profile as given by Madhusudhan & Seager (2009). Having been developed with gas planets in focus, the retrieval code does not offer the possibility to include a solid surface. Still, we mimic the location of the surface in the retrieval modeling by the pressure level above which an isotherm is assumed, set by the parameter P_3 in the formalism of Madhusudhan & Seager (2009). As a surface radiating as a blackbody is equivalent to an optically thick atmosphere of a single temperature, P_3 should in theory correspond to a limit on the surface pressure. Approximating the surface with a blackbody does not allow us to make any statements about a nonzero surface albedo, and thus we do not retrieve on the surface but merely on the atmosphere in this work.

We perform our retrievals on the spectra generated using `Pandexo` by binning (resampling) the data down to a resolution of $R = 100$ and using uncertainties based on five simulated secondary eclipse observations.

3. Results

3.1. LHS 3844b Spitzer Eclipse Depth Constraint

In the following we first explore which of the tested surface types agree with the Spitzer eclipse depth measurement without considering an atmosphere (*no-atmosphere limit*). Then, analogously, we find the atmospheric models (varying composition and thickness) that are consistent with the Spitzer measurement. Among those we determine the atmospheres with the largest spectral features that may allow for characterization with the JWST (*thick atmosphere limit*).

Whether or not a model is in agreement with the Spitzer observation is determined by comparing the `HELIOS`-generated spectrum of the planet for each surface with the observed Spitzer 4.5 μm eclipse depth. The model is considered consistent with the data point if the simulated emission over the Spitzer bandpass, obtained by convolving the model spectrum with the Spitzer IRAC channel 2 bandpass function,[14] is within a 3σ confidence interval from the observed value.

3.1.1. No-atmosphere Limit

In the no-atmosphere limit we generate planetary emission spectra including the surface albedo signal only. Each model assumes that the whole planetary dayside is covered by one of the considered surface compositions. The secondary eclipse spectra of the surface-only models are shown in Figure 2. Based on this analysis, the Spitzer data point is most consistent with a metal-rich surface (1.39σ from Spitzer measurement), followed by an iron-oxidized (2.05σ) and a basaltic (2.28σ) surface. The ultramafic (3.61σ), feldspathic (4.79σ), and granitoid (4.80σ) surfaces are excluded by the data based on the assumed confidence limit. Curiously, all explored models predict a smaller eclipse depth than Spitzer with no single model prediction within 1σ of the observation. The best-fit surface model is the metal-rich surface at 1.39σ, giving a p-value of 0.163. First, this could be due to stochastic noise (although unlikely), but it could also indicate that the error of the Spitzer observation is underestimated due to an unknown systematic bias. As the reanalysis of the raw Spitzer phase curve of Kreidberg et al. (2019) is beyond the scope of this work, we limit ourselves to taking the Spitzer measurement at its face value. Second, it should be noted that not only is the albedo affected by grain size but even more the shortwave albedo can be lowered by small amounts of minor surface constituents, e.g., carbon, like on Mercury's surface (Izenberg et al. 2014), and metallic iron produced by space weathering.

[14] We use the Spitzer IRAC.I2 filter transmission function from http://svo2.cab.inta-csic.es/theory/fps/.

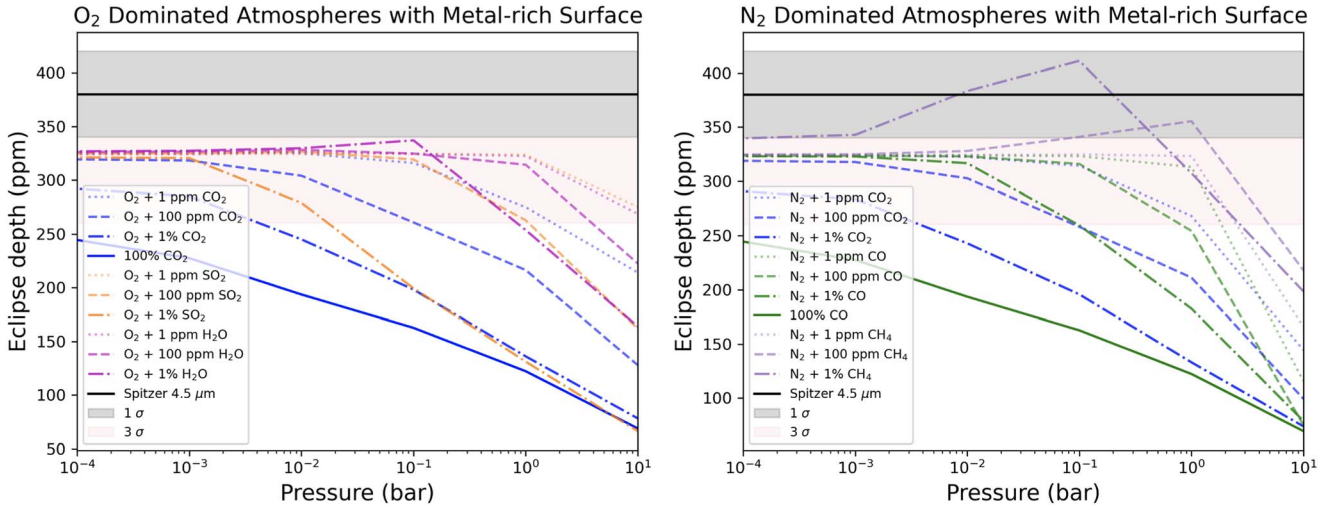

Figure 3. Predicted Spitzer 4.5 μm eclipse depth for the explored atmospheric models in combination with a metal-rich surface as a function of the surface pressure compared to the Spitzer measurement (black horizontal line). Oxidizing/O_2-dominated atmospheres are on the left and reducing/N_2-dominated atmospheres on the right. The gray shaded area corresponds to a 1σ uncertainty, and the pink shaded area to a 3σ uncertainty in the negative direction from the observed value.

Table 2
Maximum Surface Pressure (bar) Consistent with the Spitzer 4.5 μm Eclipse Depth at 3σ, for Each Atmospheric Composition Modeled

	Maximum Pressure (bar) Consistent with Spitzer Eclipse Depth					
	N_2 with CO_2	N_2 with CO	N_2 with CH_4	O_2 with CO_2	O_2 with SO_2	O_2 with H_2O
1 ppm	1	1	1	1	10	**10**
100 ppm	10^{-2}	10^{-1}	1	**10^{-1}**	**1**	1
1%	10^{-3}	10^{-2}	**1**	10^{-3}	10^{-2}	10^{-1}
100%	No Soln.	No Soln.	Not Modeled	No Soln.	Not Modeled	Not Modeled

Note. Abundances listed in the left-hand column correspond to abundances of the second-listed gas. The overall maximum surface pressure consistent with Spitzer is 10 bar among all setups and 1 bar once an infrared absorber at $\geqslant$100 ppm is included. The atmospheric models in bold possess the largest features and are used in the JWST observability analysis. If a value is listed as "No Soln"., then there was no solution consistent with Spitzer.

Thus, by increasing the absorption of light in the shortwave and consequently boosting the thermal emission in the infrared, minor modeling assumptions can in theory affect the ordering of surfaces as well as the answer to whether they are consistent with the Spitzer data point. Testing these assumptions is again beyond the scope of this work.

Compared to the previous analysis of Kreidberg et al. (2019) we obtain consistently shallower eclipse depths for the same surface compositions and also find the ultramafic crust inconsistent with the observation, which is in contrast to the modeling result in Kreidberg et al. (2019). This discrepancy is discussed in Section 4.2.

3.1.2. Including an Atmosphere

As the bottom boundary our nominal atmosphere models use the metal-rich surface. Being the most consistent with Spitzer, this choice maximizes the allowed atmospheric parameter space. Following the style of Figure 3 from Kreidberg et al. (2019), our Figure 3 shows the modeled eclipse depths for each tested atmospheric composition plotted across the range of probed surface pressures. The atmospheric temperature–pressure (T–P) profiles for the 10 bar and 1 mbar atmospheres are shown in Figure C1. The associated maximum surface pressures consistent with Spitzer for the 20 combinations of gases, and their abundances are listed in Table 2. We find that the modeled eclipse depths result from a combination of three distinct effects: day–night heat transport, greenhouse warming, and strength of gaseous absorption at 4.5 μm. Optically thick atmospheres transport more heat from the day to the nightside than thinner atmospheres and consequently lead to a smaller dayside emission (Koll 2022). This mechanism translates to a decreasing eclipse depth with increasing surface pressure and also a decreasing eclipse depth with increasing absorber abundance. The second effect is the greenhouse effect. As all of the explored absorbers are greenhouse gases, the T–P profiles increase with the pressure in optically thick atmospheric regions. (Note that the surface is smoothly connected to the first atmospheric layer via convective stability, i.e., temperature jumps at the surface are not permitted.) If the planetary emission originates at or near the surface, increasing surface pressure consequently leads to an increased thermal emission and larger eclipse depth. Similarly, increasing the greenhouse gas abundance increases the strength of the greenhouse effect, which warms the surface leading to a larger eclipse depth in spectral window regions (the total energy budget remains constant in radiative equilibrium). Finally, some species (CO_2, CO) exhibit a strong absorption band around 4.5 μm, whereas others do not (H_2O, SO_2) or even exhibit an absorption window (CH_4) that affects the eclipse depth at that wavelength location. These combined effects lead to nonmonotonic trends in the eclipse depth versus surface pressure or absorber abundance (as visible in Figure 3) and also nonmonotonic trends in the maximum surface pressure

consistent with Spitzer versus the absorber abundance (as visible in Table 2).

Overall, we find that some atmospheres with a surface pressure of 10 bar are consistent with the Spitzer measurement (see Table 2), but those possess a mixing ratio of only 1 ppm of a near-infrared absorber, which in the vast majority of cases has a marginal effect on the atmospheric extinction (1 ppm of H_2O being the exception among our models). Once absorbing species with a mixing ratio of at least 100 ppm are included, the maximum surface pressure for any tested atmosphere is 1 bar, obtained for compositions with CH_4, SO_2, or H_2O. With CO_2 or CO as the absorber, the maximum surface pressure allowed decreases to 0.1 bar.

The eclipse depths of the atmospheric models with O_2 and CO_2 deviate from the model predictions in Kreidberg et al. (2019) and also the overall maximum surface pressure is lower than the previously obtained maximum limit of 10 bar previously found. These differences are discussed in Section 4.2.

Lastly, exploring the impact of the surface on the allowed atmospheric parameter space, Figure C2 shows the analogous suite of atmospheric models as Figure 3 but with a basaltic surface. As expected, the modeled eclipse depths for each atmospheric composition are generally more consistent with Spitzer for the metal-rich surface than the basaltic surface as the latter is overall more reflective and thus less consistent with the Spitzer observation.

3.1.3. Determining the Thick Atmosphere Limit

In the following we pick for each gas pair (background and infrared absorber) the atmospheric model that provides the largest spectral features. Those models are then used for the JWST observability analysis presented in Section 3.2.3 and the retrieval analysis in Section 3.3. The eclipse spectra of the atmospheric models that are consistent with the Spitzer measurement are shown in Figure 4. As the atmospheric optical thickness depends on both the amount of absorbers in the atmosphere and the overall atmospheric density, the size of absorption features is degenerate in those parameters. As CO_2 and CO are strongly absorbing in the Spitzer bandpass, the only atmospheric models not excluded by Spitzer are those that either have no significant amount of these absorbers or low surface pressures. Hence, model spectra including those two absorbers have only weak features. The models including CO_2 appear to be very similar to each other, independent of the background gas, as the only noticeable absorption features are due to CO_2. Thus, we pick the scenario with O_2 as the background case for further analysis, in particular the O_2 with 100 ppm CO_2 model as it provides the largest CO_2 feature at 4.3 μm. No model with CO exhibits visible features and thus no CO models are used for further analysis. In contrast, H_2O, SO_2, and CH_4 have no absorption feature directly at 4.5 μm, which allows for larger atmospheric abundances of these species, still satisfying the Spitzer constraint. Among the corresponding models, we find that O_2 with 100 ppm SO_2 and N_2 with 1% CH_4 provide the largest spectral features. In particular, SO_2 possesses a large double-peaked feature at 7–9 μm and a smaller one at 4 μm, and CH_4 has strong absorption bands at 2–3 μm and from 4 to 8 μm. Lastly, the O_2 with 1 ppm H_2O model provides the largest spectral feature among all models that include H_2O. Interestingly, the striking feature at 5.5–7.5 μm is not due to H_2O alone, but also has a contribution from O_2–O_2 collision-induced absorption (CIA), which covers the small dip at 6.3 μm in the H_2O opacity.

The four atmospheric models picked for the JWST observability analysis in the thick atmosphere limit are O_2 with 100 ppm SO_2 and $P_{surface} = 1$ bar (hereafter $O_2|SO_{2(100\ ppm)}$), N_2 with 1% CH_4 and $P_{surface} = 1$ bar ($N_2|CH_{4(1\%)}$), O_2 with 100 ppm CO_2 and $P_{surface} = 10^{-1}$ bar ($O_2|CO_{2(100\ ppm)}$), and O_2 with 1 ppm H_2O and $P_{surface} = 10$ bar ($O_2|H_2O_{(1ppm)}$).

3.2. Observability with the JWST

In the following, we first explore the observability of surfaces without an overlying atmosphere (no-atmosphere limit). For that we use all six surface types shown in Figure 1 including the three surfaces excluded by the Spitzer measurement for general applicability of our results (see Section 3.1.1). In contrast, the observability of atmospheres is assessed only for the atmospheric models that are consistent with the recorded Spitzer eclipse depth and that provide the largest signal for each background composition and gas absorber (thick atmosphere limit; see Section 3.1.3).

3.2.1. No-atmosphere Limit

First, we test how many secondary eclipse observations would be necessary with the JWST in order to distinguish each surface emission from a blackbody spectrum using the MIRI LRS or the NIRSpec G395H instrument. We use a blackbody spectrum to establish a featureless reference spectrum, which allows us to test the visibility of the features stemming from the surface albedo variations.

For the two models to be considered "distinguishable," a chi-square test is conducted and a p-value is obtained for a given number of eclipses, where the p-value represents the probability that the simulated data from the model to be tested are consistent with a reference model. We begin with one eclipse and increase in integer steps until the two models are distinguishable by 3σ, given by a p-value of less than 0.0027. We declare the models indistinguishable under realistic observation times if the number of required eclipse observations exceeds 30, which is an arbitrary number chosen to be higher than what is expected to be economically feasible for the JWST to allocate for observing a single planet. The blackbody spectrum is obtained by fitting the mock JWST spectrum using SciPy's "curve_fit" function, with uncertainties on each data point at the given resolution obtained by `Pandexo`. Fitting the blackbody spectrum to the data allows us to infer a brightness temperature and corresponding planetary albedo; see Section 3.2.2 for further details. First, the mock data used as input for the blackbody fitting are downsampled to the resolution $R = 10$. Second, to determine the eclipses needed to differentiate each surface from a blackbody we use $R = 3$, analogous to distinguishing surfaces pairwise (see reasoning and Figure 5 further below).

