

The Atmospheric Composition of Sub-Neptune K2-18 b and Implications for its Formation

Gareb Fernández-Rodríguez^{1,2}, Giuseppe Morello^{3,4}, Jonathan C. Tan^{2,5}, Enric Pallé^{1,6}, Mark R. Swain⁷, Efthymios Poulourtzidis^{2,8}, Alfredo Biagini⁴, Quentin Changeat⁹, Chengzi Jiang^{1,6}, Francisco J. Pozuelos³, Pedro J. Amado³

¹ Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain

² Dept. of Space, Earth & Environment, Chalmers University of Technology, Gothenburg, Sweden

³ Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain

⁴ INAF- Palermo Astronomical Observatory, Piazza del Parlamento, 1, 90134 Palermo, Italy

⁵ Dept. of Astronomy & Virginia Institute for Theoretical Astronomy, University of Virginia, Charlottesville, VA, USA

⁶ Instituto de Astrofísica de Canarias (IAC), 38205 La Laguna, Tenerife, Spain

⁷ California Institute of Technology, NASA Jet Propulsion Laboratory, 4800 Oak Grove Dr, La Cañada Flintridge, CA 91011, USA

⁸ Department of Physics, Aristotle University of Thessaloniki, University Campus, Thessaloniki, 54124, Greece

⁹ Kapteyn Institute, University of Groningen, 9747 AD Groningen, NL

Received MONTH XX, XXXX; accepted MONTH XX, XXXX

ABSTRACT

Unlocking the atmospheres of sub-Neptune planets is one of the revolutionary accomplishments of JWST. However such observations require complex data analysis methodologies, which have a strong impact on the derived conclusions. Here, we present an independent reanalysis of the JWST transmission spectrum of the temperate sub-Neptune K2-18 b, to assess the robustness of previously claimed atmospheric detections, explore the planet’s possible parameter space, and implications for its formation. The NIRISS/SOSS and NIRSpec/G395H observations were reduced using a combination of public and customized pipelines. We produced a total of 12 different versions of the transmission spectrum by varying key steps: spectral binning, limb-darkening treatment, and the application of a novel correction for an occulted stellar spot, as well as error inflation and instrumental offsets. We then performed atmospheric retrievals using TauREx 3, comparing models of varying complexity. We robustly detect CH₄ at $>4\sigma$ significance across all reduction and retrieval setups. The evidence for CO₂ is weaker and highly model-dependent, with a typical significance of $\sim 2\sigma$. The tentative detection of dimethyl sulphide (DMS), reported in previous studies, vanishes in our most comprehensive retrieval models. We find that correcting the stellar spot in the NIRISS transit is a critical step, introducing a uniform offset that primarily drives the inference of a lighter, lower mean molecular weight atmosphere. Furthermore, the assumed complexity of the retrieval model itself introduces significant biases; including more molecules systematically increases the retrieved CH₄ abundance and atmospheric mean molecular weight, even for species without spectral features. The data are consistent with a hydrogen-rich, i.e., primordial, atmosphere with an elevated O abundance and an even more elevated C abundance, leading to a C/O ratio significantly greater than solar. We show that the physical properties of the K2-18 system planets, i.e., innermost planet K2-18 c, and K2-18 b are consistent with those expected by the *in situ* formation theory of Inside-Out Planet Formation (IOPF). Furthermore, these properties predict the temperature of K2-18 b at time of formation was $\gtrsim 500$ K, i.e., much warmer than the current equilibrium temperature and just interior to the carbon “soot” line, where an elevated C/O ratio of a primordial atmosphere is expected to be inherited from the protoplanetary disk.

Key words. Planets and satellites: individual: K2-18 b – Planets and satellites: atmospheres – Planets and satellites: composition – Planets and satellites: formation

1. Introduction

K2-18 b is a temperate sub-Neptune orbiting an M-dwarf star located at 38 pc in the constellation Leo (Montet et al. 2015). With a mass of $8.63 \pm 1.35 M_{\oplus}$ (Cloutier et al. 2019) and a radius of $2.610 \pm 0.087 R_{\oplus}$ (Benneke et al. 2019), the planet lies in a degenerate region of parameter space concerning its interior composition. K2-18 b could host a water-rich core with a modest H₂-He envelope, or a predominantly rocky interior with a thicker primordial atmosphere (Luque & Pallé 2022; Rogers et al. 2023). Intriguingly, the planet orbits within the habitable zone of its host star, receiving a stellar irradiation comparable to that of Earth. This has motivated the proposal of a “Hycean world” scenario, in which K2-18 b might sustain liquid water beneath a hydrogen-rich atmosphere (Madhusudhan et al. 2021).

Spectroscopic observations with HST/WFC3, prior to the launch of JWST, revealed a prominent absorption feature at $1.4 \mu\text{m}$. It was initially reported as a $\gtrsim 3\sigma$ detection of water vapour (Benneke et al. 2019; Tsiaras et al. 2019). Based on self-consistent 1D radiative equilibrium models, Blain et al. (2021) and Bézard et al. (2022) argued that the $1.4 \mu\text{m}$ feature was more likely due to methane. Barclay et al. (2021) suggested unocculted stellar spots as an alternative explanation. Recently, JWST solved this dilemma by obtaining a high-precision transmission spectrum of K2-18 b over $0.9\text{--}5.2 \mu\text{m}$, combining NIRISS/SOSS and NIRSpec/G395H observations. The first analysis of these observations reported CH₄ ($\sim 5\sigma$), CO₂ ($\sim 3\sigma$), and tentative dimethyl sulphide ((CH₃)₂S or DMS) detections (Madhusudhan et al. 2023). Along with the non-detection of H₂O, CO, and NH₃, these findings were initially interpreted as

indicative of a Hycean scenario for K2-18 b (Hu et al. 2021; Madhusudhan et al. 2023). However, follow-up studies noted that they are consistent with a gas-rich envelope over a magma ocean (Shorttle et al. 2024; Wogan et al. 2024). An independent reanalysis of the same JWST data confirmed the CH₄ detection, but raised doubts about the CO₂ or DMS detections, suggesting that K2-18 b is most likely an oxygen-poor mini-Neptune without a water ocean (Schmidt et al. 2025). While this manuscript was in preparation, Hu et al. (2025) reported new JWST observations of K2-18 b covering 1.7–5.2 μm , reaffirming the detection of CO₂ at 3.7 σ and, in their interpretation, lending further support to the Hycean world hypothesis.

The tentative detection of DMS in K2-18 b has sparked interest due to its potential as a biosignature gas (Seager et al. 2013), though an abiotic origin cannot be ruled out (Reed et al. 2024). A 6–12 μm transmission spectrum obtained with JWST/MIRI LRS was initially reported as further evidence for DMS or dimethyl disulphide (DMDS) (Madhusudhan et al. 2025), though this interpretation has been widely debated (Luque et al. 2025; Stevenson et al. 2025; Taylor 2025; Welbanks et al. 2025).

In this work, we present an independent reanalysis of JWST observations of K2-18 b using NIRSpec/G395H and NIRISS/SOSS data (GO 2722, PI: Madhusudhan). We carefully examine how choices in data reduction, light-curve modelling, and retrieval configurations affect the inferred atmospheric properties. In particular, at the data processing level, we investigate the effects of spectral binning at the light-curve or transmission-spectrum stage, treatment of stellar limb-darkening and spots, instrumental offsets, and error-bar rescaling. At the modelling stage, we investigate the effect the number of included molecules, temperature–pressure profiles, and the presence of clouds and haze have in the retrievals. An additional result of our investigation is a newly developed semi-empirical method for correcting occulted stellar spots in spectral light curves, that has been proven effective. This investigation underscores the importance of repeated JWST observations to assess the robustness of molecular detections and highlights the value of future missions such as Ariel (Tinetti et al. 2018), whose simultaneous broad-wavelength coverage will help mitigate instrumental offsets and astrophysical variability. Beyond K2-18 b, our analysis also aims to provide guidelines for interpreting JWST observations of sub-Neptune atmospheres from single-transit datasets, such as TOI-421 b (Davenport et al. 2025) or GJ 9827 d (Piaulet-Ghorayeb et al. 2024). Then, in the final part of our paper we examine the implications of our derived atmospheric properties for K2-18 b for its formation, concluding that evidence points to a H₂-He dominated atmosphere with a high C/O ratio that is consistent with a scenario of *in situ* formation of a volatile-poor planet.

2. Observations

Two full transits of K2-18 b were observed with the JWST as part of GO Program 2722 (PI: N. Madhusudhan). The first transit was observed from January 20 to January 21 2023 between 18:37:38 and 01:11:32 UTC for a total of 5.3 hours. This observation was made using NIRSpec (Ferruit et al. 2012; Jakobsen et al. 2022; Birkmann et al. 2022) with the G395H grating, Bright Object Time Series mode (BOTS), F290LP filter, SUB2048 subarray, and the NRSRAPID readout pattern. The resulting spectrum was measured with the NRS1 and NRS2 detectors, spanning wavelengths of 2.8–3.72 μm and 3.83–5.17 μm each, with a detector gap between 3.72 and 3.83 μm .

The second observation took place during June 1 2023 from 13:49:20 to 19:36:05 UTC for a total of 4.9 hours. It used the

NIRISS instrument (Doyon et al. 2012, 2023) in the Single Object Slitless Spectroscopy, SOSS, mode (Albert et al. 2023), with the GR700XD grism, CLEAR filter, SUBSTRIP256 subarray, and NISRAPID readout pattern, resulting in a wavelength coverage of 0.85–2.83 μm for the first spectral order and 0.6–1.1 μm for the second spectral order.

3. Data Reduction

We reduced the observations starting from the raw data making use of a combination of publicly available pipelines and a self-developed pipeline. The first 3 stages of the JWST Calibration pipeline (jwst, Bushouse et al. 2023), which includes detector-level corrections, wavelength calibration and spectral trace extraction were carried out using standard reduction pipelines. After that we used a custom-built pipeline for stages 4–6, which includes white light curve and spectral light curve fitting and extraction of the transmission spectrum. This approach allows us to test the effects of different decisions made during the reduction process on the final transmission spectrum.

3.1. NIRISS SOSS

The first three stages of NIRISS data reduction were performed using exoTDRF (Feinstein et al. 2023; Radica et al. 2023; Radica 2024). Our reduction follows the typical reduction process for sub-Neptunes and super-Earths (Lim et al. 2023; Cadieux et al. 2024; Benneke et al. 2024). Stage 1 includes flagging of saturated pixels, subtraction of the super-bias, correction of the reference pixel position, and subtraction of the dark current. In the jump detection step we used a threshold of 7 σ for the time-domain outlier rejection algorithm (Radica et al. 2024), similarly to Ahrer et al. (2025), although we use a more conservative value. A scaled model group-level background is temporally subtracted from the data in order to better remove the 1/*f* noise which affects all JWST instruments (Schlawin et al. 2020). After removal of the aforementioned noise, the background model is re-added and the ramp linearity is fitted. Background photons are recorded simultaneously with target photons, therefore the background inherits the same non-linearity effects and should be corrected after linearity calibration (Feinstein et al. 2023; Radica et al. 2023). Our implementation of Stage 2 includes the default parameters, photometric calibration, and wavelength mapping, although we skip the flat-fielding step. In Stage 3, we perform spectral extraction using a box extraction method, assuming that the trace width remains constant, with an aperture of 30 pixels. This approach has been extensively used in prior JWST analyses (Davenport et al. 2025). In general the two spectral orders in NIRISS SOSS overlap, therefore there will be some spectral contamination. More complex extraction algorithms exist that account for this effect, like the ATOCA algorithm, however, these effects can be neglected in relative flux measurements (Darveau-Bernier et al. 2022; Radica et al. 2022a).

3.2. NIRSpec G395H

Both detectors (NRS1 and NRS2) were reduced independently with identical reduction parameters making use of the Eureka! pipeline (Bell et al. 2022). In Stage 1, we flagged saturated pixels, subtracted the super-bias, corrected the reference pixel positions, and applied the ramp linearity correction to obtain count rates and subtract the dark current. Curved traces were also masked. We increased the jump detection threshold to 10 σ , in

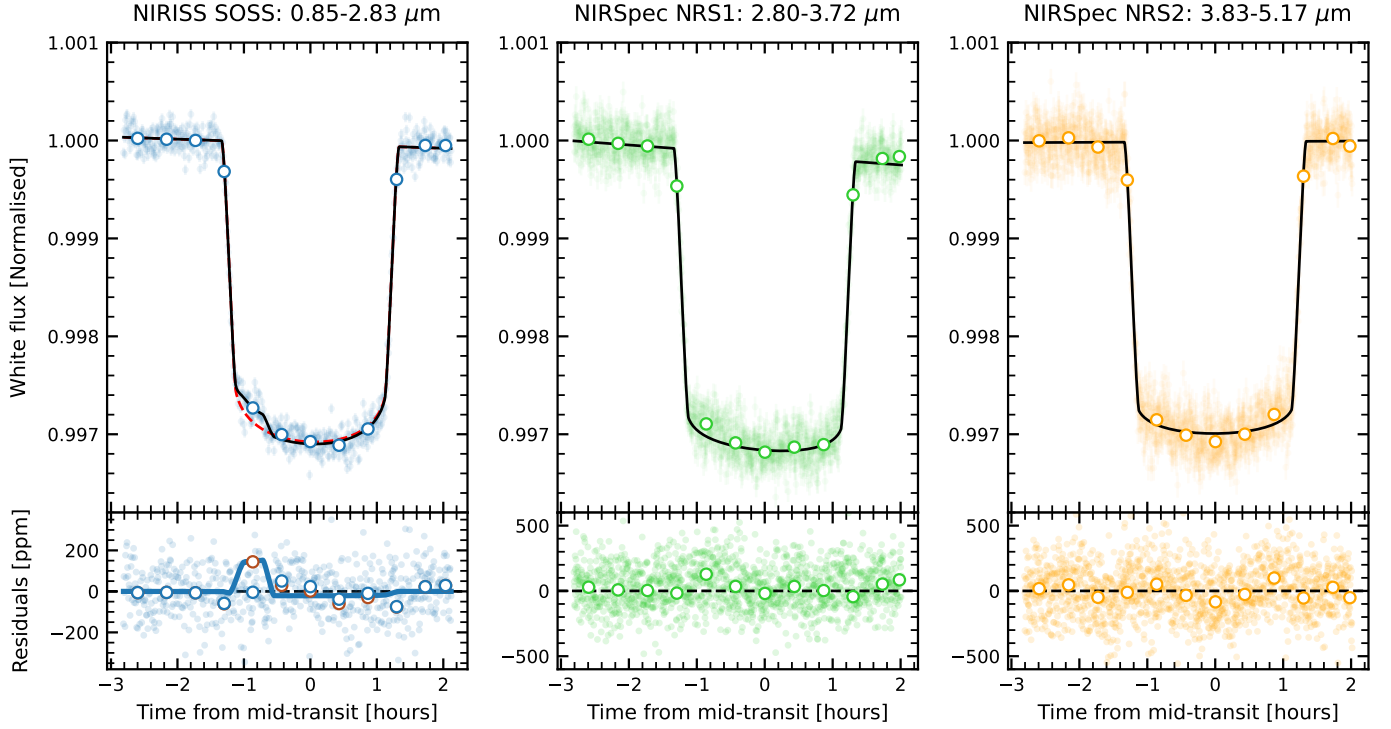


Fig. 1: White light curves for NIRISS NIRSpec NRS1 and NIRSpec NRS2 with best fit models. The blue line in the left-most panel residuals (NIRISS) indicates the resulting profile from the subtraction of the spot and spotless models. The red dashed line is the spotless model. The coloured points are 10 minute binned light curves.

line with the ERS team recommendations (Alderson et al. 2023; Rustamkulov et al. 2023). The $1/f$ noise (vertical striping pattern) is also corrected in this step by removing the column-wise median of pixels located outside the trace. Stage 2 of Eureka! is essentially a wrapper of the corresponding jwst pipeline stage (Bushouse et al. 2023). During this stage, wavelength calibration is applied. We used the default parameters for this step, but excluded photometric calibrations from the reduction process (Alderson et al. 2023). In Stage 3, we performed the spectral extraction. The region of interest was set from rows 0 to 31 along the spatial axis and columns 500 to 2042 along the spectral axis. We adopted a 10σ threshold for outlier rejection during background subtraction. The curvature of the trace was corrected by shifting the position of the pixel columns. This is done such that peak of the count distribution in the detector, which we assume to be Gaussian, lies in the centre of window, in our case row 15. The spectrum was then extracted using a 7-pixel half-width aperture centred on the trace. The background was estimated using the region 14 pixels away from the trace, well outside our extraction box.

