The Astronomical Journal, 170:205 (24pp), 2025 October https://doi.org/10.3847/1538-3881/adf198

Additional JWST/NIRSpec Transits of the Rocky M Dwarf Exoplanet GJ 1132 b Reveal a Featureless Spectrum

Katherine A. Bennett[1], Ryan J. MacDonald[2,13], Sarah Peacock[3,4,5], Junellie Perez[1,5], E. M. May[5,6], Sarah E. Moran[4,13], Lili Alderson[7,8], Jacob Lustig-Yaeger[5,6], Hannah R. Wakeford[8], David K. Sing[1,9], Kevin B. Stevenson[5,6], Natasha E. Batalha[10], Mercedes López-Morales[11], Munazza K. Alam[11], Joshua D. Lothringer[11], Guangwei Fu[9], James Kirk[12], Jeff A. Valenti[11], L. C. Mayorga[5,6], and Kristin S. Sotzen[5,6]
[1] Department of Earth & Planetary Sciences, Johns Hopkins University, Baltimore, MD 21218, USA; kbenne50@jhu.edu
[2] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
[3] University of Maryland, Baltimore County, MD 21250, USA
[4] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[5] Consortium on Habitability and Atmospheres of M-dwarf Planets (CHAMPs), Laurel, MD, USA
[6] Johns Hopkins Applied Physics Laboratory, Laurel, MD 20723, USA
[7] Department of Astronomy, Cornell University, Ithaca, NY 14853, USA
[8] School of Physics, HH Wills Physics Laboratory, University of Bristol, Bristol, BS8 1TL, UK
[9] Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[10] NASA Ames Research Center, Moffett Field, CA 94035, USA
[11] Space Telescope Science Institute, Baltimore, MD 21218, USA
[12] Department of Physics, Imperial College London, SW7 2AZ, London, UK
Received 2025 May 1; revised 2025 June 18; accepted 2025 July 16; published 2025 September 9

Abstract

As an archetypal M-dwarf rocky exoplanet, GJ 1132 b has a varied history of atmospheric measurements. At 1.13 $R_\oplus$, 1.66 $M_\oplus$, and 580 K, it orbits a bright, slowly rotating M dwarf in a 1.6 day period, making it a prime target for characterization. In this study, we combine two JWST NIRSpec/G395H transits previously reported by E. M. May et al. (2023) with two new NIRSpec/G395M transits to constrain the presence of an atmosphere. This marks the first time the G395H and G395M modes have been combined for a single target, and we report no difference in the quality of data between the two modes. For rocky M-dwarf studies, G395H may still be preferred if stacking transits to utilize the high-resolution flux-calibrated stellar spectra and assess evolving stellar heterogeneity. GJ 1132 b's coadded transmission spectrum is best fit with a flat line. A thin steam atmosphere is also consistent with the data, but this interpretation is driven almost entirely by the first transit, which suggests an increase in cool spot coverage-fraction derived from the flux-calibrated stellar spectra. This demonstrates the importance of always considering stellar heterogeneity evolution in multivisit transits, and also the importance of a "leave-one-transit-out" approach in modeling efforts of coadded transits. We combine these results with MIRI/LRS emission data to show that together, transmission and emission are consistent with only the thinnest of atmospheres. Given GJ 1132 b's age and distance from the star, a thin atmosphere is not likely stable. Therefore, the simplest explanation is that GJ 1132 b is indeed a bare rock.

Unified Astronomy Thesaurus concepts: Exoplanet astronomy (486); Exoplanet atmospheres (487); Extrasolar rocky planets (511); M dwarf stars (982); Exoplanet atmospheric composition (2021); Transmission spectroscopy (2133)

Materials only available in the online version of record: figure set

1. Introduction

Understanding whether rocky planets orbiting M dwarfs typically host atmospheres continues to be one of the most pressing questions in exoplanetary astronomy today. As the first step toward understanding how many nearby planets may be habitable, it is a critical question that nonetheless is proving a massive challenge to answer. Broadly, the community seeks to understand whether the cosmic shoreline, the hypothetical line dividing planets with and without atmospheres (K. J. Zahnle & D. C. Catling 2017), exists outside our own solar system. The shoreline is hypothesized to be due to some combination of planetary size and stellar irradiation. Planet size is important because larger planets have a higher escape velocity and thus can retain atmospheric particles more easily, whereas stellar irradiation matters because planets farther from their host stars experience lower rates of high-energy radiation over their lifetimes and thus do not undergo complete atmospheric erosion. The shoreline may be different for close-in M dwarfs, however, because they experience higher levels of irradiation, when compared to Earth, Venus, and Mars.

M dwarfs live and die slowly, meaning they experience a much more prolonged period of rapid rotation and high activity relative to other stars as they contract onto the main sequence (e.g., T. Preibisch & E. D. Feigelson 2005; E. L. Shkolnik & T. S. Barman 2014) and evolve (e.g., N. Pizzolato et al. 2003; E. K. Pass et al. 2024). Additionally (and relatedly), M dwarfs experience higher rates of stellar flares (e.g., J. R. A. Davenport 2016; M. Audard et al. 2000).

[13] NHFP Sagan fellow.

Both of these factors may contribute to atmospheric mass loss and possibly complete atmospheric erosion early on in the lifetime of close-in M-dwarf rocky exoplanets (e.g., M. L. Khodachenko et al. 2007; F. Tian 2009; J. E. Owen & A. P. Jackson 2012; S. Peacock et al. 2019; G. Van Looveren et al. 2024). There are still significant unknowns around this question due to uncertainties in high-energy stellar spectral energy distributions (K. France et al. 2016), especially the ratios between the extreme-UV (XUV), far-UV, and near-UV fluxes, which affect atmospheric chemistry (A. Segura et al. 2005; S. Rugheimer et al. 2015), precise flare rates (K. France et al. 2020) and broadband flare energy distributions (C. E. Brasseur et al. 2023; R. R. Paudel et al. 2024; K. Burton et al. 2025), the mechanisms of atmospheric loss (nonthermal versus thermal processes; H. Lammer et al. 2008; J. E. Owen 2019), and the variety of possible interior and atmospheric compositions (as well as outgassing mechanisms) dictated by initial planet formation (L. T. Elkins-Tanton & S. Seager 2008; E. S. Kite et al. 2009).

Nonetheless, M dwarfs offer the best opportunities right now to search for atmospheres on rocky exoplanets due to the optimally large planet-to-star radius ratio for planets around these small stars. This has led the community on a massive effort to detect atmospheres from the ground and space. With JWST, the community has already looked at ∼20 targets <2 $R_\oplus$ in transmission, emission, or as a phase curve: e.g., LHS 475b (J. Lustig-Yaeger et al. 2023), GJ 486 b (S. E. Moran et al. 2023; M. Weiner Mansfield et al. 2024), TRAPPIST-1b (T. P. Greene et al. 2023; E. Ducrot et al. 2024), TRAPPIST-1 c (S. Zieba et al. 2023; M. Radica et al. 2025), GJ 1132 b (E. M. May et al. 2023; Q. Xue et al. 2024), GJ 341 b (J. Kirk et al. 2024), GJ 367 b (M. Zhang et al. 2024), TOI-836 b (L. Alderson et al. 2024), 55 Cnc e (R. Hu et al. 2024), LHS 1140 b (C. Cadieux et al. 2024; M. Damiano et al. 2024), L98-59 d (A. Gressier et al. 2024), L98-59 c (N. Scarsdale et al. 2024), L168-9 b (M. K. Alam et al. 2025), L98-59 b (A. Bello-Arufe et al. 2025), TOI-776 b (L. Alderson et al. 2025), LHS 1478 b (P. C. August et al. 2025), LTT 1445A b (P. Wachiraphan et al. 2025), TOI-1685b (R. Luque et al. 2025), TOI-1468 b (E. A. Meier Valdés et al. 2025), and GJ 357 b (J. Adams Redai et al. 2025; J. Taylor et al. 2025). Nearly all of these have been inconclusive or hinted at no atmosphere. Only two planets—L98-59 d (A. Banerjee et al. 2024; A. Gressier et al. 2024) and 55 Cnc e (R. Hu et al. 2024)—have thus far hinted at signs of secondary atmospheres, but these findings have yet to be confirmed.

The focus of this study is the M-dwarf rocky exoplanet GJ 1132 b. Discovered in 2015 by Z. K. Berta-Thompson et al. (2015), GJ 1132 b is a 1.13 ± 0.056 $R_\oplus$ (J. A. Dittmann et al. 2017), 1.66 ± 0.23 $M_\oplus$ (X. Bonfils et al. 2018) planet in a 1.6 day orbital period around a 0.21 $R_\odot$, 0.18 $M_\odot$ 3270 K M dwarf. Assuming $A_b = 0$, the planet has $T_{\rm eq} \approx 580$ K. The planet's mass and radius give it a bulk density of 6.3 ± 1.3 g cm^{-3}, consistent with an Earth-like composition. Its star has a slow rotation period of 122 days (R. Cloutier et al. 2017; X. Bonfils et al. 2018) and is thought to be at least 5 Gyr old based on galactic kinematics, lack of Hα emission, and comparison with Barnard's Star and Proxima Centauri (Z. K. Berta-Thompson et al. 2015). This suggests the star is currently not highly active, though it is important to note that GJ 1132 b likely still experienced prolonged high stellar activity in its past, based on the stellar spindown rates with age for M dwarfs (e.g., E. R. Newton et al. 2016; E. K. Pass et al. 2022, 2024). From the estimated mass-dependent timescale of the saturated regime of XUV stellar radiation from E. K. Pass et al. (2024; see their Equation (2)), GJ 1132 could have remained in a heightened activity state for as long as ∼3 Gyr. Additionally, even older M dwarfs have well-documented frequent flare activity (e.g., K. Vida et al. 2017).

GJ 1132 b has two possible siblings: GJ 1132 c, a 2.64 ± 0.44 $M_\oplus$ minimum mass candidate at an 8.93 day orbital period (placing it in the habitable zone of GJ 1132), and another tentative planet with minimum mass $8.4^{+1.7}_{-2.5}$ $M_\oplus$ with a 177 day orbital period (X. Bonfils et al. 2018). However, this latter signal is just as likely to be due to stellar contamination.

The GJ 1132 system is just 12.0 pc away (Z. K. Berta-Thompson et al. 2015), and was the nearest transiting rocky exoplanet detected at the time of discovery in 2015, leaving it as one of the more promising targets for atmospheric characterization.

Indeed, GJ 1132 b has a long history of atmospheric detection claims and refutations. The first atmospheric investigation was by J. Southworth et al. (2017), who observed nine transits in the *griz* and *JHK* bandpasses using the MPG 2.2 m telescope at ESO La Silla in Chile. They claimed an atmospheric detection based on increased transit depths measured in the z and K bands, which they found hinted at H_2O or CH_4. Follow-up by H. Diamond-Lowe et al. (2018), however, found a featureless spectrum from 0.7 to 1 μm (covering the z band) using five transits with the Magellan II Telescope at Las Campanas Observatory, consistent with a high mean molecular weight atmosphere or bare rock.

The story continued with space-based observations of GJ 1132 b. W. C. Waalkes et al. (2019) reported a nondetection of Lyα using the Hubble Space Telescope (HST) STIS/G140M and placed a 2σ upper limit on the effective size of an exosphere at 7.3 R_P, which translates to an upper limit on the hydrogen mass-loss rate of $0.8–8 \times 10^8$ g s^{-1} (or 465–4650 Earth atmospheric masses per Gyr). Meanwhile, M. R. Swain et al. (2021) found evidence for an H_2-dominated atmosphere with HCN and CH_4 using five transits with HST WFC3/G141. They postulated this atmosphere originated via mantle H_2 degassing following the loss of a primary atmosphere. However, L. V. Mugnai et al. (2021) and J. E. Libby-Roberts et al. (2022) both analyzed the same data set and found a featureless spectrum, suggesting the M. R. Swain et al. (2021) results were due to differences in data reduction. Both L. V. Mugnai et al. (2021) and J. E. Libby-Roberts et al. (2022) found their data to be consistent with a high mean molecular weight atmosphere, which could be thick or tenuous, or a bare rock. They also stated that their data are consistent with a cloudy primary atmosphere but deemed this scenario unlikely, as XUV irradiation likely long ago evaporated away any primordial H_2/He atmosphere on this planet.

As part of GO Program 1981, our team observed GJ 1132 b in transmission with JWST NIRSpec/G395H in 2023 (E. M. May et al. 2023). We found two transits that told different stories: the first was consistent with either an H_2O-dominated atmosphere with ∼1% CH_4 or stellar contamination, while the second was featureless. We investigated stellar and planetary atmospheric variability as well as instrumental systematics as drivers of this difference, but concluded that it was likely an unlucky random noise instance.

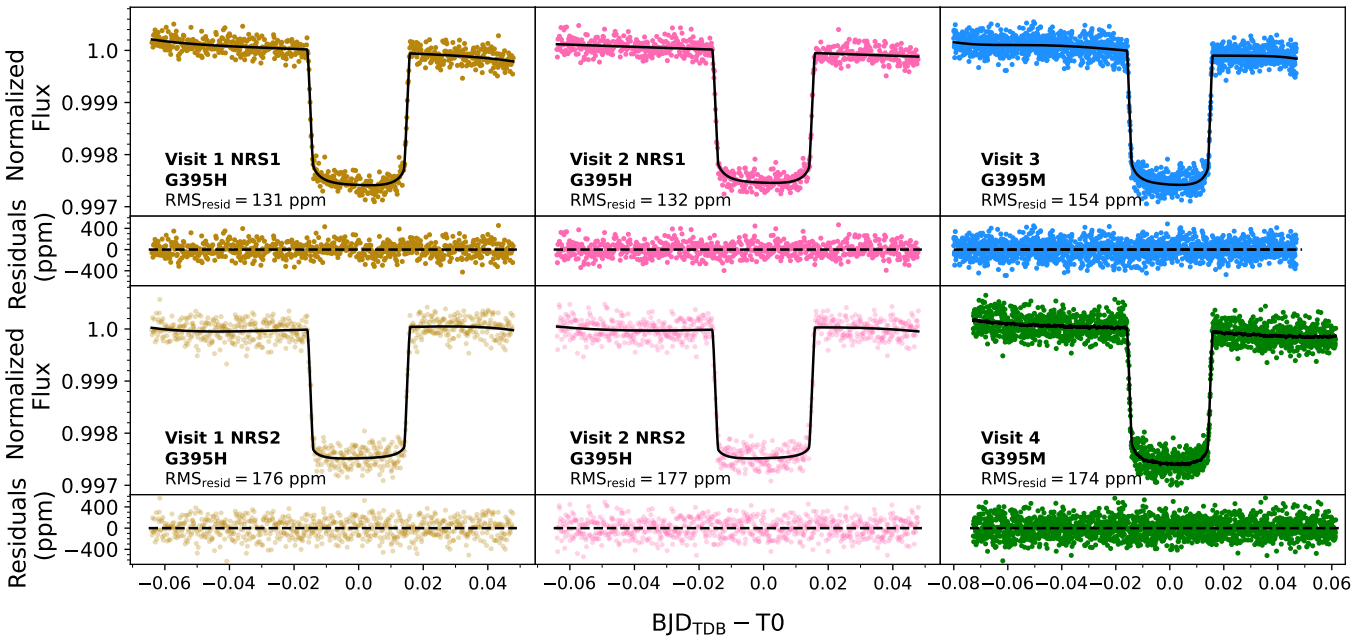

Figure 1. White light-curve data (colored points) and best-fit models (black lines) for all six light curves using the FIREFLy reduction. Residuals are shown below each respective light curve, with the mean rms of the residuals (RMS_{resid}) for each light curve given in the label.
(The complete figure set (3 images) is available in the online article.)

We do note that our analysis found that if the low mean molecular weight atmosphere detected by M. R. Swain et al. (2021) were real, there would have been large CH_4 features (on the order of 400 ppm) detectable in the NIRSpec/G395H bandpass, but we did not observe this.

GJ 1132 b was also observed by JWST in emission with GTO Program 1274 using MIRI/LRS (Q. Xue et al. 2024). One eclipse revealed a dayside temperature 1σ below that expected for a bare rock with zero albedo and no heat redistribution, with the emission spectrum consistent with a featureless blackbody or a range of atmospheric compositions with $\leqslant$10 bar. Atmospheres containing at least 1% CO_2, however, would need to be $<$10 mbar to be consistent with the data within 2σ, given the large CO_2 absorption that would otherwise be present in the MIRI/LRS range.

A more recent ground-based study using CRIRES+ on the Very Large Telescope (E. Palle et al. 2025) used the cross-correlation technique to search for HCN, CH_4, and H_2O in GJ 1132 b's atmosphere, using two models: one based on the atmosphere put forth in M. R. Swain et al. (2021), and the other based on the H_2O-dominated atmosphere postulated by E. M. May et al. (2023). None of the molecules were detected, and they largely rule out the atmospheric scenario from M. R. Swain et al. (2021), but were unable to reach sufficient sensitivity to support or refute the findings from E. M. May et al. (2023). Additionally, they did not detect 1083 nm He I and placed a 3σ upper limit on the mass-loss rate of 1.63×10^5 Earth atmospheric masses per Gyr.

Taking all of the atmospheric constraints to-date together, it is clear that that GJ 1132 b lacks a primordial H_2/He atmosphere, but it remains possible for the planet to host a thin secondary atmosphere. To follow up on our first two transit observations and determine whether or not GJ 1132 b hosts a thin atmosphere, we observe two additional GJ 1132 b transits, this time with NIRSpec/G395M instead of G395H in order to avoid the G395H detector gap. In this paper, we report on all four visits together and what they tell us, in conjunction with the JWST emission results, about the presence of an atmosphere on GJ 1132 b. In Sections 2 and 3, we describe our observations and data reduction process, respectively. We interpret the resulting spectrum in Section 4 and discuss both the technical and astrophysical implications of our findings in Section 5.

2. Observations

We observed four transits of GJ 1132 b with JWST as part of Cycle 1 GO Program 1981 (P.I.s K. Stevenson and J. Lustig-Yaeger), all using the NIRSpec Bright-Object Time Series mode (N. Espinoza et al. 2023). For the first two transits (Visits 1 and 2, hereafter), we utilized NIRSpec/G395H, as reported by E. M. May et al. (2023). For the latter two transits (Visits 3 and 4, hereafter), we utilized the NIRSpec/G395M observing mode in order to diagnose whether or not the gap between the data from the NRS1/NRS2 detectors in our G395H observations confounded our results. Specifically, we did not know whether the potential water slope seen in our first transit was due to an undiagnosed offset between the NRS1/NRS2 detectors. G395M exclusively uses the NRS1 detector and thus has no gap in the data, though it covers almost precisely the same wavelength range as G395H ($\sim$2.9–5.1 μm), albeit at a lower resolving power ($R \sim 1000$ for G395M compared to $R \sim 2700$ for G395H).