The number of eclipses needed to distinguish each surface from a blackbody is listed in the first column and the first row of Table 3, depending on the JWST instrument used. In addition, the surface spectra for all plausible (i.e., consistent with Spitzer) surfaces together with the blackbody fit are displayed in Figure 6, and those that are inconsistent with Spitzer surface compositions are displayed in Figure 7. Unfortunately, we find that not a single surface will be

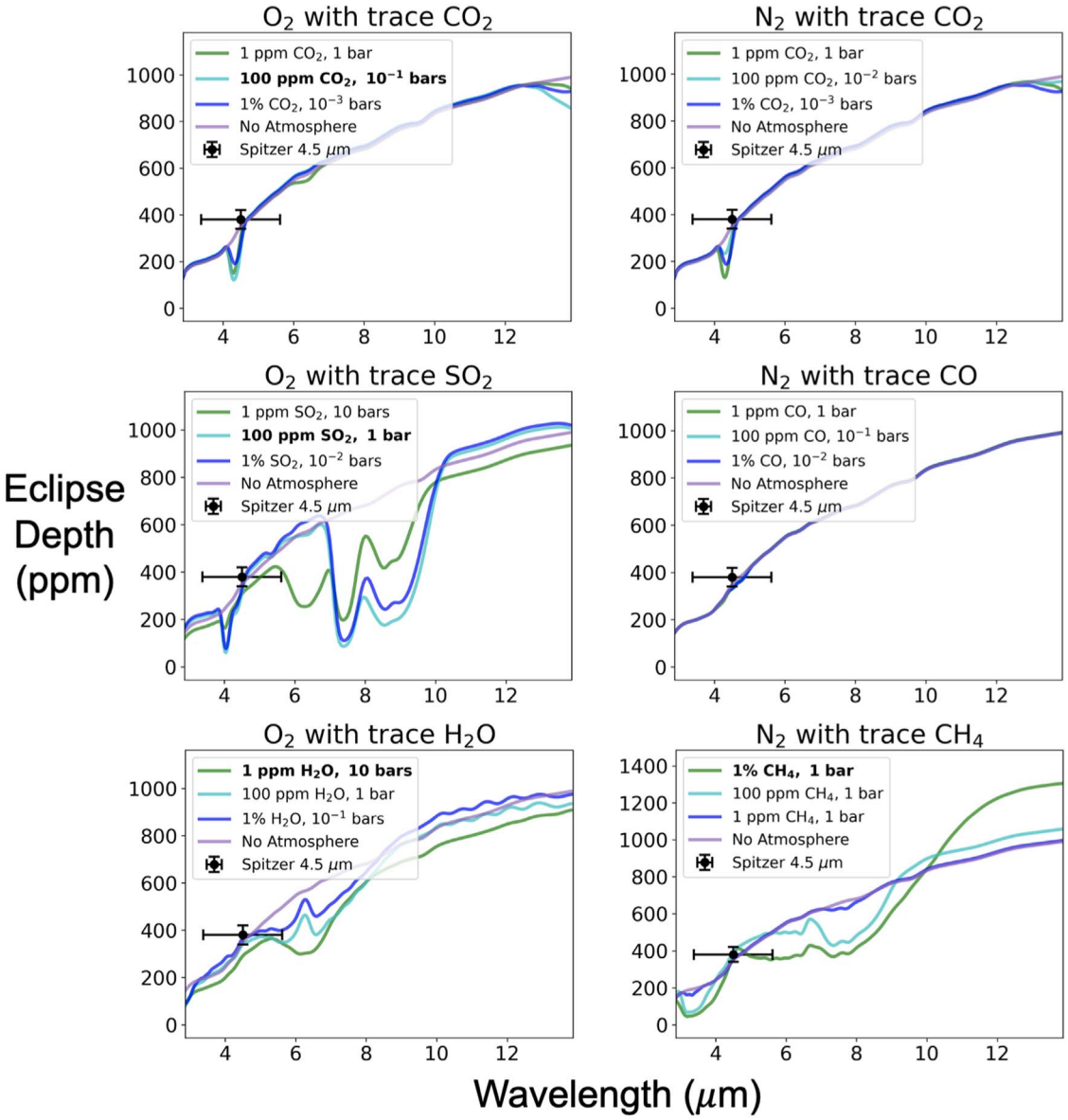

Figure 4. Secondary eclipse spectra for each atmosphere and metal-rich surface model at the maximum surface pressure consistent with the Spitzer observation. On each plot, the spectrum labeled "No Atmosphere" represents the surface-only spectrum for the metal-rich surface for comparison. The Spitzer 4.5 μm point is shown in black, where the y-axis error bar represents the uncertainty and the x-axis error bar represents the width of the bandpass. The atmospheric models with the largest features (bold) are used in the JWST observability analysis. No models are picked for the N_2 with CO_2 and the N_2 with CO cases as the former is very similar to O_2 with CO_2 and the latter does not exhibit any noticeable features.

distinguishable from a blackbody with a feasible number of eclipses. The only scenario with less than 30 eclipses is the ultramafic surface using MIRI, yet still requiring 25 eclipse observations. We conclude that the size of surface features alone is not sufficient for the tested surface to be characterized by the JWST.

As a next test, we assess the number of eclipses needed to distinguish each surface from one another, again using MIRI or NIRSpec and $R = 3$ for the data binning. This resolution is found to require the least number of eclipses for distinguishing surface models (see Figure 5 for the number of eclipses necessary to distinguish the metal-rich and granitoid surfaces dependent on the resolution of the binned mock data.) It appears that a small number of data points are sufficient, as the surface features tend to be broad and the comparison appears to be driven by the overall eclipse depth rather than the shape of individual features.

Table 3 shows the number of eclipses needed to distinguish surface pairs, and Figure C3 shows the secondary eclipse spectra for each pair of surfaces. It can be seen that many of the surfaces are more easily distinguishable from each other in the NIRSpec G395H band than in the MIRI LRS band. For example, the metal-rich and basaltic surfaces are indistinguishable in the MIRI LRS band, but are distinguishable with five secondary eclipses in the NIRSpec G395H band. Additionally, while the metal-rich and iron-oxidized surfaces are indistinguishable with MIRI LRS, they are distinguishable with 10 secondary eclipses using NIRSpec G395H. Overall, we find that less reflective materials are harder to disentangle due to their small features and eclipse depth variations than the more reflective ultramafic, granitoid, and feldspathic surfaces.

In addition to utilizing the entire instrument wavelength range and constant binning for the pairwise distinguishability,

Table 3
Number of Secondary Eclipse Observations Needed with the JWST to Distinguish Each Surface from a Blackbody Fit and from Another Surface at 3σ

Surface Detectability Using Whole Instrument Range								
	BB fit	Gray	Basaltic	Ultramafic	Metal-rich	Fe-oxidized	Granitoid	Feldspathic
BB fit		⩾30	⩾30	⩾30	⩾30	⩾30	⩾30	⩾30
gray	⩾30		**4**	1	⩾30	8	1	1
Basaltic	⩾30	⩾30		3	**5**	⩾30	1	1
Ultramafic	25	**10**	**17**		1	**2**	**3**	4
Metal-rich	⩾30	⩾30	⩾30	**11**		10	1	1
Fe-oxidized	⩾30	⩾30	⩾30	**14**	⩾30		1	1
Granitoid	⩾30	**3**	**4**	**7**	**4**	**4**		⩾30
Feldspathic	⩾30	**5**	**6**	**21**	**6**	**6**	⩾30	

Note. The number of eclipses using MIRI LRS is shown in cyan while the number of eclipses using NIRSpec G395H is shown in gray. A constant binning with $R = 3$ and the whole wavelength range of each instrument are used for this test. The detectability is not analyzed beyond 30 eclipses.

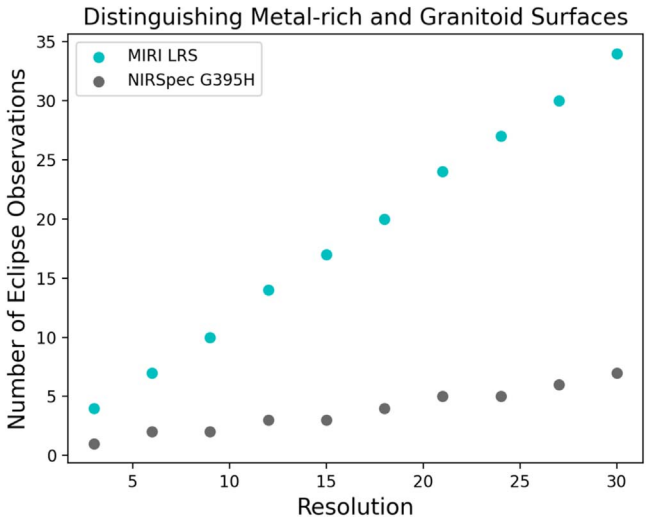

Figure 5. Number of secondary eclipse observations needed to distinguish the metal-rich and the granitoid surfaces as dependent on the binning resolution of the mock data. These two surfaces are used as an example pair to demonstrate the correlation between the resolution and eclipse number that is found to similarly hold for any surface pair. Eclipse observations with MIRI LRS are shown in cyan while observations with NIRSpec G395H are shown in dark gray.

we also conduct the same exploration focusing on the most prominent surface feature on either of the two surfaces and using optimized binning over the range of that feature. This is only done using the MIRI LRS instrument, as there are no prominent surface features within the NIRSpec G395H wavelength range (see Figure C3 on the right). Using qualitative ranking, we choose three prominent features to be included in this analysis. The double feature around 10 μm for the ultramafic, the double feature at 8.5 μm for the granitoid, and the double feature around 10 μm for the feldspathic surface, each of which is visible in its respective surface spectrum shown in Figure 7. The gray, basaltic, metal-rich, and iron-oxidized surfaces do not exhibit any sufficiently significant features for this analysis. Table 4 shows the number of eclipses needed to distinguish the ultramafic, granitoid, and feldspathic from all other surfaces using the described method, while the corresponding pairs of surface spectra are shown in Figure C4. We adopt a binning with a resolution such that each feature would encompass two data points spanning its whole width. The resolution this has resulted in is $R = 4$, 7, or 9 depending on the feature being spanned, and can be seen in the upper left corner of each plot in Figure C4.

Using this optimized binning over the most prominent surface feature, the number of eclipses needed to distinguish between three of the surface combinations was reduced, meaning that this method represented an improvement upon constant binning over the entire instrument wavelength range. More specifically, the number of eclipses was reduced from 6 to 3 for the granitoid and ultramafic surfaces, from 27 to 11 for the granitoid and feldspathic surfaces, and from 16 to 10 for the feldspathic and ultramafic surfaces. However, all other surface combinations tested using this optimized binning either remained unchanged or needed a greater number of eclipses than in the runs using the whole instrument range. As pointed out before, this result indicates that the constraining of surface compositions is largely driven by the Bond albedo of the surface (overall eclipse depth offset over the width of an instrument bandpass) rather than the spectral variation of features.

3.2.2. *Inferring Planetary Albedo and Temperature*

In the previous section we have fit the model spectra with a blackbody to determine whether surface features are detectable. However, the blackbody fit can also be used to determine the brightness temperature over a certain wavelength range. The brightness temperature in turn allows us to infer the planetary Bond albedo, which is a useful quantity that can be used to place constraints on the presence and properties of an atmosphere, as elucidated in detail in Mansfield et al. (2019). However, as pointed out in that work, inferring the Bond albedo from MIRI observations is prone to be biased due to the nongray shape of the surface albedo. In the following we reenact their assessment of this issue for LHS 3844b using the nongray radiative transfer code `HELIOS` and extend it to NIRSpec observations as well. First we assume that the brightness temperature is equivalent to the temperature of the blackbody fit for each surface. The inferred albedo is then

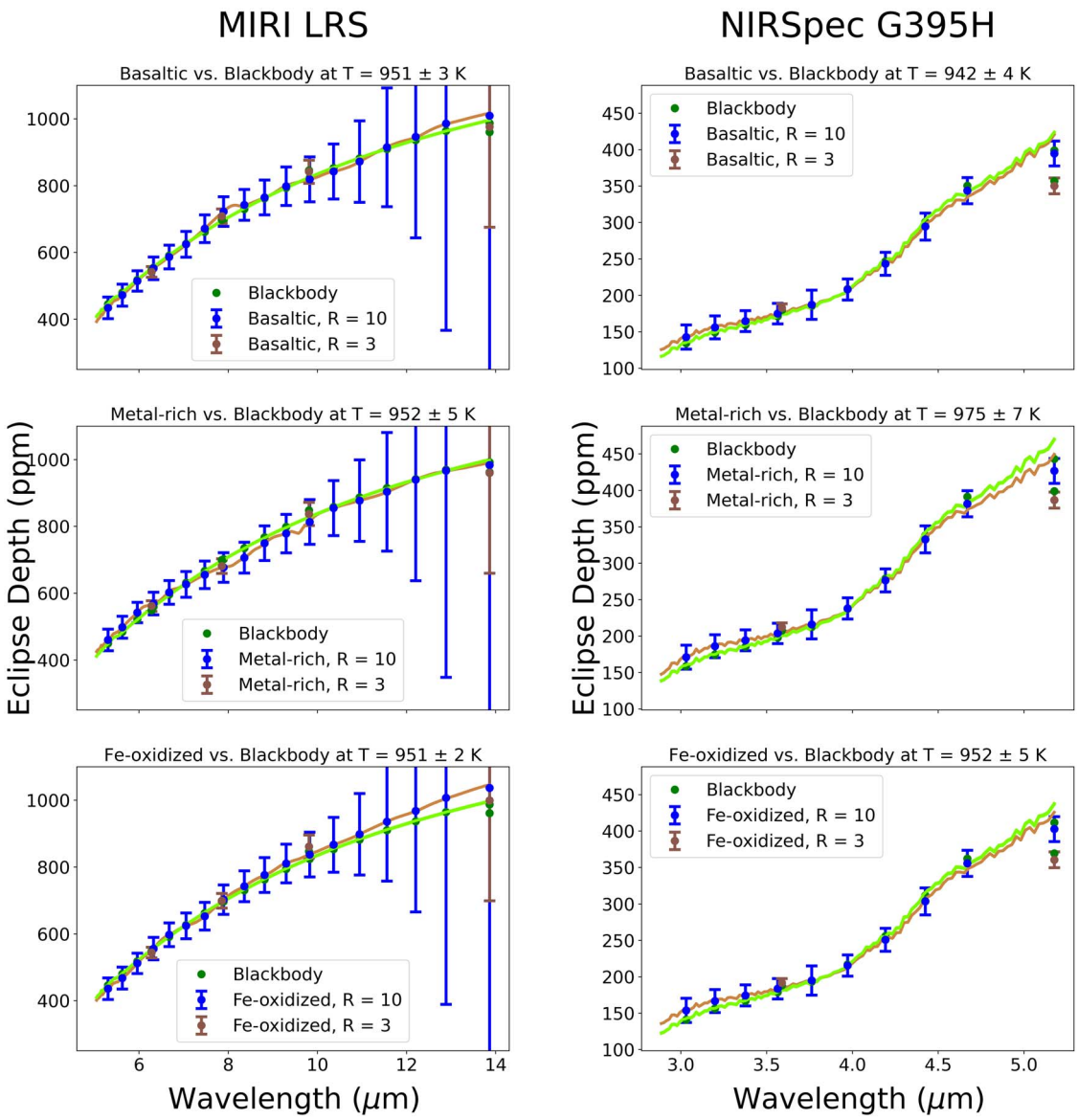

Figure 6. Secondary eclipse spectra of the surface-only models, mock data points for observations with the JWST MIRI LRS (left) and NIRSpec G395H (right), and corresponding blackbody fits. Note that the blackbody models do not appear as smooth curves because the emitted flux is divided by the stellar spectrum model to obtain the eclipse depth. Only the surfaces consistent with the Spitzer 4.5 μm eclipse depth are shown here. The mock data points are rebinned to $R = 3$ for the chi-square analysis and rebinned to $R = 10$ for the blackbody fitting. The error bars correspond to five eclipse observations. The original model spectra are downsampled from their native resolution to $R = 100$ for clarity.

calculated according to

$$\alpha_{\rm in} = 1 - \frac{\epsilon}{f}\left(\frac{T_{\rm bb}}{T_{\rm star}}\right)^4\left(\frac{a}{R_{\rm star}}\right)^2, \quad (1)$$

where $\alpha_{\rm in}$ is the inferred albedo, $T_{\rm bb}$ is the temperature from the blackbody fit, $T_{\rm star}$ is the temperature of the star, a is the semimajor axis of the planetary orbit, and $R_{\rm star}$ is the radius of the star. We set the heat redistribution factor f to 2/3 as predicted for a "vanishingly thin" atmosphere. The wavelength-integrated emissivity ϵ is a priori unknown and thus set to 1 in our reference models, which is equivalent to assuming that the surface radiates as a blackbody. We also discuss the effects of assuming $\epsilon < 1$ in Appendix B. We calculate the true Bond albedo for each surface by integrating the flux reflected from the planet over all wavelengths, and dividing by the total stellar flux received by the planet.

The upper two rows of Table 5 list the temperatures of the best fit associated with each of these blackbody models based on the MIRI or NIRSpec spectra, while the third row of Table 5 lists the true surface temperatures obtained in the HELIOS models.

The Bond albedos and inferred albedos for each surface are listed in the bottom part of Table 5. Each surface has a separate inferred albedo for each JWST instrument, in contrast to the Bond albedo, which is a single quantity resulting from the interplay between the planetary surface and the stellar irradiation.