4. Light-curve Analysis

4.1. White-light curves

For the white light curve analysis (Stage 4 of data reduction) we used a custom built pipeline. The white light curves were obtained by summing the flux within the spectral extraction aperture for each frame. In this way, we extracted three light curves, one from each detector. These time series were fitted using a transit model with a polynomial baseline (order ≤ 2) to account for instrumental systematics. The basic free transit parameters included the planet-to-star radius ratio (R_p/R_s), orbital period

(P), mid-transit time (T_0), impact parameter (b), and a transit duration parameter assuming a point-like planet (τ_0). The selection of the latter two parameters follows the recommendations of Morello (2018), where formulae relating them to orbital parameters are provided. Limb-darkening coefficients were fixed to theoretical values corresponding to a four-parameter law (Claret et al. 2020), computed using the ExoTETHyS package (Morello et al. 2020a,b). These coefficients were derived from the spherical one-dimensional (1D) PHOENIX stellar atmosphere model (Husser et al. 2013), adopting stellar parameters of $T_{\text{eff}} = 3496\text{K}$ and $\log(g/[\text{cm s}^{-2}]) = 4.858$ (Crossfield et al. 2016). Additionally, we included an error-scaling factor (γ) to rescale the uncertainties and ensure that the reduced χ^2 is close to unity.

The NIRISS SOSS white light curve integrates all flux within the $0.85\text{--}2.83\ \mu\text{m}$ range. This data set includes a spot crossing event near the end of transit ingress, as reported by Madhusudhan et al. (2023). We adopted a specialized software to simulate the planetary transits in front of a limb-darkened stellar disc with spots (Cracchiolo et al. 2021; Biagini et al. 2024). The stellar disk is divided radially and azimuthally into 1000 steps each, resulting in a grid of one million elements. The flux contribution from each element is computed at each time step during the transit, based on the planet’s position. A single spot is included in the fit, described by four free parameters: latitude (ϕ_{spot}), longitude (λ_{spot}), the intensity ratio ($f_{\text{spot}}/f_{\text{disk}}$) at the spot’s centre, and the spot radius ratio ($r_{\text{spot}}/r_{\text{disk}}$). Model selection based on the Bayesian Information Criterion (BIC; Schwarz 1978) indicates a first-order polynomial baseline with two free parameters (r_0 , r_1) as the preferred choice for the NIRISS SOSS data as opposed to constant or second-order polynomial models, which yielded higher BIC values by ~ 8 .

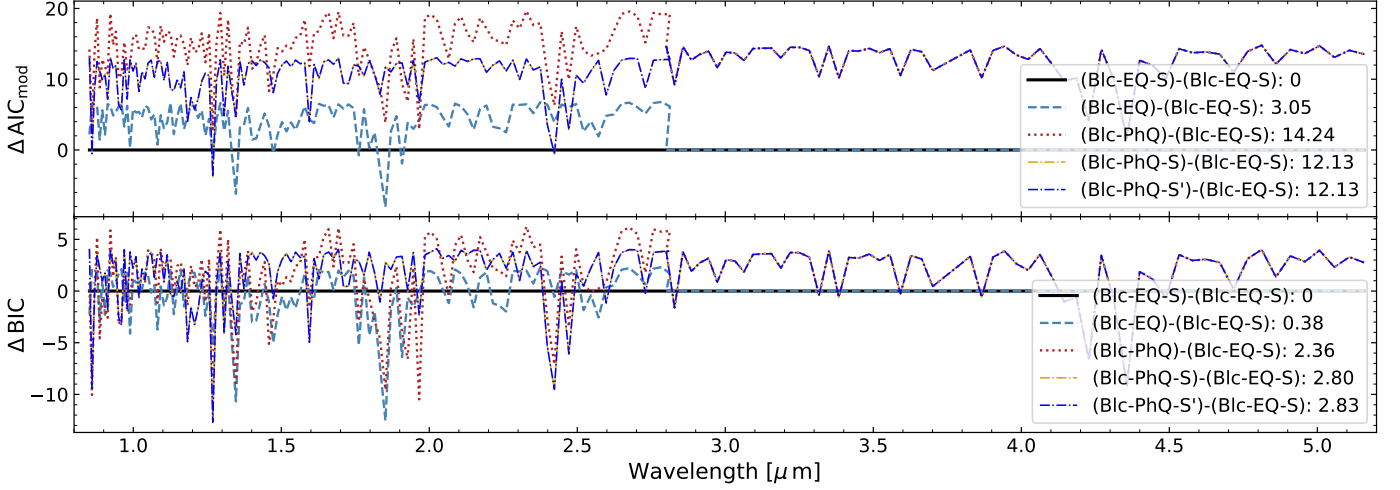


Fig. 2: Comparison between different reduction modified AICs in the top panel and BICs in the lower panel (all reduction names and their characteristics are described in Section 4.2) taking the Blc-EQ-S (binned before fitting light curves - empirical limb darkening - with spot correction) reduction as baseline. The median of the modified AIC or BIC is shown in the legend. For those cases where the NIRSpec data is shared between reductions, only the median for the NIRISS data is shown. Although there are peaks below, the majority of the points lie above the 0 ΔAIC_{mod} and ΔBIC lines.

Parameter	Prior	Posterior
NIRISS		
R_p/R_s	$\mathcal{U}(0, 1)$	$0.05400^{+0.00016}_{-0.00016}$
T_0^a	$\mathcal{N}(60096.729374, 0.000035)$	$60096.729369^{+0.000032}_{-0.000031}$
P [d]	$\mathcal{N}(32.939623, 0.0000012)$	$32.9396230^{+0.0000012}_{-0.0000013}$
τ_0 [s]	$\mathcal{U}(0, 50000)$	$8855.4^{+10.9}_{-10.1}$
b	$\mathcal{U}(0, 1)$	$0.611^{+0.016}_{-0.018}$
ϕ_{spot} [°]	$\mathcal{U}(-80, -10)$	$-31.3^{+6.3}_{-7.0}$
λ_{spot} [°]	$\mathcal{U}(-70, -15)$	$-47.9^{+2.5}_{-3.8}$
$f_{\text{disk}}/f_{\text{spot}}$	$\mathcal{U}(0, 1)$	$0.816^{+0.026}_{-0.035}$
$r_{\text{disk}}/r_{\text{spot}}$	$\mathcal{U}(0.01, 0.35)$	$0.259^{+0.058}_{-0.057}$
NIRSpec NRS1		
R_p/R_s	$\mathcal{U}(0, 1)$	$0.054254^{+0.000093}_{-0.000094}$
T_0^a	$\mathcal{N}(59964.969453, 0.000035)$	$59964.969452^{+0.000026}_{-0.000026}$
P [d]	$\mathcal{N}(32.939623, 0.0000012)$	$32.9396230^{+0.0000012}_{-0.0000012}$
τ_0 [s]	$\mathcal{U}(0, 50000)$	$8872.4^{+6.8}_{-6.9}$
b	$\mathcal{U}(0, 1)$	$0.612^{+0.010}_{-0.010}$
NIRSpec NRS2		
R_p/R_s	$\mathcal{U}(0, 1)$	$0.05402^{+0.00013}_{-0.00013}$

Table 1: Priors and posteriors for each instrument. Orbital parameters in the case of NIRSpec white light curve fits are shared.

Notes. a : Format is BJD TDB - 2400000.5

From NIRSpec, the NRS1 and NRS2 light curves integrate $2.80 - 3.72 \mu\text{m}$ and $3.83 - 5.17 \mu\text{m}$, respectively. In this case, we assumed a simpler transit model generated with PYLIGHTCURVE (Tsiasar et al. 2016). Our choice is motivated by the lack of

statistical evidence for spots, also in the original study of Madhusudhan et al. (2023). The NRS1 and NRS2 light curves were fitted simultaneously with shared orbital parameters (P , T_0 , b , and τ_0), but independent radius ratios (R_p/R_s), due to its possible wavelength-dependence. We find that the derived orbital parameter values are within uncertainties to those from the NIRISS fit, see Table 1. Similarly to NIRISS a linear ramp was preferred, as constant or quadratic alternatives resulted in higher BIC values by 5 and 7, respectively.

A summary of the most relevant priors and retrieved parameters can be found in Table 1. The priors for the orbital period, mid transit times, were taken from Madhusudhan et al. (2023), while the others are uniform. The resulting fitted light curves can be seen in Figure 1.

4.2. Spectroscopic light curves

Spectroscopic light curves were extracted from the same apertures as the white ones, using either single-pixel columns or flux integrated over wavelength bins. For all spectroscopic fits, the orbital parameters (P , T_0 , b and τ_0) were fixed to the values derived from the white light curves, while R_p/R_s , the baseline ramp parameters and error scaling factor were left free. We ran several approaches to obtain the transmission spectrum, testing the impact of varying the following procedures:

- **Spectral binning.** The transmission spectrum was derived from binned light curves (Blc) at a resolution of $R \sim 100$, following recommendations from the JWST Transiting Exoplanet Community Early Release Science team (Carter et al. 2024), or from light curves at the native pixel resolution. For comparison, we also derived the binned transmission spectrum (Bts) from that obtained at the native pixel resolution. Previous studies have suggested that $R \sim 100$ represents the lowest resolution at which JWST transmission spectra retain most of the information, providing an optimal compromise between computational efficiency and reliability of atmospheric retrievals (Davey et al. 2025);
- **Limb-darkening.** We fitted the binned light curves allowing for empirical quadratic limb-darkening coefficients (EQ),

or fixing them to the corresponding PHOENIX-model values (PhQ) computed with ExoTETHyS. For the single-column light curves, we always fixed the limb-darkening coefficients to PHOENIX-model values, considering both the quadratic (PhQ) and four-coefficient (PhC4) laws.

- **Occluded spot.** We introduce a new method to account for the spot-crossing feature in spectroscopic light-curve fitting. The spot signal is obtained as the difference between the spotted transit model from the white light-curve fit and the corresponding spotless model, created by setting $r_{\text{spot}}/r_{\text{disk}} = 0$. We then assume that the spot signal shape is consistent across wavelengths, but with a varying amplitude. Accordingly, the spectroscopic light curves are fitted as a linear combination of a spotless transit model and the spot signal (S). As a more advanced approach, we computed an ad hoc spot signal for each spectral bin by adjusting the limb-darkening coefficients to theoretical values (S'). We also ran tests without spot correction to quantify its importance.

4.3. Fiducial transmission spectrum

We performed extensive retrieval analyses on all extracted spectra to evaluate how the inferred atmospheric properties depend on choices made during data reduction, light-curve fitting, and retrieval configuration. Not all extractions are equally reliable. Previous studies have shown that JWST high-resolution modes are particularly sensitive to analysis details such as limb-darkening, likely due to non-Gaussian parameter correlations in low signal-to-noise (S/N) light curve fits (Carter et al. 2024). For this reason, we consider the Blc extractions more robust, as the binning increases the effective S/N of the fitted light curves. Within the Blc subset, those accounting for the NIRISS spot-crossing event are regarded as the most trustworthy. The use of empirical limb-darkening coefficients further minimizes model dependence and reduces potential biases from theoretical prescriptions. Accordingly, we adopt the Blc-EQ-S spectrum as our fiducial reference.

Moreover this is shown quantitatively by comparing the modified Akaike information criterion (AIC, Akaike 1974) and the BIC at each wavelength bin and deriving an average across all wavelengths, as shown in Figure 2. Both methods confirm that indeed the Blc-EQ-S reduction is moderately preferred.

5. Retrievals

We conducted a suite of free-chemistry retrievals on multiple variants of the combined transmission spectrum using TauREx 3.1 (Waldmann et al. 2015; Al-Refaie et al. 2021). The planetary atmosphere was modelled numerically with 100 layers, evenly spaced on a logarithmic pressure scale from 10^{-7} to 10^2 bar. Our retrieval setups included absorption from selected molecular species, collision-induced absorption by H_2 – H_2 and H_2 –He (Abel et al. 2011, 2012; Fletcher et al. 2018), Mie scattering following the formalism of Lee et al. (2013), and an optically thick cloud deck defined by its top pressure. The temperature–pressure profile was parametrized using the formulation of Guillot (2010), or a simpler isothermal model for comparison. For the former, we assumed an internal temperature of 60 K, which is well within the range predicted by previous studies (Blain et al. 2021; Madhusudhan et al. 2023).

We considered absorption from the most prominent molecular species expected in mini-Neptune atmospheres, including H_2O (Polyansky et al. 2018), CO (Li et al. 2015), CO_2

(Yurchenko et al. 2020), CH_4 (Yurchenko et al. 2017), NH_3 (Coles et al. 2019), HCN (Barber et al. 2014), H_2S (Azzam et al. 2016), and SO_2 (Underwood et al. 2016) (resolutions 5×10^4). Additionally, we considered potential biosignatures tested by Madhusudhan et al. (2023), though adopting newer opacity data sources for some species, such as OCS (Owens et al. 2024), CH_3Cl (Owens et al. 2018), N_2O (Yurchenko et al. 2024), CS_2 (Karlovetz et al. 2021), and $(\text{CH}_3)_2\text{S}$ or DMS. Note that for DMS and CS_2 , we extrapolated the absorption cross sections provided directly by HITRAN (Sharpe et al. 2004; Kochanov et al. 2019; Gordon et al. 2022) at 1 bar and 298 K to other temperature and pressure conditions, following the method of Madhusudhan et al. (2021, 2023). Retrievals were also carried out with selected subsets of molecules to assess the statistical significance of different contributions and evaluate how the inferred physical parameters and chemical abundances were affected by the species considered.