Our G395H observations occurred on 2023 February 25 and 2023 March 5, a little over a week apart, while our G395M observations occurred on 2024 February 8 and 2024 June 5. We acquired the target using the SUB32 subarray with the F140X filter. For the science exposures, we used the SUB2048 subarray with the NRSRAPID readout mode. Visits 1 and 2 (using G395H) used 14 groups per integration, with 814 integrations per visit (each lasting 4.62 hr including overhead). Visits 3 and 4 (using G395M) used six groups per integration,

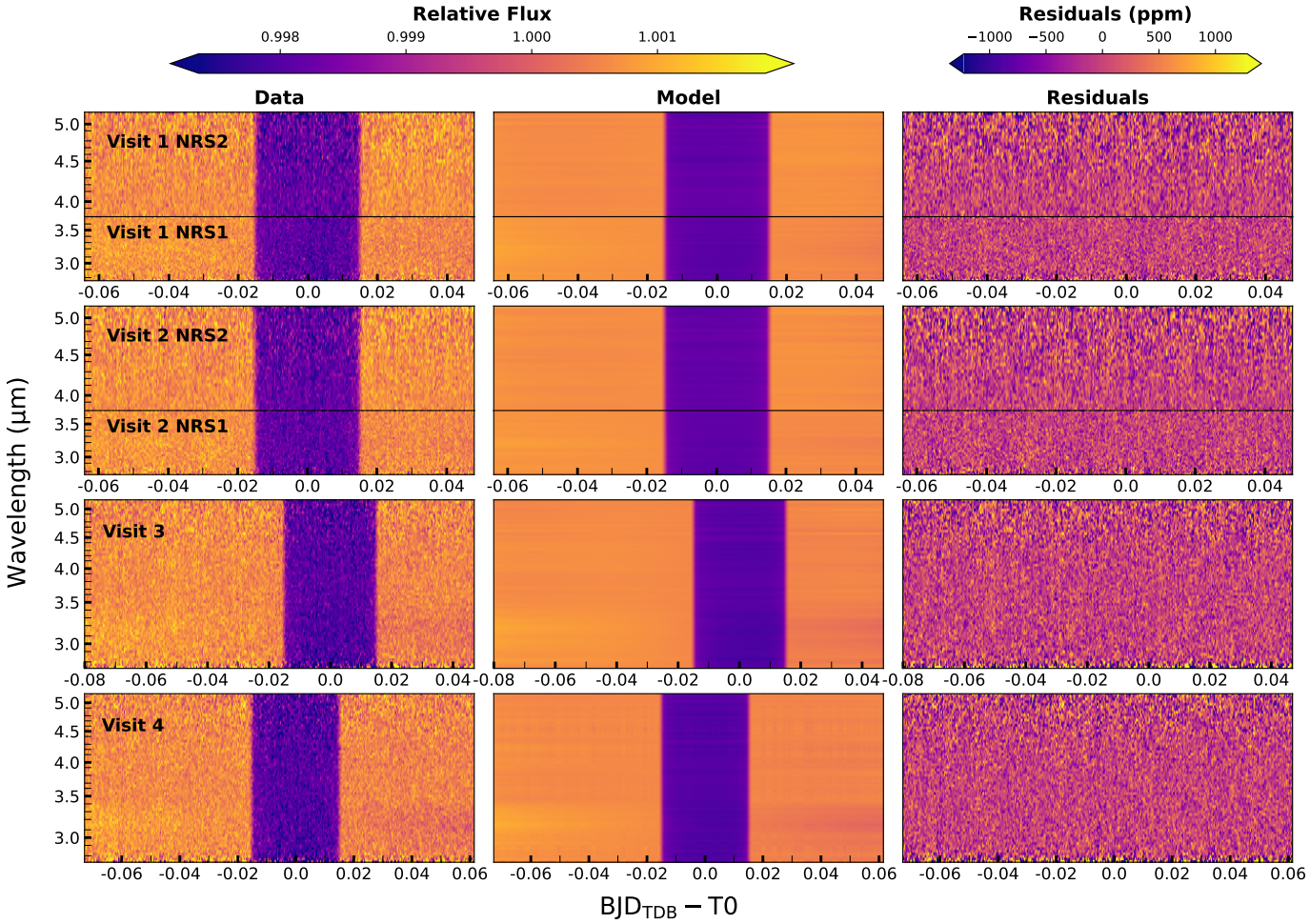

Figure 2. Spectroscopic light-curve data, model, and residuals for all six light curves using the FIREFLy reduction. Each row reflects a different observation. (The complete figure set (3 images) is available in the online article.)

with 1732 integrations (4.60 hr) for Visit 3 and 1932 integrations (4.95 hr) for Visit 4.

3. Data Reduction

As in E. M. May et al. (2023), we reduce the data with three independent reduction pipelines: FIREFLy (Z. Rustamkulov et al. 2022, 2023), Eureka! (T. Bell et al. 2022), and ExoTiC-JEDI (L. Alderson et al. 2022). Below we detail the reduction process for each pipeline. The weighted mean (at $R \sim 100$) of all four visits for each pipeline is shown in Figure 4, and the fit orbital parameters for each reduction are given in Table 1.

3.1. FIREFLy

FIREFLy begins with the uncal fits files and reduces the data through the spectroscopic light-curve fitting. In addition to reducing the G395M data, we completely re-reduced the G395H data, in order to make sure all four visits were reduced with the same calibration files from the jwst pipeline (version 1.13.4 with CRDS context (pmap) 1241) and with the same version of FIREFLy.

FIREFLy first utilizes Stages 1 (for group-level corrections) and 2 (for integration-level corrections) of the jwst reduction pipeline. As in previous reductions, the only changes we make in Stage 1 are to apply a group-level $1/f$ subtraction, and to skip the dark current step and the jump step. In Stage 1, we use the default jwst superbias file, as we found in E. M. May et al. (2023) that at this time, the most up-to-date superbias file does a better job at removing the bias than our custom bias subtraction described in S. E. Moran et al. (2023). Stage 1 ends with ramp fitting and the gain-scale step, after which in Stage 2 we only utilize the "assign WCS" step—we skip flat-fielding—and proceed to FIREFLy's custom stages.

For the stellar extraction, we first apply a cosmic-ray correction using lascosmic (P. G. van Dokkum 2001). Additionally, we apply a second $1/f$ subtraction at the integration-level and measure the x-position and y-position detector shifts over the course of the observation. We extract the 1D stellar spectra by using a fourth-order polynomial to compute the trace, and extracting the flux around an aperture full-width of ∼5.3 pixels across G395H and G395M. We determine the precise pixel width of the trace by finding the width that minimizes the out-of-transit scatter in the white light curve. This amounts to including all flux within 3.8σ of the center of the point-spread function (PSF) per column.

Before fitting the light curves, we trim the first 100 integrations for all four visits. We fit all observations separately, including the two different detectors for the G395H observations. We assume a circular orbit and fix the period to 1.628931 days (X. Bonfils et al. 2018) while fitting for the semimajor axis in units of stellar radii (a/R_s), the

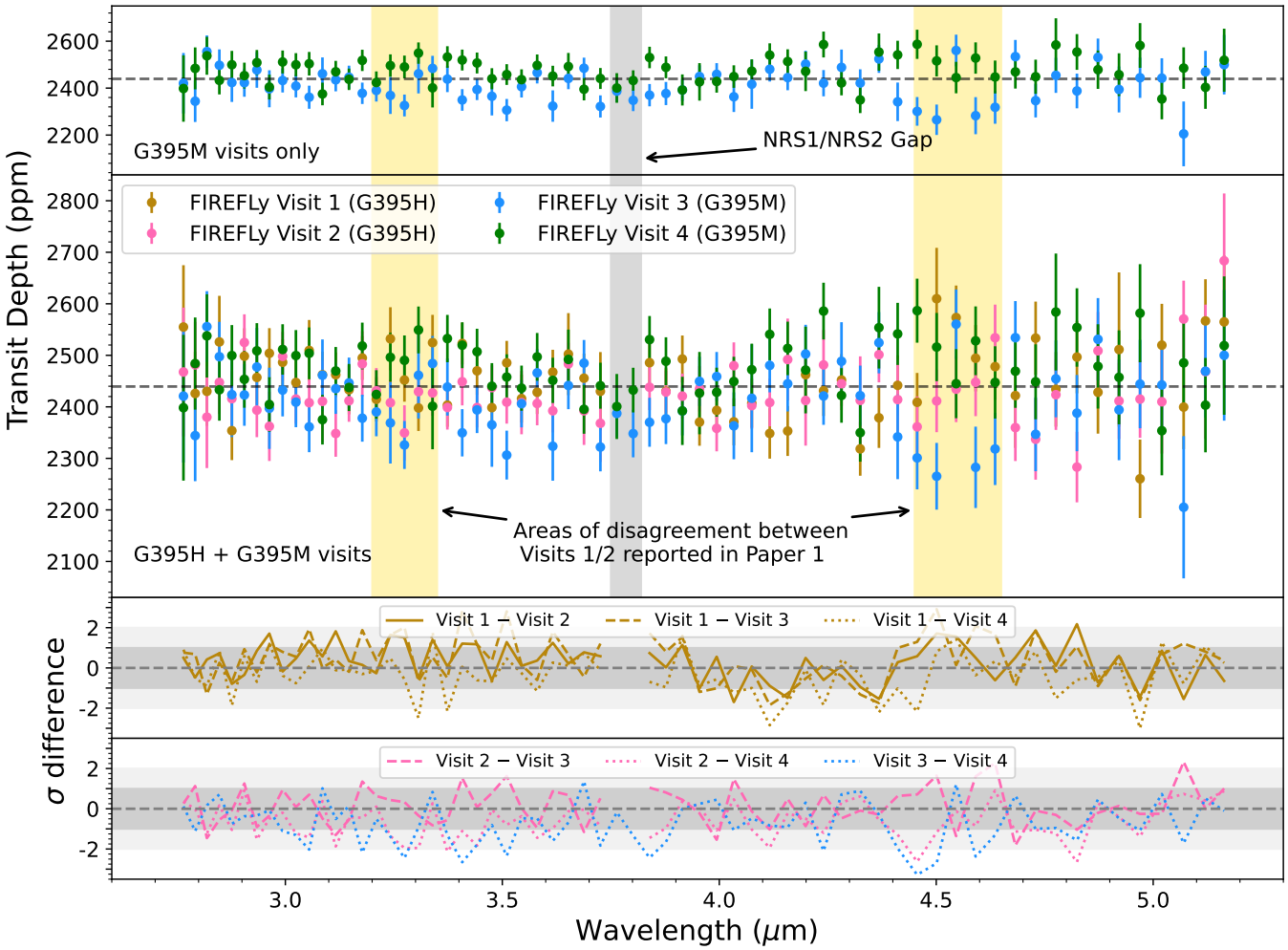

Figure 3. Top: transmission spectra of GJ 1132 b shown individually for the two G395M visits (Visits 3 and 4) for the FIREFLy reduction. Middle: the same as the top, but including the two G395H visits (Visits 1 and 2). For both panets, no offsets have been applied. The gray dashed horizontal line is the mean transit depth. The vertical gray region shows the gap between the NRS1 and NRS2 detectors, and the vertical yellow regions show the areas of the spectrum that drove differences in retrievals for the first two visits from E. M. May et al. (2023; Paper I). Bottom: the σ differences between different visits, with the 1σ and 2σ regions shaded in dark and light gray, respectively. There is strong intervisit and intermode agreement. Though there is scatter, all four visits largely show agreement to within 2σ.
(The complete figure set (3 images) is available in the online article.)

impact parameter (b), the center-of-transit time (T_0), the flux, and the transit depth ($(R_p/R_s)^2$). We fit for limb darkening using the quadratic law. Initially, we used the q_1/q_2 Kipping parameterization (D. M. Kipping 2013) but found that q_2 was consistently not fit well as it trended toward zero. Instead, we shifted to using the classic u_1/u_2 parameterization (Z. Kopal 1950). While we find in this case that u_1 still trended toward zero (which is consistent with q_2 trending toward zero, since $q_2 = 0.5u_1(u_1 + u_2)^{-1}$), the sigma differences between the individually fitted a/R_s and b decreases when switching from the q_1/q_2 to u_1/u_2 parameterization, and so we elect to stick with the latter. We set $u_1 = 0$ and fit for u_2. Once we fit all six white light curves individually and confirm that the best-fit system parameters are within 1σ of each other, we calculate the weighted mean of the fitted a/R_s, b, and u_2 values. We then go back and refit the white light curves fixing a/R_s and b to their weighted means from all six light curves, and fixing u_2 to its weighted mean within each respective observing mode.

To determine the most optimal systematics model in the light-curve fits, we examine all combinations of the possible detrending vectors (8191 total possible combinations), which includes a polynomial in time (up to order six), as well as the x- and y-position shifts, their squares (e.g., x^2 where x is the subpixel position in the spectral direction), and their product ($x \times y$). We determine the most optimal combination of detrending vectors by determining which combination has the lowest Akaike information criterion (AIC) value. When fitting the white light curves, if a given fitting parameter is not statistically significant (to $\gtrsim 3\sigma$), we then manually exclude it, except in the case where excluding it increases the red noise in the residuals. We found this to be the best method to determine what systematics model to use. If we simply relied on the model with the lowest Bayesian information criterion (BIC), there was almost always significant red noise in the residuals. We used a linear polynomial in time for Visit 2 NRS1, a third-order polynomial for Visit 2 NRS2, and a fifth-order polynomial in time for Visit 1 NRS1 and NRS2 and Visit 3. For Visit 4, we used a fourth-order polynomial plus the y-position shift. To actually fit the white light curves, we use emcee (D. Foreman-Mackey et al. 2013) with 1000 steps and 200 burn-ins. We show the individually fitted white light curves and their residuals in Figure 1. Comparing the residuals on the NRS1 detector, the rms values of the G395M residuals

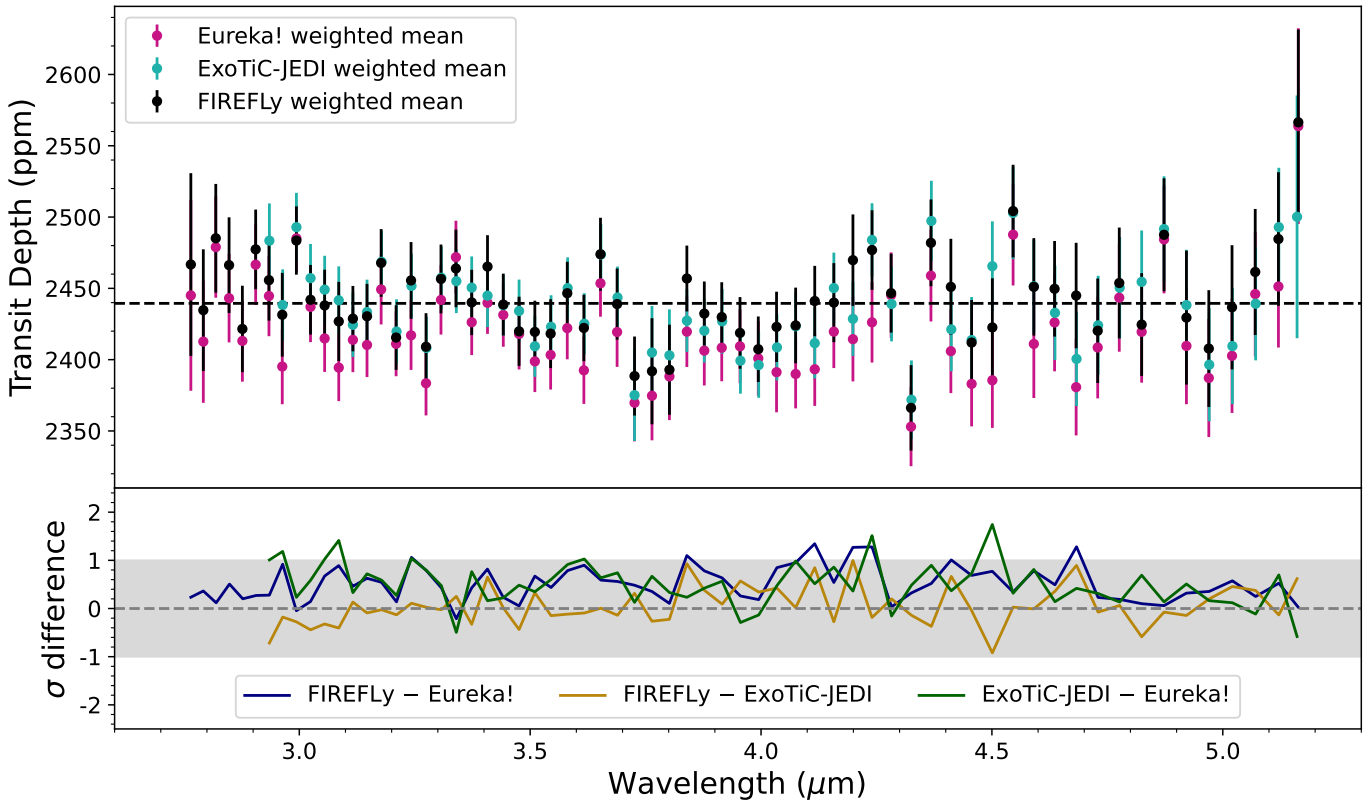

Figure 4. Top: the weighted mean of all four visits is shown for all three reductions, with the mean transit depth for `FIREFLy` shown as a horizontal dashed black line. No offsets have been applied to any data set. Bottom: the σ difference between reductions is shown, with the 1σ difference region highlighted in gray. Reductions show exquisite agreement to within 1σ for most channels.

are generally larger than those of G395H NRS1 due to the lower number of groups per integration in these visits.

For the spectroscopic fits, we use least-squares fitting with `lmfit` (M. Newville et al. 2014), as this is much more computationally efficient and has been shown to have consistent results with `emcee` fitting at the spectroscopic stage (E. M. May et al. 2023). We apply a binning scheme of $R \sim 100$, following E. M. May et al. (2023). We fix system parameters to the weighted-mean white light-curve values and apply the same systematics model as in the white light-curve fit. We fix the detrending parameter coefficients to their white light-curve values, except for the linear term, which has a wavelength dependency. We also extensively explore whether or not to fit or fix limb darkening in the spectroscopic fits. In past studies, `FIREFLy` has fixed limb darkening at this stage (E. M. May et al. 2023, S. E. Moran et al. 2023), since we are working at long wavelengths where in theory the limb darkening should not change that much. However, there was a slight wavelength dependency (on the order of 10–20 ppm) when we compared the fit-limb-darkening versus fixed limb-darkening transmission spectrum. Additionally, when plotting the weighted mean of the fitted u_2 values as a function of wavelength, there visually is a wavelength dependency shortward of $\sim$4 μm, whereas longward of $\sim$4 μm, u_2 is quite flat. This can be seen in each individual u_2 spectrum per visit as well, but is most apparent in the weighted mean spectrum. Thus, we adopt an "empirical best-fit" approach in which we fit a linear slope to the fitted weighted mean u_2 values shortward of 4 μm and fit a zero-order polynomial (i.e., just take the weighted mean) of the u_2 values longward of $\sim$4 μm. We then adopt these values as our fixed limb-darkening values in the spectroscopic fits. Our spectroscopic light curves for `FIREFLy` are illustrated in Figure 2.

At both the white and spectroscopic light-curve level, we rescale our error bars when needed so that $\chi^2_\nu = 1$ between the data and model. While this allows us to more realistically describe the scatter of the data, it does not actually take into account correlated noise, the recurring demon of exoplanet transmission spectroscopy studies. Thus, in order to properly rescale our final transmission spectrum error bars, we use the binned rms of the residuals (hereafter referred to as a "red noise diagnostic plot"; formerly referred to as an "Allan variance plot"; but see D. Kipping 2025) to measure the correlated noise across different bin sizes in time (F. Pont et al. 2006) for each spectroscopic channel. We then inflate the calculated error bar for that channel in quadrature with the median correlated noise across all bin sizes.