Figure 8 displays the inferred brightness temperatures against the true model temperatures (left panel) and the inferred albedos against the Bond albedo (right panel) for each surface based on simulated MIRI observations. We find that the surface temperature of the planet is consistently underestimated by the blackbody fitting. This is because the fitting routine assumes an emissivity of 1 (equivalent to an albedo of 0), which assumes that the planet radiates heat more efficiently

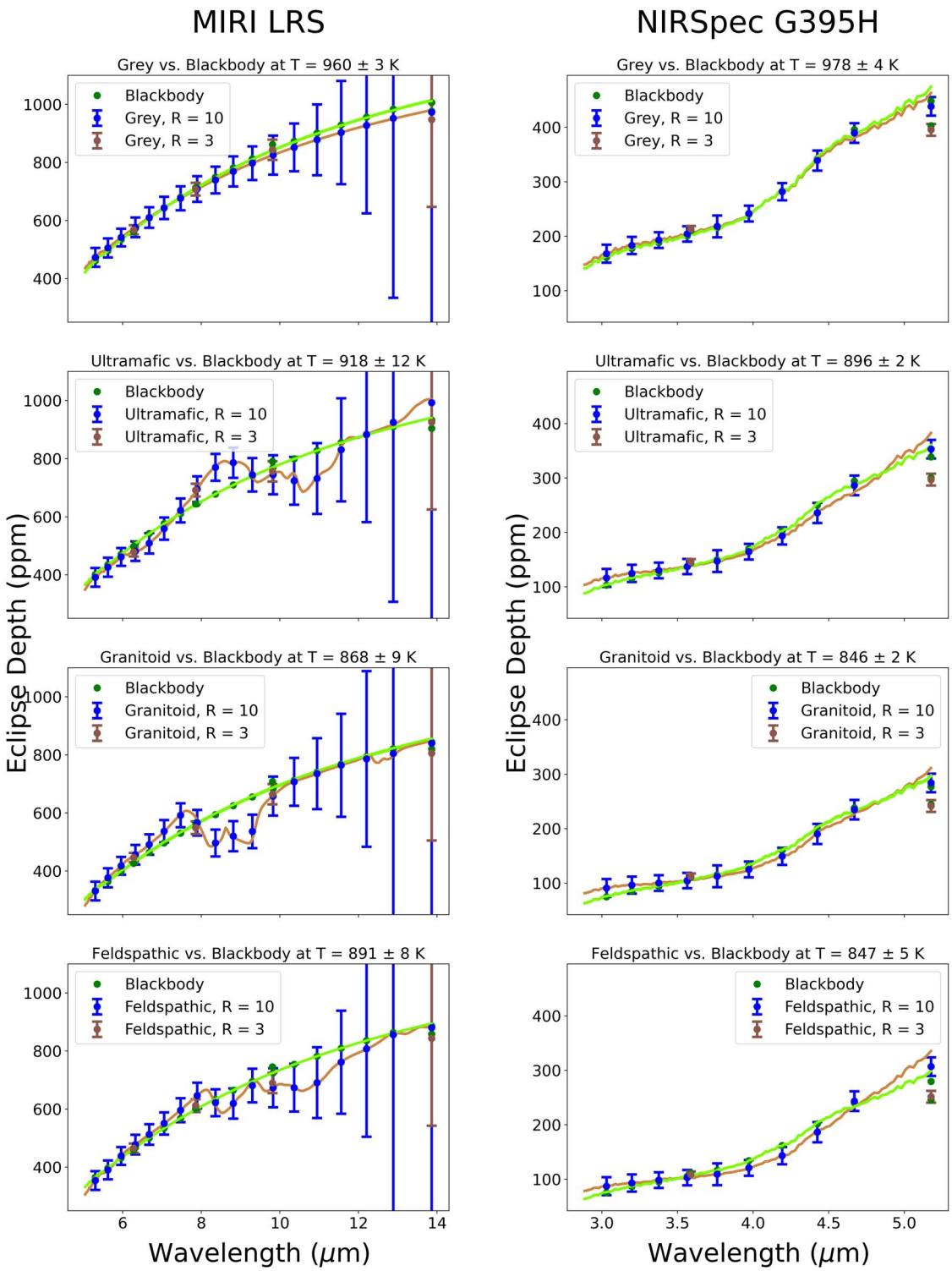

Figure 7. Secondary eclipse spectra of the surface-only models, mock data points for observations with the JWST MIRI LRS (left) and NIRSpec G395H (right), and corresponding blackbody fits. Shown here are the surfaces inconsistent with the Spitzer 4.5 μm eclipse depth. The mock data points are rebinned to $R = 3$ for the chi-square analysis and rebinned to $R = 10$ for the blackbody fitting. Error bars correspond to five eclipse observations. The model spectra are downsampled from their native resolution to $R = 100$ for clarity.

than it does in reality, finding a lower temperature for a given amount of thermal heat. The right panel shows that the majority of surfaces have an inferred albedo in the MIRI LRS band, which is lower than the true Bond albedo. This can be explained by the nongray albedo variation, discussed in detail in Mansfield et al. (2019). As a realistic, nongray surface albedo exhibits strong variations in the near to mid-infrared, the planetary emission is muted in the albedo peaks, and amplified in the albedo troughs, in accordance with Kirchhoff's law of radiation. As the majority of the surfaces tested here have a lower albedo within the MIRI bandpass than outside of it, emission within the bandpass is amplified, leading to an

Table 4
Number of Secondary Eclipse Observations Needed with the JWST to Distinguish Surfaces at 3σ, Focusing on Isolated Features with Optimal Binning

Distinguishing Surfaces with Optimal Binning of MIRI LRS						
	Gray	Basaltic	Ultramafic	Metal-rich	Fe-oxidized	Granitoid
Ultramafic	$\geqslant$30	$\geqslant$30		$\geqslant$30	25	
Granitoid	3	3	3	3	3	
Feldspathic	7	7	10	10	6	11

Note. Only MIRI LRS is used for this test as there are no sufficiently pronounced features in the NIRSpec G395H bandpass for any of the tested surface compositions.

Table 5
Top: Inferred Brightness Temperature Corresponding to the Blackbody Fit Temperature for Each Surface Based on Simulated MIRI LRS or NIRSpec G395H Observations at $R = 10$, Compared to the True Surface Temperature in Radiative Equilibrium Found by the HELIOS Radiative Transfer Code

Inferred Temperature and Albedo for Each Surface							
Inferred Temperature / Blackbody Fit Temperature (K)							
	Gray (0.1)	Basaltic	Ultramafic	Metal-rich	Fe-oxidized	Granitoid	Feldspathic
MIRI	961 ± 1	950 ± 1	918 ± 6	953 ± 2	950 ± 1	866 ± 8	891 ± 4
NIRSpec	977 ± 2	942 ± 2	895 ± 4	974 ± 3	951 ± 3	844 ± 4	846 ± 6
HELIOS Surface Temperature (K)							
	1000.78	979.831	990.324	998.682	992.145	926.571	928.914
Inferred Albedo							
	Gray (0.1)	Basaltic	Ultramafic	Metal-rich	Fe-oxidized	Granitoid	Feldspathic
MIRI	0.150 ± 0.005	0.189 ± 0.005	0.292 ± 0.018	0.179 ± 0.008	0.188 ± 0.005	0.439 ± 0.021	0.372 ± 0.012
NIRSpec	0.091 ± 0.006	0.215 ± 0.006	0.359 ± 0.010	0.103 ± 0.011	0.184 ± 0.009	0.494 ± 0.009	0.488 ± 0.014
Bond Albedo							
	0.100	0.218	0.384	0.128	0.185	0.526	0.511

Note. The uncertainties correspond to five eclipse observations. The observed brightness temperature as measured by Spitzer is 1040 ± 40 K. Bottom: Inferred albedo derived from the best-fit blackbody temperatures, listed on the top, and the true Bond albedo for each surface.

underestimation of the inferred albedo. In contrast, the albedo of the iron-oxidized surface is determined accurately and moreover the metal-rich albedo is overestimated. Here, the wavelength variation effect is countered by another bias induced by the assumption of an emissivity of unity, which leads to an overestimation of the albedo, best visible for the gray surfaces added as control cases (for more details see Appendix B). Lastly, another bias on the inferred albedo stems from the reflected stellar light being superimposed on the planetary thermal emission, mimicking a higher planetary emission and lower albedo. However, as the fraction of reflected light in the MIRI bandpass is minor, we find that this effect is relatively small (see Figure C5, left panel).

Interestingly, compared to the previous MIRI predictions in Mansfield et al. (2019) our new calculation predicts a consistently smaller observational bias for the albedo. After comparing the current and previous calculations, we have found an inconsistency in Mansfield et al. (2019) that we think is responsible for the stark difference between the current and previously obtained albedo values. In contrast to our current method, which is using blackbody fitting to infer the brightness temperature over an instrument bandpass, in Mansfield et al. (2019) we solved for the secondary eclipse depth integrated over the MIRI bandpass to infer the brightness temperature (see Equation (5) therein). For that we calculated the planetary eclipse depth using a PHOENIX stellar spectrum that is extrapolated by a blackbody of temperature 2755 K over the MIRI bandpass, but subsequently used a blackbody spectrum based on the effective temperature of LHS 3844 ($T_{\rm eff} =$ 3036 K) to infer the brightness temperature over the MIRI bandpass. This inconsistency in the stellar flux results in different baselines for the secondary eclipse depth, which directly translates to a bias in the planetary flux and in turn the inferred brightness temperature and the albedo.

With the NIRSpec instrument the albedo can be inferred more accurately than with MIRI, even though the temperature of the blackbody fit is similarly underestimated (see Figure 9). The albedo is in general, just as with MIRI, somewhat underestimated. However, here the main reason for the bias is the scattered light, the amount of which in the NIRSpec bandpass is significant for the more reflective surfaces. If the scattered light is removed from the analysis, the inferred albedo values become higher and predominantly overestimated (see Figure C5, right panel). The emissivity effect, a major source of

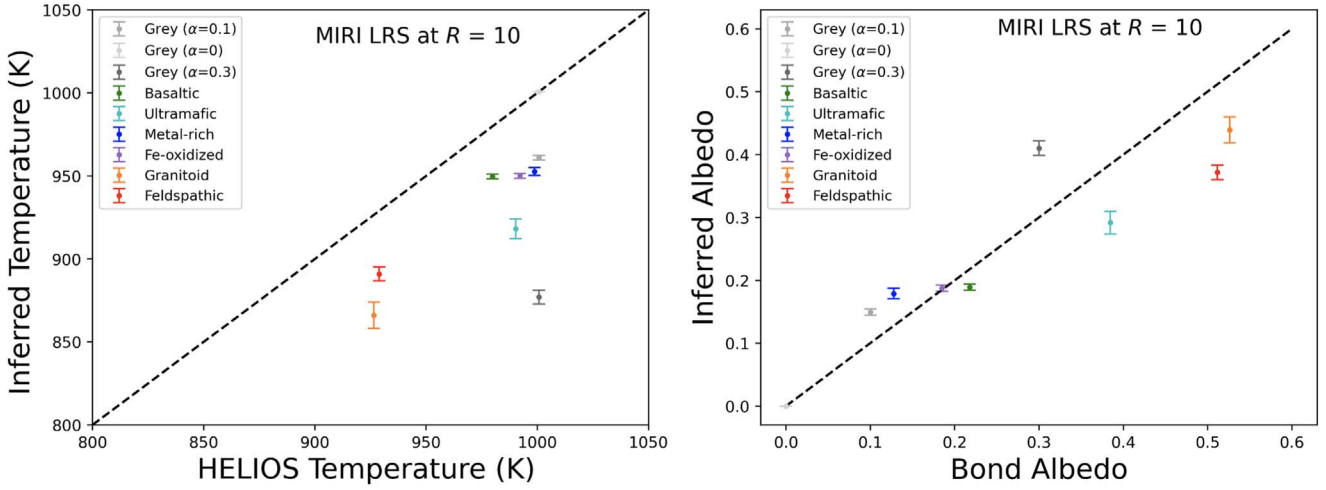

Figure 8. Left: Surface temperature obtained with HELIOS vs. inferred temperature of the blackbody fit for simulated MIRI LRS data at $R = 10$. The black dotted line represents the location where the inferred and HELIOS temperatures are equal. The error bars correspond to five eclipse observations. The choice of an emissivity of 1 when inferring the temperature causes it to be an underestimate compared to the true surface temperature. Right: Analogous to the left panel but showing the Bond albedo vs. inferred albedo. The inferred albedo is calculated with Equation (1) using the inferred temperature of the blackbody fit.

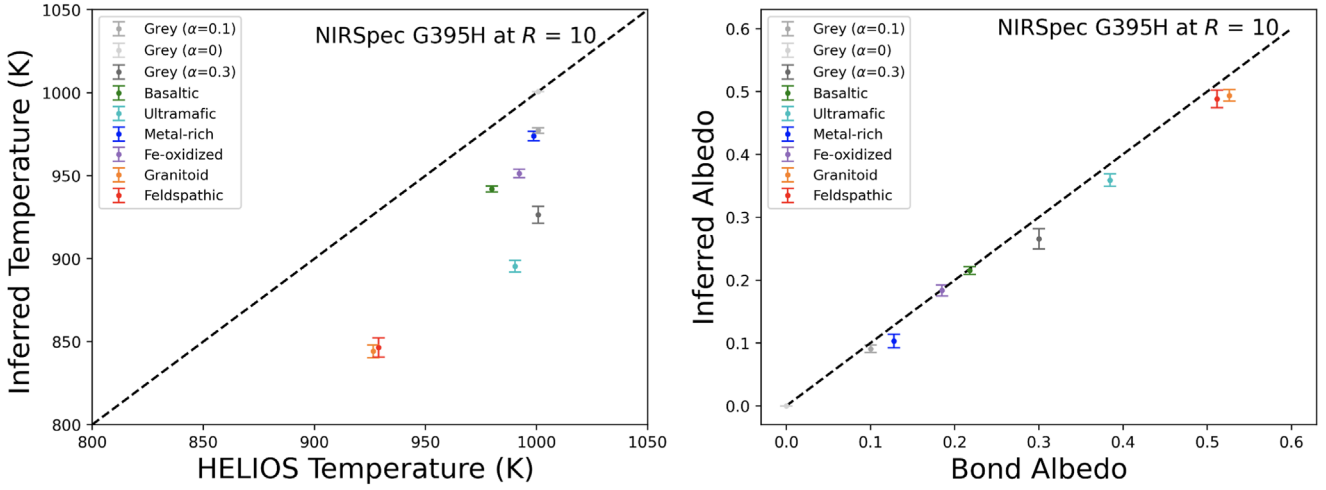

Figure 9. Same as Figure 8 but using simulated observations with the NIRSpec G395H instrument at $R = 10$.

inference bias when using MIRI, is only significant for the higher-albedo surface in the case of NIRSpec, as is further explained in Appendix B.

3.2.3. Thick Atmosphere Limit

In the following we explore whether the atmospheres in the thick limit (recall Section 3.1.3) are detectable with the JWST, i.e., if they can be distinguished from a pure surface spectrum. First, we utilize the whole instrument bandpasses of MIRI LRS and NIRSpec G395H and choose the resolution $R = 10$ as this represents a good compromise between statistical noise and sampling of the spectral features. The number of secondary eclipse observations needed to detect the atmosphere in each tested scenario is listed in Table 6 (top rows). We find that all of the atmospheres are detectable with $\leqslant 9$ eclipse observations if using the preferred instrument, with NIRSpec being preferable at finding CO_2 and CH_4, and MIRI preferable at finding SO_2 and H_2O. The detections with the lowest time requirement are $O_2|SO_{2(100\ ppm)}$ with MIRI and $N_2|CH_{4(1\%)}$ with NIRSpec, both requiring three eclipse observations. The $O_2|H_2O_{(1ppm)}$ atmosphere is detectable with MIRI and seven eclipses. As a worst case observing scenario, with MIRI the $O_2|CO_{2(100\ ppm)}$ atmosphere is indistinguishable from a surface. The spectra of the atmospheric models and the associated surface-only spectra can be seen in Figure C7.

Similar as in the surface observability analysis, we conduct another examination focusing on individual atmospheric absorption features or bands in isolation instead of taking the whole instrument range. In particular, for SO_2 we test the feature at 4 μm and the double-peaked band around 8 μm, for CH_4 we test the feature at 3.5 μm and the band around 7 μm, and for CO_2 we test the feature at 4.3 μm. There are no discernible features within the MIRI instrument range for the atmosphere with CO_2 and, analogously, no discernible features within the NIRSpec range for the atmosphere with H_2O. The resolution and binning of each spectrum are chosen to ensure that there are two to three data points sampling the feature being probed. The number of secondary eclipses needed to distinguish each atmospheric feature from the surface-only spectrum is listed in Table 6 (bottom rows). The resulting atmospheric model spectra and their no-atmosphere counterparts are shown in Figure 10, with error bars representing the

Table 6
Top: Number of Secondary Eclipses Needed to Distinguish an Atmospheric Model in the "Thick Limit" (see Section 3.1.3) from a Surface-only Spectrum at 3σ using MIRI and NIRSpec Mock Observations at $R = 10$

Eclipses Needed to Distinguish Atmospheres from the Surface-only Case				
	O_2 with 100 ppm CO_2	O_2 with 100 ppm SO_2	O_2 with 1 ppm H_2O	N_2 with 1% CH_4
Utilizing Whole Instrument Bandpass				
MIRI LRS	⩾30	3	7	6
NIRSpec G395H	9	7	10	3
Focusing on Individual Gas Features				
MIRI LRS	No Discernible Features	1	2	2
NIRSpec G395H	3	3	No Discernible Features	1

Note. The metal-rich surface is used for this test. The distinguishability of models is not analyzed beyond 30 secondary eclipses. Bottom: Same as above but focusing on the strongest gas features and absorption bands in isolation. No discernible features are found within the MIRI LRS range for the O_2 with 100 ppm CO_2 case and within the NIRSpec G395H range for the O_2 with 1 ppm H_2O case.

uncertainty after five secondary eclipse measurements. We find that isolating the gaseous features increases the detectability significantly to the extent that all explored atmospheres can be distinguished from a surface with at most three eclipses if using the preferred instrument for each case. Most promisingly, the CH_4 and SO_2 cases can be detected with only a single eclipse using NIRSpec or MIRI, respectively. The $O_2|H_2O_{\,(1ppm)}$ atmosphere is detectable with MIRI and two eclipses and the $O_2|CO_{2(100\ ppm)}$ atmosphere with NIRSpec and three eclipses.