The atmospheric parameter space was sampled using the nested sampling algorithm PyMultiNest (Feroz et al. 2009; Buchner et al. 2014), with 3000 live points to ensure a thorough exploration. We adopted broad (log-)uniform parameter priors, which are largely uninformative. In particular, the molecular mixing ratios were allowed to vary across 17 orders of magnitude, ranging from negligible quantities to as much as 99% of the total. Along with the posterior distributions of atmospheric parameters, PyMultiNest provides the Bayesian evidence for each retrieval model, which is particularly useful for comparing nested models (Trotta 2008; Tsiraras et al. 2018). For instance, the detection significance of a specific molecule can be inferred from the difference in log-evidence between otherwise identical models, differing only by the inclusion of that molecule, a method widely used in recent studies (e.g., Guilluy et al. 2021; Madhusudhan et al. 2023; Powell et al. 2024).

6. Results

6.1. Retrievals from the Fiducial Transmission Spectrum

The atmospheric retrievals on our fiducial transmission spectrum confirm CH_4 and CO_2 as the most likely absorbers, although their absolute abundances and detection significances vary depending on the retrieval setup (see Tables 2 and B.1).

When including all molecular species in an H_2 –He-dominated atmosphere, a cloud deck, Mie scattering, and assuming a Guillot temperature–pressure profile, we obtain $\log(\text{CH}_4) = -0.96^{+0.45}_{-0.57}$ and $\log(\text{CO}_2) = -3.63^{+1.75}_{-5.13}$. Notably, the posterior for CO_2 is significantly skewed, favouring higher abundances with a peak around $\log(\text{CO}_2) \sim -1.88$, but with a long tail extending toward much lower values. Indeed, Bayesian model comparison using a leave-one-out approach indicates that only CH_4 is detected with strong statistical significance (4.16σ). Retrievals excluding CO_2 yield comparable Bayesian evidence, suggesting no robust detection of CO_2 . For other species, we report only broad 95th-percentile upper limits, such as $\log(\text{H}_2\text{O}) < -2.29$ and $\log(\text{NH}_3) < -4.41$ (see Table 4 for the full list).

Given the lack of evidence for molecules beyond CH_4 and CO_2 , we explored simplified retrievals including one or both of these species, with or without an additional molecule. A model with only CH_4 and CO_2 yields nearly the highest Bayesian evidence, with additional molecules offering negligible or no improvement. The corresponding abundances are $\log(\text{CH}_4) = -1.73^{+0.66}_{-0.75}$ and $\log(\text{CO}_2) = -3.94^{+1.68}_{-4.20}$. This solution is only marginally favoured over a pure CH_4 scenario with $\log(\text{CH}_4) = -1.78^{+0.64}_{-0.76}$, corresponding to a $\sim 2\sigma$ detection of CO_2 . Across

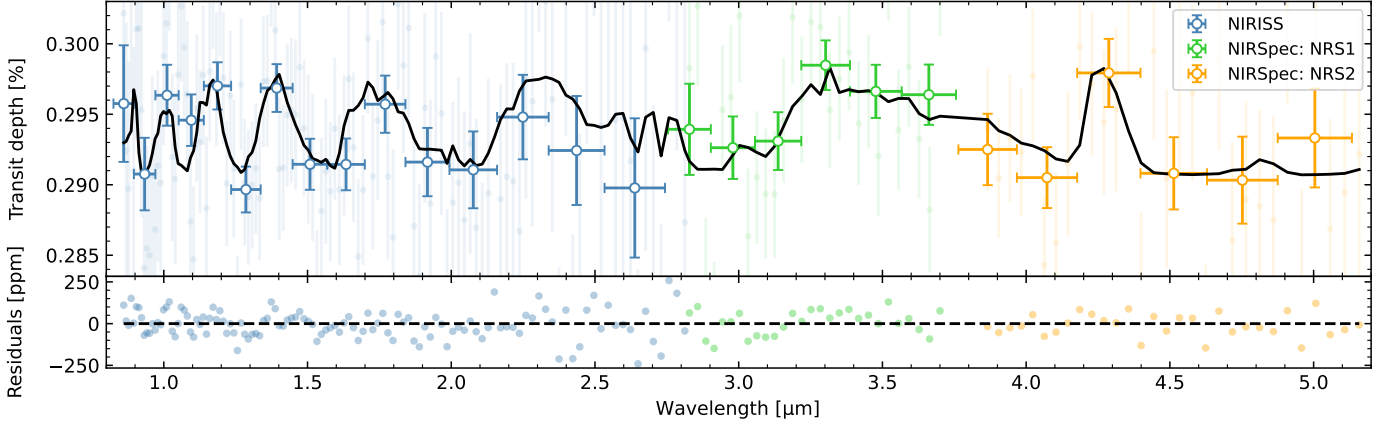


Fig. 3: Fiducial reduction and retrieval which includes: spot correction, 2 coefficient empirical limb darkening law, light curves binned to a resolution of 100 before the fit (Blc-EQ-S). The data (lighter shaded points) are binned to a resolution of $R \sim 15$ (darker shaded points) to highlight the spectral features. The bottom panel shows the residuals, at $R \sim 100$, which is the resolution at which the retrieval was done.

Table 2: Selected parameters from retrievals on the fiducial transmission spectrum (Blc-EQ-S).

Molecules	χ^2_0	log ev	log (CO ₂)	log (CH ₄)	$T_{10\text{ mbar}}$ (K)	MMW	log ($P_{\text{cloud}}/\text{bar}$)
Flat ^a	1.26	1440.86	–	–	–	–	–
CO ₂	1.34	–0.34	N	–	–	–	–
CH ₄	1.19	+7.66	–	$-1.78^{+0.64}_{-0.76}$ (4.32 σ)	124^{+44}_{-38}	$2.53^{+0.75}_{-0.19}$	$-1.30^{+1.24}_{-0.60}$
DMS	1.28	+0.61	–	–	–	–	–
CO ₂ +CH ₄ ^b	1.17	+8.62	$-3.94^{+1.68}_{-4.20}$ (2.00 σ)	$-1.73^{+0.66}_{-0.73}$ (4.62 σ)	138 ⁺⁴⁸ _{–39}	2.64 ^{+1.16} _{–0.28}	-1.14 ^{+1.35} _{–0.59}
CO ₂ +CH ₄ +DMS	1.18	+8.52	$-4.05^{+1.83}_{-4.86}$	$-1.65^{+0.65}_{-0.77}$	126^{+47}_{-38}	$2.71^{+1.31}_{-0.34}$	$-1.02^{+1.53}_{-0.75}$
All ^c	1.25	+7.49	$-3.63^{+1.75}_{-5.13}$ (N)	$-0.96^{+0.45}_{-0.57}$ (4.16 σ)	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$	$-0.88^{+1.59}_{-1.10}$

Notes. Retrieval setups including Guillot T-P profiles, an optically-thick cloud deck, and Mie scattering. The significance of detection for molecular species are reported in parenthesis. ^(a) Forcedly flat spectrum obtained by imposing the highest cloud deck, but adjusting the radius to better match the data ; ^(b) Simplified retrieval setup, including molecules with at least a 2σ detection ; ^(c) Comprehensive retrieval setup, including the 13 molecular species considered in this study.

retrieval setups, the inferred abundances are consistent within 1σ , though CH₄ in particular tends to shift toward higher values when including more molecules. This higher CH₄ abundance, together with the high upper limits for certain species, leads to a systematic increase of the mean molecular weight, ranging from $MMW = 2.53^{+0.75}_{-0.19}$ (CH₄ only) up to $MMW = 4.32^{+3.02}_{-1.45}$ (all molecules).

The retrieved temperature-pressure profiles are nearly isothermal above the 0.01 bar level, with median posterior temperatures ranging from $T_{10\text{ mbar}} \sim 124 - 161$ K across configurations, and 1σ credible intervals spanning 86 – 251 K. In our reference setup including CO₂ and CH₄, we obtain $T_{10\text{ mbar}} = 138^{+48}_{-39}$ K. All retrievals provide consistent hints of a deeper cloud layer, with the preferred setup indicating a cloud top pressure of $\log(P_{\text{cloud}}/\text{bar}) = -1.14^{+1.35}_{-0.59}$. The Mie scattering parameters remain poorly constrained.

We also tested retrieval variants assuming an isothermal profile and/or excluding Mie scattering and cloud opacity (see Table B.1). These simpler setups yield comparable Bayesian evidences, indicating that neither a thick cloud deck nor Mie scattering is statistically required by the data. The retrieved molec-

ular abundances and temperatures are consistent across setups well within 1σ .

6.2. Retrievals from Alternative Spectral Extractions

All our data processing methods yield visually similar transmission spectra for K2-18 b, with most points agreeing within 1σ (see Appendix A). Nonetheless, even these modest differences can notably affect the inferred atmospheric properties. We applied the same suite of retrieval setups to each version of the transmission spectrum and ultimately report the “simplified solution,” including only molecules with a $\geq 2\sigma$ detection, alongside the “most comprehensive” model. Key results are summarized in Table 3 and shown in Figure 4.

In all cases, CH₄ is the most abundant species and is always detected at $> 3\sigma$. However, its absolute abundance varies substantially across spectral extractions and retrieval setups, with 1σ intervals spanning $\log(\text{CH}_4) \in [-2.79, -0.23]$. The comprehensive retrievals yield systematically higher CH₄ abundances than the simplified chemical setups, though they remain consistent within 1σ . CO₂ is likely the second most abundant molecule, but not always detected with statistical significance. Consequently,

Table 3: Selected parameters from retrievals on different spectral extractions, including the simplified and comprehensive retrieval setups.

Spectrum	Molecules	χ^2_0	log ev	log (CO ₂)	log (CH ₄)	$T_{10\text{mbar}}$ (K)	MMW
Blc-EQ-S	CO ₂ +CH ₄	1.17	1449.47	-3.94^{+1.68}_{-4.20} (2.00σ)	-1.73^{+0.66}_{-0.73} (4.62σ)	138⁺⁴⁸₋₃₉	2.64^{+1.16}_{-0.28}
	All	1.25	1448.35	-3.63 ^{+1.75} _{-5.13} (N)	-0.96 ^{+0.45} _{-0.57} (4.16 σ)	161 ⁺⁹⁰ ₋₅₄	4.32 ^{+3.02} _{-1.45}
Blc-PhQ-S	CH ₄	1.31	1450.59	–	-1.95 ^{+0.67} _{-0.75} (4.56 σ)	121 ⁺⁴⁴ ₋₃₅	2.46 ^{+0.56} _{-0.13}
	All	1.39	1449.88	-4.76 ^{+2.56} _{-5.30} (N)	-0.95 ^{+0.45} _{-0.60} (4.06 σ)	146 ⁺⁸⁹ ₋₄₈	4.26 ^{+3.05} _{-1.44}
Blc-PhQ-S'	CH ₄	1.30	1450.94	–	-1.94 ^{+0.68} _{-0.76} (4.56 σ)	123 ⁺⁴³ ₋₃₇	2.46 ^{+0.59} _{-0.13}
	All	1.39	1450.20	-4.98 ^{+2.74} _{-5.24} (N)	-0.95 ^{+0.46} _{-0.59} (4.01 σ)	148 ⁺⁹⁶ ₋₄₉	4.25 ^{+3.20} _{-1.44}
Blc-EQ	CO ₂ +CH ₄	1.19	1446.63	-3.29 ^{+1.73} _{-3.44} (1.99 σ)	-1.84 ^{+0.93} _{-0.79} (4.52 σ)	155 ⁺⁶⁷ ₋₄₆	2.63 ^{+3.06} _{-0.28}
	All	1.27	1446.09	-3.00 ^{+1.56} _{-5.58} (N)	-0.80 ^{+0.41} _{-0.59} (3.75 σ)	195 ⁺¹¹⁸ ₋₇₄	5.63 ^{+3.74} _{-2.42}
Blc-PhQ	CO ₂ +CH ₄	1.36	1446.61	-1.74 ^{+0.81} _{-4.56} (2.30 σ)	-1.00 ^{+0.64} _{-1.79} (4.96 σ)	213 ⁺¹⁵⁵ ₋₇₉	6.73 ^{+5.15} _{-4.40}
	All	1.45	1447.18	-2.17 ^{+1.07} _{-5.46} (N)	-0.53 ^{+0.30} _{-0.51} (3.59 σ)	283 ⁺¹⁷⁵ ₋₁₃₂	8.83 ^{+3.92} _{-4.27}
Bts-PhQ	CO ₂ +CH ₄ +DMS ^a	1.50	1443.52	-5.54 ^{+4.45} _{-6.07} (3.89 σ)	-1.57 ^{+0.67} _{-0.89} (5.94 σ)	134 ⁺⁸⁰ ₋₃₇	2.77 ^{+5.31} _{-0.41}
	All	1.58	1444.07	-1.18 ^{+0.32} _{-1.60} (2.33 σ)	-0.64 ^{+0.33} _{-0.44} (3.83 σ)	328 ⁺¹⁸⁴ ₋₁₄₂	10.14 ^{+3.21} _{-3.39}
Bts-PhC4	CO ₂ +CH ₄ +DMS ^b	1.50	1443.63	-6.79 ^{+3.96} _{-5.23} (3.00 σ)	-1.80 ^{+0.62} _{-0.78} (5.83 σ)	125 ⁺⁴² ₋₃₂	2.58 ^{+0.84} _{-0.23}
	All	1.58	1444.01	-1.21 ^{+0.34} _{-2.11} (2.17 σ)	-0.62 ^{+0.32} _{-0.45} (3.85 σ)	321 ⁺¹⁸⁷ ₋₁₄₁	10.20 ^{+3.21} _{-3.59}
Bts-Mad23	CO ₂ +CH ₄	1.56	1423.53	-1.17 ^{+0.24} _{-1.96} (3.14 σ)	-1.03 ^{+0.43} _{-1.36} (6.13 σ)	206 ⁺⁸⁷ ₋₆₃	7.23 ^{+2.63} _{-4.81}
	All	1.66	1423.74	-1.40 ^{+0.35} _{-1.44} (2.13 σ)	-0.70 ^{+0.30} _{-0.42} (3.80 σ)	248 ⁺¹⁴⁴ ₋₉₅	7.68 ^{+2.62} _{-2.69}
Orig-Mad23	CO ₂ +CH ₄ +DMS ^c	1.08	28820.93	-1.38 ^{+0.32} _{-0.87} (3.64 σ)	-1.02 ^{+0.41} _{-0.63} (6.24 σ)	204 ⁺⁹⁷ ₋₇₃	5.94 ^{+2.78} _{-2.55}
	All	1.09	28820.72	-1.44 ^{+0.31} _{-0.57} (2.90 σ)	-0.84 ^{+0.34} _{-0.40} (4.19 σ)	254 ⁺¹²⁹ ₋₉₄	6.48 ^{+2.35} _{-1.83}

Notes. ^(a) log (DMS) = -4.60^{+0.95}_{-1.70} (2.13 σ) ; ^(b) log (DMS) = -4.46^{+0.78}_{-0.81} (2.31 σ) ; ^(c) log (DMS) = -6.29^{+2.06}_{-5.60} (2.25 σ) .