Fitted system parameters for each visit are shown in Table 1. We show each individual visit reduction for `FIREFLy` in Figure 3, and the final weighted mean spectrum (coadding all four visits) in Figure 4.

3.2. *Eureka!*

The `Eureka!`[14] (T. Bell et al. 2022) package begins with the `uncal` JWST files. Stages 1 and 2 are wrappers of the `jwst` pipeline (H. Bushouse et al. 2024; version 1.14.0 with crds context (pmap) 1253 for M visits and crds context (pmap) 1241 for H visits) with some modifications, primarily custom group-level background subtraction. This improves the

[14] For this work, we use an untagged development version after v0.10 that addressed bug fixes prior to the large v1.0 release.

Table 1
Best-fit Orbital Parameters from White Light-curve Fitting

Parameter		FIREFLy	Eureka!	ExoTiC-JEDI
$(R_p/R_s)^2$	(ppm)	2428 ± 8	2448 ± 13	2439 ± 8
Visit 1 T_0—2460000	(BJD_{TDB})	0.97654 ± 0.00002	0.97652 ± 0.00001	0.97653 ± 0.00002
Visit 2 T_0—2460009	(BJD_{TDB})	0.12115 ± 0.00002	…	0.12115 ± 0.00002
Visit 3 T_0—2460349	(BJD_{TDB})	0.56744 ± 0.00002	…	0.56746 ± 0.00002
Visit 4 T_0—2460466	(BJD_{TDB})	0.85035 ± 0.00002	…	0.85036 ± 0.00002
Period	(days)	1.628931 (fixed)	1.62892962 ± 7.15E-8	1.628931 (fixed)
a/R_s	(unitless)	15.53 ± 0.22	15.88 ± 0.19	15.51 ± 0.17
b	(unitless)	0.44 ± 0.03	0.39 ± 0.03	0.44 ± 0.02
i	(°)	88.37 ± 0.10	88.61 ± 0.11	88.38 ± 0.09

Note. FIREFLy and ExoTiC-JEDI report $(R_p/R_s)^2$, a/R_s, and b (or i) as the weighted mean from all six independent white light-curve fits. Eureka! reports values from a joint white light-curve fit, and therefore, only directly fits for the first center of transit time and the period. Eureka! and ExoTiC-JEDI use time units of $BMJD_{TDB}$, which is converted to BJD_{TDB} here. b is measured in the FIREFLy light-curve fits and then converted to i for this table, whereas the opposite is true for ExoTiC-JEDI and Eureka!. All fitted parameters agree to within 1.7σ across reductions.

precision of ramp fitting by removing the striping resulting from $1/f$ noise at the group level. It consists of identifying the center of light in each column and masking the values 6 pixels within each side of the center. For this study, we estimate the background as the median of all other pixels in a column with a 3σ outlier threshold. Lastly, for Stage 1, we skip the jump step detection to avoid more noise being introduced in the light curves. For Stage 2, we skip both the flat-field and photon steps similar to E. M. May et al. (2023) since these steps give us absolute flux calculations that we do not need for this reduction.

After running these stages of the Eureka! pipeline, we continue running Stage 3, which first corrects the curvature of the trace. We identify the center of light in each column, then roll the pixels within that column so that the stellar spectrum aligns with the center of the detector. We then run an additional background subtraction and remove the median of all pixels more than 8 pixels away from the center of the trace with a 5σ outlier threshold. We utilize optimal spectral extraction (K. Horne 1986) to extract the 1D stellar spectra time series. The median frame is assembled with a 5σ outlier rejection threshold, smoothing over 20 frames. For Visit 3 with G395M, the 1D stellar spectra are extracted with a 20σ outlier rejection threshold relative to the median frame and a 3 pixel half-width aperture size. For Visit 4, we use a 30σ outlier rejection threshold and a 3 pixel half-width aperture size. We choose these thresholds and aperture sizes based on the combination of these parameters that result in the lowest median absolute deviation (MAD) in the white light curves.

We also reanalyze both G395H visits presented in E. M. May et al. (2023) following the same steps outlined within, while using updated NIRSpec reference files. For the G395H observations, white light curves are created by summing all light between 2.752 and 3.721 μm for the NRS1 detector and 3.820–5.177 μm for the NRS2 detector. For the two new G395M visits presented in this work, white light curves are created by summing all light between 2.726 and 5.181 μm. Spectroscopic light curves for all visits are extracted on the same $R \sim 100$ wavelength grid as FIREFLy.

We perform a joint white light-curve fit with emcee (D. Foreman-Mackey et al. 2013) of all six white light curves (two for each G395H visit, one for each G395M visit) using batman for the astrophysical signals and a linear temporal ramp for each light curve. For the temporal ramp, we initially consider zeroth- to third-order polynomials in time, and any combination thereof (e.g., third order with no quadratic term). We select the model with the lowest BIC, which corresponds to the standard linear ramp. In retrospect, we note that the BIC might not always be the best model selection tool to use, as there remains a higher level of red noise in the Eureka! white light curves compared to those of FIREFLy, which uses the AIC to arrive at a more complex systematics model. The planetary orbital period, T_0, a/R_s, and orbital inclination (i) are required to be the same for all visits, while the transit depth is allowed to vary between visits and detectors. Quadratic limb-darkening parameters are fixed independently to the MPS2 grid (N. M. Kostogryz et al. 2022) for each white light curve and calculated with ExoTIC-LD (D. Grant & H. R. Wakeford 2022; D. Grant & H. Wakeford 2024) using the same stellar parameters as in E. M. May et al. (2023), namely $T_{\rm eff} = 3261$ K, $\log g = 5.02$ (both from K. G. Stassun et al. 2019), and a metallicity $[{\rm Fe/H}] = -0.12$ (Z. K. Berta-Thompson et al. 2015). All chains are run to at least 50× the autocorrelation time to ensure convergence. Specifically, for these joint white light-curve fits, we use 25,000 steps, 250 walkers, with a burn-in of 5000 steps. Best-fit orbital parameters are given in Table 1.

We then hold orbital parameters fixed at our best-fit white light values for the spectroscopic light curves and only fit for the transit depth and temporal ramp parameters. Limb darkening is once again held to the values from the MPS2 grid with ExoTIC-LD for the given channel. Final error bars are inflated following the correlated-noise approach described in the FIREFLy section. Some type of error inflation like this is critical to do; although it only increases the error bars in the final weighted mean spectrum in select channels by ∼5 ppm, it changed the precise weighted mean spectrum enough to alter interpretation of the spectrum in our retrievals. Our final weighted mean spectrum is shown in Figure 4.

3.3. ExoTiC-JEDI

For our ExoTiC-JEDI (L. Alderson et al. 2022) reduction of the G395M data, we follow the same procedures as for the G395H data presented in E. M. May et al. (2023), with minor adjustments to account for the differences between the two modes. In summary, the ExoTiC-JEDI reduction begins with a modified version of stage 1 of the jwst pipeline (v.1.14.0, H. Bushouse et al. 2024), using the custom bias correction available within ExoTiC-JEDI (e.g., L. Alderson et al. 2023, 2024; N. Scarsdale et al. 2024), before cleaning the

2D images and extracting 1D stellar spectra. The reduction of the G395M data uses the same reduction parameters as outlined in E. M. May et al. (2023) other than the full-width of the aperture, which for both Visit 3 and 4 is 3.5× the trace FWHM, equivalent to approximately 5 pixels. With the extracted 1D stellar spectra, we produce white light curves spanning the full G395M range (2.814–5.1 μm), as well as spectroscopic light curves at $R \sim 100$.

We additionally re-reduce the G395H data to take into account lessons learned from other `ExoTiC-JEDI` analyses as well as updates to the `jwst` pipeline. This analysis also followed the same procedures as presented in E. M. May et al. (2023); however, the `ExoTiC-JEDI` custom bias subtraction was used as opposed to the default within the JWST pipeline. With our re-reduced 1D stellar spectra, we produced new white light curves spanning the full G395H range (2.814–3.717 μm for NRS1 and 3.824–5.111 μm for NRS2) and spectroscopic light curves at $R \sim 100$. Our pipeline does not extract flux blueward of 2.814 μm due to the limb-darkening prescription used, as `ExoTiC-LD` (D. Grant & H. Wakeford 2024) is limited to the official throughput of the NIRSpec instrument.

To account for the full information provided by the four total observed transits, we fit the four individual G395H white light curves (one each for NRS1 and NRS2 per visit) along with the two G395M white light curves (one per visit) before taking the weighted average of the system parameters to ensure consistency when obtaining the final transmission spectra. In all cases, we fit the light curves using a least-squares optimizer with an L. Kreidberg (2015) transit model (`batman`) and a systematic model and calculated limb-darkening coefficients with `ExoTiC-LD` (D. Grant & H. Wakeford 2024) using the nonlinear law (A. Claret 2000) and a Phoenix stellar model (T. O. Husser et al. 2013; $T_{\rm eff} = 3300$ K, $\log(g) = 5.0$, $[{\rm Fe/H}] = 0.0$). For all six sets of light curves (two for G395M and four for G395H), we used the same systematics model, $S(\lambda)$, as in E. M. May et al. (2023), correcting for a linear trend in time, t, plus the change in x-position, x_s, multiplied by the absolute magnitude of the y-positional change, $|y_s|$, i.e., $S(\lambda) = s0 + (s1 \times x_s|y_s|) + (s2 \times t)$, (where $s0$, $s1$, $s2$ are coefficient terms). We also removed any data points that were greater than 4σ outliers in the residuals, and rescaled the flux errors using the beta value (F. Pont et al. 2006) to account for any remaining red noise. In the case of the G395H white light curves, we additionally removed the first 15 integrations from the time series to remove a settling ramp seen in the light curves.

For our initial white light-curve fits, we fitted for R_p/R_s, a/R_s, i and T_0, holding the period fixed to 1.628931 days and the eccentricity fixed to 0 (X. Bonfils et al. 2018). We then calculated the weighted average of the a/R_s and i values from the two G395M and four G395H white light curves and refit them holding a/R_s and i fixed to these values (see Table 1). Finally, we fit the spectroscopic light curves for each visit, fitting only for R_p/R_s, fixing T_0 to the value obtained from each respective white light-curve refit, and fixing all other parameters to the values as outlined in Table 1. The final weighted mean spectrum can be seen alongside the `FIREFLy` and `Eureka!` spectra in Figure 4.

3.4. Intervisit and Reduction Agreement

Figure 3 shows the `FIREFLy` transmission spectra for the two new G395M visits (we refer the reader to Figure 1 from E. M. May et al. (2023) for an illustration of the two G395H visits), as well as all four visits overplotted. Importantly, no offsets have been applied to any visits nor between the NRS1/NRS2 data sets in the G395H visits. The G395M data fill the gap between the NRS1/NRS2 detectors smoothly, suggesting that offsets between NRS1/NRS2 may not prevent useful data interpretation, at least for observations with a large number of groups per integration ($\gtrsim$7). For more details, see Section 5.1.

In Figure 3, all four visits agree well, with the majority of points agreeing within 2σ. Across the four visits, there is more visual scatter at the red end, but this is largely due to the decrease in stellar flux at these wavelengths, which increases the overall noise in the spectrum. To illustrate this, note that though the scatter increases visually in terms of ppm, the scatter as measured in σ remains consistent across the spectrum.

With the addition of two more visits, it appears by simple visual inspection that the disagreements between Visits 1 and 2 reported in E. M. May et al. (2023), highlighted in yellow in the figure, may be due to random noise. There is not inherently more scatter in the region around 3.3 μm compared to elsewhere in the spectrum when looking across all four visits. We note, however, that there does appear to be a disagreement between Visit 3 and the other visits in the region around 4.5 μm. The culprit here is unlikely to be planetary in nature, as the fluctuations are on the order of 200 ppm, too large to be explainable by atmospheric variation. Instrumental artifacts remain a possibility, though there is no evidence for increased red noise or scatter in these light curves. Intriguingly, this region aligns with the fundamental band of CO centered around 4.5 μm, suggesting that an astrophysical explanation is also worth speculating: the discrepancy could perhaps arise from spatially variable CO on the star itself. While the exoplanet community has been focused on stellar contamination driven by cooler water-dominated spots in the stellar atmosphere (e.g., B. V. Rackham et al. 2018, S. E. Moran et al. 2023), stellar contamination by spatially variable CO has not been explored as deeply, though recent magnetohydrodynamic modeling efforts illustrate that excluding CO and other diatomic molecules can impact the stellar (specifically, the umbral and penumbral) spectra of early M dwarfs longward of ~4 μm (H. N. Smitha et al. 2025). This suggests that stellar contamination at longer wavelengths cannot be ruled out, though of course random noise instances are also likely to play a role in the Visit 3 discrepancies. We will further discuss stellar variability in Section 4.2.

In addition to good intervisit agreement, we also have exquisite agreement between reductions, as shown in Figure 4. This figure shows the weighted mean spectra from all three reductions, which agree to largely within 1σ. Again, no offsets have been applied to any data set. The slightly lower transit depths for `Eureka!` at the blue end are an artifact of limb darkening. When `Eureka!` uses 3D stellar models from Z. Magic et al. (2015) instead of the MPS2 grid, `Eureka!` actually has slightly *higher* transit depth values at the blue end compared to `FIREFLy` or `ExoTiC-JEDI`. This demonstrates the importance of careful limb-darkening selection/reduction comparison so as to not attribute limb-darkening systematics to astrophysical phenomena. Finally, Table 1 demonstrates the agreement in fitted orbital parameters across the three reductions, all of which agree to within 1.7σ.

4. Analysis

4.1. Gaussian Detection Test Favors a Flat Line

Following previous works (e.g., E. M. May et al. 2023; S. E. Moran et al. 2023; J. Kirk et al. 2024), we begin our analysis of the transmission spectrum by simply determining whether it rejects the null hypothesis. In this case, the null hypothesis is represented by a flat, featureless spectrum, indicative of no evidence for spectroscopic absorption features from an atmosphere. We perform two common statistical tests.

In our first test, we assess the consistency of the data with the null hypothesis by computing the cumulative distribution function of the observed chi-squared under the null model. We then convert this probability, p, into a Gaussian equivalent significance using $z = \sqrt{2}\,\mathrm{erf}^{-1}(p)$, representing the number of standard deviations by which the observation deviates from the null expectation. Results of this test for each data reduction are given in Table 2. Small values of z (e.g., $z < 2$) indicate a high degree of consistency with the null hypothesis, while larger values (e.g., $z > 3$) suggest increasing tension that may justify rejection of the null at high confidence.

In our second test, we directly compare the null hypothesis model to a more complex, albeit still quite simple, alternative model with a Gaussian absorption feature added to the flat model. We elect to use a Gaussian model to approximate molecular absorption in a transmission spectrum (e.g., J. Taylor 2025); we allow the feature to be as broad as the full G395H wavelength range and centered outside of this range so as to also approximate a spectral slope (e.g., from H_2O, as in S. E. Moran et al. 2023) should that be preferred. Since the models are nested, they enable a comparison of their respective Bayesian evidence terms estimated using the `dynesty` Nested Sampling code (J. Skilling 2004; J. S. Speagle 2020). We calculate Bayes factors from the ratio of the two model evidences and then report equivalent σ values in Table 2 for the detection significance of the Gaussian model over the null hypothesis (R. Trotta 2008). We use a negative sign in this case to flip the interpretation and denote the detection of the flat line over the Gaussian feature.

Table 2 reports results from both our null hypothesis and Gaussian feature detection tests for all three independent data reductions using the weighted mean of the four visits (Figure 4). As previously discussed, the three reductions are in excellent agreement, and this agreement holds for the statistical properties of our null hypothesis tests. All three reductions are consistent with a featureless spectrum to within approximately 1σ of the expected range in χ^2_ν values given the null hypothesis and no significant evidence is found to favor a Gaussian feature in the data.

Table 2
Is It Flat?

Reduction	χ^2_ν	ν	Dev. from Null	Gaussian Detection
`Eureka!`	1.09	64	1.04σ	1.71σ
`FIREFLy`	0.88	64	0.32σ	-2.00σ
`ExoTiC-JEDI`	1.13	58	1.20σ	1.49σ

Note. χ^2_ν is the reduced chi-squared resulting from the best-fitting featureless fit (null hypothesis) to the observed spectrum. "Dev. from Null" (z) refers to how many σ the χ^2_ν deviates from expectation under the null. "Gaussian Detection" shows the equivalent detection significance of an agnostic Gaussian absorption feature in the spectrum over the flat null hypothesis from a Bayes factor model comparison, where the negative sign denotes preference instead for the flat model over the Gaussian.

4.2. Models Hint at Evolution of Stellar Heterogeneity between Visits

To explore potential changes in stellar heterogeneity, we forward modeled the out-of-transit, flux-calibrated stellar spectra for all four visits. This analysis followed the same procedures presented in E. M. May et al. (2023), but using the `FIREFLy` reductions and adjusting each spectrum to match the lower resolution and wavelength sampling of the G395M visits (for consistency in determining inhomogeneities between visits). Additionally, we use a wider aperture for the flux-calibrated spectrum than the transmission spectrum, because we want to ensure we capture the entirety of the stellar flux. We include all flux within 9.7σ of the center of the column-level PSF (the maximum available width before exceeding the bounds of the detector), which equates to an aperture full-width of 13.35 pixels for G395H NRS1, 15.15 pixels for G395H NRS2, and 14.12 pixels for G395M. We confirm that the flux-calibrated spectra do not differ between reductions. Because the JWST pipeline-estimated error bar size is underestimated (P. Jakobsen 2022), we calculate empirical error bars. We estimate an initial error bar size based on the mission-required flux precision from JDox,[15] then inflate/deflate the error bars until the χ^2_ν between the data and best-fit model (see below) is 1. Assuming the best-fit model is true, we achieve a precision of ~2.2% and ~1.3% for the visits to G395H and G395M, respectively. This is consistent with the absolute flux calibration level (~2%) of the NIRSpec instrument.

We utilize multicomponent stellar models based on the F. Allard et al. (2012) `PHOENIX` models, aiming to detect any rotational or evolutionary modulation of surface features. Using a weighted combination of three `PHOENIX` models, we modeled the background photosphere, cooler regions representing starspots ($T_{\rm eff} \leqslant T_{\rm eff,\ photosphere} - 100$ K), and warmer regions representing faculae ($T_{\rm eff} \geqslant T_{\rm eff,\ photosphere} + 100$ K). Each component shared the same $\log(g)$ and metallicity (held constant at $\log(g) = 4.5\ \mathrm{cm\ s^{-2}}$ and $Z = 0$; consistent with E. M. May et al. 2023), and the coverage of the cooler/warmer regions was each allowed to represent up to 50% of the surface. The model grid ranged from $T_{\rm eff} = 2500$–4500 K.

For direct comparison with the observed spectra, we normalized the models to the observed flux at 3 μm and interpolated them to match the resolution and wavelength sampling of Visit 4 before computing χ^2_ν. Each χ^2_ν was calculated across 1273 wavelength points and five free parameters: $T_{\rm eff}$ of the photosphere, cool spots and warm spots, and coverage fractions of both feature types.

All four visits favored a photosphere temperature of 3100 K, which is consistent within 2σ of the literature value of 3270 ± 140 K (X. Bonfils et al. 2018). They also all favor the same 2900 K for the cooler spots and 3300 K for warmer spots. The visits do differ, however, in the percent coverage, where Visit 1 is an outlier, suggesting ~10% more cool spots and ~5% less warm spots than the other three visits. The best-fit parameters and their uncertainties (computed from models with $\Delta\chi^2 \leqslant 6$ above the minimum, which, given the five

[15] https://jwst-docs.stsci.edu/jwst-calibration-status/jwst-absolute-flux-calibration

model parameters, yields a 1σ confidence interval) are given in Table 3. All best-fit model spectra (including those captured in the uncertainties) are plotted in Figure 5.