3.3. Atmospheric Retrieval

In addition to the chi-square tests presented in the last section, we further run atmospheric retrieval models in order to determine how JWST observations may help constrain the atmospheric composition of LHS 3844b using a Bayesian framework. To this end, we use the simulated JWST data from both the NIRSpec G395H and MIRI LRS instruments with errors equivalent to five eclipses as input for the retrieval modeling. Note that, as described in Section 2.5, our code generating the forward models, `HELIOS`, adds the nongray surface to the spectrum, but the retrieval code `PLATON` does not include the surface and only assumes an isotherm extending to the infinite, which should mimic retrieving the pressure where the surface is expected. Hence this analysis further allows us to explore the impact of the realistic surface albedo on the results of the retrieval modeling.

Instead of immediately using the forward models that include a nongray surface, we first test the retrieval performance on pure atmosphere models, i.e., models with an atmosphere and a zero albedo surface. This test is indicative of whether the atmospheric signal alone is sufficient to make any statements about the gas inventory of the planetary envelope. Second, we then use the same atmospheric models and include the albedo signal of various surfaces. We include the crusts that are consistent with the Spitzer measurement, namely, basaltic, metal-rich, and iron-oxidized surfaces with the addition of ultramafic surfaces, in order to have an example of a more reflective surface. Finally, the last set of retrievals uses the surface spectrum only in order to explore whether the surface albedo variation itself can be interpreted as a false atmospheric signal by the retrieval algorithm in the case of the planet having no atmosphere at all.

The results of the atmospheric retrievals with and without a surface contribution are listed in Table 7 and the surface-only retrievals in Table 8. The striking results, discussed in detail in the main text and presented in Figures 11 and 12, are highlighted in magenta. This includes the gas volume mixing ratios retrieved from the no-surface models, a comparison between retrieved mixing ratios from the $O_2|H_2O_{\,(1ppm)}$ model when including an ultramafic surface and no surface, and H_2O mixing ratios retrieved for the surface-only models. Note that the species O_2 and N_2, not explicitly listed in the tables, are treated as background gases in the retrieval, and their mixing ratios always add up to unity with the other absorbers.

In the atmosphere-only setup the retrieval model manages to recover the absorbing species in each case it is present (see Figure 11, top row). Curiously, only the true H_2O mixing ratio of 1 ppm lies within the 1σ region of the retrieved posterior, $\log f_{H_2O} = -6.67^{+3.38}_{-2.80}$. The mixing ratios of the other absorbers, CO_2, SO_2, and CH_4, although the true values lie close to the 1σ range of the respective posteriors, are somewhat overestimated by the retrieval algorithm. This is likely a consequence of the degeneracy between the mixing ratio of an absorber, the atmospheric temperature, and the mean molecular weight of the atmosphere; all parameters that impact the optical depth of the atmosphere and different combinations of which can lead to a similar planetary spectrum. In this context, a systematic bias in any of these parameters could be caused by the fact that `HELIOS` and `PLATON` have radiative transfer solvers and ingredients that are not identical, e.g., they use different opacity line lists and scattering prescriptions. Furthermore, being limited by its simplified numerical scheme that relies on parameterized temperature profiles, `PLATON` cannot fully reproduce the radiative equilibrium state found by `HELIOS`. A second general finding is that, just as the retrieval algorithm succeeds at recovering present species, it also correctly disfavors absent species, as the retrieved mean mixing ratios of the latter do not exceed ∼0.1 ppm in any of the tested cases (see Figure C8 for a full grid of atmosphere-only model posteriors).

We find that the addition of a realistic surface crust generally appears not to impede our ability to detect absorbing species that are present in the atmosphere and sufficiently visible in the planetary spectrum. In all setups the posterior distributions of absorbers that are present are consistent between the cases with and without a nongray surface signal. However, we find a tendency that a nongray surface makes it harder for the retrieval code to reject species that are absent in the forward model. This phenomenon is best visible in the case of the $O_2|H_2O_{\,(1ppm)}$ atmosphere (see Figure 12). While the posteriors of the atmosphere-only setup appear to be well-behaved, with a clear upper limit on the mixing ratio (see CO_2, SO_2, and CH_4 panels of Figure 12), the posterior distributions in the case with the ultramafic surface remain broad and undetermined. Technically, it appears that the retrieval algorithm tries to match the

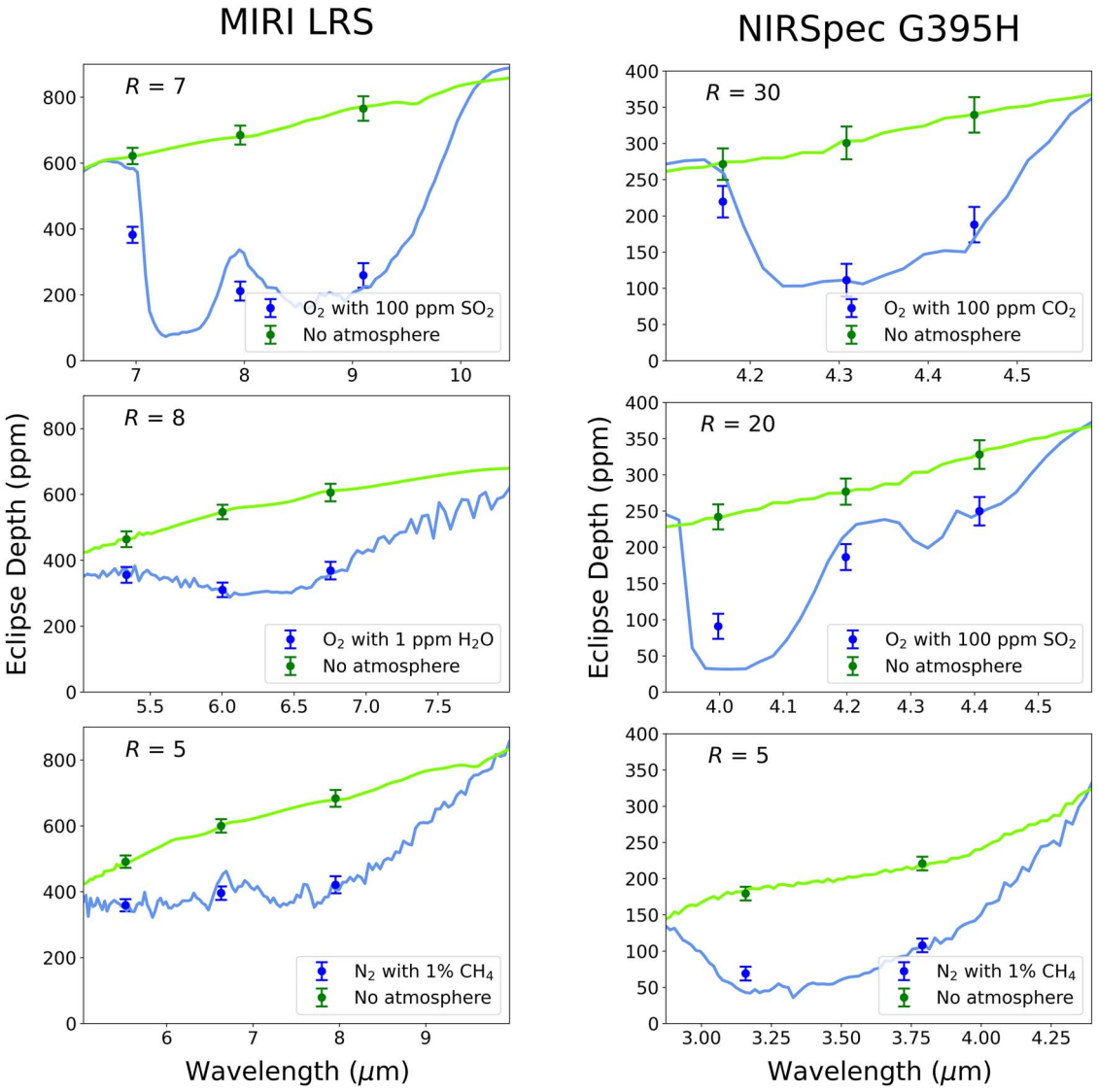

Figure 10. Secondary eclipse spectra of atmospheric models vs. surface-only models over the range of the strongest absorption bands in the respective models, overlaid with MIRI LRS (left) and NIRSpec G395H (right) simulated data points. The metal-rich surface is used for this test. The error bars show the uncertainties for five eclipse observations. The model spectra are downsampled from the native resolution to $R = 100$ for clarity.

surface features in the planetary spectrum with combinations of small amounts of atmospheric absorbers leading to many degenerate solutions. For instance, for SO_2 higher mixing ratios of $\gtrsim$1 ppm are strongly disfavored in the no-surface case, ($\log f_{SO_2} = -9.17^{+2.54}_{-1.68}$), but once the ultramafic signal is added ($\log f_{SO_2} = -5.75^{+3.26}_{-3.95}$) even mixing ratios of $\sim$1 % cannot be ruled out.

This last finding is further supported by our results from retrieving on the surface-only models. In general we find that the surface signatures imprinted on the planetary spectrum mislead the retrieval algorithm to falsely allow for the presence of atmospheric gases, with possible mixing ratios given by broad posterior distributions with a mean around $\sim$1 ppm. The extreme case is H_2O for which the retrieved mixing ratios are higher with a mean around $\sim$100 ppm in the cases of the basaltic, iron-oxidized, and ultramafic crusts (see Figure 11, bottom row). Also, in these cases the posterior distributions are more clearly defined and could misleadingly point to a weak H_2O detection. Figure C9 shows the full grid of surface-only model posteriors.

4. Discussion and Conclusions

4.1. New Constraints on the Surface and Atmosphere of LHS 3844b and their Observability

In this study we have explored the feasibility of characterizing the atmosphere and surface of rocky super-Earth LHS 3844b with the JWST. To find the parameter space of atmospheres and surface types that are plausible for LHS 3844b, we have modeled the planetary emission of LHS 3844b, including the spectral signal of both atmosphere and surface, and exhaustively explored all scenarios that are consistent with the existing Spitzer 4.5 μm measurement of Kreidberg et al. (2019). For the surface we have assumed six crust compositions that are common in the solar system and found that the surfaces that are consistent with Spitzer are metal-rich, iron-oxidized, and basaltic crusts (see Figure 2). In contrast, inconsistent with the data are ultramafic, granitoid, and feldspathic surfaces, whose high albedos lead to a planetary thermal emission too far below the recorded one. Our atmospheric models consist of O_2-, N_2-, CO_2-, and CO-dominated

Table 7
Logarithm of Input and Retrieved Gas Volume Mixing Ratios for the Atmospheres in the "Thick" Limit (see Section 3.1.3) without and with a Realistic Surface Crust

Retrieval Results for Atmosphere Models with and without Surface								
	CO_2		SO_2		H_2O		CH_4	
No Surface, O_2|CO $_{2(100\ ppm)}$	−4	$-2.33^{+1.17}_{-1.38}$	N Inc	$-8.24^{+2.24}_{-2.35}$	N Inc	$-7.46^{+2.91}_{-2.81}$	N Inc	$-8.45^{+2.70}_{-2.33}$
No Surface, O_2|SO $_{2(100\ ppm)}$	N Inc	$-8.19^{+2.25}_{-2.32}$	−4	$-1.92^{+1.05}_{-1.25}$	N Inc	$-8.30^{+2.64}_{-2.35}$	N Inc	$-8.13^{+2.44}_{-2.36}$
No Surface, O_2|H_2O $_{(1ppm)}$	N Inc	$-9.49^{+3.55}_{-1.56}$	N Inc	$-9.17^{+2.54}_{-1.68}$	−6	$-6.67^{+3.38}_{-2.80}$	N Inc	$-8.40^{+2.12}_{-2.05}$
No Surface, N_2|CH $_{4(1\%)}$	N Inc	$-7.16^{+1.93}_{-3.15}$	N Inc	$-6.79^{+2.66}_{-2.90}$	N Inc	$-7.02^{+3.23}_{-3.28}$	−2	$-0.95^{+0.58}_{-0.84}$
Basaltic, O_2|CO $_{2(100\ ppm)}$	−4	$-2.31^{+1.21}_{-1.63}$	N Inc	$-8.17^{+2.02}_{-2.25}$	N Inc	$-7.28^{+3.23}_{-2.80}$	N Inc	$-7.65^{+2.72}_{-2.65}$
Basaltic, O_2|SO $_{2(100\ ppm)}$	N Inc	$-8.27^{+2.58}_{-2.40}$	−4	$-1.66^{+0.97}_{-1.20}$	N Inc	$-7.77^{+2.76}_{-2.58}$	N Inc	$-7.23^{+2.49}_{-2.85}$
Basaltic, O_2|H_2O $_{(1ppm)}$	N Inc	$-6.62^{+4.00}_{-3.60}$	N Inc	$-6.31^{+3.52}_{-3.55}$	−6	$-5.93^{+3.56}_{-3.75}$	N Inc	$-6.12^{+3.37}_{-3.51}$
Basaltic, N_2|CH $_{4(1\%)}$	N Inc	$-8.12^{+2.26}_{-2.03}$	N Inc	$-7.75^{+3.10}_{-2.78}$	N Inc	$-7.54^{+2.97}_{-2.82}$	−2	$-1.21^{+0.65}_{-0.92}$
Metal-rich, O_2|CO $_{2(100\ ppm)}$	−4	$-2.33^{+1.19}_{-1.56}$	N Inc	$-8.17^{+2.28}_{-2.40}$	N Inc	$-7.73^{+3.05}_{-2.66}$	N Inc	$-7.81^{+2.71}_{-2.55}$
Metal-rich, O_2|SO $_{2(100\ ppm)}$	N Inc	$-8.41^{+2.51}_{-2.25}$	−4	$-2.10^{+1.24}_{-1.27}$	N Inc	$-8.05^{+2.70}_{-2.22}$	N Inc	$-8.15^{+2.37}_{-2.30}$
Metal-rich, O_2|H_2O $_{(1ppm)}$	N Inc	$-7.54^{+4.99}_{-2.99}$	N Inc	$-7.53^{+4.52}_{-2.87}$	−6	$-6.86^{+3.70}_{-3.24}$	N Inc	$-7.36^{+4.13}_{-2.87}$
Metal-rich, N_2|CH $_{4(1\%)}$	N Inc	$-7.80^{+2.19}_{-2.33}$	N Inc	$-7.47^{+2.53}_{-2.34}$	N Inc	$-6.68^{+2.79}_{-3.35}$	−2	$-1.16^{+0.61}_{-0.75}$
Fe-oxidized, O_2|CO $_{2(100\ ppm)}$	−4	$-2.06^{+1.10}_{-1.68}$	N Inc	$-7.86^{+2.42}_{-2.55}$	N Inc	$-7.35^{+3.07}_{-2.92}$	N Inc	$-7.75^{+2.73}_{-2.50}$
Fe-oxidized, O_2|SO $_{2(100\ ppm)}$	N Inc	$-7.92^{+2.19}_{-2.47}$	−4	$-1.71^{+0.97}_{-1.34}$	N Inc	$-8.32^{+2.94}_{-2.40}$	N Inc	$-7.67^{+2.57}_{-2.59}$
Fe-oxidized, O_2|H_2O $_{(1ppm)}$	N Inc	$-6.31^{+3.58}_{-3.62}$	N Inc	$-5.99^{+3.47}_{-3.55}$	−6	$-5.77^{+3.40}_{-3.91}$	N Inc	$-6.09^{+3.64}_{-3.56}$
Fe-oxidized, N_2|CH $_{4(1\%)}$	N Inc	$-8.23^{+1.88}_{-2.08}$	N Inc	$-7.77^{+2.59}_{-2.10}$	N Inc	$-8.13^{+3.83}_{-2.69}$	−2	$-1.39^{+0.71}_{-0.98}$
Ultramafic, O_2|CO $_{2(100\ ppm)}$	−4	$-2.62^{+1.45}_{-2.01}$	N Inc	$-8.30^{+2.45}_{-2.12}$	N Inc	$-6.09^{+3.24}_{-3.71}$	N Inc	$-7.14^{+2.42}_{-2.96}$
Ultramafic, O_2|SO $_{2(100\ ppm)}$	N Inc	$-8.04^{+2.61}_{-2.31}$	−4	$-1.64^{+0.90}_{-1.30}$	N Inc	$-7.61^{+2.84}_{-2.68}$	N Inc	$-7.67^{+3.03}_{-2.66}$
Ultramafic, O_2|H_2O $_{(1ppm)}$	N Inc	$-6.83^{+3.56}_{-3.47}$	N Inc	$-5.75^{+3.26}_{-3.95}$	−6	$-6.03^{+3.47}_{-3.70}$	N Inc	$-6.36^{+3.76}_{-3.68}$
Ultramafic, N_2|CH $_{4(1\%)}$	N Inc	$-8.40^{+2.39}_{-2.35}$	N Inc	$-7.44^{+2.57}_{-2.40}$	N Inc	$-6.31^{+2.82}_{-2.97}$	−2	$-1.18^{+0.65}_{-0.84}$

Note. "N Inc" means that the species is not included in the forward model. Individual results discussed in Section 3.3 and presented in Figures 11 and 12 are highlighted in magenta. Note that O_2 and N_2, treated as background gases in the retrieval modeling, are not directly retrieved, and thus their mixing ratios are not listed here.