Blc: transmission spectrum derived from binned light curves at R~100; Bts: binned transmission spectrum obtained from the native pixel resolution; EQ, PhQ, PhC4: light-curve fits using empirical quadratic, PHOENIX quadratic, or PHOENIX four-coefficient limb-darkening laws, respectively; S, S': spot correction adopting either a single spot signal with a free scaling factor for each wavelength bin, or a wavelength-dependent signal accounting for limb darkening.

Table 4: 95th-percentile upper limits for the volume mixing ratios of undetected species.

Molecule	95 th Upper Limit
H ₂ O	<-2.29 / -2.83
CO	<-2.80 / -4.03
NH ₃	<-4.41 / -4.96
HCN	<-2.54 / -3.22
H ₂ S	<-2.85 / -3.36
SO ₂	<-3.20 / -3.90
DMS	<-4.39 / -4.98
OCS	<-4.61 / -5.77
CH ₃ Cl	<-3.72 / -4.84
N ₂ O	<-3.14 / -3.49
CS ₂	<-5.32 / -5.79

Notes. Derived from the Blc-EQ-S spectrum, adopting the comprehensive or simplified retrieval setups, respectively.

the corresponding 1 σ intervals span an extremely broad range, log (CO₂) $\in$ [-12, -0.93].

Retrieval parameters from the Blc spectra agree within fractions of a sigma, particularly when applying spot correction or empirical limb-darkening. The Blc spectra without spot correction exhibit a bimodal solution, with peaks corresponding to H₂-He and high-MMW scenarios. Applying the spot correction suppresses the high-MMW peak, particularly in the simplified chemical setup. The Bts spectra indicate a light or heavy atmosphere, depending on whether the simplified or comprehensive setup is used. These are the only spectra suggesting a > 3 σ detection of CO₂ and showing hints of DMS ($\gtrsim$ 2 σ) in the optimal retrievals.

Interestingly, we find a general trend across our results in which higher MMWs correlate with higher retrieved temperatures (see Figure 5). This trend is opposite to the known degeneracy for H₂-He-dominated atmospheres, where increasing abundances or temperatures both enhance the spectral features, and likely arises from the plausibility of both light- and heavy-atmosphere scenarios for this planet. In the latter, increasing MMW can have the opposite effect on spectral features by lowering the atmospheric scale height.

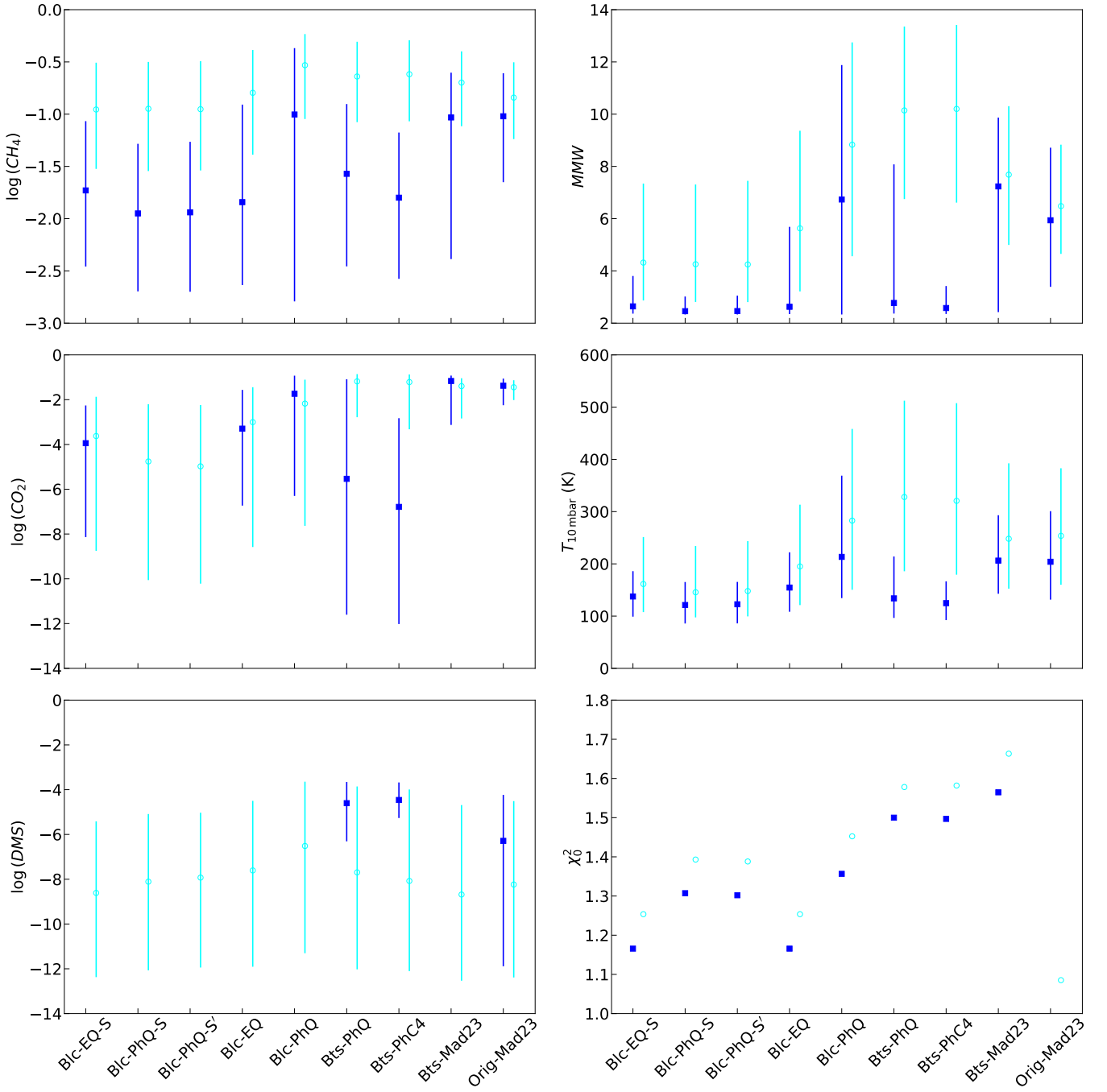


Fig. 4: Left: Main molecular abundances obtained from different spectral extractions, using simplified (blue) and comprehensive (cyan) retrieval setups. Note that CO_2 was not included in the Blc-PhQ-S and Blc-PhQ-S' simplified retrievals, as its significance was below 2σ . Similarly, DMS was not included in the majority of simplified retrievals. Right: Retrieved mean molecular weights, temperature at a pressure of 10 mbar, and reduced chi-squared values.

6.3. Retrievals from a Literature Spectrum

We also tested our retrieval setups on the transmission spectrum published by Madhusudhan et al. (2023) (hereafter “Orig-Mad23”)¹, as well as to a rebinned version using the same wavelength bins as our extractions (“Bts-Mad23”). As in the original study, the TauREx retrievals find tighter constraints on the CO_2 abundance when using the Orig-Mad23 and Bts-Mad23 spectra,

compared to any of our own extractions. The corresponding CO_2 detection significances are $> 3\sigma$ with the simplified chemical setups, and $2\text{--}3\sigma$ with the comprehensive setup. DMS is tentatively detected at 2.25σ from the Orig-Mad23 only using the simplified chemical setup, while its significance drops to 1.90σ for the Bts-Mad23 spectrum. Notably, the TauREx retrievals converge to higher CH_4 and CO_2 abundances (still consistent within 1σ) than those reported in the original study, shifting toward a high-MMW scenario, similar to our spectra without spot correction. These discrepancies are partially due to the larger priors

¹ Downloaded from: <https://osf.io/36djh/>

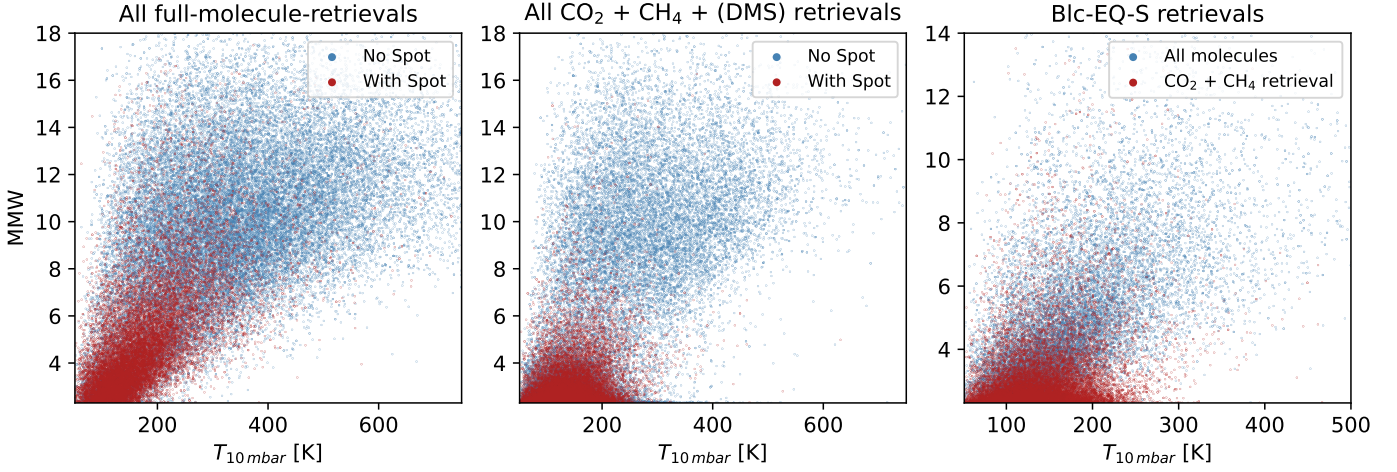


Fig. 5: Posterior distribution correlations between mean molecular weight (MMW) and temperature at 10 mbar. The left-most panel shows all retrievals which include the full suite of molecules. The middle panel shows the same but only including CO₂, CH₄ and DMS with blue and red indicating no spot and spot correction reductions respectively. The right-most panel shows all retrievals for the fiducial spectrum (Blc-EQ-S, binned before fitting light curves - empirical limb darkening - with spot correction) with blue indicating those with all molecules and red those with only CO₂ and CH₄.

adopted in this study, effectively allowing for purely metallic atmospheres and higher temperatures than might be physically expected. Nevertheless, retrieval tests using more constrained temperature and abundance priors, analogous to Madhusudhan et al. (2023), still recover a bimodal solution that includes the high-MMW mode. Additional subtle differences may arise from the use of different retrieval frameworks and/or opacity cross sections, underscoring the importance of careful code comparisons.

6.4. Retrieval Performances

The χ^2_0 values of our retrievals fall in the range 1.17–1.58, comparable to those reported in other exoplanet atmospheric studies with JWST (e.g., Banerjee et al. 2024), thus indicating a satisfactory fit. Importantly, differences in χ^2_0 across extractions are driven mainly by the size of the spectral error bars rather than a closer match to theoretical models; swapping error bars between spectra produces corresponding changes in χ^2_0 values.

The comprehensive retrieval setups systematically yield slightly higher χ^2_0 values, by 0.08–0.09, compared to the simplified chemical setups. The latter also have higher or comparable Bayesian evidences, supporting their statistical preference. Even slightly better metrics are obtained with an isothermal T-P profile and/or without Mie scattering and clouds (see Table B.1).

TauREx retrievals on the Orig-Mad23 spectrum yield $\chi^2_0 = 1.08$ –1.09, matching the AURA fit quality reported by Madhusudhan et al. (2023). For the Bts-Mad23 spectrum, we obtain $\chi^2_0 = 1.56$ –1.66, slightly higher than the average from our own spectral extractions. Even replacing the error bars in the Bts-Mad23 spectrum with those from our Blc-EQ-S, the resulting χ^2_0 would remain higher (1.42–1.51), potentially suggesting a better agreement between our fiducial spectrum and theoretical models.

6.5. Retrievals with Inflated Error Bars

As an additional test, we performed retrievals with inflated spectral error bars, scaled to enforce $\chi^2_0 = 1$ (Benneke & Seager 2013; Lueber et al. 2025). In this case, the overall atmospheric

picture remains largely unchanged, with CH₄ detected at $> 3\sigma$, while the significance of CO₂ and DMS drops below 3σ even in the Bts retrievals (see Table D.1).

7. Implications for Atmospheric Retrievals

7.1. Stellar Spot and Spectral Offsets

At zeroth order, the spot correction primarily introduces a uniform offset in the NIRISS spectrum: +15 ppm (Blc-EQ-S relative to Blc-EQ) or +24 ppm (Blc-PhQ-S/Blc-PhQ-S' relative to Blc-PhQ), with a standard deviation of 9 ppm in the corresponding difference spectra (see Figure A.1). For comparison, the mean spectral error bar at $R = 100$ is ~ 70 ppm.

In Appendix C, we show that the mean offset between the NIRISS and NIRSpec spectra largely explains the different interpretations found when including spot corrections or not. In short, spectra offsets are the main driver of the variations in the retrieved CH₄ abundance and atmospheric MMW for this K2-18 b data. We also find that the TauREx retrievals statistically prefer our Blc-EQ-S and Blc-PhQ-S spectra as-is—i.e., without additional offsets (see Table C.1)—suggesting that the spot correction might be performing effectively. Moreover, a comparison between multiple retrieval configurations can be seen in Figure E.1.

However, we note that offsets in transmission and eclipse spectra in multi-epoch and/or instrument observations can arise from multiple sources, including instrumental systematics and inter-visit variability of any kind, and can influence the retrieved molecular abundances and atmospheric properties. This effect can be mitigated if each dataset has resolvable spectral features and some overlap between them (Edwards & Changeat 2024; Edwards et al. 2024). Nevertheless, a marginally better retrieval fit does not constitute conclusive evidence that such offsets are absent.