We note that the results from Visit 1 and 2 presented here differ from those computed in E. M. May et al. (2023), where both visits were previously favored by a 3200 K photosphere and ∼40% 2900 K spots and ∼25% 3500 K faculae. This is due to myriad factors including different aperture sizes used in the reduction (which drive a 1% difference in the spectra at native resolution), using a lower wavelength resolution in the current analysis, and allowing the cool and warm region coverage to now go up to 50%, where before it was capped at 45%. We made this change here to allow for the possibility of a two-component best-fit model.

The results presented here suggest that the visible surface of the star could have had a large starspot near the limb during Visit 1 that rotated off before Visit 2, 8 days later. Interestingly, in the `FIREFLy` reduction, the fit spectroscopic limb-darkening coefficients are slightly different in Visit 1, compared to the rest of the visits. Assuming a rotation period of the star of $122.31^{+6.03}_{-5.04}$ days (R. Cloutier et al. 2017), there is a 7% rotation between Visits 1 and 2 and a 95% rotation between Visits 3 and 4. There are 340.43 days between the two sets of visits, corresponding to a 78% rotation (when phase-folding). Therefore, we are looking at similar faces of the star for Visits 1 and 2 and for Visits 3 and 4, but between the two, we are seeing slightly different portions of the stellar disk. The total time period between Visit 1 and Visit 4 is 466 days, which allows for the additional possibility that the purported spot from Visit 1 evolved before Visits 3 and 4.

4.3. Forward Models Favor Featureless Spectrum or Thin Steam Atmosphere

To further interpret the combined transmission spectrum, we follow our previous works (J. Lustig-Yaeger et al. 2023; E. M. May et al. 2023; S. E. Moran et al. 2023; J. Kirk et al. 2024) and compare our observations of GJ 1132 b to a series of atmospheric forward models generated for GJ 1132 b's planetary mass, radius, and equilibrium temperature. We include the same series of models from our original GJ 1132 b study (E. M. May et al. 2023), as summarized briefly below, as well as an expanded model set with additional atmospheric pressures given the increased precision enabled by our four-visit combined spectrum. All models were produced using the `PICASO` (N. E. Batalha et al. 2019) radiative transfer code, using molecular opacities derived from `PICASO`'s Zenodo v2 database of Resampled Opacities at $R = 60{,}000$ (N. Batalha et al. 2022). We rebinned each model to the resolution of the combined transmission spectrum for each reduction to compute a goodness-of-fit in terms of the χ^2_ν assuming 64 degrees of freedom ($\nu = 64$). We show these results in Figure 6, where the best-fit results ($\chi^2_\nu < 1.18$) have p-values corresponding to low ($<1\sigma$) significance levels for rejecting the models, indicating no significant disagreement with the data.

We include one model using abundances of a 1000× solar metallicity, solar C/O ratio atmosphere generated using the `CHIMERA` (M. R. Line & Y. L. Yung 2013; M. R. Line et al. 2014) thermochemical equilibrium model. This model includes opacities from H_2 collision-induced absorption (CIA), H_2, He, H_2O, CH_4, CO, CO_2, NH_3, N_2, HCN, and H_2S with a T. Guillot (2010) temperature–pressure profile. Of the models explored in E. M. May et al. (2023), this 1000× solar metallicity model was not ruled out when considering Visits 1 and 2 independently. Now, when considering the weighted mean of all four visits, all three reductions robustly reject this atmospheric scenario by over 4σ.

We next compare to a series of either single- or two-component atmospheric models at 1 bar in surface/cloud-top pressure with an isothermal temperature set to GJ 1132 b's equilibrium temperature of ∼580 K. Two of these models—either clear, pure CH_4 or CO_2 atmospheres—are strongly rejected by each reduction at $>10\sigma$ and 3σ, respectively. These models can be rejected because the transmission spectrum lacks significant absorption in either the 3.35 μm or 4.3 μm regions where these molecules have spectral features. The CH_4 model was previously ruled out by our original two visits of GJ 1132 b (E. M. May et al. 2023), but we could not rule out the CO_2 model without stacking multiple transits. Finally, motivated by the potential water-rich, methane-containing atmosphere we obtained from Visit 1 in E. M. May et al. (2023), we include water-rich atmospheres with varying amounts of CH_4. The pure 1 bar water atmosphere is a better fit (χ^2_ν of ∼1.2–1.3 for each reduction) than that of either 10% or 1% CH_4 in water (χ^2_ν of ∼1.5–2.7 across reductions). All of these clear atmospheric forward models are shown in the top panel of Figure 6. Finally, we also include a flat line model representative of either no atmosphere or a very high altitude cloud deck. Compared to the clear, 1 bar atmosphere models discussed above, the flat line model is the preferred forward model fit with the lowest χ^2_ν for all three reductions ($\chi^2_\nu \sim 0.9$–1.1).

Next, we examine how changing the pressure of our H_2O atmosphere affects our interpretation, generating models with an opaque pressure level from 1 bar to 0.01 mbar in five logarithmic steps. This opaque pressure level can be interpreted either as that of the pressure of an opaque cloud-top or haze layer or the surface pressure. The χ^2_ν values do decrease with decreasing pressure levels, but all thin atmosphere water scenarios remain consistent with the data. Though it is not a significant difference, the water atmosphere models fit the data better than the flat line because they simultaneously

Table 3
Best-fit Multicomponent Stellar Models to Each Visit

Parameter	Visit 1	Visit 2	Visit 3	Visit 4
$T_{\rm photo}$ [K]	3100	3100	3100	3100
$f_{\rm cooler}$ [%]	31 ± 2	21^{+3}_{-4}	24 ± 1	22^{+2}_{-1}
$T_{\rm cooler}$ [K]	2900 ± 0	2900^{+0}_{-100}	2900 ± 0	2900 ± 0
$f_{\rm warmer}$ [%]	22^{+3}_{-4}	27^{+6}_{-3}	26 ± 2	28 ± 2
$T_{\rm warmer}$ [K]	3300 ± 0	3300 ± 0	3300 ± 0	3300 ± 0
$\chi^2_{\nu,\rm min}$	1.0	1.0	1.0	1.0
n [a]	25	46	9	10
δt [b] [days]	0	8.14	348.57	465.86
φ [c]	0	0.07 ± 0.003	$0.85^{+0.13}_{-0.12}$	$0.81^{+0.18}_{-0.16}$

Notes. All observations were reduced to the resolution of G395M, and the photospheric temperatures are held constant at 3100 K. Uncertainties are computed from models with $\Delta\chi^2 \leqslant 6$.
[a] Number of models with $\Delta\chi^2 \leqslant 6$.
[b] Time (in days) from the first observation.
[c] Phase-folded phase difference from the first observation assuming a rotation period of $122.31^{+6.03}_{-5.04}$ days (R. Cloutier et al. 2017).

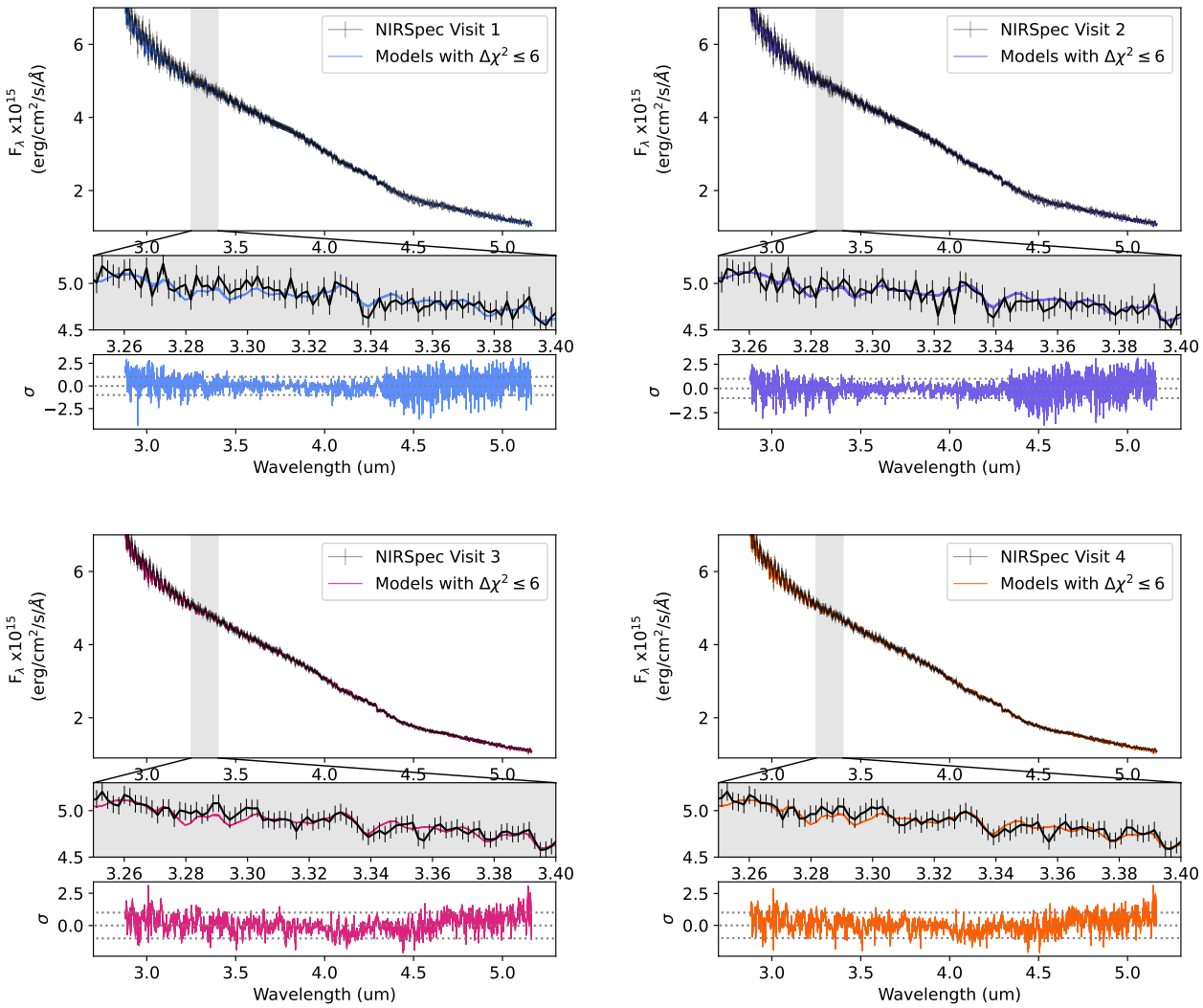

Figure 5. Best-fit multicomponent stellar models for each visit. Each of the four main panels corresponds to one visit, and contains three subpanels. In every subpanel, all models with $\Delta\chi^2 \leqslant 6$ from the best-fit are shown (25 models for Visit 1, 46 for Visit 2, 9 for Visit 3, and 10 for Visit 4). The models are nearly indistinguishable due to their close agreement. The observed spectra, smoothed to the resolution of G395M, are overplotted in black. The residuals (in units of σ) are shown in each bottom subpanel, with dotted lines indicating the $\pm 1\sigma$ range.

maintain the slight visual slope from 2.8–3.5 μm while flattening the spectrum from 3.5–5.0 μm. In contrast, the higher-pressure water atmosphere has more structure at these longer wavelengths. The results of these water atmosphere, pressure-varying fits are shown in the middle panel of Figure 6.

Given a pure water atmosphere at the 580 K temperature of GJ 1132 b, there are few if any cloud species (e.g., P. Gao et al. 2020) likely to form to produce the low-pressure cloud tops favored by our forward models. Hazes could potentially form in enough abundance at these pressures (e.g., S. E. Moran et al. 2018; C. He et al. 2024) to fully mute out water absorption at deeper layers, or the planet could simply have a very tenuous atmosphere dominated by water vapor.

Finally, we examine whether we can refute or confirm the best-fit atmospheric models allowed by the emission study of GJ 1132 b reported in Q. Xue et al. (2024). Their best-fit atmospheric models (based on the lowest χ^2_ν values reported in their Figure 5) included (1) a 1 bar O_2 atmosphere with 100 ppm H_2O and (2) a 10 bar O_2 atmosphere with 1 ppm CO_2. Both of these models fit their emission spectrum with a $\chi^2_\nu < 1$. We reproduce these conditions and compute transmission spectrum models (with an isothermal temperature–pressure profile set to 580 K) to compare to our transmission data. Their 100 ppm H_2O in an O_2 atmosphere is fully consistent with our data and similarly fit with $\chi^2_\nu \leqslant 1$ for each of our `Eureka!`, `ExoTiC-JEDI`, and `FIREFLy` reductions. However, we are able to rule out their 10 bar, 1 ppm CO_2 in O_2 atmosphere with moderate confidence, rejecting this possibility by over 3σ from both the `Eureka!` and `ExoTiC-JEDI` reductions and to 1.6σ from the `FIREFLy` reduction. These models are shown in the bottom panel of Figure 6. The combination of emission and transmission spectra together can clearly winnow down the allowed potential atmospheric scenarios beyond either set of observations individually (M. Weiner Mansfield et al. 2024), with transmission observations more sensitive to more tenuous atmospheres. See Section 5.3 for a more complete discussion of this.

4.4. Atmospheric Retrievals Favor Thin or No Atmosphere and Reveal the First Visit Dominates Interpretation

We also interpret GJ 1132 b's transmission spectrum via Bayesian atmospheric retrievals. We use the open source retrieval

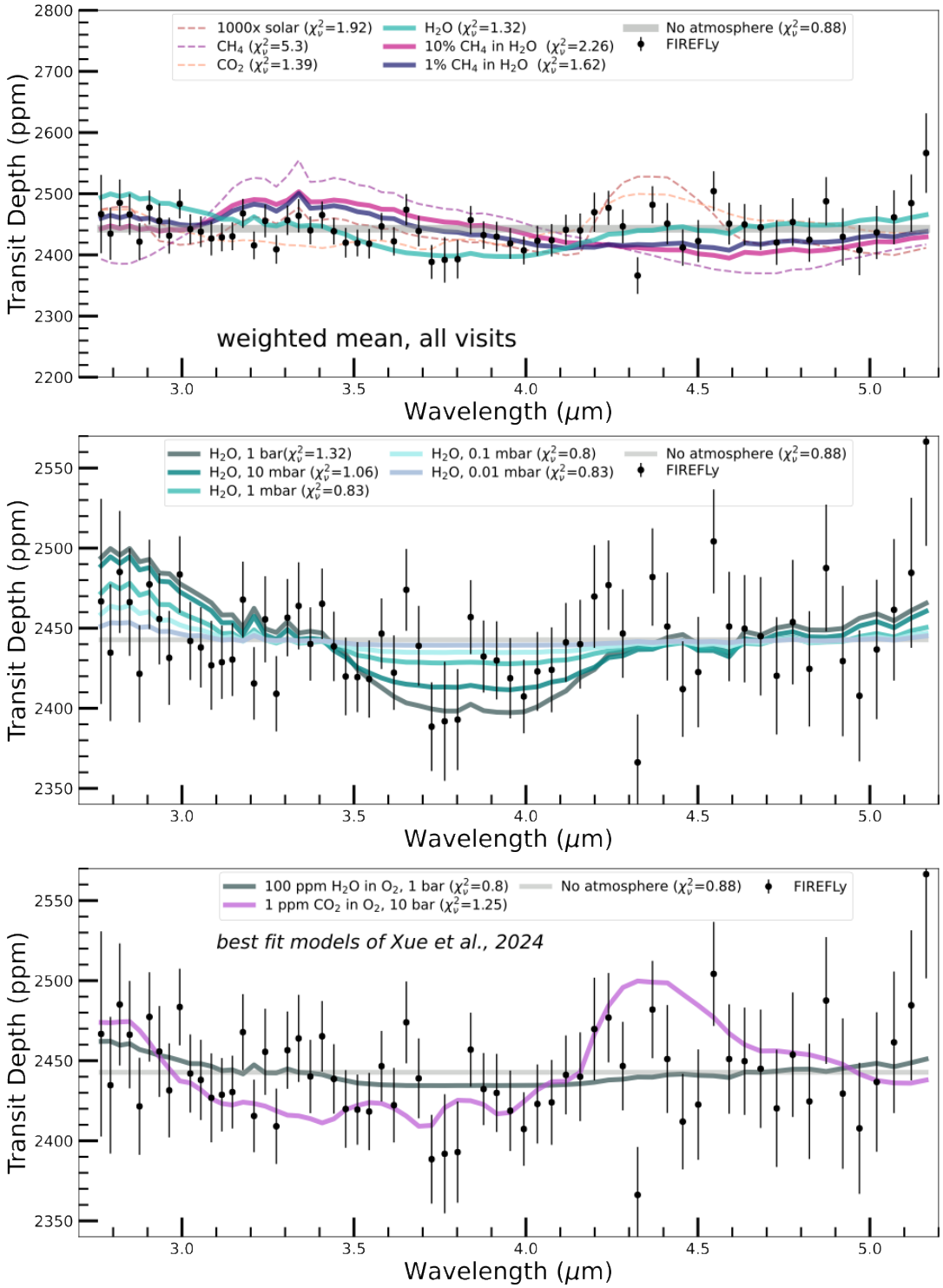

Figure 6. Forward model fits compared to the weighted mean of all four visits are shown for the FIREFLy reduction. χ^2_ν values, shown in the legend, are calculated assuming 64 degrees of freedom. Top panel: forward model fits for a variety of clear atmospheres, similar to those shown in E. M. May et al. (2023). The best-fit of those shown is a flat line model. Middle panel: forward model fits for a pure water atmosphere with varying opaque pressure levels, which can represent either a surface pressure or that of a cloud-top. A water atmosphere approximately 1 mbar in pressure is more likely than a flat line model (see the text for details). Bottom panel: the two best-fitting atmospheric forward models from the MIRI emission study of Q. Xue et al. (2024) compared to our transmission spectra. We rule out their 1 ppm CO_2 in O_2 atmosphere to $\geqslant 3\sigma$ with our data while their 100 ppm H_2O case is consistent with our transmission spectrum.
(The complete figure set (3 images) is available in the online article.)

code POSEIDON (R. J. MacDonald & N. Madhusudhan 2017; R. J. MacDonald 2023), following a similar approach to recent studies (E. M. May et al. 2023; S. E. Moran et al. 2023) whereby we explore planetary atmospheres, stellar contamination, and flat spectra as potential explanations. Given the high-precision of our multivisit GJ 1132 b transmission spectrum, we additionally

Table 4
Atmospheric Retrieval Parameters and Priors

Parameter	Prior
Atmosphere Model	
Reference Radius ($R_{p,ref}$)	$\mathcal{U}(0.9605, 1.2995)\ R_\oplus$
Surface Pressure ($\log_{10}(P_{surf}$ / bar))	$\mathcal{U}(-7, 2)$
"Ghost" Mean Molecular Weight (μ_{back})	$\mathcal{U}(2.3, 100)$
Atmospheric Temperature (T)	$\mathcal{U}(200, 900)$ K
Molecule Volume Mixing Ratios ($\log_{10} X_i$)	$\mathcal{CLR}(-12, 0)$
Stellar Contamination Model	
Planetary Radius (R_p)	$\mathcal{U}(0.9605, 1.2995)\ R_\oplus$
Photosphere Temperature (T_{phot})	$\mathcal{N}(3270, 140^2)$ K
Heterogeneity Temperature (T_{het})	$\mathcal{U}(2300, 3924)$ K
Heterogeneity Covering Fraction (f_{het})	$\mathcal{U}(0, 0.6)$
Flat Line Model	
Planetary Radius (R_p)	$\mathcal{U}(0.9605, 1.2995)\ R_\oplus$

Note. The mixing ratios of H_2O, CH_4, CO_2, CO, SO_2, N_2, and H_2 are free parameters, with the mixing ratio of the "ghost" gas satisfying the summation to unity constraint. "CLR" refers to the centered log-ratio prior, which allows any gas to have equal probability a priori to be the background gas (e.g., B. Benneke & S. Seager 2012). The Gaussian prior for T_{phot} is expressed as $\mathcal{N}(\mu, \sigma^2)$, where μ and σ are the mean and standard deviation, respectively, for the stellar effective temperature from X. Bonfils et al. (2018).

explore joint constraints on the mean molecular weight and (apparent) surface pressure for allowable atmospheres.