Table 8
Logarithm of Retrieved Gas Volume Mixing Ratios for Surface Models without Including an Atmosphere

Retrieval Results for Surface-only Models				
	CO_2	SO_2	H_2O	CH_4
Basaltic, No Atmo.	$-6.79^{+3.58}_{-3.11}$	$-6.11^{+3.20}_{-3.53}$	$-4.07^{+2.56}_{-5.10}$	$-6.57^{+3.41}_{-3.18}$
Metal-rich, No Atmo.	$-6.67^{+3.83}_{-3.31}$	$-6.86^{+3.55}_{-3.15}$	$-5.53^{+3.43}_{-4.18}$	$-6.32^{+3.55}_{-3.55}$
Fe-oxidized, No Atmo.	$-6.57^{+3.77}_{-3.40}$	$-6.95^{+3.62}_{-3.06}$	$-3.54^{+2.21}_{-4.42}$	$-6.99^{+3.85}_{-2.98}$
Ultramafic, No Atmo.	$-7.03^{+3.62}_{-3.12}$	$-6.04^{+3.70}_{-3.91}$	$-4.05^{+2.49}_{-4.85}$	$-6.10^{+3.82}_{-3.64}$

Note. Individual results discussed in Section 3.3 and presented in Figure 11 are highlighted in magenta. Note that O_2 and N_2, treated as background gases in the retrieval modeling, are not directly retrieved, and thus their mixing ratios are not listed here.

atmospheres with H_2O, CO_2, CO, CH_4, and SO_2 as additionally included near-infrared absorbers of various mixing ratios. We have found that in order to be consistent with Spitzer the maximum surface pressure is 10^{-1} bars for the models including CO_2 or CO and 1 bar for the models including H_2O, CH_4, or SO_2 if the mixing ratio of the included infrared absorber $\geqslant 100$ ppm (see Figure 3 and Table 2).

Next, we have conducted a JWST observability analysis exploring two limits; on one end we have assumed there is no atmosphere, and on the other end we have investigated the atmospheric scenarios that provide the largest gas features while being consistent with the Spitzer data. In the no-atmosphere limit our analysis predicts that it will be very difficult to disentangle specific surface types from a blackbody null hypothesis. However, making use of not only the feature strength but also the total offset of the surface emission (or reflection) over an instrument bandpass, we have found that among the plausible surfaces the basaltic and metal-rich

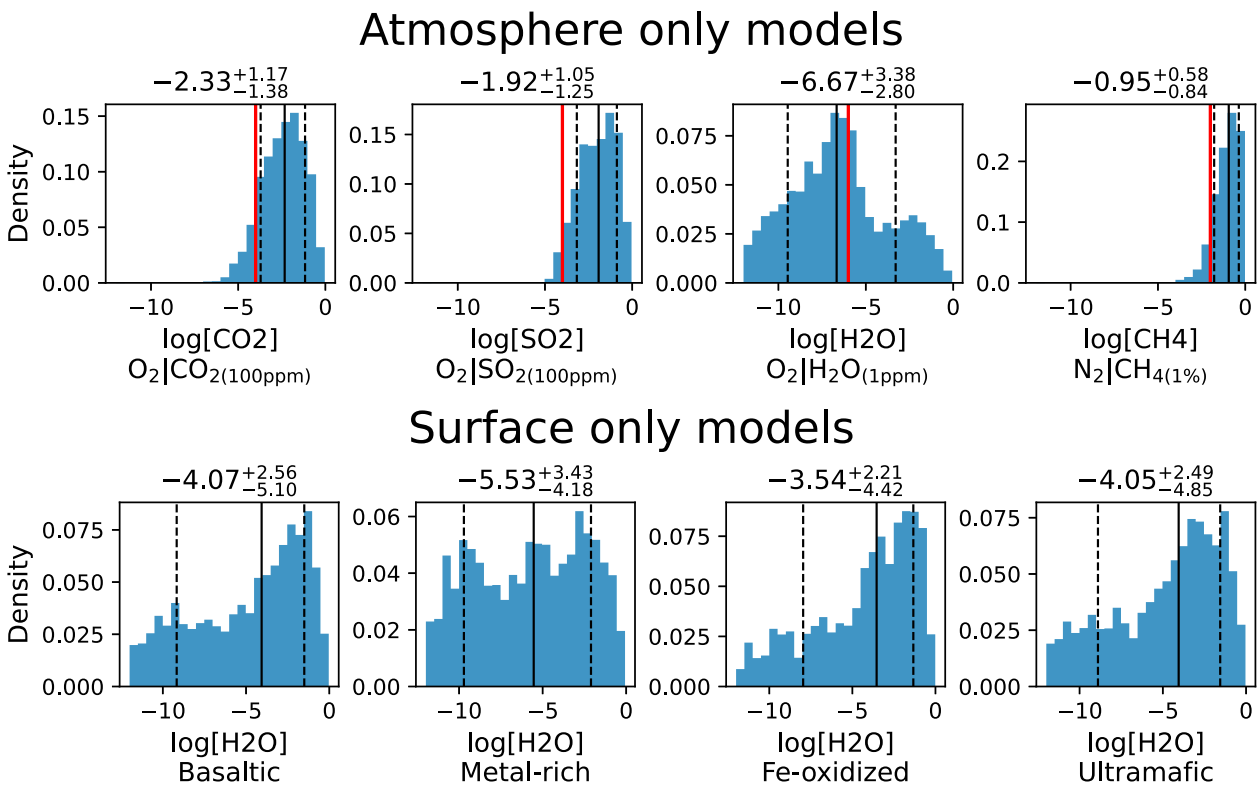

Figure 11. Top: Posterior distributions of the retrieved volume mixing ratios of CO_2, SO_2, H_2O, and CH_4 using atmospheric forward models without a nongray surface. The retrieved mean and 1σ width is given on top of each panel and shown with black solid and dashed vertical lines. The input mixing ratio (the "true" value) is shown as a red vertical line. The atmospheric composition used in the forward model is listed below the horizontal axis label. All atmospheric species are recovered in the retrieval, although the retrieved mixing ratio tends to be somewhat overpredicted for CO_2, SO_2, and CH_4. Bottom: Posterior distributions of the retrieved H_2O volume mixing ratios from the surface-only (no atmosphere) forward models. The surface crust used in the forward model is listed below the horizontal axis label. The basaltic, iron-oxidized, and ultramafic surfaces lead to a relatively pronounced posterior distribution that may erroneously be interpreted as a weak detection of atmospheric H_2O.

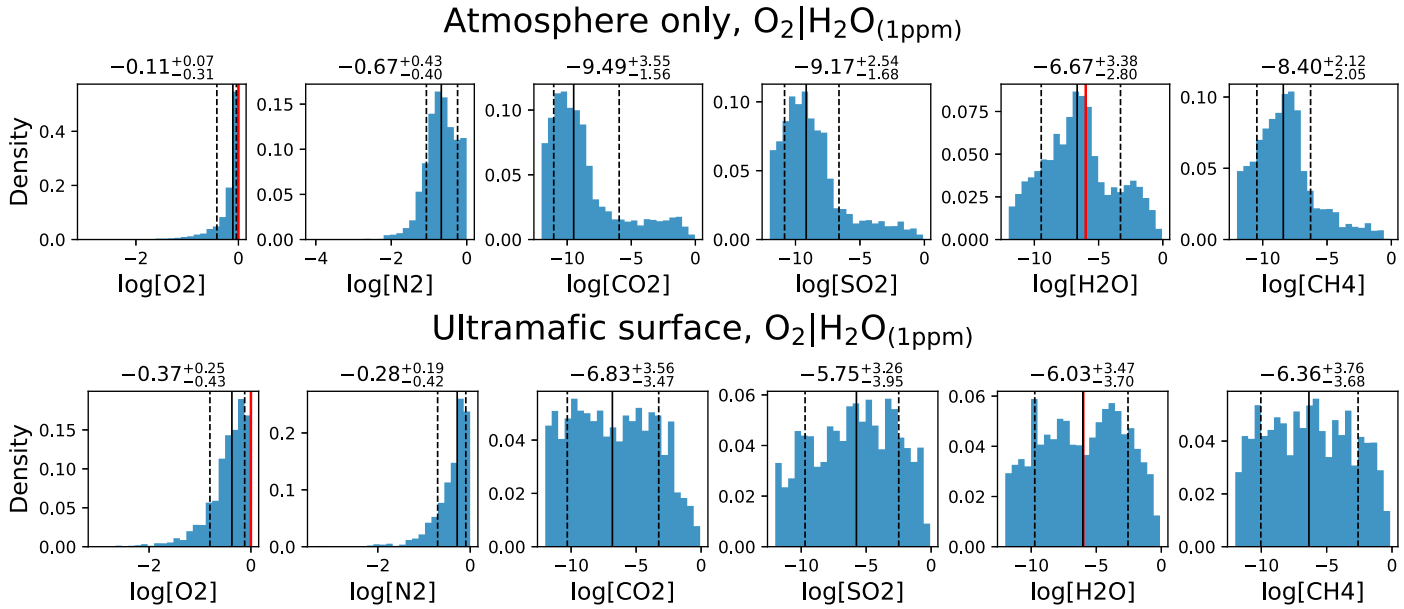

Figure 12. Posterior distributions of volume mixing ratios using the O_2 with 1 ppm H_2O atmosphere as input for the retrieval modeling, once without a nongray surface (top) and once with an ultramafic surface (bottom). The retrieved mean and 1σ width is given on top of each panel and shown with black solid and dashed vertical lines. The input mixing ratio (the "true" value) is shown as a red vertical line. Compared to the atmosphere-only case, adding the ultramafic signal leads to "washed-out" posteriors making it harder to place limits on the mixing ratios. Note that O_2 and N_2 are treated as background gases in the retrieval modeling and are not directly retrieved.

surfaces are distinguishable with five eclipse measurements and the metal-rich and iron-oxidized surfaces are distinguishable with 10 eclipse measurements using NIRSpec. The more reflective surfaces disfavored by Spitzer, the granitoid, feldspathic, and ultramafic crusts, would be distinguishable from most of the other surfaces with 1–4 eclipses using NIRSpec (as seen in Table 3).

Exploring the limit with an atmosphere, we have found that each atmospheric model with discernible features is distinguishable from a surface-only spectrum in under three eclipse observations if the better-suited JWST instrument, MIRI or NIRSpec, is used in each case. The most amenable cases are the atmospheres with SO_2 and CH_4, each requiring only a single eclipse to be constrained using MIRI or NIRSpec, respectively (see Figure 10 and Table 6).

Lastly, we have run a suite of atmospheric retrieval models to determine how JWST may help constrain the atmospheric composition of LHS 3844b with a Bayesian framework. We have also explored whether surface albedo variations could bias the retrieval algorithm or even be mistaken for atmospheric signatures. Using uncertainties based on five eclipse observations, we have been able to detect all four of the included gas absorbers, albeit three of the retrieved mixing ratios are somewhat overpredicted (see Figure 11). Furthermore, we have found that the nongray surface has negligible effect on an atmospheric species' retrieved mixing ratio if the species' spectral signature is sufficiently large. However, the surface contamination of the atmospheric retrieval may lead to a false weak detection of atmospheric H_2O and may also make it harder to disfavor the presence of gas species (see Figure 12 and Tables 1 and 8).

LHS 3844b will be observed during three secondary eclipses with the JWST MIRI/LRS as part of Cycle 1 of the General Observer program. First, these observations will extend our knowledge about the presence and thickness of an atmosphere and also verify the existing Spitzer data point, testing the measured eclipse depth at 4.5 μm and its derived error. Furthermore, according to our predictions these MIRI observations will indicate whether the planet's surface is granitoid in nature, but the allocated time will not suffice to place any constrains on the other potential surface crusts. Also, if present, any significant amounts of H_2O, SO_2, or CH_4 should be detected, provided the planetary atmosphere is thick enough.

4.2. Comparison to Kreidberg et al. (2019) and Implications for Surface and Atmosphere

Our current study expands the LHS 3844b phase curve analysis of Kreidberg et al. (2019; herein after K19) on multiple fronts. First, we have extended the number of tested surface types from ultramafic, feldspathic, basaltic, and granitoid to further include a metal-rich and an iron-oxidized surface. Additionally, while K19 assumed atmospheres made up of O_2, N_2, and CO_2, we have tested a wider range of atmospheres by including CO, CH_4, SO_2, and H_2O as absorbers. Finally, our modeling of LHS 3844b is done using a radiative–convective two-stream radiative transfer code that additionally accounts for energy balance at the planet's solid surface. The analysis of LHS 3844b in K19 was performed against atmosphere models using the techniques described in Morley et al. (2017), which utilize a simplified temperature prescription and do not natively account for a surface.

While our results are in broad agreement with the LHS 3844b phase curve analysis of K19, we also find differences between our and their model predictions that are worth discussion. First, the wavelength-varying secondary eclipse depths and consequently the modeled Spitzer 4.5 μm eclipse depths of the bare-rock models in our work are noticeably lower than predicted by K19 for the same conditions. For example, our predicted eclipse depths for the basaltic and ultramafic models are 79 ppm (1.98σ) and 40 ppm (1σ) lower than predicted by K19. Second, analogously to the bare-rock models, the eclipse depths of the models with (optically) thin atmospheres are also noticeably lower than found by K19 for the same conditions. In contrast, the optically thick atmospheres provide eclipse depths that are larger than in K19. For example, K19's calculated eclipse depths of the O_2 + 1 ppm CO_2 atmospheres are approximately 15–50 ppm (0.4–1.3σ) larger than ours for all surface pressures $\leqslant$ 10 bar. However, K19's models with 100% CO_2 provide an approximately 40–60 ppm (1–1.5σ) shallower eclipse depth throughout almost all surface pressures compared to ours (apart from 1 bar where the eclipse depth is roughly the same).

We believe the discrepancy in the atmosphere models can be at least partially attributed to the sophistication of the atmospheric radiative transfer modeling as we utilize self-consistent radiative–convective equilibrium models, whereas they relied on parameterized temperature profiles that do not take the nongray radiative feedback of gas absorbers into account. Yet, the modeling treatment of the atmosphere can only explain the cases with optically thick atmospheres, as otherwise the atmosphere has a marginal impact on the planetary spectrum. We believe that the discrepancy in the optically thin atmosphere and no-atmosphere models is due to the treatment of the host star radiation. In K19, they scaled the stellar spectrum model to match the measured stellar flux density over the Spitzer 4.5 μm bandpass. However, this absolute flux measurement is not consistent with the parameters of LHS 3844 derived from spectral energy distribution fitting (Kreidberg et al. 2019; Vanderspek et al. 2019). In the present work, we have chosen to follow the literature and use a PHOENIX stellar spectrum for the parameters given in K19 without any additional spectrum scaling, as it remains unclear whether the eclipse modeling should be guided by the single-band photometric measurement (that could be prone to an unknown systematic error) or the parameters derived from the overall stellar fit. Ultimately, if multiple flux measurements of the star cannot be reconciled, the correct modeling procedure remains ambiguous. This highlights the need for accurate stellar flux measurements as any uncertainty in the treatment of the star directly affects the secondary eclipse depth.

A shallower eclipse depth prediction for the bare-rock models compared to K19 directly translates to tighter constraints on the type of surfaces that are possible for this planet when taking the Spitzer data point at face value. We recover their result that a basaltic surface is consistent with Spitzer at 3σ, but our ultramafic model is outside of this confidence interval in contrast to their modeling.

In terms of atmospheric thickness, our new modeling sets the top limit on the surface pressure at $\sim$1 bar down from the previous limit of 10 bar, if an infrared absorbing gas is included at $\geqslant$100 ppm. Their best-fit models ($< 1\sigma$ deviation from observation) that include a nonmarginal amount of near-infrared absorber at $\geqslant$10 ppm require a thin atmosphere with a

surface pressure < 0.1 bar. Interestingly, our models that fall within 1σ allow atmospheres with up to 1 bar surface pressure if CH_4 is included. That is because CH_4 is not only a strong greenhouse gas, thus warming the deep atmosphere and surface, but also CH_4 is weakly absorbing in a spectral window around 4.5 μm, enabling a higher thermal emission from the deep atmosphere in the Spitzer bandpass.