7.2. Understanding Trends with Retrieval Complexity

A clear outcome of our thorough analysis of K2-18 b is that assumed model complexity plays a critical role in shaping atmo-

spheric inferences. In particular, at low SNR such as with this K2-18 b data, retrieval conclusions are dependent on the number of molecules included in the fit. Our simplified retrieval setups systematically lead to higher molecular detection significances than the comprehensive ones, based on direct model comparisons with the leave-one-out approach. This effect is particularly important for establishing the presence of certain absorbers. For instance, CO₂ and DMS are marginally supported only in the simplified setups (at 2–3 σ), with DMS further restricted to a few specific spectral extractions. Their vanishing significance in more comprehensive setups indicates that alternative explanations can account for the data. Differences in log-evidence alone may lead to fake detections if the range of plausible models is not adequately explored. This might be the case for the initial claims of H₂O vapour in K2-18 b from HST/WFC3 spectra (Benneke et al. 2019; Blain et al. 2021), as well as for other controversial results in the exoplanet literature (e.g., Evans et al. 2016; Merritt et al. 2020; Welbanks et al. 2025).

Beyond detection significances, we noted that the number of molecules affects the retrieved abundances, with comprehensive setups yielding higher CH₄ and atmospheric MMW (see Sections 6.1 and 6.2). It may be surprising that molecules lacking notable spectral signatures can influence retrieval results, despite broad priors allowing them to remain minimal. A careful examination shows that most undetected molecules are not strongly ruled out by the data, permitting only loose upper bounds on their abundances. We hypothesize that these molecules contribute spectral noise and/or raise the MMW (thus reducing the atmospheric scale height), thereby pushing retrievals toward higher CH₄ abundances to match the observed features.

In summary, the choice of molecules in a retrieval setup can bias results in multiple ways. Including too few species may lead to spurious detections and bias the inferred parameters as the model compensates for missing inputs. Conversely, including too many species can also introduce biases, particularly when their abundances are poorly constrained by the data. Our approach of systematically comparing many retrieval configurations provides a practical compromise, capturing the key atmospheric signals while minimizing biases from both extremes.

7.3. Comparison with Previous Studies

Our reference analysis (Blc-EQ-S extraction + simplified retrieval) recovers a CH₄ abundance very similar to the three initial values reported by Madhusudhan et al. (2023), while our CO₂ posterior distribution has a median two orders of magnitude smaller and a long tail toward negligible values, though still consistent within 1 σ . Considering all our retrieval results, the conservative 1 σ range for CH₄ is slightly broader than that obtained by Schmidt et al. (2025), whereas the 1 σ upper limit for CO₂ exceeds their 2 σ upper limit. The latest study by Hu et al. (2025), incorporating new datasets, reports CH₄ and CO₂ abundances that fall in the upper halves of our ranges (most closely matching the CH₄ range from our comprehensive retrievals), while preserving a CO₂-to-CH₄ ratio similar to ours.

Regarding DMS, we find that its detection is marginal and limited to specific extraction and retrieval configurations, consistent with the results reported in Schmidt et al. (2025).

8. Implications for Planet Formation

8.1. Overview of Formation Theories

The origin of the close-in Super-Earth / Sub-Neptune planets remains under debate. Formation theories can be divided into two broad classes: 1) Formation in the outer disk, followed by tidal migration to the inner region (e.g., Kley & Nelson 2012; Bitsch et al. 2015; Emsenhuber et al. 2021); 2) Formation *in situ* in the inner region (e.g., Hansen & Murray 2012, 2013; Chiang & Laughlin 2013; Chatterjee & Tan 2014).

The composition of the planetary “core” and atmosphere are likely to depend on the location in the protoplanetary disk where formation occurred. Note, here we define “core” as the solid/liquid component of the planet that is not part of the atmosphere. Planetary cores that are assembled, either by planetesimal or pebble accretion, in the cooler, outer disk, especially beyond the water ice line at ~ 170 K, are expected to have bulk compositions that are rich in volatile species. This would lead to the core having significant amounts of water such that a global planetary ocean is a significant fraction of the mass budget.

On the other hand, if the planetary core is assembled from small solids, i.e., pebbles or planetesimals, in the warm, inner regions interior to the water ice line, then its bulk composition is expected to be mainly composed of metals and metal-rich silicates (e.g., Cevallos Soto et al. 2022). Refractory interstellar dust also contains a carbonaceous component, often referred to as “soot”, which is significant in terms of the global carbon budget, although likely to be sub-dominant in terms of mass compared to the metals/silicates (e.g., Li et al. 2025). This soot is expected to sublimate at temperatures near ~ 500 K, with this location in the protoplanetary disk referred to as the “soot line”. Thus, planetary cores are expected to become relatively carbon poor if formed interior to the soot line, although the overall impact on bulk density may be relatively modest.

If core masses become sufficiently massive ($\gtrsim 1 M_{\oplus}$) while the planet is still embedded in a gas-rich disk, then accretion of a “significant” H/He-rich “primordial” atmosphere becomes possible. The ability of the planetary core to accrete H/He gas depends on local environmental conditions and the ability of the gas to cool. For example, even at temperatures of ~ 1000 K, planetary cores of $\sim 1 M_{\oplus}$ may accrete a few percent by mass of H/He (Thomopoulos et al., in prep.). Such an atmosphere, while a very minor component of the total mass, causes the size of the planet, i.e., defined by its optically thick photosphere, to be much larger than the core surface. Under cooler conditions and for more massive cores, the mass fraction of the primordial atmosphere is expected to increase. Finally, we note that for core masses greater than a certain critical value, $\sim 10 M_{\oplus}$, runaway gas accretion leading to gas giant formation is expected (e.g., Pollack et al. 1996; Piso & Youdin 2014).

If dominated by H/He, a primordial atmosphere will have a low MMW, i.e., 2.33 amu in the limit of pure H₂ and He with $n_{\text{He}} = 0.2n_{\text{H}_2}$. The primordial atmosphere’s abundance of heavier elements will also reflect the composition of gas in the local zone of the protoplanetary disk. For example, in astrochemical models of protoplanetary disks that including radially drifting pebbles, enhanced gas phase water is expected interior to the water ice line (e.g., Cevallos Soto et al. 2022). Indeed, elevated abundances of gas phase water have been reported in the inner regions of some disks (Banzatti et al. 2023), which is interpreted as indirect evidence for pebble drift. However, in addition to gas accretion, the composition of the primordial atmosphere could be affected by continued pebble and/or planetesimal ac-

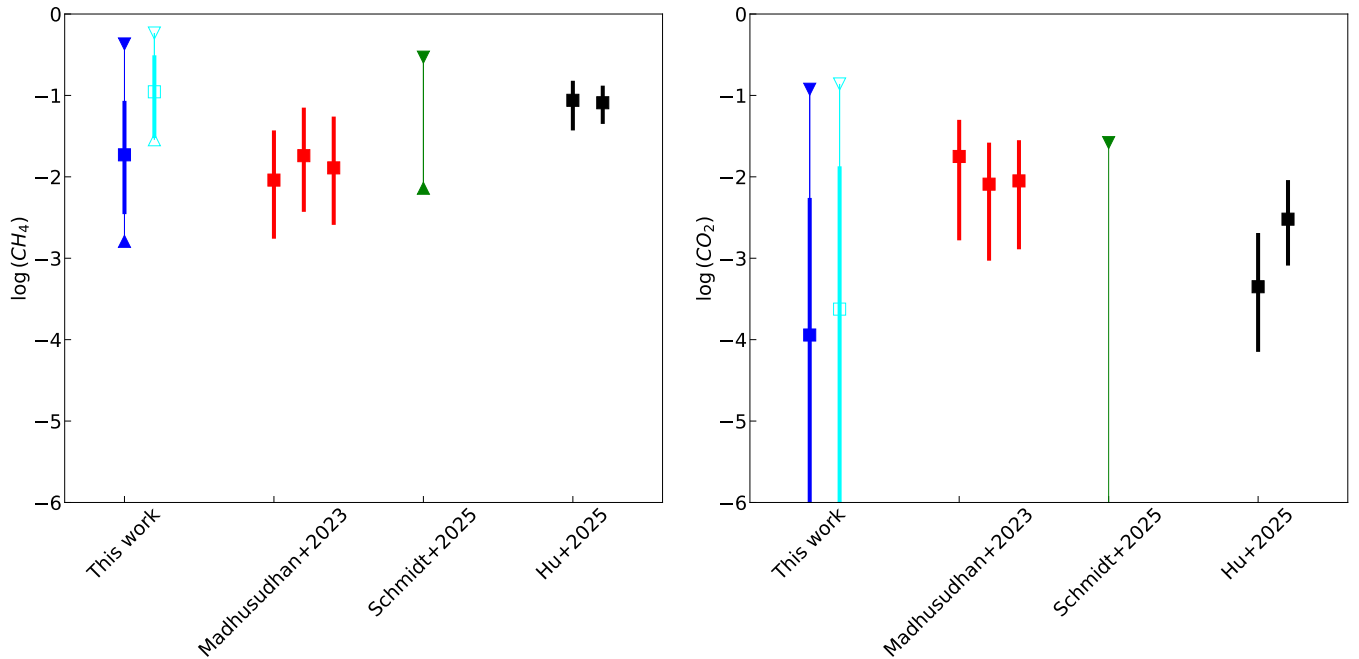


Fig. 6: Atmospheric abundances of CH_4 and CO_2 obtained in this work, compared with those reported by other papers. This work: blue denotes simplified retrieval, cyan is for comprehensive retrieval results; the thicker error lines are derived from the fiducial Blc-EQ-S spectrum, the thinner lines denote the 1σ range covered by all spectral extractions. From Madhusudhan et al. (2023): three values from their Table 2, obtained with zero, one or two free offsets between detectors. From Schmidt et al. (2025): total 1σ range for CH_4 and 2σ upper limit across multiple reductions, as reported in their Section 4.2.2. From Hu et al. (2025): Results from their ExoTR retrievals (Table 5, Shifted Average) and Aura retrievals (Table 7, Hazes). Notably, all results are consistent within about $1-2\sigma$.

cretion, which would lead to enhanced abundances of refractory elements.

Once a primordial atmosphere has formed, its composition may be further altered by mass loss due to photoevaporation, especially if this tends to preferentially remove lighter species, i.e., H and He. Indeed, Rogers & Owen (2021) have modelled the Super-Earth / Sub-Neptune population as having common properties of $\sim$ few Earth-mass rocky/iron-rich cores with initial primordial atmospheres of a few percent by mass, with the Super-Earths, i.e., below the “radius valley”, being revealed by photoevaporation of their atmospheres. Interaction with the planetary core could also alter the composition of a primordial atmosphere. For example, atmospheric material may dissolve in a global ocean. And/or volatile materials may be outgassed from the core (e.g., Heng et al. 2025). In the limiting case where this outgassed component dominates, then this “secondary” atmosphere would no longer reflect the original gaseous composition of the local protoplanetary disk.

Given the above possibilities, it is in general difficult to make definitive predictions of atmospheric composition that are unique indicators of a given formation location in the protoplanetary disk. On the other hand, for a given formation scenario, where both the location and mode of accretion are specified, we may attempt to make predictions for the atmospheres of its resulting planets. In the next sub-section we outline such predictions for the *in situ* model Inside-Out Planet Formation (Chatterjee & Tan 2014).

8.2. K2-18 b and c in the context of Inside-Out Planet Formation

Inside-Out Planet Formation (IOPF) (Chatterjee & Tan 2014) is a model for the *in situ* formation of close-in Super-Earth/Sub-Neptune multi-planet systems. Planets form from pebble-rich rings that are trapped at the pressure maximum associated with the dead zone inner boundary (DZIB) with an interior magneto-rotational-instability (MRI) active zone, with this location first set by thermal ionization of alkali metals at about 1,200 K. The location of the DZIB, i.e., where the disk midplane temperature reaches about 1,200 K, is estimated to be:

$$r_{\text{DZIB}} = 0.0653 \phi_{\text{DZIB},0.5}^{-2/9} \gamma_{1.4}^{2/9} \kappa_{10}^{-2/9} \alpha_{-3}^{1/3} m_{*,1}^{1/3} (f_r \dot{m}_{-9})^{4/9} \text{ au}, \quad (1)$$

where $\phi_{\text{DZIB},0.5} \equiv \phi_{\text{DZIB}}/0.5$ is a dimensionless parameter of order unity accounting for potential differences from a pure viscous disk model (especially expected to be somewhat less than unity due to energy extraction by a disk wind), $\gamma \equiv 1.4\gamma_{1.4}$ with fiducial normalization to a value of 1.4 for H_2 with rotational modes excited, $\kappa_{10} \equiv \kappa/10 \text{ cm}^2 \text{ g}^{-1}$ is the disk mean opacity with fiducial normalization appropriate for inner disk conditions, $\alpha_{-3} \equiv \alpha/10^{-3}$ is the Shakura-Sunyaev viscosity parameter with normalization guided by numerical simulations of dead zone regions of protoplanetary disks, $m_{*,1} \equiv m_*/M_\odot$ is the stellar mass with fiducial normalization of $1 M_\odot$, $f_r \equiv 1 - \sqrt{r_*/r_{\text{DZIB}}}$, r_* is the stellar radius, and $\dot{m}_{-9} \equiv \dot{m}_*/10^{-9} M_\odot \text{ yr}^{-1}$ is the disk accretion rate.