4.4.1. Retrieval Configuration

We consider three retrieval model categories with `POSEIDON`: (i) a planetary atmosphere with no stellar contamination; (ii) stellar contamination from unocculted stellar heterogeneities; and (iii) a flat line. We also consider a sub-model of the atmosphere category with a pure H_2O composition (i.e., a "steam world") motivated by recent studies (E. M. May et al. 2023; C. Piaulet-Ghorayeb et al. 2024). We apply each retrieval model to the three data reductions described above, and for each reduction, we consider separate retrievals for the four-visit combined data set and three-visit spectra with one visit excluded. The latter "leave-one-out" retrievals allow us to assess the robustness of results derived from all four visits combined. All of our retrieval models use 10,000 `PyMultiNest` (F. Feroz et al. 2009; J. Buchner et al. 2014) live points and calculate model transmission spectra at a resolving power of $R = 20{,}000$ from 2.6–5.3 μm.

Our atmosphere models consider a variety of potential compositions. We consider seven specific gases (H_2, H_2O, CO_2, CO, CH_4, N_2, and SO_2) alongside a "ghost" gas with a freely fit mean molecular weight. Any of these eight gases can dominate the atmospheric composition, since we use the background-gas agnostic centered log-ratio prior (e.g., B. Benneke & S. Seager 2012; J. Lustig-Yaeger et al. 2023). The molecular cross sections used in the retrieval forward model's radiative transfer uses the opacity database corresponding to `POSEIDON` v1.2 (E. Mullens et al. 2024), which uses the following ExoMol (J. Tennyson & S. N. Yurchenko 2018; J. Tennyson et al. 2024) line lists: H_2O (O. L. Polyansky et al. 2018), CO_2 (S. N. Yurchenko et al. 2020), CO (G. Li et al. 2015), CH_4 (S. N. Yurchenko et al. 2024), and SO_2 (D. S. Underwood et al. 2016). Collision-induced absorption from H_2 to H_2 and pairs with N_2 is also included from HITRAN (T. Karman et al. 2019). Each model atmosphere spans 10^{-7}–100 bar, split into 100 layers spaced uniformly in log-pressure, with a freely fit effective surface pressure (representing either a solid surface or cloud-top) and is assumed to be isothermal with pressure. We additionally fit for the atmospheric radius at the 10 bar reference pressure. The priors for this 11-parameter model are listed in Table 4.

Table 5
Retrieval Model Statistics

Data + Model	N_{param}	$\ln \mathcal{Z}$	χ^2	dof	χ^2_ν
`FIREFLy`					
Four Visits					
Atmosphere	11	575.03	49.0	53	0.92
Stellar Contam.	4	575.44	50.8	60	0.85
Thin Steam World	3	576.36	50.6	61	0.83
Flat	1	575.45	56.1	63	0.89
Three Visits (No Visit 1)					
Atmosphere	11	563.59	53.0	53	1.00
Stellar Contam.	4	562.23	58.8	60	0.98
Thin Steam World	3	563.37	58.0	61	0.95
Flat	1	563.47	60.7	63	0.96
`Eureka!`					
Four Visits					
Atmosphere	11	570.62	56.3	53	1.06
Stellar Contam.	4	571.68	61.4	60	1.02
Thin Steam World	3	572.50	60.4	61	0.99
Flat	1	569.83	69.5	63	1.10
Three Visits (No Visit 1)					
Atmosphere	11	561.76	59.5	53	1.12
Stellar Contam.	4	561.36	63.5	60	1.06
Thin Steam World	3	562.52	62.3	61	1.02
Flat	1	562.51	65.0	63	1.03
`ExoTiC-JEDI`					
Four Visits					
Atmosphere	11	519.63	52.3	47	1.11
Stellar Contam.	4	521.89	52.3	54	0.99
Thin Steam World	3	521.74	55.3	55	1.01
Flat	1	518.59	65.7	57	1.15
Three Visits (No Visit 1)					
Atmosphere	11	514.76	48.8	47	1.04
Stellar Contam.	4	513.98	53.5	54	0.99
Thin Steam World	3	514.75	53.7	55	0.98
Flat	1	514.89	56.2	57	0.99

Note. The models are ordered by complexity (N_{param}). $\mathcal{Z}$ is the Bayesian evidence, dof is the degrees of freedom ($N_{data} - N_{param}$), and χ^2_ν is the reduced chi-squared.

Our stellar contamination retrieval model uses a one-heterogeneity model (e.g., A. D. Rathcke et al. 2021). We initially explored a two-heterogeneity model (i.e., including both cold and hot active regions; e.g., M. Fournier-Tondreau et al. 2024), but found no need for a second heterogeneity population. Therefore, we focus on the one-heterogeneity model in what follows. This four-parameter model is parameterized by the planetary radius, the stellar photosphere temperature, the heterogeneity covering fraction, and the stellar heterogeneity temperature (corresponding to spots if the heterogeneity is cooler than the photosphere, or faculae for the converse). The stellar spectra of the photosphere and heterogeneities are calculated by interpolating `PHOENIX` models (T. O. Husser et al. 2013) using the `PyMSG` package

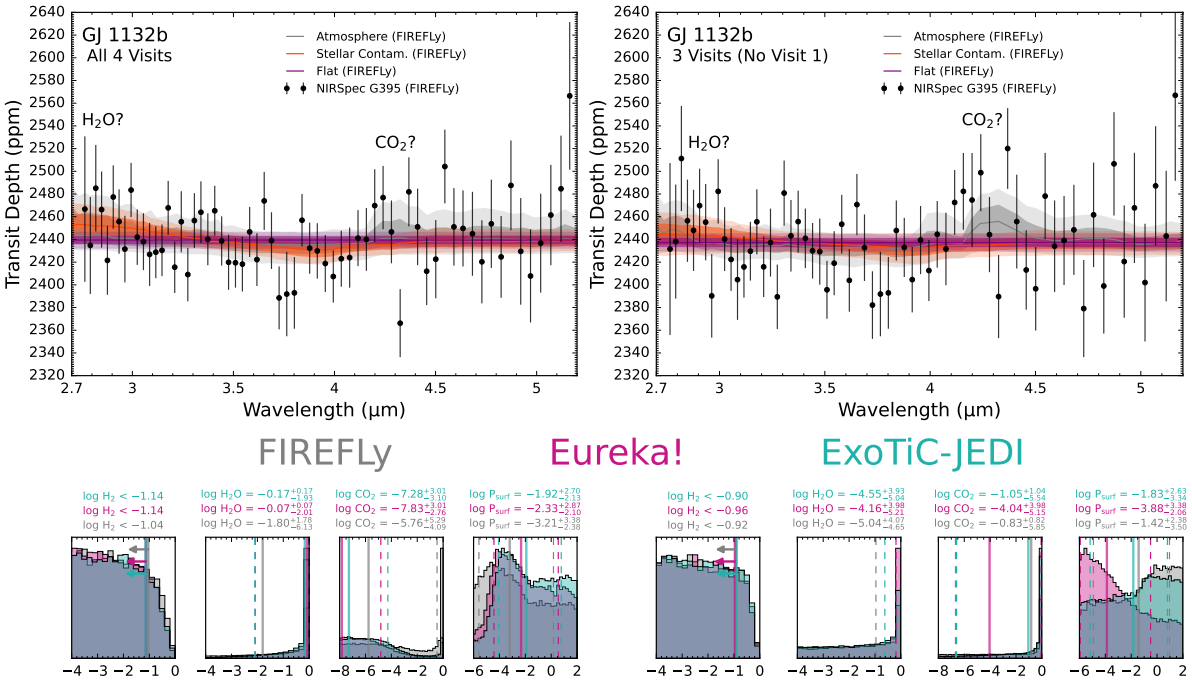

Figure 7. Retrieval results for GJ 1132 b. Top panels: comparison between the retrieved transmission spectra for the `FIREFLy` reduction with all four visits (top-left panel) and for three visits with Visit 1 excluded (top-right panel). Three `POSEIDON` retrieval models are shown: (i) an atmosphere (gray), (ii) stellar contamination (orange), and (iii) a flat line (purple). The corresponding median retrieved spectra (solid lines) and 1σ and 2σ model confidence intervals (dark and light contours) are overlaid. Labels indicate H_2O and CO_2 absorption bands (neither are detected, but CO_2 is allowed by `FIREFLy`). Bottom panels: atmosphere posterior histograms for the three data reductions: `FIREFLy` (gray), `Eureka!` (pink), and `ExoTiC-JEDI` (cyan). The four-visit combined analysis weakly favors a thin H_2O-rich atmosphere (driven by a short-wavelength slope), but this solution is not favored when Visit 1 is excluded.

(R. Townsend & A. Lopez 2023). We adopt the stellar radius and effective temperature from X. Bonfils et al. (2018; $R_* = 0.2105\,R_\odot$; $T_{\rm eff} = 3270$ K), the stellar surface gravity from J. Southworth et al. (2017; $\log_{10} g_* = 4.88$), and the stellar metallicity from Z. K. Berta-Thompson et al. (2015; $[{\rm Fe/H}] = -0.12$). The priors are also summarized in Table 4.

4.4.2. Retrieval Results

Our GJ 1132 b retrieval analysis favors a thin H_2O-rich atmosphere when using all four visits, in agreement with our forward modeling analysis (Section 4.3.) A thin ($\sim 10^{-4}$ bar) pure H_2O atmosphere maximizes the Bayesian evidence and is preferred over both a flat line and stellar contamination for all three data reductions (Table 5). The Bayes factors for the simple three-parameter pure H_2O model compared to a flat line are 2.5 (`FIREFLy`), 14 (`Eureka!`), and 23 (`ExoTiC-JEDI`). The lower significance from `FIREFLy` arises from the slightly larger error bars in certain channels in this reduction, which also manifests in the smaller reduced-χ^2 values compared to `Eureka!` and `ExoTiC-JEDI` (Table 5). The improvement in the Bayesian evidence between the pure H_2O model and the stellar contamination model is not statistically significant, so these NIRSpec G395 data cannot distinguish between stellar contamination and an atmosphere (as in previous studies, e.g., S. E. Moran et al. 2023).

Figure 7 shows the retrieved spectra fits to the `FIREFLy` data and the posterior distributions for the atmosphere model for all three data reductions. We show the results from the "full" atmosphere model, rather than the simple pure H_2O model, since the full model encompasses the best-fitting pure H_2O solution while also marginalizing over other potential atmosphere components. While our retrievals identify H_2O as the most likely background gas for all three data reductions, we see that the H_2O abundance posterior displays a long tail to lower abundances without a clear lower limit. Compared to the results from E. M. May et al. (2023; for Visit 1 only), the weaker H_2O abundance constraints here are due to our four-visit average spectrum being flatter, which requires a low apparent surface pressure peaking at $\sim 10^{-4}$ bar. In addition to the preference for H_2O, the `FIREFLy` data tentatively allow for the presence of CO_2 (see Figure 7), but the `Eureka!` and `ExoTiC-JEDI` data do not indicate any evidence for CO_2—this likely arises from small differences in the data scatter near 4.3 μm. Our data also rules out H_2-rich atmospheres, with 2σ upper limits of $\approx 10\%$ H_2 (Figure 7), even after marginalizing over surface pressure, for all three data reductions.

Given that our four-visit retrieval analysis finds weaker evidence for an atmosphere than the Visit-1-only results from E. M. May et al. (2023), we repeated our retrievals with three-visit average spectra excluding one visit in turn. This leave-one-out test allows us to assess whether a single visit drives our inferences. We find generally similar results when excluding Visits 2, 3, or 4 to the results described above, so we do not discuss these further. However, the exclusion of Visit 1 significantly alters the retrieval results. Therefore, in what follows, we present the retrieval results for both the four-visit and three-visit (with Visit 1 excluded) data sets on equal grounds.

With Visit 1 excluded, we find that GJ 1132 b's transmission spectrum is statistically flat with no evidence for atmospheric absorption. Table 5 demonstrates that the thin pure H_2O model that was preferred for the four-visit spectrum has a similar or slightly lower Bayesian evidence than the flat model for all three data reductions, indicating that a flat model

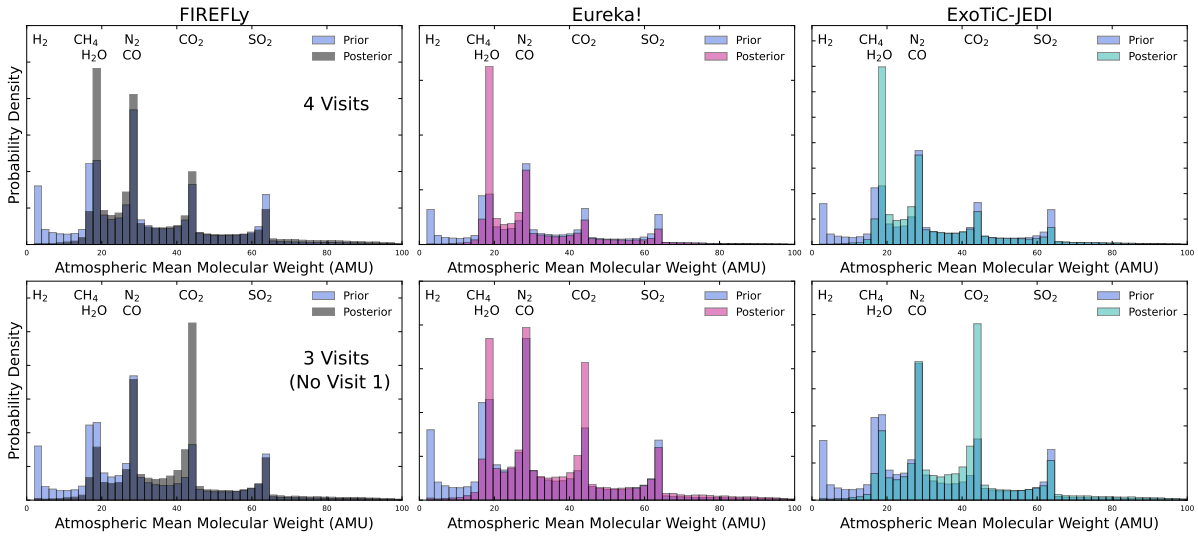

Figure 8. Mean molecular weight constraints for GJ 1132 b's atmosphere. Top panels: constraints from all four visits combined, comparing the prior distribution (blue) with the posteriors from `FIREFLy` (gray; left), `Eureka!` (pink; middle), and `ExoTiC-JEDI` (cyan; right). Bottom panels: constraints from three visits with Visit 1 excluded. The peaks in the prior at ∼100% abundances of each considered molecule (annotated) arise from the use of the centered log-ratio prior (e.g., B. Benneke & S. Seager 2012), which formally encodes that all gases are equally likely to be the background gas. Higher relative prior amplitudes arise from multiple molecules overlapping in a given mean molecular weight bin (e.g., N_2 and CO both have $\mu = 28$). The four-visit results favor an H_2O-dominated composition for all three reductions, but without Visit 1, this solution is disfavored relative to higher mean molecular weights (e.g., CO_2-dominated atmospheres). H_2-dominated atmospheres are excluded in all cases.

sufficiently describes the data (see Figure 7). Should GJ 1132 b possess an atmosphere, the transmission spectrum can be rendered flat by a very low surface pressure (`Eureka!`), or a high mean molecular weight dominated by CO_2 (`FIREFLy` and `ExoTiC-JEDI`). Nevertheless, atmospheres rich in H_2 remain ruled out with a 2σ upper limit of $\approx 10\%$ H_2 for the three visits without Visit 1 (see Figure 7).

Our atmospheric mean molecular weight constraints for GJ 1132 b are shown in Figure 8. The prior probability distribution for the mean molecular weight displays distinct "spikes" at the mean molecular weight (given in AMU, atomic mass units) of specific molecules ($H_2 = 2$, $CH_4 = 17$, $H_2O = 18$, $N_2 = CO = 28$, $CO_2 = 44$, $SO_2 = 64$), corresponding to atmospheres dominated by one of the considered gases. Atmospheres dominated by the molecular weight of a single gas are a manifestation of the centered log-ratio prior, which allows every gas an a priori equal prior probability of being the dominant background gas. Since at least one gas must have a significant abundance in order for the mixing ratios to sum to unity, the prior probability for a mean molecular weight corresponding to one of the considered gases is higher than intermediate mean molecular weights produced by a combination of different gases. We note that the prior distribution for $\mu_{\rm atm} > 64$ arises from the "ghost" gas in our retrieval model, which we allow up to 100. While the ghost gas also follows the CLR prior, since its mean molecular weight is also a free parameter, it produces a diffuse prior density across the parameter space. The prior density falls toward the upper limit ($\mu_{\rm atm} = 100$) because for high values of $\mu_{\rm atm}$, there are fewer samples where the ghost gas is both the dominant gas and has a high randomly sampled mean molecular weight. In summary, the prior distribution in Figure 8 is a reflection of our prior knowledge that not all mean molecular weights are equally likely a priori, since there are characteristic molecular weights corresponding to the specific molecules that can dominate an atmosphere.

We find robust lower limits on GJ 1132 b's atmospheric mean molecular weight. When considering all four visits, we find 99.7% credible interval lower limits of $\mu_{\rm atm} > 7.5$ (`FIREFLy`), 7.6 (`Eureka!`), and 8.8 (`ExoTiC-JEDI`). The equivalent limits excluding Visit 1 are $\mu_{\rm atm} > 5.9$ (`FIREFLy`), 6.2 (`Eureka!`), and 6.5 (`ExoTiC-JEDI`). Therefore, an H_2-rich atmosphere is ruled out to high significance across all three data reductions, regardless of the inclusion of Visit 1. While a specific atmospheric composition is not uniquely identified for our atmospheric retrieval model considering multiple gases, Figure 8 demonstrates that the most likely composition from the four-visit spectrum is an H_2O-dominated atmosphere. The `Eureka!` and `ExoTiC-JEDI` reductions yield the highest posterior probability for an H_2O-dominated atmosphere, but the `FIREFLy` data allows for a significant contribution from N_2. However, the three-visit spectrum without Visit 1 produces significantly different solutions. Figure 8 shows that the relative posterior probability shifts from the H_2O-rich peak to atmospheres dominated by CO_2 (`FIREFLy` and `ExoTiC-JEDI`) or N_2 (`Eureka!`). In essence, the flatter transmission spectrum without Visit 1 (Figure 7) requires a higher mean molecular weight than even a pure H_2O atmosphere can provide, resulting in a probability density shift toward heavier background gases (in particular, CO_2).