4.3. General Implications for the Characterization of Rocky Exoplanets

Expanding the study of exoplanet atmospheres into the terrestrial planet regime is a major goal in the JWST era. Yet such planets will still be challenging atmospheric characterization targets, even with the improved observing capabilities of the JWST. It is therefore necessary to establish robust strategies for extracting meaningful constraints on terrestrial planet atmospheres. Among such targets, hot rocky planets orbiting bright stars, such as LHS 3844b present the best opportunities for detailed characterization.

A zeroth-order question to answer for rocky exoplanet targets is whether they possess atmospheres at all, and if so, how thick said atmospheres are (Koll et al. 2019; Kreidberg et al. 2019). The next natural follow-up question is to establish the composition of the planet's atmosphere and surface.

In this work, we have demonstrated how to go about addressing questions of surface and atmosphere composition for the thick (or thin) and no-atmosphere cases, as applied to the most amenable rocky target for thermal emission characterization, LHS 3844b. We have quantified the amount of observing time required and the limitations to which types of surfaces and atmospheres can be uniquely distinguished. We have also identified shortcomings in our current retrieval capabilities for measuring the properties of terrestrial planet atmospheres. The approach that we have outlined here can be applied, in principle, to any rocky planet target that presents sufficiently high signal-to-noise thermal emission.

One particularly subtle challenge that we have discussed at length in this work is that of measuring a planet's albedo and its surface temperature. When interpreting the measured planetary emission one should be aware that due to the less-than-unity emissivity of the realistic surfaces, the inferred brightness temperature is substantially lower than the true surface temperature for both NIRSpec and MIRI observations. However, while inferring the planetary albedo (using the techniques described in this paper and in Mansfield et al. 2019) with NIRSpec is relatively accurate, using MIRI the inferred albedo is somewhat underestimated for the higher-albedo surfaces and overestimated for the lower-albedo surfaces.

The Kreidberg et al. (2019) Spitzer phase curve observations and our interpretation in this paper have already provided a phenomenal degree of insight into the properties of LHS 3844b. If the planet possesses an atmosphere at all, it is thinner than the Earth's. Yet this conclusion is based only on photometric measurements in a single bandpass. JWST will, for the first time, open the door to the spectroscopic characterization over a wide wavelength range of LHS 3844b as well as a wealth of other rocky planets, building on previous efforts that started with the spectroscopic analysis of terrestrial exoplanets with HST/WFC3 and Spitzer, like the observations of the TRAPPIST-1 system and GJ 1132b (de Wit et al. 2016, 2018; Mugnai et al. 2021; Swain et al. 2021; Libby-Roberts et al. 2022). Ultimately, efforts like these will provide much deeper insight into the fundamental properties of terrestrial planet surfaces and atmospheres. We look forward to the era of rocky planet characterization that is just beginning.

This study has been conducted as part of the NASA Exoplanets Research Program, grant No. 80NSSC20K0269 (PI: M.Malik). E.M.-R.K. and J.I. acknowledge support from the National Science Foundation under CAREER grant No.1931736 and from the AEThER program, funded by the Alfred P. Sloan Foundation under grant G202114194. M.M. acknowledges support from the NASA Hubble Fellowship grant HST-HF2-51485.001-A awarded by STScI. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This research has made use of the SVO Filter Profile Service (http://svo2.cab.inta-csic.es/theory/fps/) supported from the Spanish MINECO through grant AYA2017-84089.

Software: `astropy` (Price-Whelan et al. 2018), `CUDA` (Nickolls et al. 2008), `HELIOS` (Malik et al. 2017, 2019a, 2019b), `HELIOS-K` (Grimm & Heng 2015; Grimm et al. 2021), `Matplotlib` (Hunter 2007), `NumPy` (van der Walt et al. 2011), `Pandexo` (Batalha et al. 2017), `PLATON` (Zhang et al. 2019, 2020) `PyCUDA` (Klöckner et al. 2012) `SciPy` (Oliphant 2007).

Appendix A
Updated Numerical Iteration in `HELIOS` With Added Surface

For this work we have updated the numerical forward stepping algorithm of `HELIOS` that iterates toward radiative–convective equilibrium, and, in contrast to the previous implementation where the surface temperature was calculated directly from the energy balance across the surface boundary (see Equation (4) in Malik et al. 2019a), we have also included the surface temperature in the numerical iteration.

The original expression for the temperature iteration (see Equation (24) in Malik et al. 2017), required the knowledge of local atmospheric properties, such as the density and heat capacity. Furthermore, the radiative timescale was used (see Equation (27) in Malik et al. 2017), to make the forward stepping independent of local atmospheric inertia, allowing for a faster and more stable convergence toward equilibrium. However, as the steady-state radiative equilibrium merely requires a vanishing net flux divergence across each atmospheric layer, we have found the inclusion and calculation of such a large number of atmospheric properties not necessary. That is why we have simplified the original formalism to the following expressions. Note that we denote the wavelength-integrated "bolometric" flux with $\mathcal{F}$ ($[\mathcal{F}]$= erg s^{-1} cm^{-2}) and the spectral flux with F ($[F]$ = erg s^{-1} cm^{-3}). The change in the temperature of atmospheric layer i, ΔT_i, between successive iteration steps is calculated as

$$\Delta T_i = -f_{\mathrm{pre},i}\,\frac{P_i}{\Delta P_i}\Delta\mathcal{F}_{\mathrm{net},i}, \tag{A1}$$

where P_i is the pressure in the center of layer i, ΔP_i is the pressure difference across layer i, $\Delta\mathcal{F}_{\mathrm{net},i}$ is the net flux difference across layer i, and $f_{\mathrm{pre},i}$ is a dimensional prefactor

given by

$$f_{\mathrm{pre},i} = \frac{\mathrm{K\ s\ cm^2\ erg^{-1}}}{[|(\Delta\mathcal{F}_{\mathrm{net},i})|/(\mathrm{erg\ s^{-1}\ cm^{-2}})]^{0.9}}. \quad \text{(A2)}$$

The goal of this prefactor is first to make the iteration more stable by dampening the impact of $\Delta\mathcal{F}_{\mathrm{net},i}$, thus making the temperature iteration smoother between neighboring layers. Second, $f_{\mathrm{pre},i}$ is used to adapt the temperature step during the iteration. If the temperature is close to equilibrium, $f_{\mathrm{pre},i}$ becomes increasingly smaller. This prevents numerical oscillations around the equilibrium temperature without ever reaching it (details on the exact evolution of $f_{\mathrm{pre},i}$ are given in Malik et al. 2017). Analogous to the atmospheric layer temperatures, the surface temperature starts from the planetary dayside-averaged temperature at the beginning of the run and then advances with each iteration step as

$$\Delta T_{\mathrm{surf}} = -f_{\mathrm{pre,\ surf}} \frac{P_0}{\Delta P_0} \Delta\mathcal{F}_{\mathrm{net,\ surf}}, \quad \text{(A3)}$$

with

$$\Delta\mathcal{F}_{\mathrm{net,surf}} = \mathcal{F}_{\mathrm{net},0} - \mathcal{F}_{\mathrm{intern}}, \quad \text{(A4)}$$

where $\mathcal{F}_{\mathrm{intern}}$ is the internal heat flux. The index "0" denotes the first atmospheric layer or interface from the bottom. The prefactor $f_{\mathrm{pre,\ surf}}$ is calculated with Equation (A2), but using $\Delta\mathcal{F}_{\mathrm{net,surf}}$ in the denominator. The net flux at the surface boundary is given by $\mathcal{F}_{\mathrm{net},0} = \mathcal{F}_{\uparrow,0} - \mathcal{F}_{\downarrow,0}$. One caveat of including the surface in the numerical iteration of the atmosphere is that the surface and the first atmospheric layer sometimes become "stuck," oscillating back and forth between iterations, prone to happen when the near-surface atmosphere is optically thick. We employ a numerical trick to solve that issue. As long as the first atmospheric layer is not in radiative equilibrium, we use $\mathcal{F}_{\mathrm{net},1}$ in place of $\mathcal{F}_{\mathrm{net},0}$ in Equation (A4), which essentially includes the first atmospheric layer in the energy balance of the surface. Once the first atmospheric layer has converged, we switch back to the correct energy balance expression for the surface, as given by Equation (A4).

Finally, the wavelength-varying surface albedo is included in the upward flux from the surface. Namely, the upward pointing spectral flux at the surface is calculated as

$$F_{\uparrow,0,\lambda} = A_{\mathrm{surf},\lambda} F_{\downarrow,0,\lambda} + (1 - A_{\mathrm{surf},\lambda})\pi B_\lambda(T_{\mathrm{surf}}), \quad \text{(A5)}$$

where $A_{\mathrm{surf},\lambda}$ is the wavelength-dependent surface albedo, B_λ is the Planck function, and T_{surf} is the surface temperature.

Appendix B
Role of the Emissivity on the Inferred Albedo

The effect of assuming different emissivity values on the inferred albedo is shown in Figure C6, where we plot three different emissivity values: 1.0, 0.9, and 0.7, as seen in the upper left of each plot. When fitting a blackbody curve to a spectrum, the total area under the curve integrated from zero to infinity, analogous to the total energy emitted, is not conserved. This discrepancy is exacerbated if the assumed emissivity of the blackbody model significantly differs from the one of the realistic surface spectrum. Moreover, the farther the wavelength region for the fitting is separated from the peak of the planetary thermal emission, the larger is the discrepancy in the areas below the two curves. In the MIRI case, if the blackbody fit assumes an emissivity higher than the one used in the physical model the total energy of the blackbody emission is smaller than in the physical model, resulting in a too low blackbody temperature and consequently a too high inferred albedo. This is best visible for the three gray albedo models with $\alpha = 0$, $\alpha = 0.1$, and $\alpha = 0.3$. When the fit for the gray albedo models is conducted with the correct emissivity the inferred albedo matches the real bond albedo.

When using NIRSpec, this emissivity effect is somewhat smaller and also nonmonotonic. This is because in this case the instrument bandpass overlaps with the peak planetary emission making the areas between the blackbody curve and the planetary spectrum more similar. Furthermore, there appears to be a dependence on individual albedo variation. The surfaces showing an inverse trend have an overall higher albedo and a more strongly varying albedo within the NIRSpec bandpass.

Appendix C
Additional Figures

Figure C1 displays temperature-pressure profiles for all relevant models of atmospheric constituents at surface pressures of 10^{-3} bars (top) and 10 bars (bottom). Figure C2 is analogous to Figure 3, but for a basaltic rather than a metal-rich surface. Figure C3 displays secondary eclipse spectra for all possible model pairs for both the MIRI LRS and NIRSpec G395H instruments over their entire wavelength range. These spectra help visualize the surface detectability values listed in Table 3. The spectra used for distinguishing surfaces with optimal binning of MIRI LRS (see Table 4) are shown in Figure C4. Figure C5 shows the bond albedo versus the inferred albedo obtained for each modeled surface, including albedo inferred from the full spectrum (emission and reflection) as well as the albedo inferred using only the emission from the planet. Similarly, Figure C6 shows the bond albedo versus inferred albedo for each surface using the full spectrum, but this time the emissivity assumed is varied. Figure C7 shows the eclipse spectra of the atmospheric models vs. surface-only models, where here the full wavelength range of each instrument is covered. Figure C8 shows the posterior distributions of all atmosphere-only models retrieved on by `PLATON`, while Figure C9 shows posterior distributions for all surface-only models retrieved on by `PLATON`.

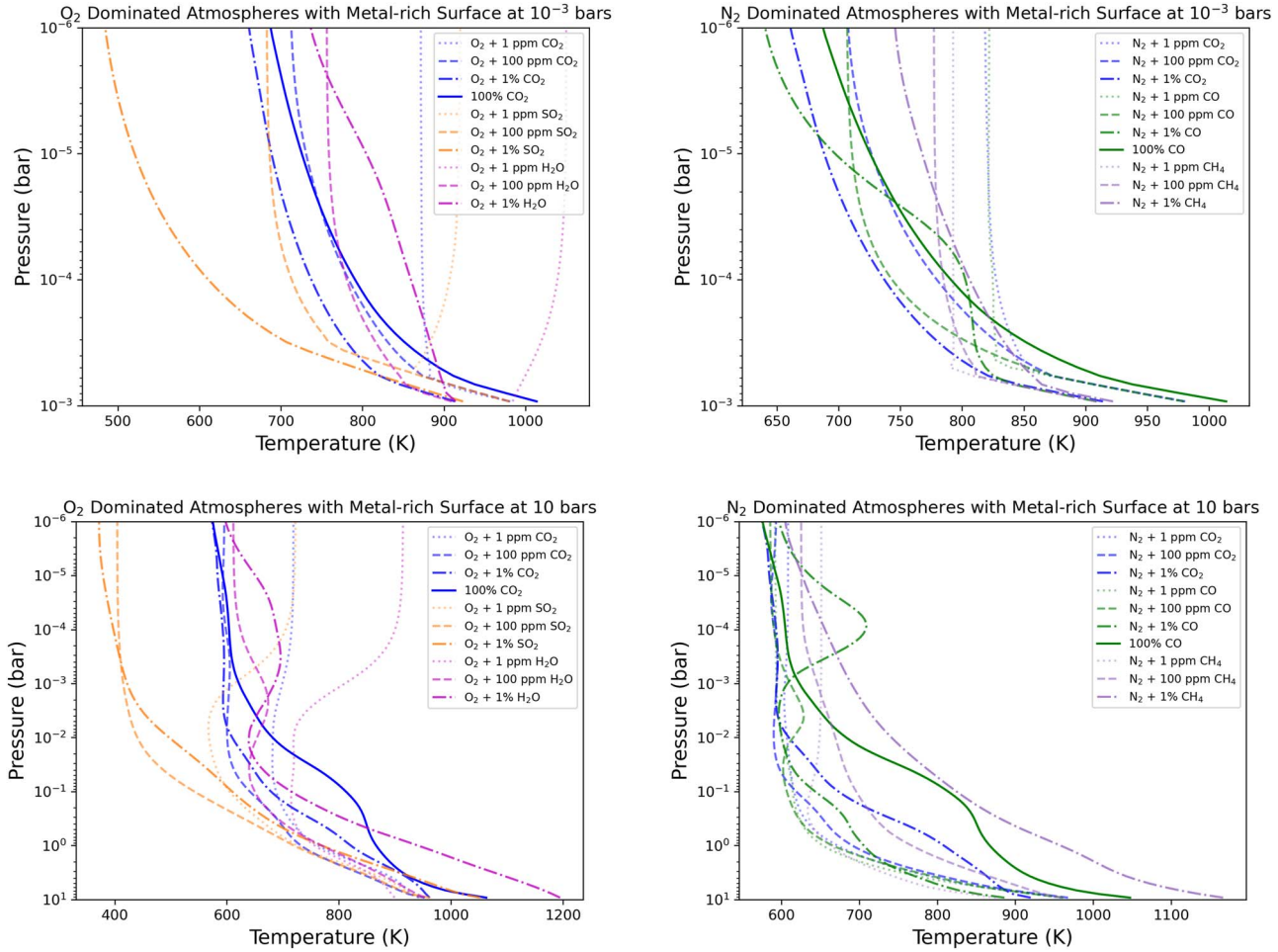

Figure C1. Temperature–pressure (T–P) profiles for oxidizing/O_2-dominated atmosphere models (left) and reducing/N_2-dominated atmosphere models (right) assuming a surface pressure of 10^{-3} bars (top) and 10 bars (bottom).