The innermost planet in the K2-18 system is K2-18 c. It is a non-transiting Super-Earth detected via radial velocity measurements (Cloutier et al. 2017, 2019). A later analysis by Radica

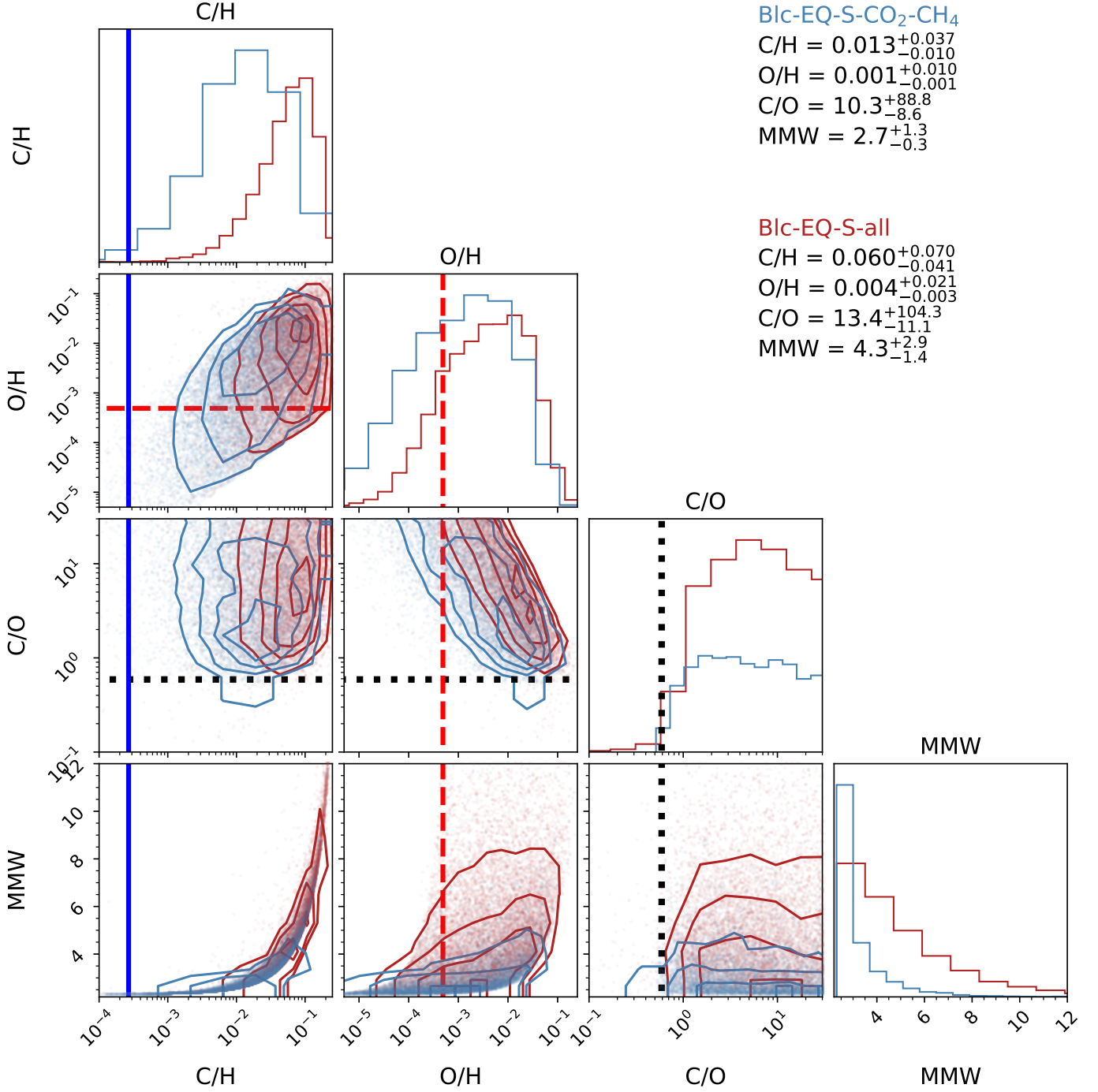


Fig. 7: Derived posterior distributions of carbon and oxygen abundances with respect to hydrogen, elemental C/O ratio and mean molecular weight (MMW) for the fiducial reduction. Red shows the comprehensive retrieval and blue the simplified retrieval. The plots are zoomed in to highlight regions of interest such as C/O $\sim 0.1 - 30$. The solid blue line is the solar C/H, the dashed red line the solar O/H, and the black dotted line the solar C/O. On the top right, the median and 1σ uncertainties for both reductions are shown.

et al. (2022b) using the line-by-line framework confirmed the presence of the planet and revised the mass and period. Nevertheless, Radica et al. (2022b) do not provide values for semimajor axis so we opt to use the values from Cloutier et al. (2019) for semimajor axis. These are a minimum mass of $6.99^{+0.96}_{-0.99} M_{\oplus}$,

semi-major axis of 0.0670 ± 0.0002 au and equilibrium temperature of 409 ± 8 K assuming a bond albedo of 0.3.

If the current location of K2-18 c, i.e., $r = 0.067$ au, is the location of the DZIB, then, given the host star mass of $m_* = 0.495 M_{\odot}$, this implies $\dot{m}_{-9} = 1.80 \phi_{\text{DZIB},0.5}^{-9/4} \alpha_{-3}^{1/2}$, where f_r and other factors have been assumed to be unity. We note that this accretion rate, i.e., $\sim 10^{-9} M_{\odot} \text{ yr}^{-1}$, is typical of those observed in transition disk systems (Manara et al. 2014), i.e., where dust

² The value obtained by Cloutier et al. (2019) is $5.62 \pm 0.84 M_{\oplus}$

appears to be depleted from the innermost regions. This derived condition on accretion rate and viscosity in the vicinity of the DZIB also places a constraint on the size of the pebbles that are trapped at this location, i.e., pebble radius $a_p > 0.15$ cm (Hu & Tan 2025).

The planet mass in the IOPF theory is assumed to be the gap opening mass. From IOPF Paper IV (Hu et al. 2018), this is given by

$$M_{p,1} = 11.1 \phi_{G,D,1.44} \phi_{DZIB,0.5}^{-9/8} \gamma_{1.4}^{5/4} \alpha_{-3}^{1/2} m_{*,1}^{-1/4} r_{0.1\text{au}}^{5/4} M_{\oplus}, \quad (2)$$

where $\phi_{G,D,1.44} \equiv \phi_{G,D}/1.44$ is a dimensionless factor with fiducial value of unity that has been calibrated from numerical simulations. Again, evaluating for K2-18 c parameters, yields $M_{p,1} = 8.00 \phi_{DZIB,0.5}^{-9/8} \alpha_{-3}^{1/2} M_{\oplus}$. Comparing to the observed mass of K2-18 c, i.e., $M_{p,1,\text{obs}} \simeq 6.99_{0.99}^{0.96} M_{\oplus}$ (Radica et al. 2022b) (note, we assume $\sin i \simeq 1$), we see that the fiducial IOPF predicted mass is consistent with K2-18 c's properties if $\alpha \simeq 10^{-3}$.

The predicted mass of K2-18 b can be estimated from that of the innermost planet given the scaling of gap opening mass:

$$M_{G,D} = 6.86 \phi_{G,D,1.44} \gamma_{1.4} \kappa_{10}^{1/4} \alpha_{-3}^{1/4} m_{*,1}^{1/8} (f, \dot{m}_{-9})^{1/2} r_{0.1\text{au}}^{1/8} M_{\oplus}, \quad (3)$$

i.e., the gap opening mass scales as $r^{1/8}$. Thus, given the properties of K2-18 c and the location of K2-18 b at 0.159 au, the latter's mass is predicted to be $(6.99 \pm 0.98) \times (0.159/0.067)^{1/8} M_{\oplus} = 7.8 \pm 1.1 M_{\oplus}$. This is consistent with the observed value of $8.63 \pm 1.35 M_{\oplus}$.

Next we consider the implications of an IOPF formation scenario for the composition of the core and atmosphere of K2-18 b. In IOPF, the solid core of the planet would be built up by pebbles that have survived to the local disk temperature at K2-18 b's location. For a viscously heated disk, the midplane temperature is predicted to scale as $T \propto \gamma_{1.4}^{-1/5} \kappa_{10}^{1/5} \alpha_{-3}^{-1/5} \dot{m}^{2/5} r^{-9/10}$. The location of K2-18 c sets the place in the disk where the midplane temperature was $\sim 1,200$ K. Thus, the local disk temperature of K2-18 b during formation is expected to be ~ 550 K. This temperature, which we note is much warmer than the current equilibrium temperature, is expected to be just interior to the “soot line”, which has fiducial expected temperature of ~ 500 K (e.g., Li et al. 2021). Thus the core of the planet is expected to be volatile poor, i.e., mainly metal/rock-rich composition, with low water and carbon content. However, because the planet is embedded in a gas-rich disk, it has the chance to accrete a primordial atmosphere, with typical mass fractions of ~ 1 to 10%. These conditions for the solid cores and initial primordial atmospheres have been inferred for the general Super-Earth/Sub-Neptune population by the photo-evaporation modelling study of Rogers & Owen (2021). In this context, for K2-18 b if the core has a mass of $\sim 8 M_{\oplus}$ average density of $\sim 6 \text{ g cm}^{-3}$, then its radius would be $1.94 R_{\oplus}$. Given the observed radius of K2-18 b of $2.61 \pm 0.09 R_{\oplus}$, the atmosphere needs to have a thickness of about 35% of the core radius. Such a primordial atmosphere, retained since the time of formation of the planet, would be in the regime of being a low mean molecular weight atmosphere.

Detailed predictions for the atmospheric composition of the innermost planets forming via IOPF have not yet been made. IOPF Paper VII (Cevallos Soto et al. 2022) presented chemodynamical models of protoplanetary disks, coupling a gas-grain astrochemical network with pebbles undergoing radial drift. They found enhanced abundances of gas-phase H_2O were delivered to regions inside the water ice line at $T \sim 170$ K. This could boost the gas-phase O abundance by factors of several, leading to gas-phase C/O ratios of ~ 0.1 . In IOPF the primordial atmospheric

composition is expected to reflect that of the gas-phase abundances in the disk, since pebble accretion would be truncated by the DZIB that has retreated further out in the disk. In this case, planets just interior to the water-ice line would have low C/O ratios. In the case of the K2-18 system, scaling from the innermost planet that is assumed to be at 1,200 K, the location of the 170 K water ice line is $r_{170\text{K}} = 0.587$ au.

However, as mentioned above, at a temperature of about 500 K, the carbonaceous “soot” component of dust is expected to sublimate. The location of this soot line in the K2-18 system is at $r_{500\text{K}} = 0.177$ au, so K2-18 b is expected to be just interior to this soot line. From stoichiometry, the abundance of “soot”, assuming it dominates the refractory C content of dust, is expected to be about $[\text{C}_{\text{soot}}/\text{H}] \sim 1.6 \times 10^{-4}$, i.e., about $10\times$ greater than that of water ice (Cevallos Soto et al. 2022), so this would boost the C/O ratio to values $\gtrsim 1$. Thus, the prediction of the IOPF model for the atmospheric composition of K2-18 b is a primordial, i.e., low MMW, atmosphere with an elevated C/O ratio.

In Figure 7 we highlight several metrics derived from our atmospheric retrievals of K2-18 b: C/H; O/H; C/O; and MMW. We see that low MMW solutions are preferred, favouring the scenario of a primordial H/He dominated atmosphere. An abundance of O that is elevated above solar, typically by a factor of 10, is the best solution, which would be consistent with gas that has been enriched in water inside the water ice line (Cevallos Soto et al. 2022). However, the retrieved C abundance is even more enhanced above solar, i.e., by factors of ~ 100 , which is consistent with the expected composition of protoplanetary disk gas interior to the soot line.

The above conclusions are consistent with a number of recent studies of K2-18 b. For example, Leconte et al. (2024) presented a detailed model of the structure of K2-18 b's atmosphere, concluding it was unlikely to be able to host a liquid water ocean below a H/He dominated atmosphere. Liu et al. (2025) have recently presented results of atmospheric retrievals of K2-18 b based on JWST NIRISS/NIRSpec and MIRI LRS data, and allowing for the presence of hazes. Their preferred solutions are low MMW (~ 2.4 amu) atmospheres with the presence of hydrocarbon hazes, whose formation would tend to indicate an elevated C/O ratio.

9. Conclusions

In this work we re-reduced and analysed the original NIRISS and NIRSpec JWST observations of the sub-Neptune K2-18 b. We have made use of publicly available pipelines in combination with custom-made pipelines producing a total of 12 transmission spectra, exploring the effects of different treatments of limb darkening, spectral binning, error inflation, instrumental offsets, and spot correction. Notably, we introduced a new semi-empirical method for correcting stellar spot crossing events. We also performed retrievals on the transmission spectrum from Madhusudan et al. (2023) and our binned version of it.

We recommend extracting a robust transmission spectrum by using binned light curves ($R \sim 100$) to mitigate correlated-parameter biases in the low-S/N regime, adopting empirical limb-darkening coefficients, and, where applicable, applying our novel semi-empirical spot correction method to minimize model-imposed biases. The validity of this approach is further supported by our statistical tests and is consistent with lessons learned in previous studies by the community (Carter et al. 2024; Davey et al. 2025). Binning is particularly effective to enhance the definition of broad features, such as those of CH_4 and CO_2

present in this atmosphere. Nevertheless some narrow or small-amplitude features may be lost when binning in wavelength, depending on the atmospheric composition and S/N of the data, therefore the optimal resolution or choice of bins may be case-dependent.

We also propose a suite of retrieval configurations to test atmospheric model dependencies, including different clouds, hazes, and temperature–pressure parametrizations, both comprehensive and minimal sets of chemical species, and criteria for selecting a simplified retrieval setup. We detect CH₄ with a significance of above 3–4 σ in all configurations. The detection of CO₂ is configuration dependent and, in most cases, around 2 σ . The absolute abundances of both species are broadly constrained and slightly dependent on the configuration, encompassing both scenarios with a primordial H₂-He atmosphere or a heavier secondary atmosphere. However our fiducial solution favours the first scenario.

Finally, by computing the C/H, O/H, and C/O ratios, we find that K2-18 b exhibits super-solar abundances of both carbon and oxygen, with a much stronger enhancement in carbon, resulting in a significantly super-solar C/O ratio. These atmospheric properties are in line with expectations of the Inside-Out Planet Formation (IOPF) model, if K2-18 b formed interior to the “soot” line. We have also shown that the physical properties (masses and orbital locations) of the innermost planet, K2-18 c, and K2-18 b are consistent with the IOPF scenario. This is the first test of IOPF via examination of both physical planetary properties and atmospheric chemical composition. Clearly larger samples of planets with measured atmospheric compositions are needed, ideally in multi-planet systems that span a range of disk conditions that straddle the soot and water ice lines.

Acknowledgements. We thank Arturo Cevallos Soto, Kevin Heng, Greg Houlihan, Xiao Hu and Antonino Petralia for helpful discussions. G.F.R. and E.Po. acknowledge support from Chalmers Astrophysics and Space Sciences Summer (CASSUM) fellowships. G.M. acknowledges financial support from the Severo Ochoa grant CEX2021-001131-S and from the Ramón y Cajal grant RYC2022-037854-I funded by MCIN/AEI/1144 10.13039/501100011033 and FSE+. J.C.T. acknowledges support from the Chalmers Initiative on Cosmic Origins (CICO) and the Virginia Institute for Theoretical Astrophysics (VITA), supported by the College and Graduate School of Arts and Sciences at the University of Virginia. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute (STScI). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127.