Given the significant sensitivity of GJ 1132 b's transmission spectrum and retrieval results to the inclusion of Visit 1, we briefly consider whether time-variable stellar contamination could offer an explanation, as in Section 4.2. We compare the flux-calibrated stellar spectra comparison approach to the retrieval approach to determine if the retrieval approach can also tease out meaningful information on the cooler/warmer spot coverage fractions and temperatures on a visit-to-visit basis. We would also like to understand if the constraints in the retrieval approach differ between the G395H and G395M modes. For this, we test a single G395H visit (Visit 1) and a

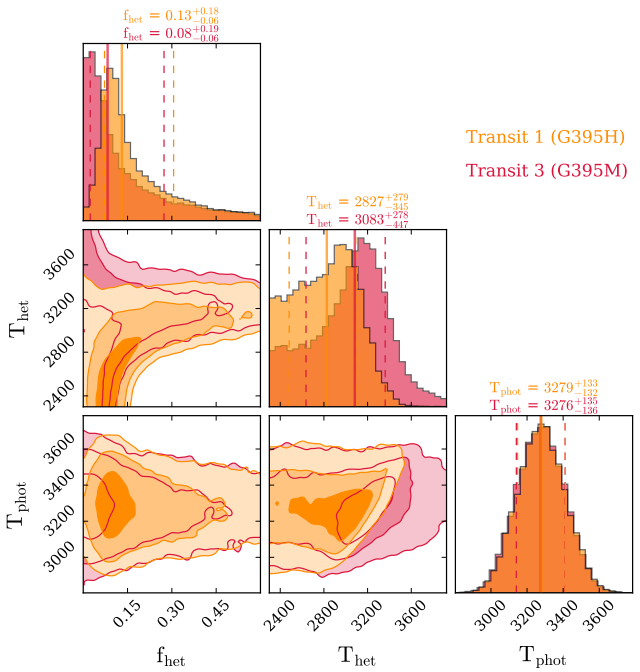

Figure 9. Retrieved stellar properties from GJ 1132 b's FIREFLy transmission spectra from Visit 1 (NIRSpec G395H) and Visit 3 (NIRSpec G395M). These stellar contamination retrievals assume all spectral features arise from unocculted heterogeneities (i.e., no planetary absorption). The slope seen in Visit 1 can be explained with a moderate starspot covering fraction ($f_{\rm het} \sim 10\%$), but the relatively flat Visit 3 spectrum does not require unocculted starspots ($f_{\rm het} \rightarrow 0$).

single G395M visit (Visit 3). Figure 9 shows retrieval results only using a stellar contamination retrieval model. We see that the spectral slope exhibited in the Visit 1 transmission spectrum requires $13^{+18}_{-6}\%$ starspot coverage to be explained by stellar contamination, but Visit 3 does not require stellar contamination ($f_{\rm het} \rightarrow 0$), since this visit is statistically flat. However, the two posterior distributions are formally consistent within 1σ, given the strong $f_{\rm het}$–$T_{\rm het}$ degeneracy, so the presence of time-variable stellar contamination cannot be established from the individual visit spectra alone using this technique, illustrating that analyzing the flux-calibrated stellar spectrum (Section 4.2) may be more information-rich.

Finally, since different surface pressures can affect the viability of atmospheric compositions differently, we present the joint posteriors for surface pressure and mean molecular weight in Figure 10. This visual allows for a direct comparison between our GJ 1132 b retrieval results and the solar system terrestrial atmospheres (Venus, Mars, Earth, and Titan), three exoplanets with well-measured atmospheric mean molecular weights heavier than H_2+He (B. Benneke et al. 2024; C. Piaulet-Ghorayeb et al. 2024; S. P. Schmidt et al. 2025), and previous atmospheric constraints for GJ 1132 b from JWST emission spectroscopy (Q. Xue et al. 2024). In Figure 10, the CLR prior spikes seen in Figure 8 translate to columns in the $P_{\rm surf}$–$\mu_{\rm atm}$ parameter space.

Figure 10 demonstrates that the H_2O-rich atmospheric solution for the four-visit combined spectrum is localized in a thin pressure region centered near 10^{-4} bar. This solution is consistent with the GJ 1132 b surface pressure upper limit from Q. Xue et al. (2024) for a pure H_2O composition ($P_{\rm surf} \lesssim 10^{-1}$ bar) and with the "steam world" atmosphere detected for GJ 9827 d (C. Piaulet-Ghorayeb et al. 2024). However, the removal of Visit 1 shifts the solution space to favor heavier atmospheres. For three visits, the Eureka! reduction favors a thin ($P_{\rm surf} < 10^{-3}$ bar) N_2-rich atmosphere, while the FIREFLy and ExoTiC-JEDI reductions favor thick ($P_{\rm surf} > 10^{-1}$ bar) CO_2-rich atmospheres. Given the featureless spectrum without Visit 1, small differences between the reductions alter the relative probability between the thin N_2-rich and thick CO_2-rich solutions. We note that the GJ 1132 b MIRI LRS emission spectrum from Q. Xue et al. (2024) strongly disfavors a thick CO_2-rich atmosphere, so our transmission spectra results should not be interpreted as evidence for such an atmosphere, despite the low significance possibility of 4.3 μm CO_2 absorption shown in Figure 7. Our most robust result is that GJ 1132 b cannot have an atmosphere with a mean molecular weight <6, for any surface pressure $>10^{-7}$ bar, which rules out both H_2-rich atmospheres and a miscible supercritical envelope similar to that favored for TOI-270d by B. Benneke et al. (2024; $\mu_{\rm atm} \approx 5$).

The extreme sensitivity of our retrieval results to the inclusion of Visit 1 suggests that this Visit is a statistical outlier, and therefore that the most reliable constraints on GJ 1132 b's atmosphere should be drawn from the three-visit transmission spectrum with Visit 1 excluded. For this data set, we find a nondetection of atmospheric absorption, in particular a lack of H_2O absorption, which can result from GJ 1132 b being either a bare rock or possessing a thin atmosphere dominated by heavy molecules such as N_2 or CO_2.

5. Discussion

In this section, we discuss both the technical implications of utilizing the NIRSpec/G395M mode and stacking multiple transits as well as the astrophysical implications for GJ 1132 b.

5.1. Mind the Gap: NIRSpec G395M versus G395H

For transmission spectroscopy with JWST targeting the 3–5 μm range, the exoplanet community by and large has utilized NIRSpec/G395H (or NIRSpec/PRISM for fainter targets). To date, only two other studies have observed with G395M (A. W. Mayo et al. 2025; B. Davenport et al. 2025). (Note that there are multiple ongoing/upcoming programs intending to use this mode, but none have been published as of this writing.)

Importantly, our observations are the first time both the G395H and G395M mode have been utilized for a single target, offering a critical opportunity to compare the two modes. G395M is an attractive alternative to the higher-resolution mode primarily as a way to avoid the notorious detector gap with G395H, which splits the incoming light between the NRS1 and NRS2 detectors with a gap between ~3.75 and 3.82 μm (see Figure 3). This allows for greater spectral resolution, but there has not actually yet been evidence that suggests the higher resolution afforded by G395H improves abundance constraints for detected molecules compared to G395M (A. Guzmán-Mesa et al. 2020). Additionally, the use of G395H comes at the price of needing to characterize any potential offsets between the data sets due to this gap, which can confound results (e.g., N. Madhusudhan et al. 2023). With the G395M mode, on the other hand, light falls solely on the NRS1 detector while still covering the full ~2.9–5.1 μm wavelength range, allowing us to avoid the gap.

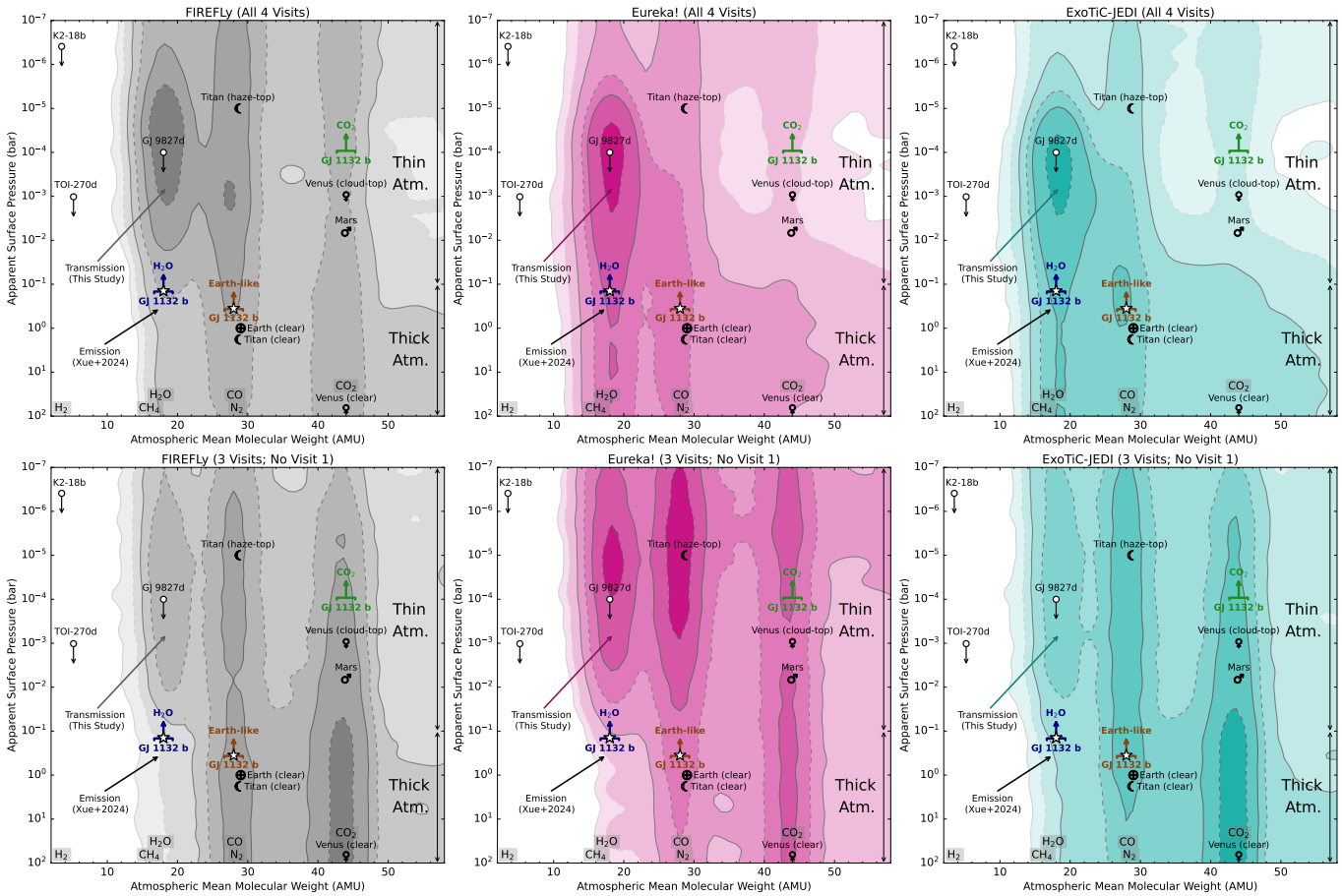

Figure 10. Surface pressure vs. mean molecular weight constraints for GJ 1132 b. Top panels: posterior probability density (colored contours) from four NIRSpec G395H/M transits combined from `FIREFLy` (gray; left), `Eureka!` (pink; middle), and `ExoTiC-JEDI` (cyan; right). Bottom panels: the same as the top panels, but excluding Visit 1. The 1σ, 2σ, and 3σ credible regions have solid line boundaries, while the 0.5σ, 1.5σ, and 2.5σ credible regions have dashed boundaries. Pressure upper limits from emission spectroscopy of GJ 1132 b (2σ limits from Q. Xue et al. 2024) are annotated for 100% H_2O (blue arrow), an Earth-like composition (brown arrow), and 100% CO_2 (green arrow). Three exoplanets with detected atmospheres and well-constrained mean molecular weights heavier than pure H_2+He (TOI-270d; B. Benneke et al. 2024; GJ 9827d; C. Piaulet-Ghorayeb et al. 2024; and K2-18b; S. P. Schmidt et al. 2025) are annotated (circles with arrows, located at the 2σ lower limit on the surface / cloud pressure), alongside terrestrial atmospheres in the solar system. GJ 1132 b's transmission spectrum favors a thin $\sim 10^{-4}$ bar H_2O-dominated atmosphere when using all four visits, but the exclusion of Visit 1 allows for a wider range of heavier and/or thin atmospheres (or no atmosphere) as viable ways to produce a featureless three-visit transmission spectrum (see Figure 7).

In Figure 3, we see that the data for Visits 3 and 4 (the G395M visits) bridge the detector gap and "fill in the blank" very smoothly. Visually, there is no apparent need for an offset between NRS1/NRS2 in the first two visits (the G395H visits). Though large offsets have been seen in G395H data ($\sim$10 s of ppm), this is frequently seen in data sets observing bright targets with few ($\leqslant$3) groups per integration (S. E. Moran et al. 2023, L. Alderson et al. 2024; N. L. Wallack et al. 2024; A. Bello-Arufe et al. 2025). Our G395H visits, on the other hand, use 14 groups per integration. This hints that this detector offset is driven by poor constraints on the up-the-ramp fit due to this low number of groups. This could subsequently lead to increased systematics in the light curves, necessitating the need for more complex models and thus inducing offsets. However, offsets are also seen in observations with as many as six or seven groups per integration (e.g., A. Gressier et al. 2024; L. Alderson et al. 2025). It could be, therefore, that (1) approximately >7 groups for G395H data are required in order to mitigate detector offsets, or (2) the offsets are unrelated to the number of groups and instead are a result of some other underlying instrumental effect. Overall, coupling the G395M mode provides an important check on whether an offset is needed in G395H data.

Beyond detector gaps, we can evaluate the usefulness of G395H versus G395M based on noise levels seen in the two modes. As far as the light-curve noise levels, we see comparable levels of photon noise reached between G395M and G395H for both the white and spectroscopic light curves. Additionally, the standard deviation and mean absolute deviation in the white light-curve fits are comparable between modes, as illustrated in Table 6, which illustrates these statistics for Visits 1 (G395H) and 3 (G395M). Comparable values are seen in the other visits for each mode. We note that this may be somewhat reduction dependent, as `FIREFLy` and `Eureka!` have marginally larger SD/MAD values for G395M relative to G395H/NRS1 (but marginally smaller values compared to G395H/NRS2), but `ExoTiC-JEDI` has G395M values slightly worse than both NRS1 and NRS2. Nevertheless, there are no glaring differences in white light noise levels between the observing modes.

Additionally, we observe no meaningful differences in correlated noise between the detectors, as visualized by the

Table 6
White Light Curve and Transmission Spectrum Statistics for G395H vs. G395M

	Photon Noise	SD (ppm)	MAD (ppm)	Spectroscopic Error (ppm)
G395H (NRS1)	1.4×	131	88	48
G395H (NRS2)	1.6×	177	123	58
G395M	1.4×	154	103	60

Note. Visit 1 (G395H) and Visit 3 (G395M) are shown for the `FIREFLy` reduction. The first three columns refer to the white light-curve fits. The level of photon noise is calculated by dividing the observed scatter by the predicted pipeline scatter, SD is standard deviation, and MAD is mean absolute deviation. Median spectroscopic error is shown in the final column. Statistics are comparable between G395H and G395M.

white light red noise diagnostic plots shown in the top panel of Figure 11. All four visits bin down close to photon noise. Likewise, the spectroscopic red noise diagnostic plots are similar, with a similar number of wavelength channels requiring error inflation between G395H and G395M, as described in Section 3.1. We note, however, that this is also somewhat reduction dependent. `FIREFLy` sees a similar number of channels requiring error inflation, but `Eureka!` requires more channels for error inflation in G395M, although this is likely due to the simpler systematics model employed by `Eureka!`.

Still, at the end of the day, this error inflation technique does not meaningfully impact the average error bars between the visits, which is also comparable across the observing modes, regardless of reduction. Median error bar sizes for Visits 1 and 3 (binned to the same resolution wavelength grid, as in Figure 3) are shown in Table 6. G395M sees a very small increase in median (and mean) error bar size compared to G395H—on the order of 5 ppm per visit for all reductions—but this is well within the per-visit error (∼50–60 ppm).

One positive benefit of utilizing G395H is the possibility of doing isotopologue investigations using cross-correlation with the pixel-resolution data (e.g., E. Esparza-Borges et al. 2023). However, this is a new strategy that is likely only applicable to targets with large signals with JWST, such as brown dwarfs and gas giants. Rocky planets have such small and noisy transit depths in the pixel-resolution data that the signal is insufficient for this type of science.

Our last remaining point in evaluating the two modes is that G395H does allow for more information-rich flux-calibrated stellar spectra. As shown in Section 4.2, examining the flux-calibrated spectra is a powerful tool to assess underlying stellar heterogeneity that is missed in the transmission spectrum. The precision of spot coverage/temperature constraints is comparable between the modes. However, the increased spectral detail for G395H allows for greater flexibility in model matching. Higher resolution provides more data points and finer structure in the observed spectrum, which can accommodate a broader range of models that capture different physical conditions. This effectively reduces degeneracies and makes it easier for multiple models to achieve a statistically acceptable fit. In contrast, the G395M spectra contain fewer constraints, making it harder for many models to simultaneously match the observed features. Whether or not the stellar models are optimally extracting this possible source of extra information

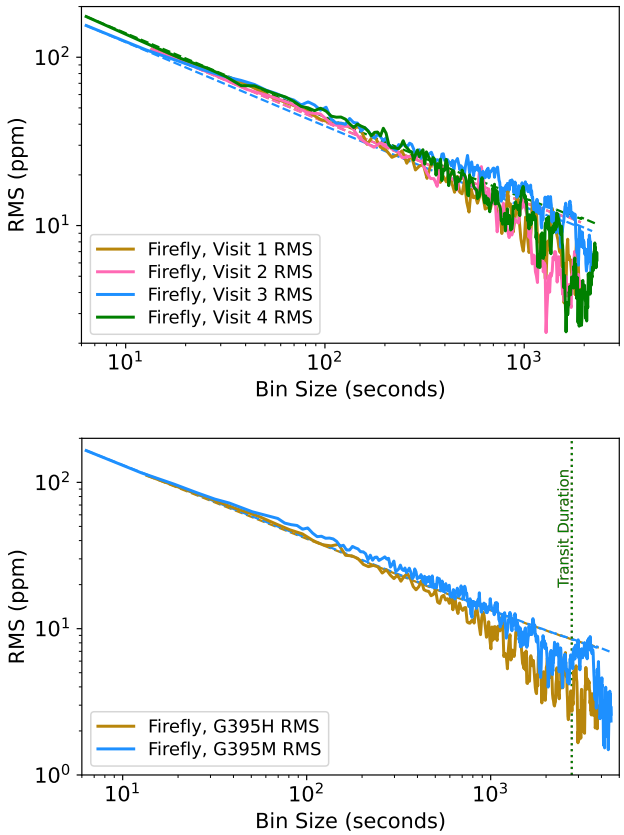

Figure 11. Red noise diagnostic plot (depicting the rms of the residuals as a function of bin size in time) for the residuals from the four `FIREFLy` visits (top panel) and each observing mode with residuals from visits concatenated (bottom panel). rms as a function of bin size is shown, with the dotted lines representing the expected decrease in noise with the square root of bin size ($\sqrt{N}$). Bin size is in seconds, not number of integrations, in order to appropriately compare G395H and G395M. There is no apparent difference in correlated noise between the G395H (Visits 1 and 2) and G395M (Visits 3 and 4) observations, nor is there an indication that an instrumental noise floor is being reached.

from the higher-resolution spectra is a topic for a separate investigation.