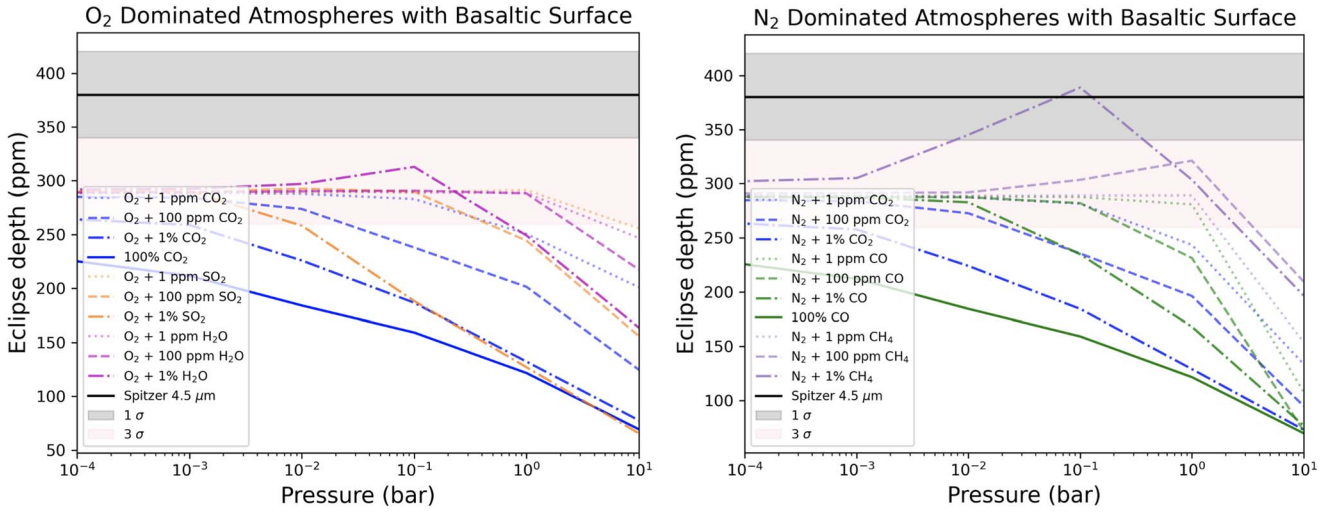

Figure C2. Predicted Spitzer 4.5 μm eclipse depth for the explored atmospheric models in combination with a basaltic surface as a function of the surface pressure compared to the Spitzer measurement (black horizontal line). Oxidizing/O_2-dominated atmospheres are on the left and reducing/N_2-dominated atmospheres on the right. The gray shaded area corresponds to a 1σ uncertainty, and the pink shaded area to a 3σ uncertainty in the negative direction from the observed value.

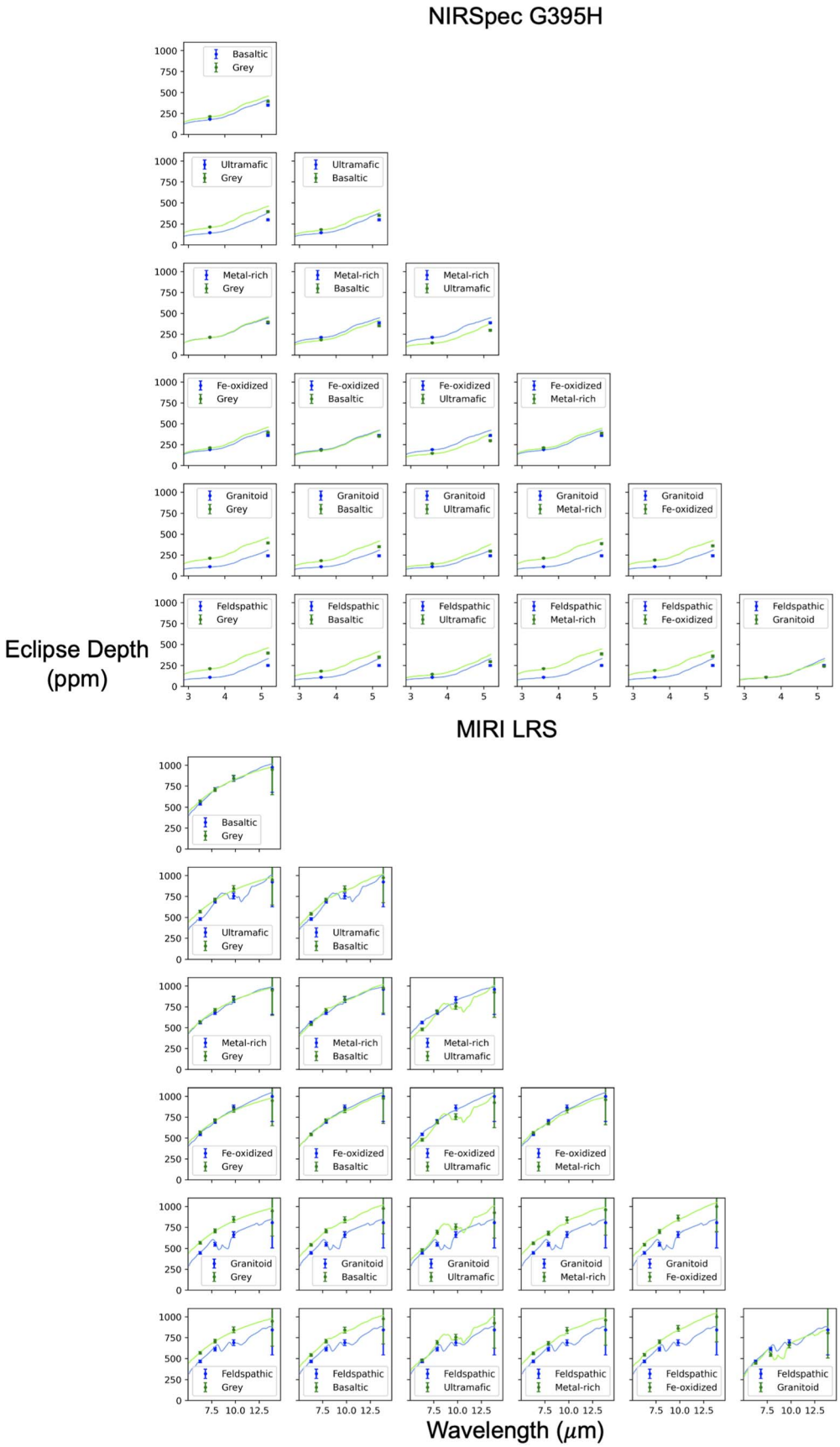

Figure C3. Secondary eclipse spectra of surface-only model pairs over the full range of the NIRSpec G395H (top) and MIRI LRS (bottom) bandpasses with overlaid mock data points at $R = 3$. The error bars correspond to five secondary eclipse observations. The model spectra are downsampled to $R = 100$ for clarity.

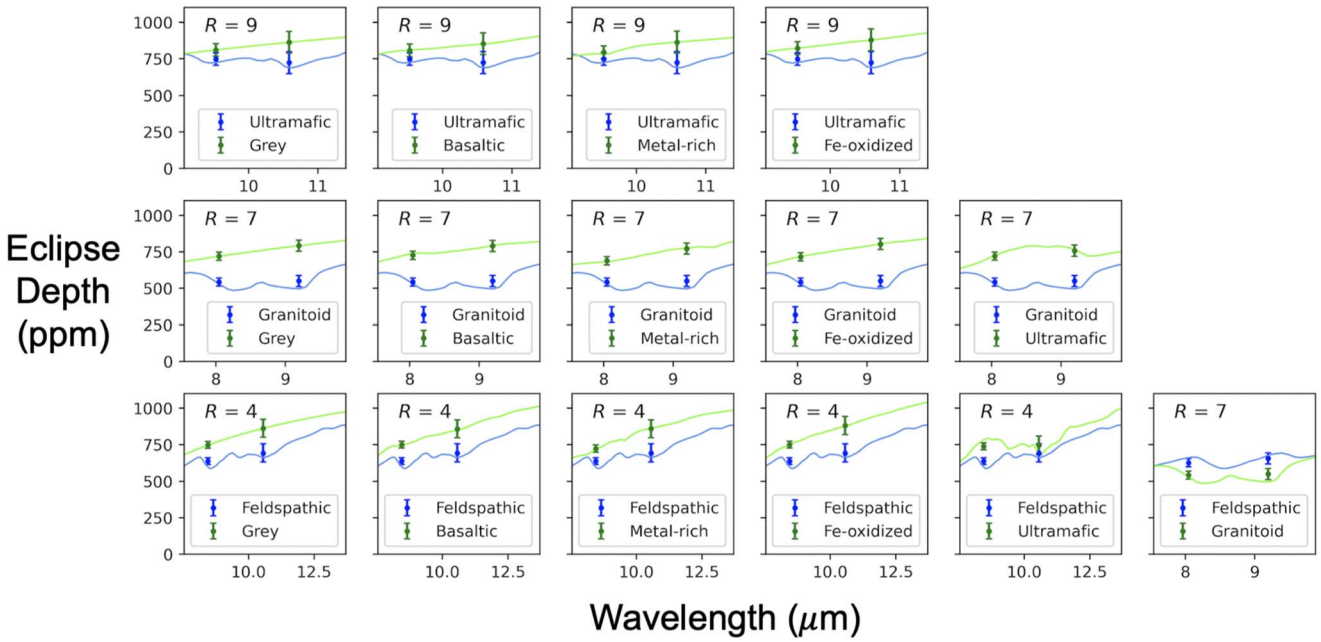

Figure C4. Secondary eclipse spectra of surface-only model pairs over the range of selected surface features within the MIRI LRS bandpass with overlaid mock data points. Each case uses a resolution (shown in the top left corner) to optimally match the spectral features. The error bars correspond to five secondary eclipse observations. The model spectra are downsampled to $R = 100$ for clarity.

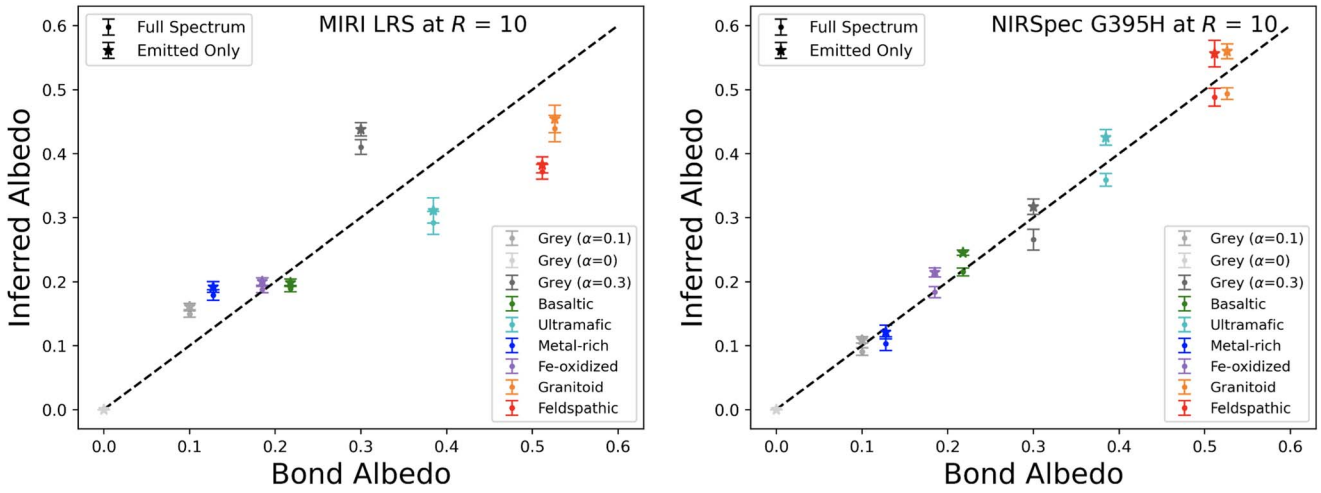

Figure C5. Left: Bond albedo vs. inferred albedo for simulated MIRI LRS data at $R = 10$. The black dotted line represents the location where the inferred and Bond albedos are equal. The inferred albedo is determined using once the full spectrum (i.e., both planetary emission and reflected light contributions) and once the emitted radiation only. The nominal case uses the full spectrum. Right: analogous to the left panel but using simulated NIRSpec G395H data at $R = 10$.

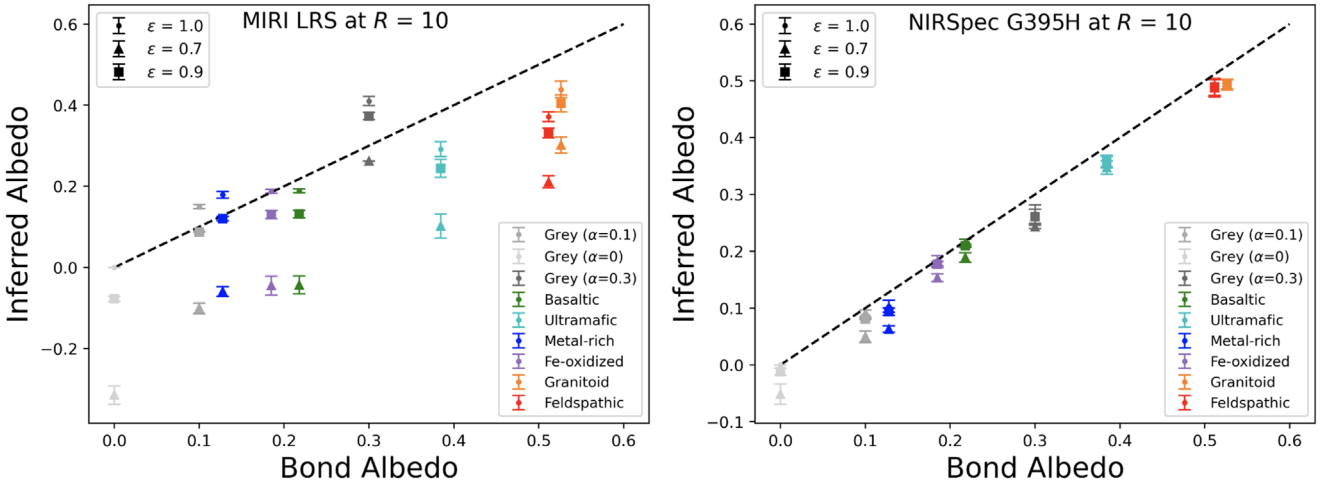

Figure C6. Left: Bond albedo vs. inferred albedo for simulated MIRI LRS data at $R = 10$. The black dotted line represents the location where the inferred and Bond albedos are equal. Three values of emissivity, 1.0, 0.9, and 0.7, are assumed for the planetary emission when inferring the temperature and the albedo. The nominal value used is 1.0. Right: analogous to the left panel but using simulated NIRSpec G395H data at $R = 10$.

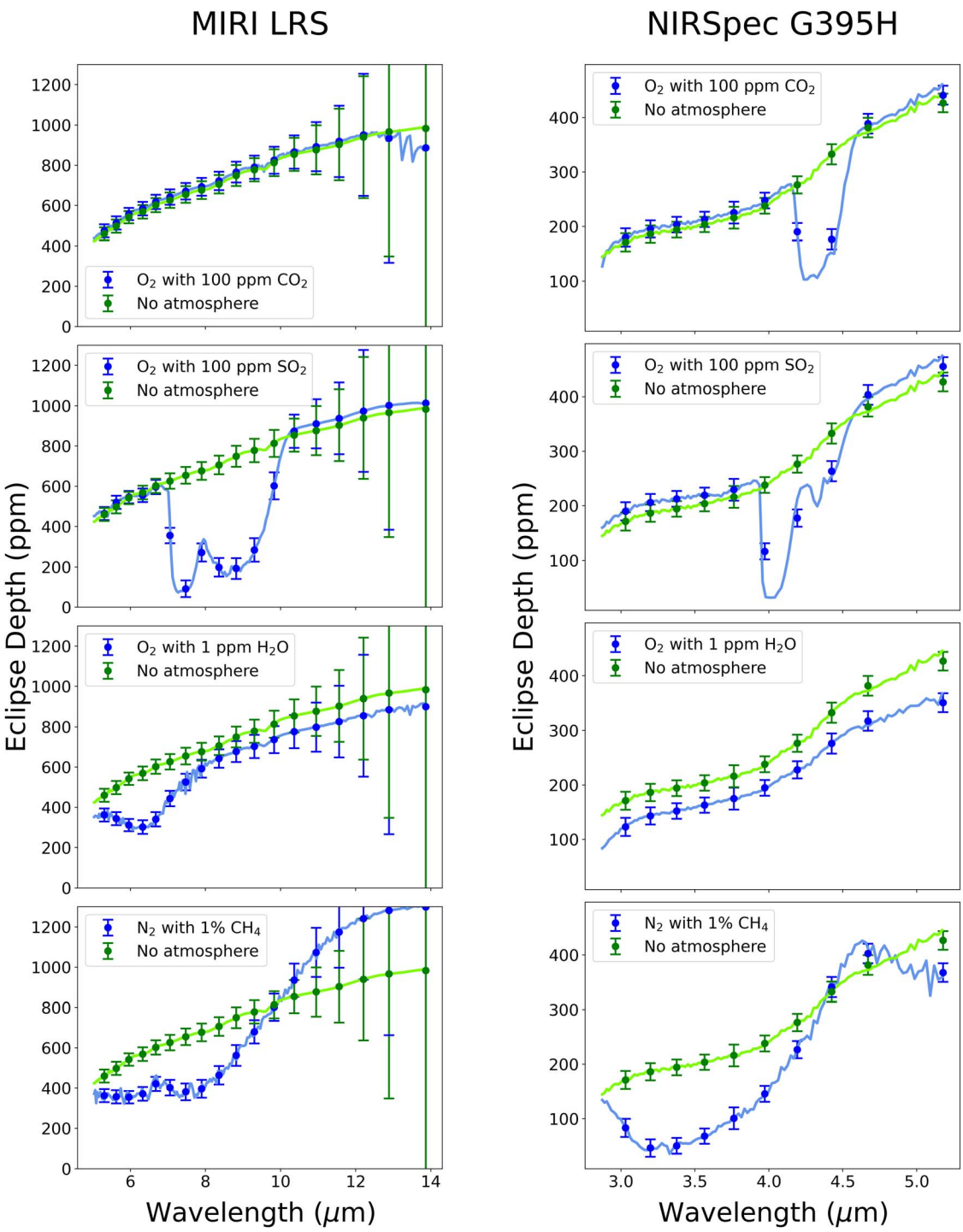

Figure C7. Secondary eclipse spectra of atmospheric models vs. surface-only models, overlaid with MIRI LRS (left) and NIRSpec G395H (right) simulated data points at $R = 10$ over the whole instrument bandpass. The metal-rich surface is used for this test. The error bars show the uncertainties for five eclipse observations. The model spectra are downsampled from the native resolution to $R = 100$ for clarity.