References

Abel, M., Frommhold, L., Li, X., & Hunt, K. L. C. 2011, *J. Chem. Phys.*, 115, 6805
 Abel, M., Frommhold, L., Li, X., & Hunt, K. L. C. 2012, *J. Chem. Phys.*, 136, 044319
 Ahrer, E.-M., Radica, M., Piaulet-Ghorayeb, C., et al. 2025, *ApJ*, 985, L10
 Akaike, H. 1974, *A New Look at the Statistical Model Identification* (Springer New York), 215–222
 Al-Refaie, A. F., Changeat, Q., Waldmann, I. P., & Tinetti, G. 2021, *ApJ*, 917, 37
 Albert, L., Lafrenière, D., René, D., et al. 2023, *PASP*, 135, 075001
 Alderson, L., Wakeford, H. R., Alam, M. K., et al. 2023, *Nature*, 614, 664–669
 Azzam, A. A. A., Tennyson, J., Yurchenko, S. N., & Naumenko, O. V. 2016, *MNRAS*, 460, 4063
 Banerjee, A., Barstow, J. K., Gressier, A., et al. 2024, *ApJ*, 975, L11
 Banzatti, A., Pontoppidan, K. M., Carr, J. S., et al. 2023, *ApJ*, 957, L22
 Barber, R. J., Strange, J. K., Hill, C., et al. 2014, *MNRAS*, 437, 1828
 Barclay, T., Kostov, V. B., Colón, K. D., et al. 2021, *AJ*, 162, 300
 Bell, T., Ahrer, E.-M., Brande, J., et al. 2022, *JOSS*, 7, 4503
 Benneke, B., Roy, P.-A., Coulombe, L.-P., et al. 2024, *arXiv e-prints*, arXiv:2403.03325
 Benneke, B. & Seager, S. 2013, *ApJ*, 778, 153

Benneke, B., Wong, I., Piaulet, C., et al. 2019, *ApJ*, 887, L14
 Bézard, B., Charnay, B., & Blain, D. 2022, *Nature Astronomy*, 6, 537
 Biagini, A., Cracchiolo, G., Petralia, A., et al. 2024, *MNRAS*, 530, 1054
 Birkmann, S. M., Ferruit, P., Giardino, G., et al. 2022, *A&A*, 661, A83
 Bitsch, B., Lambrechts, M., & Johansen, A. 2015, *A&A*, 582, A112
 Blain, D., Charnay, B., & Bézard, B. 2021, *A&A*, 646, A15
 Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *A&A*, 564, A125
 Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2023, *JWST Calibration Pipeline*
 Cadieux, C., Doyon, R., MacDonald, R. J., et al. 2024, *ApJ*, 970, L2
 Carter, A. L., May, E. M., Espinoza, N., et al. 2024, *Nature Astronomy*, 8, 1008
 Cevallos Soto, A., Tan, J. C., Hu, X., Hsu, C.-J., & Walsh, C. 2022, *MNRAS*, 517, 2285
 Chatterjee, S. & Tan, J. C. 2014, *ApJ*, 780, 53
 Chiang, E. & Laughlin, G. 2013, *MNRAS*, 431, 3444
 Claret, A. et al. 2020, *A&A*, 634, A93
 Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2017, *A&A*, 608, A35
 Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2019, *A&A*, 621, A49
 Coles, P. A., Yurchenko, S. N., & Tennyson, J. 2019, *MNRAS*, 490, 4638
 Cracchiolo, G., Micela, G., Morello, G., & Peres, G. 2021, *MNRAS*, 507, 6118
 Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, *ApJS*, 226, 7
 Darveau-Bernier, A., Albert, L., Talens, G. J., et al. 2022, *PASP*, 134, 094502
 Davenport, B., Kempton, E. M. R., Nixon, M. C., et al. 2025, *ApJ*, 984, L44
 Davey, J. J., Yip, K. H., Al-Refaie, A. F., & Waldmann, I. P. 2025, *MNRAS*, 536, 2618
 Doyon, R., Hutchings, J. B., Beaulieu, M., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8442, *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. Oschmann, Jacobus M., 84422R
 Doyon, R., Willott, C. J., Hutchings, J. B., et al. 2023, *PASP*, 135, 098001
 Edwards, B. & Changeat, Q. 2024, *The Astrophysical Journal Letters*, 962, L30
 Edwards, B., Tsirias, A., Changeat, Q., & Yip, K. H. 2024, *RAS Techniques and Instruments*, 3, 415
 Emsenhuber, A., Mordasini, C., Burn, R., et al. 2021, *A&A*, 656, A69
 Evans, T. M., Sing, D. K., Wakeford, H. R., et al. 2016, *ApJ*, 822, L4
 Feinstein, A. D., Radica, M., Welbanks, L., et al. 2023, *Nature*, 614, 670
 Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, 398, 1601
 Ferruit, P., Bagnasco, G., Barho, R., et al. 2012, in *SPIE Conference Series*, Vol. 8442, *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, 84422O
 Fletcher, L. N., Gustafsson, M., & Orton, G. S. 2018, *ApJS*, 235, 24
 Gordon, I. E., Rothman, L. S., Hargreaves, R. J., et al. 2022, *J. Quant. Spectr. Rad. Transf.*, 277, 107949
 Guillot, T. 2010, *A&A*, 520, A27
 Guilluy, G., Gressier, A., Wright, S., et al. 2021, *AJ*, 161, 19
 Hansen, B. M. S. & Murray, N. 2012, *ApJ*, 751, 158
 Hansen, B. M. S. & Murray, N. 2013, *ApJ*, 775, 53
 Heng, K., Owen, J. E., & Tian, M. 2025, *arXiv e-prints*, arXiv:2504.02499
 Hu, R., Bello-Arufe, A., Tokadjian, A., et al. 2025, *arXiv e-prints*, arXiv:2507.12622
 Hu, R., Damiano, M., Scheucher, M., et al. 2021, *ApJ*, 921, L8
 Hu, X. & Tan, J. C. 2025, *arXiv e-prints*, arXiv:2508.10755
 Hu, X., Tan, J. C., Zhu, Z., et al. 2018, *ApJ*, 857, 20
 Husser, T. et al. 2013, *A&A*, 553, A6
 Jakobsen, P., Ferruit, P., Alves de Oliveira, C., et al. 2022, *A&A*, 661, A80
 Karlovets, E. V., Gordon, I. E., Hashemi, R., et al. 2021, *J. Quant. Spectr. Rad. Transf.*, 258, 107275
 Kley, W. & Nelson, R. P. 2012, *ARA&A*, 50, 211
 Kochanov, R. V., Gordon, I. E., Rothman, L. S., et al. 2019, *J. Quant. Spectr. Rad. Transf.*, 230, 172
 Leconte, J., Spiga, A., Clément, N., et al. 2024, *A&A*, 686, A131
 Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, *ApJ*, 778, 97
 Li, G., Gordon, I. E., Rothman, L. S., et al. 2015, *ApJS*, 216, 15
 Li, J., Bergin, E. A., Blake, G. A., Ciesla, F. J., & Hirschmann, M. M. 2021, *Science Advances*, 7, eabd3632
 Li, J., Bergin, E. A., Hirschmann, M. M., et al. 2025, *arXiv e-prints*, arXiv:2508.16781
 Lim, O., Benneke, B., Doyon, R., et al. 2023, *ApJ*, 955, L22
 Liu, R., Lavvas, P., Tinetti, G., et al. 2025, *arXiv e-prints*, arXiv:2509.10947
 Lueber, A., Karchev, K., Fisher, C., et al. 2025, *ApJ*, 984, L32
 Luque, R. & Pallé, E. 2022, *Science*, 377, 1211
 Luque, R., Piaulet-Ghorayeb, C., Radica, M., et al. 2025, *A&A*, 700, A284
 Madhusudhan, N., Constantinou, S., Holmberg, M., et al. 2025, *ApJ*, 983, L40
 Madhusudhan, N., Piette, A. A. A., & Constantinou, S. 2021, *ApJ*, 918, 1
 Madhusudhan, N., Sarkar, S., Constantinou, S., et al. 2023, *ApJ*, 956, L13
 Manara, C. F., Testi, L., Natta, A., et al. 2014, *A&A*, 568, A18
 Merritt, S. R., Gibson, N. P., Nugroho, S. K., et al. 2020, *A&A*, 636, A117
 Montet, B. T., Morton, T. D., Foreman-Mackey, D., et al. 2015, *ApJ*, 809, 25
 Morello, G. 2018, *AJ*, 156, 175

- Morello, G., Claret, A., Martin-Lagarde, M., et al. 2020a, *The Journal of Open Source Software*, 5, 1834
- Morello, G., Claret, A., Martin-Lagarde, M., et al. 2020b, *AJ*, 159, 75
- Owens, A., Yachmenev, A., Thiel, W., et al. 2018, *MNRAS*, 479, 3002
- Owens, A., Yurchenko, S. N., & Tennyson, J. 2024, *MNRAS*, 530, 4004
- Piaulet-Ghorayeb, C., Benneke, B., Radica, M., et al. 2024, *ApJ*, 974, L10
- Piso, A.-M. A. & Youdin, A. N. 2014, *ApJ*, 786, 21
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icarus*, 124, 62
- Polyansky, O. L., Kyuberis, A. A., Zobov, N. F., et al. 2018, *MNRAS*, 480, 2597
- Powell, D., Feinstein, A. D., Lee, E. K. H., et al. 2024, *Nature*, 626, 979
- Radica, M. 2024, *The Journal of Open Source Software*, 9, 6898
- Radica, M., Albert, L., Taylor, J., et al. 2022a, *PASP*, 134, 104502
- Radica, M., Artigau, É., Lafrenière, D., et al. 2022b, *MNRAS*, 517, 5050
- Radica, M., Coulombe, L.-P., Taylor, J., et al. 2024, *ApJ*, 962, L20
- Radica, M., Welbanks, L., Espinoza, N., et al. 2023, *MNRAS*, 524, 835
- Reed, N. W., Shearer, R. L., McGlynn, S. E., et al. 2024, *ApJ*, 973, L38
- Rogers, J. G. & Owen, J. E. 2021, *MNRAS*, 503, 1526
- Rogers, J. G., Schlichting, H. E., & Owen, J. E. 2023, *ApJ*, 947, L19
- Rustamkulov, Z., Sing, D. K., Mukherjee, S., et al. 2023, *Nature*, 614, 659–663
- Schlawin, E., Leisenring, J., Misselt, K., et al. 2020, *AJ*, 160, 231
- Schmidt, S. P., MacDonald, R. J., Tsai, S.-M., et al. 2025, *arXiv e-prints*, arXiv:2501.18477
- Schwarz, G. 1978, *Annals of Statistics*, 6, 461
- Seager, S., Bains, W., & Hu, R. 2013, *ApJ*, 777, 95
- Sharpe, S. W., Johnson, T. J., Sams, R. L., et al. 2004, *Applied Spectroscopy*, 58, 1452
- Shorttle, O., Jordan, S., Nicholls, H., Lichtenberg, T., & Bower, D. J. 2024, *ApJ*, 962, L8
- Stevenson, K. B., Lustig-Yaeger, J., May, E. M., et al. 2025, *arXiv e-prints*, arXiv:2508.05961
- Taylor, J. 2025, *Research Notes of the American Astronomical Society*, 9, 118
- Tinetti, G., Drossart, P., Eccleston, P., et al. 2018, *Experimental Astronomy*, 46, 135
- Trotta, R. 2008, *Contemporary Physics*, 49, 71
- Tsiaras, A., Waldmann, I. P., Rocchetto, M., et al. 2016, *ApJ*, 832, 202
- Tsiaras, A., Waldmann, I. P., Tinetti, G., Tennyson, J., & Yurchenko, S. N. 2019, *Nature Astronomy*, 3, 1086
- Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, *AJ*, 155, 156
- Underwood, D. S., Tennyson, J., Yurchenko, S. N., et al. 2016, *MNRAS*, 459, 3890
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, *ApJ*, 802, 107
- Welbanks, L., Nixon, M. C., McGill, P., et al. 2025, *arXiv e-prints*, arXiv:2504.21788
- Wogan, N. F., Batalha, N. E., Zahnle, K. J., et al. 2024, *ApJ*, 963, L7
- Yurchenko, S. N., Amundsen, D. S., Tennyson, J., & Waldmann, I. P. 2017, *A&A*, 605, A95
- Yurchenko, S. N., Mellor, T. M., Freedman, R. S., & Tennyson, J. 2020, *MNRAS*, 496, 5282
- Yurchenko, S. N., Mellor, T. M., & Tennyson, J. 2024, *MNRAS*, 534, 1364

Appendix A: Comparison of Transmission Spectra

Figure A.1 presents a corner plot of all pairwise differences between the transmission spectra from various methods in this work (see Sections 4.2 and 4.3) and the spectrum from Madhusudhan et al. (2023), binned to a common wavelength grid. Transmission spectra from all our Blc setups agree well within 1σ , though stellar spot and limb-darkening treatments may lead to subtle systematic differences. For instance, the spot correction shifts the average NIRISS transit depth upward by 15 ppm or 24 ppm compared to the uncorrected cases when using empirical or PHOENIX quadratic limb-darkening coefficients, respectively, with maximum differences reaching 45 ppm or 62 ppm. Accounting for the wavelength-dependent spot signal shape due to limb-darkening alters transit depths by at most 5 ppm. The difference spectra between extractions using empirical and PHOENIX limb-darkening are more randomly scattered, though both instruments show a consistent mean offset of $\sim +15$ ppm when using empirical rather than PHOENIX coefficients.

Differences between Blc and Bts transmission spectra exhibit greater scatter, with some data points deviating by more than 1σ . The largest discrepancies occur near the detector edges, consistent with previous technical reports (e.g., Carter et al. 2024). In particular, the $2.5\text{--}4\mu\text{m}$ region, corresponding to the NIRISS–NIRSpec transition, is critical for interpreting atmospheric composition. This range includes CO_2 and DMS features, which are likely driving the different retrieval outcomes (see Section 6.2).

We also compare our transmission spectra with the first reduction published by Madhusudhan et al. (2023), rebinned to wavelength bins adopted in this study (hereafter “Bts-Mad23”). The Bts-Mad23 spectrum is morphologically closer to our Blc spectra, with standard deviations of the differences of ~ 43 and 31 ppm for the NIRISS and NIRSpec segments, respectively. In comparison, the analogous values with our Bts spectra are 69 and 36 ppm. Additionally, offsets between our spectral segments and Bts-Mad23 range from -14 to $+25$ ppm; these offsets are similar across instruments when no spot correction is applied, (Blc-EQ, Blc-PhQ, Bts-PhQ and Bts-PhC4) but differ more when our spot correction is included.