Overall, the decision to utilize G395M versus G395H should be made on a case-by-case basis. A balance must be struck between concerns about systematics and concerns of underlying stellar heterogeneity when stacking multiple transits. If observing sources with at least ∼8 groups per integration, and stacking multiple visits, it may be worth utilizing G395H, as it is less likely that NRS1/NRS2 offsets will appear in the data and more likely to need robust stellar heterogeneity measurements to ensure minimal visit-to-visit stellar heterogeneity. On the other hand, if transits are not being stacked and the source can be observed with G395M without saturating, G395M will suffice. We note that in practice, any target that requires few (less than ∼8) groups with G395H will likely saturate (or come close to saturating) G395M, so one may still be constrained to a given observing mode by the saturation limits of NIRSpec.

5.2. *Stacking Multiple Transits: A Treasure or a Curse?*

This study marks the second time as many as four transits for a single planet have been coadded together using JWST data, after L98-59b (A. Bello-Arufe et al. 2025). The hope of the rocky exoplanet community for many years has been to use

the power of stacking to bring down transmission spectra error bar size and reach the regime where tiny, 10–20 ppm signals are discernible from a secondary atmosphere composed of high mean molecular weight gases (e.g., L. Kaltenegger & W. A. Traub 2009; J. K. Barstow & P. G. J. Irwin 2016; C. V. Morley et al. 2017; J. Lustig-Yaeger et al. 2019). Additionally, coadding may be better suited for small planets, because even if the atmosphere shows some variability (e.g., changes in average cloud coverage), the amplitude of the changes in the spectra introduced by atmospheric variability are smaller than the error bars on individual transits. The same may not be true for giant planets.

Our study demonstrates that reaching these small signals is technically possible. The median error for all four individual transits is 55 ppm, which bins down to a median error of 28 ppm in the weighted mean. We also find that red noise in individual wavelength channels becomes less important when stacking visits, as those channels with inflated error bars due to red noise simply become weighted less.

Additionally, coadding transits still shows promise because there is still no indication that we are hitting an instrumental noise floor. We test this by concatenating the residuals from each observing mode and computing a red noise diagnostic plot for each, as shown in the bottom panel of Figure 11. We do not concatenate the G395M and G395H residuals, because for every integration of G395H (with 14 groups), there are 2.3 integrations of G395M (with six groups), so we cannot simply bin an integer number of G395M residuals to match G395H. Figure 11 illustrates that the residuals are still binning down with increasing bin size down to the duration of the transit itself, which tells us no noise floor is hit down to 8 ppm. Again, this shows promise for the potential of stacking transits.

However, the community is also now coming to terms with the reality of stacking transits in the face of stellar contamination that may impact the transmission spectrum visit-to-visit (e.g., O. Lim et al. 2023; M. Radica et al. 2025). Typically, the presence of stellar contamination is discerned by unocculted cold or hot spots that present themselves as unexpected slopes and/or water features in the spectrum itself (this was first postulated by F. Pont et al. 2008 and popularized as the "transit light source effect" by B. V. Rackham et al. 2018). Attempts to mitigate this include incorporating stellar contamination into retrievals (e.g., C. Cadieux et al. 2024) and using a transmission spectrum from another (presumably bare rock) planet in the same system as an empirical correction factor (e.g., A. D. Rathcke et al. 2025).

However, if there is no obvious sign of stellar variability from the weighted mean transmission spectrum, as we demonstrate in our study, there is no clear reason to suspect that there might be a problem stacking multiple transmission spectra together. However, as we see in Sections 4.2 and 4.4, Visit 1 is an outlier in terms of its stellar heterogeneity and outsized influence on retrieval results. By simply looking at the visit-by-visit transmission spectrum (see Section 3.4 and Figure 3), it appears that the apparent discrepancies between Visits 1 and 2 discussed in E. M. May et al. (2023) are likely random noise instances based on the scatter of the data seen in all four visits, with the notable exception of the curious disagreement at 4.5 μm between Visit 3 and the other visits (see Section 3.4). However, further investigation revealed that the random noise interpretation is not favored. Rather, the star may be changing in more subtle ways than previously assumed.

This exercise demonstrates that single visits can still dominate interpretation of planetary spectra when stacking multiple visits, even in the apparent face of little-to-no obvious stellar variability in the transmission spectrum. It is unclear if a single visit could have such a dominating influence if, say, 10 transits were stacked instead of four. There should be a limit in which a single transit, no matter how contaminated, becomes "diluted" in the face of many transits. In any case, moving forward, we recommend a "leave-one-out" approach in atmospheric retrievals and forward models to ensure that one visit is not dominating the results.

5.3. Comparison with MIRI/LRS Findings

We examine the transmission and emission constraints together by combining our retrieval results (using the three-visit case, excluding Visit 1) with the atmospheric constraints from Q. Xue et al. (2024). We compare our results to their emission spectroscopy constraints, which are more stringent compared to their atmospheric constraints from their white light eclipse depth.

In Figure 10, we overplot the Q. Xue et al. (2024) 2σ lower limit (see their Figure 5) on atmospheric pressures allowable for atmospheres of three end-member compositions: pure H_2O, Earth-like, and pure CO_2. With transmission alone (considering the case where we exclude Visit (1)), our data are equally consistent with thin (<100 mbar) atmospheres that are 14 AMU or greater, or thick (>100 mbar) atmospheres that are 25 AMU or greater. The latter scenario, however, is rejected by the emission data.

The emission 2σ lower limit of either a ~0.4 bar Earth-like atmosphere or a ~0.1 mbar CO_2 atmosphere agrees with our findings across reduction pipelines. These limits help contextualize our findings that, for example, our transmission data are consistent with a thick pure CO_2 atmosphere, but given the emission lower limit of 0.1 mbar CO_2, we can exclude this as a possibility. Thus, the emission limits for atmospheres >25 AMU generally place more stringent constraints on lower pressures than we can with transmission alone. (This is illustrated in the figure by the fact that the emission lower limit lies within or above our 1σ contour.) The only exception to this is in the case of the `Eureka!` reduction, which favors a very thin N_2-dominated Earth-like atmosphere. However, as discussed in Section 4.4, this demonstrates that different noise properties across reductions lead to slightly different preferences in retrieval results in the case of a featureless spectrum (as also seen in GJ 341b; J. Kirk et al. 2024), as opposed to implying that an N_2-dominated atmosphere is more likely than a CO_2 atmosphere.

In the case of the pure H_2O atmosphere, however, results from transmission spectroscopy are more constraining than those from emission (2σ limit of ~0.2 bar), as our results suggest only a very tenuous steam atmosphere is possible. To further test the possibility of a steam atmosphere in conjunction with emission data, we estimate the nightside temperature from the white light eclipse MIRI results (Q. Xue et al. 2024) to determine if a steam atmosphere is possible or if it would condense out on the nightside. Assuming $A_b = 0$ and the 1σ upper and lower limits from Q. Xue et al. (2024) on the heat redistribution efficiency, ϵ, we use Equation (5) from

N. B. Cowan & E. Agol (2011),

$$T_n = T_{\rm eff}\left(\frac{R_s}{a}\right)^{1/2}(1 - A_b)^{1/4}\left(\frac{\epsilon}{4}\right)^{1/4}, \quad (1)$$

to calculate the 1σ limits of the maximum nightside temperature, T_n. Q. Xue et al. (2024) reported a 1σ upper limit of $\epsilon = 0.52$ and 1σ lower limit of $\epsilon = 0.06$ (see their Figure 4), giving us $287 < T_n < 492$ K. At these temperatures, therefore, it is still possible from an energy transport perspective to have water vapor in the upper atmosphere.

This exercise demonstrates the utility of combining results from transmission and emission to provide a complementary and constraining picture of any possible atmosphere on a rocky exoplanet. Together, our results show that any atmosphere on GJ 1132 b is tenuous at best.

5.4. On the Nature of GJ 1132 b

Combining the emission and transmission spectroscopy data together tells us that GJ 1132 b has a thin atmosphere, or no atmosphere at all. Which of these scenarios is more likely? Is it possible for GJ 1132 b to retain a thin high mean molecular weight atmosphere?

Early attempts to predict the atmosphere composition of GJ 1132 b by and large found that the most likely outcome, given that the planet is interior to the runaway greenhouse limit (A. P. Ingersoll 1969) and assuming that it started with some amount of water, was that GJ 1132 b has a tenuous atmosphere with at most a few bars of O_2 (L. Schaefer et al. 2016). This is a result of dissociation of H_2O and subsequent preferential loss of H_2. A steam atmosphere on this (at least) 5 Gyr (Z. K. Berta-Thompson et al. 2015), 1.6 day orbital period planet would require that a magma ocean still be present on the planet's surface (as a source for the outgassed steam; L. Schaefer et al. 2016), a scenario that is highly implausible based on predicted solidification timescales of magma oceans. The only mechanism by which a magma ocean could have endured is if the initial water content of the planet was >10% by weight. L. Schaefer et al. (2016) could not rule out this high level of initial water content due to the imprecise mass from the discovery paper ($1.62 \pm 0.55\ M_\oplus$; Z. K. Berta-Thompson et al. 2015). However, using the updated radius from J. A. Dittmann et al. (2017; $1.13 \pm 0.056\ R_\oplus$) and updated mass from X. Bonfils et al. (2018; $1.66 \pm 0.23\ M_\oplus$), we use the mass–radius interior composition curves from L. Zeng et al. (2016) to argue that a 10% H_2O by weight composition can be rejected at $\sim 2.8\sigma$, and that the composition is most consistent (within 1σ) with an Earth-like interior. Therefore, it does not seem feasible for current-day GJ 1132 b to host a surface magma ocean and have any significant interior water fraction capable of maintaining a steam atmosphere. This supports our conclusion that the tentative hints of a steam atmosphere, driven largely by Visit 1, are erroneous.

We acknowledge, however, that the other scenario postulated by L. Schaefer et al. (2016)—an O_2 atmosphere leftover from preferential atmospheric escape of H_2—cannot so easily be ruled out. A 1 bar O_2-dominated atmosphere with 100 ppm H_2O (see Figure 6) is consistent with both the emission and transmission data. However, we are able to rule out the 10 bar O_2 atmosphere with 100 ppm of CO_2 that is consistent with the emission data (see the lower panel of Figure 6), as our data should be sensitive enough to detect the CO_2 feature at this atmospheric pressure. We note, however, that thinner O_2 atmospheres with some CO_2 would likely be more consistent with our data, as the lower pressures would decrease the expected feature size, analogous to our lower-pressure H_2O scenario illustrated in the middle panel of Figure 6.

In theory, the lack of an O_3 detection at 4.7 μm could be used to rule out an oxygen-dominated atmosphere, as O_3 would readily be produced by photochemistry (A. P. Lincowski et al. 2018). However, additional retrieval tests (with O_3 included) failed to place an upper limit on the species, suggesting that the precision of our transmission spectrum is insufficient to help rule out any meaningful amounts of O_3 in an oxygen-dominated atmospheres within the NIRSpec wavelength range. We conclude, therefore, that though we cannot entirely rule out an O_2 atmosphere, we find maintenance of a remnant steam atmosphere that has become O_2-dominated to be unlikely.

What of the other two scenarios highlighted in Section 4.4 and Figure 10? Is it plausible for GJ 1132 b to host a thin "Earth-like" (<0.4 bar) or CO_2 (<0.1 mbar) atmosphere? To answer this, we must consider atmosphere generation (via volcanic outgassing), stability (considering the radiative timescale and collapse), and stripping (via irradiation and stellar wind). In terms of atmospheric generation, volcanism wanes over geologic time, especially for smaller planets that deplete their mantle volatiles more quickly (C. Dorn et al. 2018). Assuming GJ 1132 is ~5 Gyr old places it right at the cusp of whether or not volcanism could be possible (E. S. Kite et al. 2009). We note this is a conservative estimate; GJ 1132 b is likely *at least* 5 Gyr old; if it is much older than this, the possibility of present-day volcanism is drastically lower. If volcanism is present, many theoretical studies predict atmospheres much thicker and/or with clouds than what our analysis suggests (E. S. Kite et al. 2009; L. Noack et al. 2017; C. Dorn et al. 2018), assuming the planet is in a stagnant-lid regime. For example, one study predicted that a volcanically outgassed CO_2 atmosphere for a planet of GJ 1132 b's mass ($1.66\ M_\oplus$) and temperature (588 K) should have surface pressures between that of Mars (~7 mbar) and Venus (~92 bar; see C. Dorn et al. 2018, their Figure 17). Our lower limit is clearly less than this.

Atmospheric stability must also be considered. We can first consider the radiative versus convective timescale, which tells us how fast the wind speeds must be to propagate the incident stellar flux from day to night before the heat is radiated away. Using a back-of-the-envelope calculation (Equations (10.1) and (10.2) from S. Seager 2010), we estimate the radiative timescale for a very thin (0.1 mbar) CO_2 atmosphere to be on the order of 10–20 s, which translates to wind speeds on the order of hundreds of meters per second. This is an order-of-magnitude greater than typical wind speeds measured on our own terrestrial planets. This suggests that in the case of a very thin atmosphere, heat is radiated away before it can be propagated to the nightside.

This also has important implications for atmospheric stability, as atmospheric collapse is a cause for concern on tidally locked rocky exoplanets when day–night heat transport is inefficient (R. Wordsworth 2015), as the nightside can condense and cold-trap atmospheric species (R. Wordsworth & L. Kreidberg 2022). R. Wordsworth (2015) estimated that close-in tidally locked exoplanets collapse at $\lesssim$0.03–0.1 bar

(see their Figure 12), although they did not extrapolate their findings to planets as irradiated as GJ 1132 b.

Lastly, we must also consider the fact that GJ 1132 b lies only 0.0153 au from its star (X. Bonfils et al. 2018), and that atmospheric stripping may occur. E. S. Kite & M. N. Barnett (2020) predicted that M-dwarf super-Earths that started out with primordial H_2 atmospheres are unlikely to retain secondary atmospheres if they are $T_{eq} > 500$ K. This is because high mmw species get lost to space alongside H_2 before the underlying magma ocean (which can absorb and protect the volatile high mmw species from loss) crystallizes. This would leave only a tenuous atmosphere that would be difficult to replenish via volcanic outgassing. Hot planets such as GJ 1132 b could theoretically have atmospheres if they formed without an H_2 envelope or if they began with a high volatile content. However, the latter, as previously discussed, is unlikely for GJ 1132 b. Finally, stellar winds may play a significant role in stripping thin atmospheres for planets interior to the habitable zone (e.g., C. Dong et al. 2017, 2018).

Taken together, our results demonstrate that though both the transmission and emission data still allow several thin atmospheres on GJ 1132 b, this scenario is disfavored from a first-principles modeling perspective. We conclude, therefore, that our featureless spectrum is more likely a reflection of a true bare rock planet, the simplest explanation that is consistent with the existing data. However, further and more detailed planetary evolution modeling is warranted to fully understand (1) the possible states for this planet considering the many remaining uncertainties (e.g., the interior composition and escape history) and (2) how the stringent constraints from the emission and transmission measurements in turn help constrain these uncertain unobservable properties of GJ 1132 b.

6. Conclusions

In this study, we follow up on the two apparently discrepant NIRSpec/G395H transits of GJ 1132 b reported in E. M. May et al. (2023). We observe two additional transits using NIRSpec/G395M in this case, and use all four transits and the recent MIRI/LRS emission spectroscopy results (Q. Xue et al. 2024) to place strict limits on any atmosphere around GJ 1132 b.

This is the first time G395H and G395M observations have studied a single target, offering a critical opportunity for comparison between modes. NIRSpec/G395M functions very well for rocky exoplanet transmission spectroscopy, with no meaningful differences in scatter, red noise, or final error bar size between the two modes. However, utilizing the G395M mode does forfeit the opportunity to gain high-resolution flux-calibrated stellar spectra, which can be critical for uncovering hidden stellar heterogeneity that is not obvious in the transmission spectrum. Characterizing this possible hidden heterogeneity is critical in the case of coadding multiple transits. In the case of rocky M-dwarf studies, we therefore recommend that G395M be utilized only if not stacking multiple visits.

For GJ 1132 b, we find good intervisit agreement in the transmission spectra, with the G395M data cleanly filling in the NRS1/NRS2 gap from the G395H data with no offsets required. Simply by looking at the transmission spectrum, it would appear that the differences in spectra reported in E. M. May et al. (2023) are simply due to random noise. However, upon further investigation, there is evidence that underlying stellar heterogeneity in Visit 1 (relative to the other visits) is driving these differences. Using the flux-calibrated stellar spectra, we find that Visit 1 likely has a higher coverage fraction of cool spots (and lower coverage fraction of warm spots). Because Visit 1 and 2 were taken only 8 days apart, this suggests (along with the slight different in fitted limb-darkening values for Visit 1) that spots on the limb may have rotated completely or partially out of view in this timeframe.

The fact that Visit 1 is an outlier is strongly corroborated by the retrieval analysis, which finds a steam atmosphere is preferred in the weighted mean of all visits. However, a featureless spectrum is preferred if Visit 1 is excluded. This illustrates the challenge of stacking multiple transits and attempting to interpret small deviation from flatness. In the future, we strongly recommend the rocky exoplanet community adopt this "leave-one-out" approach to test if a single visit is dominating one's interpretation.

What of the planet's atmosphere? Forward models prefer a featureless spectrum over most 1 bar atmospheres, whereas a thinner (<10 mbar) steam atmosphere is equally consistent with the data. However, this preference for a steam atmosphere disappears with the exclusion of Visit 1. We further explore possibilities with retrievals that use an agnostic "ghost gas" to more broadly explore the surface pressure versus mmw parameter space allowed by the data. These show that our data confidently exclude atmospheres with mmw <6 AMU (this includes H_2 atmospheres, as well as "miscible envelope" atmospheres as in B. Benneke et al. 2024) at all pressures, but are consistent with (within 1σ of) thin, <100 mbar atmospheres that are at least 14 AMU or any atmosphere that is 25 AMU or greater.

We combine these results with the MIRI/LRS emission spectroscopy results (Q. Xue et al. 2024), which are more sensitive to thicker atmospheres, which tell us that a thick atmosphere >25 AMU is unlikely. For a thinner, steam atmosphere (18 AMU), transmission results are more stringent; thin ($\lesssim$10 mbar) steam atmospheres remain consistent with the data. We further note that certain thin O_2 atmospheres remain consistent with both data sets. Due to GJ 1132 b's age and proximity to its host star, however, it is unlikely to be able to retain such a thin atmosphere, regardless of composition. At this point, the simplest explanation is in fact that GJ 1132 b is a bare rock.

If this is true, this would push our endeavor to locate the M-dwarf cosmic shoreline toward larger (>1.13 $R_\oplus$) and/or cooler ($T_{eq} < 580$ K) rocky planets. Importantly, we note that at face value, GJ 1132 was always a promising M-dwarf target for transmission spectroscopy, with its slow rotation period and old age. If GJ 1132 b is indeed a bare rock, this demonstrates that even optimal targets from an observability standpoint may not make optimal targets for atmospheres. Importantly, GJ 1132 b is in a very similar XUV instellation/escape velocity parameter space as the first two targets (LTT 1445Ac and GJ 3929b) selected for the STScI 500 hr Rocky World Director's Discretionary Time program.[16] If none of these planets show hints of atmospheres, we as a community may want to strive toward targets that are at lower instellations and therefore are more challenging observationally, but are more likely to host atmospheres.