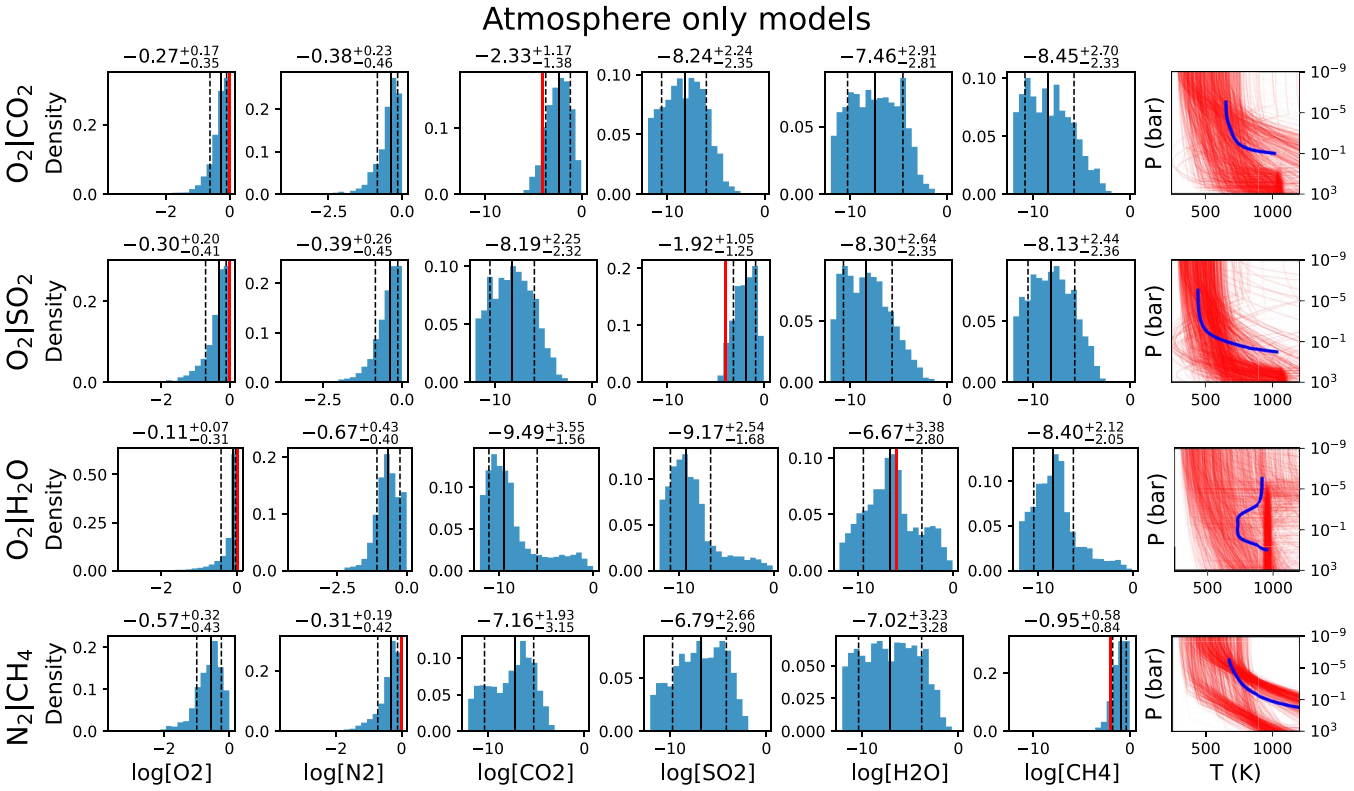

Figure C8. Posterior distributions of retrieved volume mixing ratios using atmospheric forward models without a nongray surface. The retrieved mean and 1σ width are given on top of each panel and shown with black solid and dashed vertical lines. The input mixing ratio (the "true" value) is shown as a red vertical line. The corresponding T–P profiles are shown on the right with the input profile (from the forward model) in blue, and the profiles realizations from the retrieval modeling are shown as red lines.

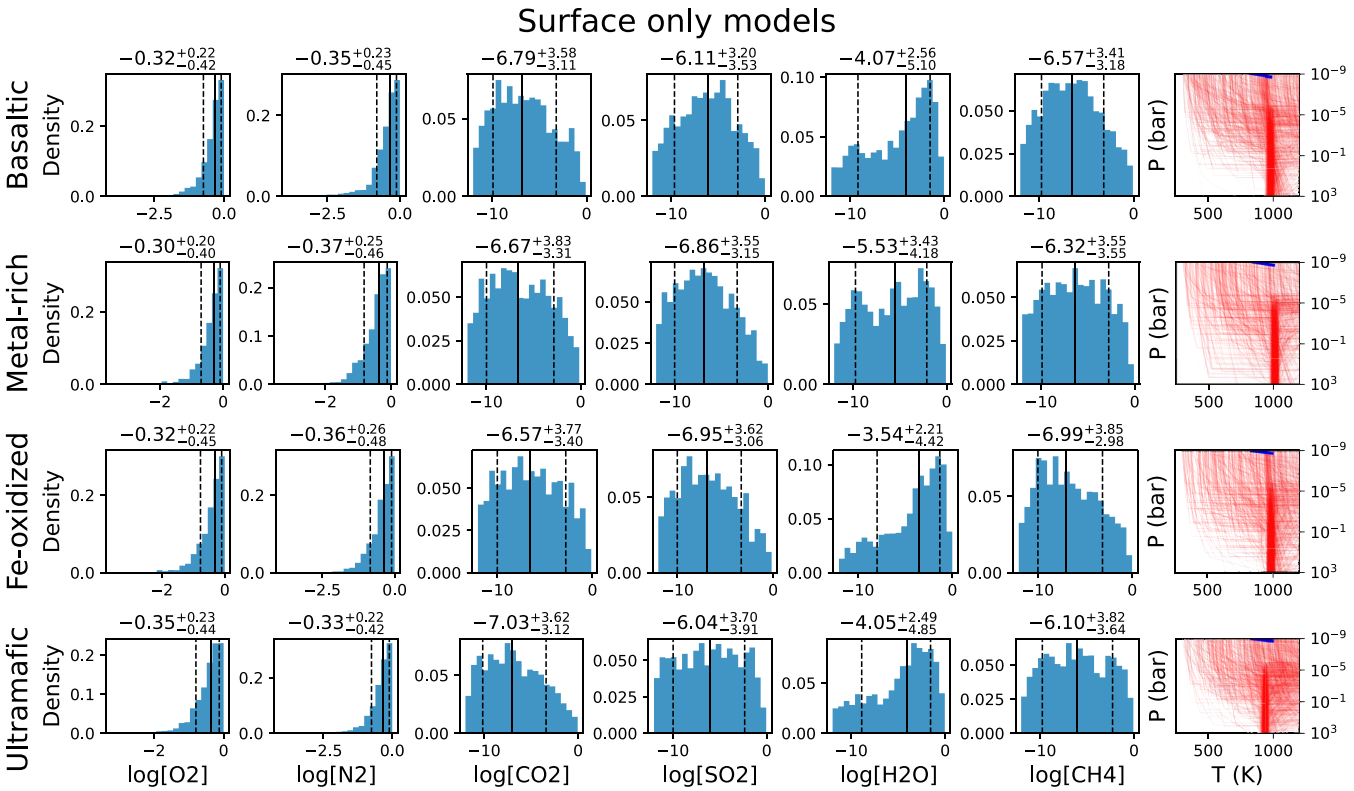

Figure C9. Same as Figure C8 but showing here the retrieval results using the surface-only (no atmosphere) forward models with different surface crusts.

ORCID iDs

Emily A. Whittaker https://orcid.org/0000-0002-1518-7475
Matej Malik https://orcid.org/0000-0002-2110-6694
Jegug Ih https://orcid.org/0000-0003-2775-653X
Eliza M.-R. Kempton https://orcid.org/0000-0002-1337-9051
Megan Mansfield https://orcid.org/0000-0003-4241-7413
Jacob L. Bean https://orcid.org/0000-0003-4733-6532
Edwin S. Kite https://orcid.org/0000-0002-1426-1186
Daniel D. B. Koll https://orcid.org/0000-0002-9076-6901
Renyu Hu https://orcid.org/0000-0003-2215-8485

References

Airapetian, V. S., Glocer, A., Gronoff, G., Hébrard, E., & Danchi, W. 2016, NatGe, 9, 452
Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, PASP, 129, 064501
Bishop, J. L. B., James, F. I., & Moersch, J. E. 2019, Remote Compositional Analysis: Techniques for Understanding Spectroscopy, Mineralogy, and Geochemistry of Planetary Surfaces, Vol. 1 (Cambridge: Cambridge Univ. Press), 18
Blewett, D. T., Lucey, P. G., Hawke, B., Ling, G., & Robinson, M. S. 1997, Icar, 129, 217
Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436
Carli, C., Serventi, G., & Sgavetti, M. 2015, GSLSP, 401, 139
Chameides, W., & Walker, J. 1981, OrLi, 11, 291
Cox, A. N. 2000, Allen's Astrophysical Quantities (New York: AIP Press)
de Wit, J., Wakeford, H. R., Gillon, M., et al. 2016, Natur, 537, 69
de Wit, J., Wakeford, H. R., Lewis, N. K., et al. 2018, NatAs, 2, 214
Diamond-Lowe, H., Charbonneau, D., Malik, M., Kempton, E. M.-R., & Beletsky, Y. 2020, AJ, 160, 188
Elkins-Tanton, L. T., & Seager, S. 2008, ApJ, 685, 1237
Gaillard, F., & Scaillet, B. 2014, E&PSL, 403, 307
Gao, P., Hu, R., Robinson, T. D., Li, C., & Yung, Y. L. 2015, ApJ, 806, 249
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Natur, 542, 456
Gordon, I., Rothman, L., Hill, C., et al. 2017, JQSRT, 203, 3
Grimm, S. L., & Heng, K. 2015, ApJ, 808, 182
Grimm, S. L., Malik, M., Kitzmann, D., et al. 2021, ApJS, 253, 30
Hansen, B. M. S. 2008, ApJS, 179, 484
Hu, R., Ehlmann, B. L., & Seager, S. 2012, ApJ, 752, 7
Hunten, D. M., Pepin, R. O., & Walker, J. C. G. 1987, Icar, 69, 532
Hunter, J. D. 2007, CSE, 9, 90
Husser, T.-O., von Berg, S. W., Dreizler, S., et al. 2013, A&A, 553, A6
Ih, J., & Kempton, E. M. R. 2021, AJ, 162, 237
Izenberg, N. R., Klima, R. L., Murchie, S. L., et al. 2014, Icar, 228, 364
Kane, S. R., Kopparapu, R. K., & Domagal-Goldman, S. D. 2014, ApJL, 794, L5
Kempton, E. M.-R., Bean, J. L., Louie, D. R., et al. 2018, PASP, 130, 114401
Kite, E. S., & Barnett, M. N. 2020, PNAS, 117, 18264
Kite, E. S., & Schaefer, L. 2021, ApJL, 909, L22
Klöckner, A., Pinto, N., Lee, Y., et al. 2012, ParC, 38, 157
Koll, D. D. B. 2022, ApJ, 924, 134
Koll, D. D. B., Malik, M., Mansfield, M., et al. 2019, ApJ, 886, 140
Kreidberg, L., Koll, D. D. B., Morley, C., et al. 2019, Natur, 573, 87
Lammer, H., Sproß, L., Grenfell, J. L., et al. 2019, AsBio, 19, 927
Li, G., Gordon, I. E., Rothman, L. S., et al. 2015, ApJS, 216, 15
Libby-Roberts, J. E., Berta-Thompson, Z. K., Diamond-Lowe, H., et al. 2022, AJ, 164, 59
Liggins, P., Jordan, S., Rimmer, P. B., & Shorttle, O. 2022, JGRE, 127, e07123
Luger, R., & Barnes, R. 2015, AsBio, 15, 119
Lupu, R. E., Zahnle, K., Marley, M. S., et al. 2014, ApJ, 784, 27
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, AJ, 153, 56
Malik, M., Kempton, E. M.-R., Koll, D. D. B., et al. 2019a, ApJ, 886, 142
Malik, M., Kitzmann, D., Mendonça, J. M., et al. 2019b, AJ, 157, 170
Mansfield, M., Kite, E. S., Hu, R., et al. 2019, ApJ, 886, 141
Matsuo, T., Greene, T. P., Johnson, R. R., et al. 2019, PASP, 131, 124502
McCord, T., Hansen, G., Buratti, B., et al. 2006, P&SS, 54, 1524
McCord, T. B., & Westphal, J. A. 1971, ApJ, 168, 141
Mikhail, S., & Sverjensky, D. A. 2014, NatGe, 7, 816
Morley, C. V., Kreidberg, L., Rustamkulov, Z., Robinson, T., & Fortney, J. J. 2017, ApJ, 850, 121
Mugnai, L. V., Modirrousta-Galian, D., Edwards, B., et al. 2021, AJ, 161, 284
Nickolls, J., Buck, I., Garland, M., & Skadron, K. 2008, Queue, 6, 40
Nittler, L. R., Starr, R. D., Weider, S. Z., et al. 2011, Sci, 333, 1847
Oliphant, T. E. 2007, CSE, 9, 10
Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105
Owen, J. E., & Wu, Y. 2017, ApJ, 847, 29
Pieters, C. M. 1978, LPSC, 3, 2825
Pieters, C. M. 1986, RvGeo, 24, 557
Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, MNRAS, 480, 2597
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Richard, C., Gordon, I. E., Rothman, L. S., et al. 2012, J. Quant. Spec. Radiat. Transf., 113, 1276
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, J. Quant. Spec. Radiat. Transf., 111, 2139
Schaefer, L., Lodders, K., & Fegley, B. 2012, ApJ, 755, 41
Schaefer, L. F., & Bruce, J. 2011, ApJ, 729, 6
Schlawin, E., Leisenring, J., McElwain, M. W., et al. 2021, AJ, 161, 115
Schlawin, E., Leisenring, J., Misselt, K., et al. 2020, AJ, 160, 231
Schlichting, H. E., Sari, R., & Yalinewich, A. 2015, Icar, 247, 81
Showman, A. P., Wordsworth, R. D., Merlis, T. M., & Kaspi, Y. 2013, in Atmospheric Circulation of Terrestrial Exoplanets, ed. S. J. Mackwell et al. (Tucson, AZ: Univ. of Arizona Press), 277
Sneep, M., & Ubachs, W. 2005, J. Quant. Spec. Radiat. Transf., 92, 293
Swain, M. R., Estrela, R., Roudier, G. M., et al. 2021, AJ, 161, 213
Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2020, JQSRT, 255, 107228
Thalman, R. J., Zarzana, K., Tolbert, M., & Volkamer, R. 2014, JQSRT, 147, 171
Underwood, D. S., Tennyson, J., Yurchenko, S. N., et al. 2016, MNRAS, 459, 3890
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Vanderspek, R., Huang, C. X., Vanderburg, A., et al. 2019, ApJ, 871, L24
Wagner, W., & Kretzschmar, H.-J. 2008, International Steam Tables—Properties of Water and Steam Based on the Industrial Formulation IAPWS-IF97 (Berlin: Springer)
Wordsworth, R., & Pierrehumbert, R. 2014, ApJ, 785, L20
Wordsworth, R. D., Schaefer, L. K., & Fischer, R. A. 2018, AJ, 155, 195
Yurchenko, S. N., Amundsen, D. S., Tennyson, J., & Waldmann, I. P. 2017, A&A, 605, A95
Zahnle, K. J., & Catling, D. C. 2017, ApJ, 843, 122
Zahnle, K. J., Lupu, R., Catling, D. C., & Wogan, N. 2020, PSJ, 1, 11
Zhang, M., Chachan, Y., Kempton, E. M. R., & Knutson, H. A. 2019, PASP, 131, 034501
Zhang, M., Chachan, Y., Kempton, E. M. R., Knutson, H. A., & Chang, W. H. 2020, ApJ, 899, 27
Zhang, Z., Zhou, Y., Rackham, B. V., & Apai, D. 2018, AJ, 156, 178