[16] https://outerspace.stsci.edu/pages/viewpage.action?pageId=257035126

The quest to determine which M-dwarf planets have atmospheres is nowhere near over. What's more, the finding of a bare rock is just as critical a finding as an atmosphere, as it helps us refine our understanding of the cosmic shoreline as well as opens up a world of possible exogeologic studies (R. Hu et al. 2012; E. A. Whittaker et al. 2022; E. C. First et al. 2025). Regardless of the ubiquity or lack thereof of nearby M-dwarf rocky exoplanet atmospheres, we are thus taking the first observational steps toward mapping our rocky neighbors and placing the solar system rocky planets in their true Galactic context.

Acknowledgments

The authors thank the reviewer for providing insightful and helpful feedback on this work. K.B. thanks N. Allen for helpful conversations about stellar contamination. This work was done based on observations using the NASA/ESA/CSA JWST. The data were obtained via the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA). Support for this work was provided by NASA grants from the JWST Cycle 1 GO Program 1981. S.P. acknowledges support from NASA under award No. 80GSFC24M0006. H.R.W was funded by UK Research and Innovation (UKRI) framework under the UK government's Horizon Europe funding guarantee for an ERC Starter grant (grant No. EP/Y006313/1). S.E.M. is supported by NASA through the NASA Hubble Fellowship grant HST-HF2-51563 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. L.A. acknowledges funding from UKRI STFC Consolidated grant ST/V000454/1 and is supported by the Klarman Fellowship. J.K. acknowledges financial support from Imperial College London through an Imperial College Research Fellowship grant.

Data Availability

All data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI). The four transit observations can be accessed via doi:10.17909/e97r-g212.

Data products are available on Zenodo: doi:10.5281/zenodo.15318917.

Facilities: JWST (NIRSpec).

Software: astropy (Astropy Collaboration et al. 2013, 2018, 2022), `batman` (L. Kreidberg 2015), `CHIMERA` (M. R. Line & Y. L. Yung 2013; M. R. Line et al. 2014), `emcee` (D. Foreman-Mackey et al. 2013), `Eureka!`(T. Bell et al. 2022), `ExoTiC-JEDI` (H. R. Wakeford et al. 2016; I. Laginja & H. Wakeford 2020), `ExoTIC-LD` (D. Grant & H. R. Wakeford 2022; D. Grant & H. Wakeford 2024), `FIREFLy` (Z. Rustamkulov et al. 2022, 2023), `jwst` (H. Bushouse et al. 2022), `lacosmic` (P. G. van Dokkum 2001), `lmfit` (M. Newville et al. 2014), matplotlib (J. D. Hunter 2007), numpy (C. R. Harris et al. 2020), pandas (W. McKinney 2010), `PHOENIX` (F. Allard et al. 2012; T. O. Husser et al. 2013), `PICASO` (N. E. Batalha et al. 2019), `POSEIDON` (R. J. MacDonald & N. Madhusudhan 2017; R. J. MacDonald 2023), `PyMSG` (R. Townsend & A. Lopez 2023), `PyMultiNest` (F. Feroz et al. 2009; J. Buchner et al. 2014), scipy (P. Virtanen et al. 2020).

ORCID iDs

Katherine A. Bennett https://orcid.org/0000-0002-9030-0132
Ryan J. MacDonald https://orcid.org/0000-0003-4816-3469
Sarah Peacock https://orcid.org/0000-0002-1046-025X
Junellie Perez https://orcid.org/0000-0002-9032-8530
E. M. May https://orcid.org/0000-0002-2739-1465
Sarah E. Moran https://orcid.org/0000-0002-6721-3284
Lili Alderson https://orcid.org/0000-0001-8703-7751
Jacob Lustig-Yaeger https://orcid.org/0000-0002-0746-1980
Hannah R. Wakeford https://orcid.org/0000-0003-4328-3867
David K. Sing https://orcid.org/0000-0001-6050-7645
Kevin B. Stevenson https://orcid.org/0000-0002-7352-7941
Natasha E. Batalha https://orcid.org/0000-0003-1240-6844
Mercedes López-Morales https://orcid.org/0000-0003-3204-8183
Munazza K. Alam https://orcid.org/0000-0003-4157-832X
Joshua D. Lothringer https://orcid.org/0000-0003-3667-8633
Guangwei Fu https://orcid.org/0000-0002-3263-2251
James Kirk https://orcid.org/0000-0002-4207-6615
Jeff A. Valenti https://orcid.org/0000-0003-3305-6281
L. C. Mayorga https://orcid.org/0000-0002-4321-4581
Kristin S. Sotzen https://orcid.org/0000-0001-7393-2368

References

Adams Redai, J., Wogan, N., Wallack, N., et al. 2025, AJ, in press
Alam, M. K., Gao, P., Adams Redai, J., et al. 2025, AJ, 169, 15
Alderson, L., Batalha, N. E., Wakeford, H. R., et al. 2024, AJ, 167, 216
Alderson, L., Grant, D., & Wakeford, H. 2022, Exo-TiC/ExoTiC-JEDI: v0.1-beta-release, v0.1, Zenodo, doi:10.5281/zenodo.7185855
Alderson, L., Moran, S. E., Wallack, N. L., et al. 2025, AJ, 169, 142
Alderson, L., Wakeford, H. R., Alam, M. K., et al. 2023, Natur, 614, 664
Allard, F., Homeier, D., & Freytag, B. 2012, RSPTA, 370, 2765
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Audard, M., Güdel, M., Drake, J. J., & Kashyap, V. L. 2000, ApJ, 541, 396
August, P. C., Buchhave, L. A., Diamond-Lowe, H., et al. 2025, A&A, 695, A171
Banerjee, A., Barstow, J. K., Gressier, A., et al. 2024, ApJL, 975, L11
Barstow, J. K., & Irwin, P. G. J. 2016, MNRAS, 461, L92
Batalha, N., Freedman, R., Gharib-Nezhad, E., & Lupu, R. 2022, Resampled Opacity Database for PICASO, v2.0, Zenodo, doi:10.5281/zenodo.6928501ss
Batalha, N. E., Marley, M. S., Lewis, N. K., & Fortney, J. J. 2019, ApJ, 878, 70
Bell, T., Ahrer, E.-M., Brande, J., et al. 2022, JOSS, 7, 4503
Bello-Arufe, A., Damiano, M., Bennett, K. A., et al. 2025, ApJL, 980, L26
Benneke, B., Roy, P.-A., Coulombe, L.-P., et al. 2024, arXiv:2403.03325
Benneke, B., & Seager, S. 2012, ApJ, 753, 100
Berta-Thompson, Z. K., Irwin, J., Charbonneau, D., et al. 2015, Natur, 527, 204
Bonfils, X., Almenara, J. M., Cloutier, R., et al. 2018, A&A, 618, A142
Brasseur, C. E., Osten, R. A., Tristan, I. I., & Kowalski, A. F. 2023, ApJ, 944, 5
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Burton, K., MacGregor, M. A., Osten, R. A., et al. 2025, ApJ, 982, 43
Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2022, JWST Calibration Pipeline, v1.8.2, Zenodo, doi:10.5281/zenodo.7325378

Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2024, JWST Calibration Pipeline, v1.8.2, Zenodo, doi:10.5281/ZENODO.10870758
Cadieux, C., Doyon, R., MacDonald, R. J., et al. 2024, ApJL, 970, L2
Claret, A. 2000, A&A, 363, 1081
Cloutier, R., Doyon, R., Menou, K., et al. 2017, AJ, 153, 9
Cowan, N. B., & Agol, E. 2011, ApJ, 729, 54
Damiano, M., Bello-Arufe, A., Yang, J., & Hu, R. 2024, ApJL, 968, L22
Davenport, B., Kempton, E. M. R., Nixon, M. C., et al. 2025, ApJL, 984, L44
Davenport, J. R. A. 2016, ApJ, 829, 23
Diamond-Lowe, H., Berta-Thompson, Z., Charbonneau, D., & Kempton, E. M. R. 2018, AJ, 156, 42
Dittmann, J. A., Irwin, J. M., Charbonneau, D., Berta-Thompson, Z. K., & Newton, E. R. 2017, AJ, 154, 142
Dong, C., Jin, M., Lingam, M., et al. 2018, PNAS, 115, 260
Dong, C., Lingam, M., Ma, Y., & Cohen, O. 2017, ApJL, 837, L26
Dorn, C., Noack, L., & Rozel, A. B. 2018, A&A, 614, A18
Ducrot, E., Lagage, P.-O., Min, M., et al. 2024, NatAs,
Elkins-Tanton, L. T., & Seager, S. 2008, ApJ, 685, 1237
Esparza-Borges, E., López-Morales, M., Adams Redai, J. I., et al. 2023, ApJL, 955, L19
Espinoza, N., Úbeda, L., Birkmann, S. M., et al. 2023, PASP, 135, 018002
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
First, E. C., Mishra, I., Gazel, E., et al. 2025, NatAs, 9, 370
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Fournier-Tondreau, M., MacDonald, R. J., Radica, M., et al. 2024, MNRAS, 528, 3354
France, K., Duvvuri, G., Egan, H., et al. 2020, AJ, 160, 237
France, K., Loyd, R. O. P., Youngblood, A., et al. 2016, ApJ, 820, 89
Gao, P., Thorngren, D. P., Lee, E. K. H., et al. 2020, NatAs, 4, 951
Grant, D., & Wakeford, H. 2024, JOSS, 9, 6816
Grant, D., & Wakeford, H. R. 2022, Exo-TiC/ExoTiC-LD: ExoTiC-LD v3.0.0, Zenodo, doi:10.5281/zenodo.7437681
Greene, T. P., Bell, T. J., Ducrot, E., et al. 2023, Natur, 618, 39
Gressier, A., Espinoza, N., Allen, N. H., et al. 2024, ApJL, 975, L10
Guillot, T. 2010, A&A, 520, A27
Guzmán-Mesa, A., Kitzmann, D., Fisher, C., et al. 2020, AJ, 160, 15
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
He, C., Radke, M., Moran, S. E., et al. 2024, NatAs, 8, 182
Horne, K. 1986, PASP, 98, 609
Hu, R., Bello-Arufe, A., Zhang, M., et al. 2024, Natur, 630, 609
Hu, R., Ehlmann, B. L., & Seager, S. 2012, ApJ, 752, 7
Hunter, J. D. 2007, CSE, 9, 90
Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Ingersoll, A. P. 1969, JAtS, 26, 1191
Jakobsen, P. 2022, ESA Instrument Report 2022-03, ESA https://jwst-docs.stsci.edu/files/154687782/154687784/1/1731088468115/quantum_yield_noise_v1.0.pdf
Kaltenegger, L., & Traub, W. A. 2009, ApJ, 698, 519
Karman, T., Gordon, I. E., van der Avoird, A., et al. 2019, Icar, 328, 160
Khodachenko, M. L., Ribas, I., Lammer, H., et al. 2007, AsBio, 7, 167
Kipping, D. 2025, arXiv:2504.13238
Kipping, D. M. 2013, MNRAS, 435, 2152
Kirk, J., Stevenson, K. B., Fu, G., et al. 2024, AJ, 167, 90
Kite, E. S., & Barnett, M. N. 2020, PNAS, 117, 18264
Kite, E. S., Manga, M., & Gaidos, E. 2009, ApJ, 700, 1732
Kopal, Z. 1950, HarCi, 454, 1
Kostogryz, N. M., Witzke, V., Shapiro, A. I., et al. 2022, A&A, 666, A60
Kreidberg, L. 2015, PASP, 127, 1161
Laginja, I., & Wakeford, H. 2020, JOSS, 5, 2281
Lammer, H., Kasting, J. F., Chassefière, E., et al. 2008, SSRv, 139, 399
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
Lim, O., Benneke, B., Doyon, R., et al. 2023, ApJL, 955, L22
Lincowski, A. P., Meadows, V. S., Crisp, D., et al. 2018, ApJ, 867, 76
Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014, ApJ, 783, 70
Line, M. R., & Yung, Y. L. 2013, ApJ, 779, 3
Luque, R., Park Coy, B., Xue, Q., et al. 2025, AJ, 170, 49
Lustig-Yaeger, J., Fu, G., May, E. M., et al. 2023, NatAs, 7, 1317
Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, AJ, 158, 27
MacDonald, R. J. 2023, JOSS, 8, 4873
MacDonald, R. J., & Madhusudhan, N. 2017, MNRAS, 469, 1979
Madhusudhan, N., Sarkar, S., Constantinou, S., et al. 2023, ApJL, 956, L13
Magic, Z., Chiavassa, A., Collet, R., & Asplund, M. 2015, A&A, 573, A90
May, E. M., MacDonald, R. J., Bennett, K. A., et al. 2023, ApJL, 959, L9
Mayo, A. W., Fortenbach, C. D., Louie, D. R., et al. 2025, AJ, 170, 50
McKinney, W. 2010, in, in Proc. 9th Python in Science Conf., ed. Stéfan van der Walt & J. Millman (Austin, TX: SciPy), 56
Meier Valdés, E. A., Demory, B. O., Diamond-Lowe, H., et al. 2025, A&A, 698, A68
Moran, S. E., Hörst, S. M., Batalha, N. E., Lewis, N. K., & Wakeford, H. R. 2018, AJ, 156, 252
Moran, S. E., Stevenson, K. B., Sing, D. K., et al. 2023, ApJL, 948, L11
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
Mullens, E., Lewis, N. K., & MacDonald, R. J. 2024, ApJ, 977, 105
Newton, E. R., Irwin, J., Charbonneau, D., et al. 2016, ApJ, 821, 93
Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, v0.8.0, Zenodo, doi:10.5281/zenodo.11813
Noack, L., Rivoldini, A., & Van Hoolst, T. 2017, PEPI, 269, 40
Owen, J. E. 2019, AREPS, 47, 67
Owen, J. E., & Jackson, A. P. 2012, MNRAS, 425, 2931
Palle, E., Yan, F., Morello, G., et al. 2025, A&A, 697, A31
Pass, E. K., Charbonneau, D., Irwin, J. M., & Winters, J. G. 2022, ApJ, 936, 109
Pass, E. K., Charbonneau, D., Latham, D. W., et al. 2024, ApJ, 966, 231
Paudel, R. R., Barclay, T., Youngblood, A., et al. 2024, ApJ, 971, 24
Peacock, S., Barman, T., Shkolnik, E. L., Hauschildt, P. H., & Baron, E. 2019, ApJ, 871, 235
Piaulet-Ghorayeb, C., Benneke, B., Radica, M., et al. 2024, ApJL, 974, L10
Pizzolato, N., Maggio, A., Micela, G., Sciortino, S., & Ventura, P. 2003, A&A, 397, 147
Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, MNRAS, 480, 2597
Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, MNRAS, 385, 109
Pont, F., Zucker, S., & Queloz, D. 2006, MNRAS, 373, 231
Preibisch, T., & Feigelson, E. D. 2005, ApJS, 160, 390
Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122
Radica, M., Piaulet-Ghorayeb, C., Taylor, J., et al. 2025, ApJL, 979, L5
Rathcke, A. D., Buchhave, L. A., de Wit, J., et al. 2025, ApJL, 979, L19
Rathcke, A. D., MacDonald, R. J., Barstow, J. K., et al. 2021, AJ, 162, 138
Rugheimer, S., Kaltenegger, L., Segura, A., Linsky, J., & Mohanty, S. 2015, ApJ, 809, 57
Rustamkulov, Z., Sing, D. K., Liu, R., & Wang, A. 2022, ApJL, 928, L7
Rustamkulov, Z., Sing, D. K., Mukherjee, S., et al. 2023, Natur, 614, 659
Scarsdale, N., Wogan, N., Wakeford, H. R., et al. 2024, AJ, 168, 276
Schaefer, L., Wordsworth, R. D., Berta-Thompson, Z., & Sasselov, D. 2016, ApJ, 829, 63
Schmidt, S. P., MacDonald, R. J., Tsai, S.-M., et al. 2025, arXiv:2501.18477
Seager, S. 2010, Exoplanet Atmospheres: Physical Processes (Princeton, NJ: Princeton Univ. Press)
Segura, A., Kasting, J. F., Meadows, V., et al. 2005, AsBio, 5, 706
Shkolnik, E. L., & Barman, T. S. 2014, AJ, 148, 64
Skilling, J. 2004, in AIP Conf. Ser. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th Int. Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint (Melville, NY: AIP), 395
Smitha, H. N., Shapiro, A. I., Witzke, V., et al. 2025, ApJL, 978, L13
Southworth, J., Mancini, L., Madhusudhan, N., et al. 2017, AJ, 153, 191
Speagle, J. S. 2020, MNRAS, 493, 3132
Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, AJ, 158, 138
Swain, M. R., Estrela, R., Roudier, G. M., et al. 2021, AJ, 161, 213
Taylor, J. 2025, RNAAS, 9, 118
Taylor, J., Radica, M., Chatterjee, R. D., et al. 2025, MNRAS,
Tennyson, J., & Yurchenko, S. N. 2018, Atoms, 6, 26
Tennyson, J., Yurchenko, S. N., Zhang, J., et al. 2024, JQSRT, 326, 109083
Tian, F. 2009, ApJ, 703, 905
Townsend, R., & Lopez, A. 2023, JOSS, 8, 4602
Trotta, R. 2008, ConPh, 49, 71
Underwood, D. S., Tennyson, J., Yurchenko, S. N., et al. 2016, MNRAS, 459, 3890
van Dokkum, P. G. 2001, PASP, 113, 1420
Van Looveren, G., Güdel, M., Boro Saikia, S., & Kislyakova, K. 2024, A&A, 683, A153
Vida, K., Kővári, Z., Pál, A., Oláh, K., & Kriskovics, L. 2017, ApJ, 841, 124
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261

Waalkes, W. C., Berta-Thompson, Z., Bourrier, V., et al. 2019, AJ, 158, 50
Wachiraphan, P., Berta-Thompson, Z. K., Diamond-Lowe, H., et al. 2025, AJ, 169, 311
Wakeford, H. R., Sing, D. K., Evans, T., Deming, D., & Mandell, A. 2016, ApJ, 819, 10
Wallack, N. L., Batalha, N. E., Alderson, L., et al. 2024, AJ, 168, 77
Weiner Mansfield, M., Xue, Q., Zhang, M., et al. 2024, ApJL, 975, L22
Whittaker, E. A., Malik, M., Ih, J., et al. 2022, AJ, 164, 258
Wordsworth, R. 2015, ApJ, 806, 180
Wordsworth, R., & Kreidberg, L. 2022, ARA&A, 60, 159
Xue, Q., Bean, J. L., Zhang, M., et al. 2024, ApJL, 973, L8
Yurchenko, S. N., Mellor, T. M., Freedman, R. S., & Tennyson, J. 2020, MNRAS, 496, 5282
Yurchenko, S. N., Owens, A., Kefala, K., & Tennyson, J. 2024, MNRAS, 528, 3719
Zahnle, K. J., & Catling, D. C. 2017, ApJ, 843, 122
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127
Zhang, M., Hu, R., Inglis, J., et al. 2024, ApJL, 961, L44
Zieba, S., Kreidberg, L., Ducrot, E., et al. 2023, Natur, 620, 746