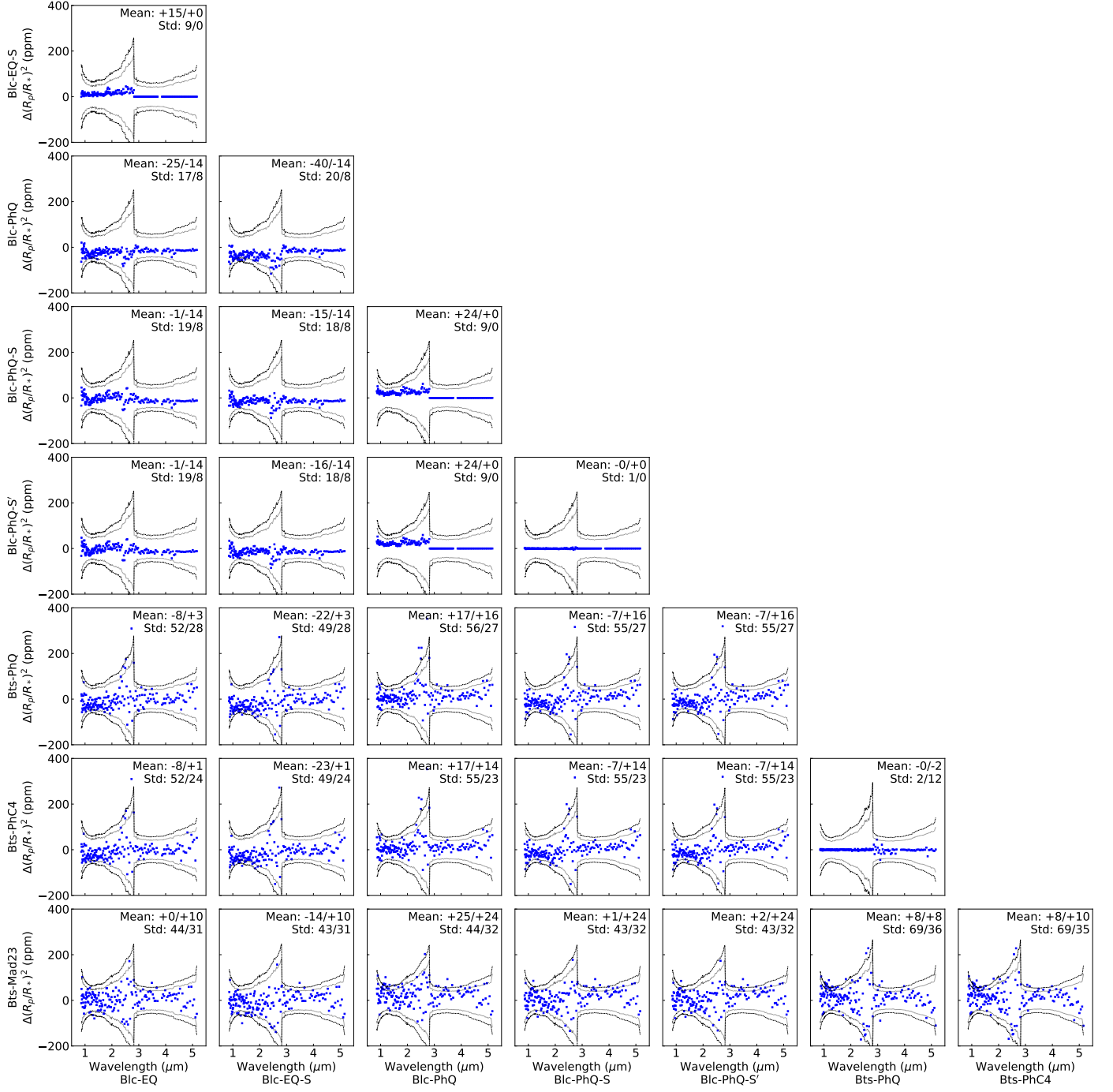


Fig. A.1: Cornerplot of difference between transmission spectra extracted with various methods (see Section 4.2).

Appendix B: Retrievals with Alternative T–P Profiles, Clouds and Haze Treatments

Table B.1 summarises the results from different retrievals on the fiducial spectrum using various temperature profiles with and without clouds/hazes.

Table B.1: Selected parameters from retrievals on the fiducial transmission spectrum (Blc-EQ-S).

Setup	χ_0^2	log ev	log (CO ₂)	log (CH ₄)	$T_{10\text{ mbar}}$ (K)	log (P_{cloud} /bar)
Gui-Cl-Mie	1.17	1448.35	-3.94^{+1.68}_{-4.20} (2.00σ)	-1.73^{+0.66}_{-0.73} (4.62σ)	138⁺⁴⁸₋₃₉	-1.14^{+1.35}_{-0.59} (N)
Gui-Cl-0	1.14	1449.34	-4.30 ^{+1.84} _{-4.90} (1.97 σ)	-1.86 ^{+0.80} _{-0.82} (4.63 σ)	131 ⁺⁵⁰ ₋₄₁	-1.12 ^{+1.42} _{-0.60} (1.96 σ)
Gui-0-Mie	1.16	1449.02	-3.92 ^{+1.91} _{-4.13} (2.11 σ)	-1.89 ^{+0.79} _{-0.97} (4.52 σ)	108 ⁺⁴² ₋₃₂	–
Gui-0-0	1.14	1448.43	-3.56 ^{+1.84} _{-3.92} (2.32 σ)	-1.99 ^{+1.01} _{-1.07} (5.53 σ)	103 ⁺⁵² ₋₃₀	–
Iso-Cl-Mie	1.15	1449.63	-4.18 ^{+1.75} _{-4.97} (2.05 σ)	-1.84 ^{+0.68} _{-0.78} (4.59 σ)	141 ⁺³⁹ ₋₃₁	-1.04 ^{+1.68} _{-0.69} (N)
Iso-Cl-0	1.12	1449.39	-3.81 ^{+2.04} _{-5.53} (2.19 σ)	-1.48 ^{+0.83} _{-0.96} (4.62 σ)	162 ⁺⁹⁵ ₋₄₇	-1.11 ^{+1.79} _{-0.82} (N)
Iso-0-Mie	1.14	1449.41	-4.10 ^{+1.87} _{-4.10} (2.22 σ)	-2.05 ^{+0.89} _{-0.90} (4.77 σ)	122 ⁺⁴³ ₋₂₅	–
Iso-0-0	1.17	1448.92	-2.84 ^{+1.46} _{-4.25} (2.44 σ)	-1.39 ^{+0.93} _{-1.33} (5.30 σ)	142 ⁺¹⁴⁰ ₋₄₁	–

Notes. Varying between Guillot or Isothermal T-P profile, inclusion of a thick cloud layer, and/or Mie scattering.

Appendix C: Retrieval Sensitivity to Instrumental Offsets

Table C.1 summarises the results from different retrievals on spot and spot-less reduction configurations including uniform offsets.

Table C.1: Selected parameters from retrievals on additional spectral extractions, analogous to Table 3.

Spectrum	Molecules	χ_0^2	log ev	log (CO ₂)	log (CH ₄)	$T_{10\text{ mbar}}$ (K)	MMW
Blc-PhQ-S	CH ₄	1.31	1450.59	–	-1.95 ^{+0.67} _{-0.75} (4.56 σ)	121 ⁺⁴⁴ ₋₃₅	2.46 ^{+0.56} _{-0.13}
	All	1.39	1449.88	-4.76 ^{+2.56} _{-5.30} (N)	-0.95 ^{+0.45} _{-0.60} (4.06 σ)	146 ⁺⁸⁹ ₋₄₈	4.26 ^{+3.05} _{-1.44}
Blc-PhQ-S -24 / 0	CO ₂ +CH ₄	1.36	1444.74	-1.11 ^{+0.33} _{-3.89} (2.42 σ)	-0.91 ^{+0.55} _{-1.82} (5.02 σ)	243 ⁺¹³⁷ ₋₉₈	9.19 ^{+4.06} _{-6.85}
	All	1.45	1445.70	-1.88 ^{+0.85} _{-4.95} (1.82 σ)	-0.47 ^{+0.27} _{-0.44} (3.48 σ)	359 ⁺¹⁸⁰ ₋₁₆₂	9.99 ^{+3.69} _{-4.09}
Blc-PhQ	CO ₂ +CH ₄	1.36	1446.61	-1.74 ^{+0.81} _{-4.56} (2.30 σ)	-1.00 ^{+0.64} _{-1.79} (4.96 σ)	213 ⁺¹⁵⁵ ₋₇₉	6.73 ^{+5.15} _{-4.40}
	All	1.45	1447.18	-2.17 ^{+1.07} _{-5.46} (N)	-0.53 ^{+0.30} _{-0.51} (3.59 σ)	283 ⁺¹⁷⁵ ₋₁₃₂	8.83 ^{+3.92} _{-4.27}
Blc-EQ-S	CO ₂ +CH ₄	1.17	1449.47	-3.94 ^{+1.68} _{-4.20} (2.00 σ)	-1.73 ^{+0.66} _{-0.73} (4.62 σ)	138 ⁺⁴⁸ ₋₃₉	2.64 ^{+1.16} _{-0.28}
	All	1.25	1448.35	-3.63 ^{+1.75} _{-5.13} (N)	-0.96 ^{+0.45} _{-0.57} (4.16 σ)	161 ⁺⁹⁰ ₋₅₄	4.32 ^{+3.02} _{-1.45}
Blc-EQ-S -26 / 0	CO ₂ +CH ₄	1.23	1443.64	-1.36 ^{+0.47} _{-2.94} (2.62 σ)	-1.17 ^{+0.68} _{-1.65} (4.90 σ)	211 ⁺¹²⁰ ₋₇₄	6.76 ^{+4.49} _{-4.42}
	All	1.31	1444.50	-1.71 ^{+0.66} _{-4.24} (1.99 σ)	-0.51 ^{+0.29} _{-0.45} (3.65 σ)	331 ⁺¹⁷² ₋₁₄₁	9.38 ^{+3.64} _{-3.62}
Bts-PhQ	CO ₂ +CH ₄ +DMS	1.50	1443.52	-5.54 ^{+4.45} _{-6.07} (3.89 σ)	-1.57 ^{+0.67} _{-0.89} (5.94 σ)	134 ⁺⁸⁰ ₋₃₇	2.77 ^{+5.31} _{-0.41}
	All	1.58	1444.07	-1.18 ^{+0.32} _{-1.60} (2.33 σ)	-0.64 ^{+0.33} _{-0.44} (3.83 σ)	328 ⁺¹⁸⁴ ₋₁₄₂	10.14 ^{+3.21} _{-3.39}
Bts-PhQ +26 / 0	CO ₂ +CH ₄	1.44	1449.16	-4.20 ^{+1.81} _{-5.12} (1.98 σ)	-2.05 ^{+0.62} _{-0.65} (5.33 σ)	132 ⁺⁴¹ ₋₃₄	2.48 ^{+0.55} _{-0.14}
	All	1.53	1448.15	-4.26 ^{+2.26} _{-5.51} (N)	-1.19 ^{+0.49} _{-0.55} (4.46 σ)	143 ⁺⁹³ ₋₄₆	3.34 ^{+3.17} _{-0.89}

Notes. Testing the impact of offsets between the NIRISS SOSS and NIRSpec G395H spectral intervals. We report difference in average transit depth between NIRISS and NIRSpec, from top to bottom: -4, -28, -28 ppm || -2, -28 ppm || -29, -3 ppm. Spectra with similar differences lead to more consistent retrieval results, with manual offsets compensating for the effects of spot correction and treatment of limb-darkening. This behaviour holds, only in part, when comparing Bts and Blc spectra with offsets.

Appendix D: Retrieval Tests with Inflated Error Bars

Figure D.1 shows the comparison between all different reductions with simplified and comprehensive retrievals for a number of parameters. Table D.1 shows the effect of inflated error bars on selected parameters and reductions.

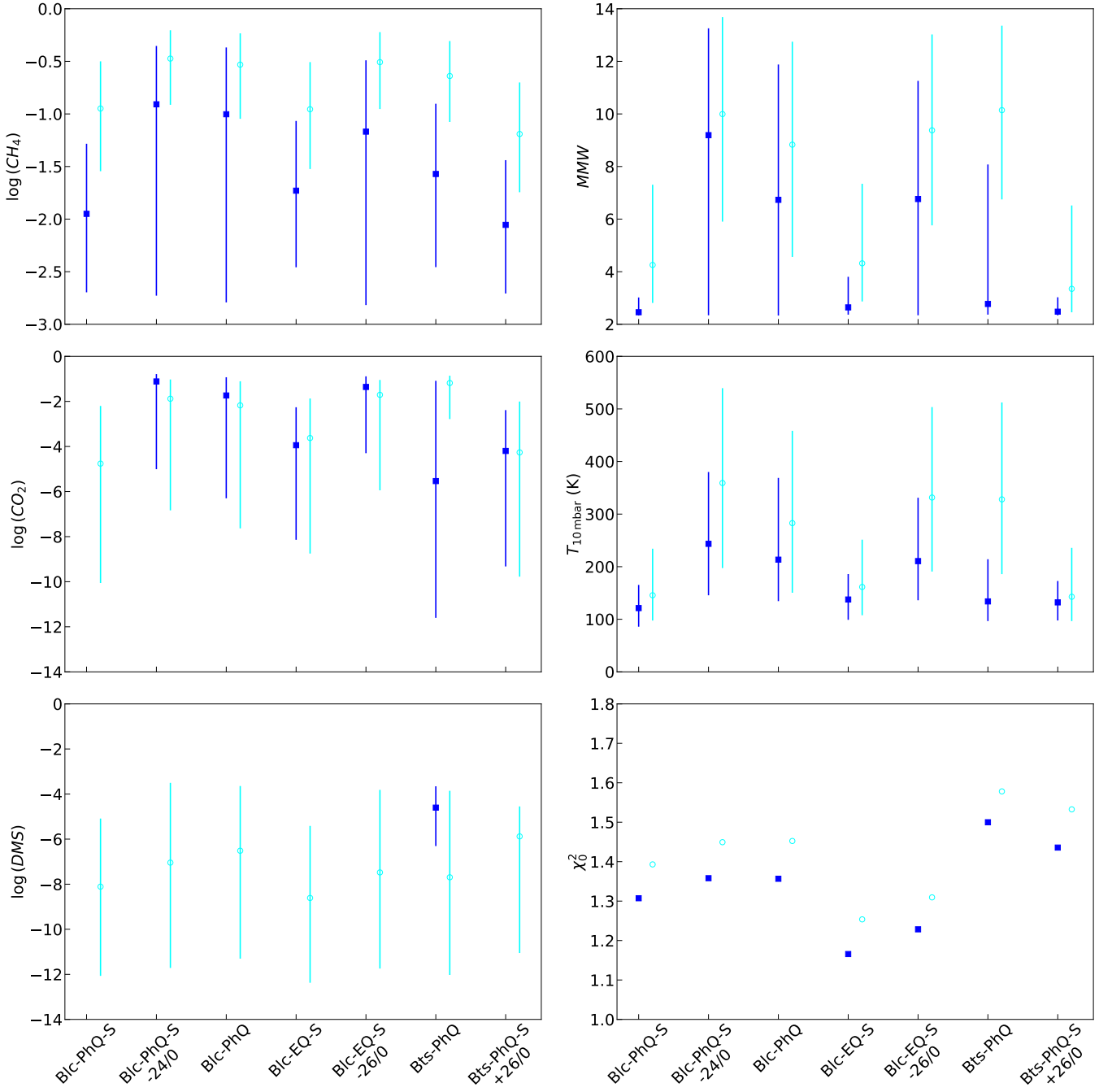


Fig. D.1: Analogous to Figure 4, but comparing spectra with and without offsets (see also Table C.1).

Table D.1: Selected parameters from retrievals on additional spectral extractions, analogous to Table 3.

Spectrum	Molecules	χ_0^2	log ev	log (CO ₂)	log (CH ₄)	$T_{10\text{ mbar}}$ (K)	MMW
Blc-EQ-S	CO ₂ +CH ₄	1.17	1449.47	$-3.94^{+1.68}_{-4.20}$ (2.00 σ)	$-1.73^{+0.66}_{-0.73}$ (4.62 σ)	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
	All	1.25	1448.35	$-3.63^{+1.75}_{-5.13}$ (N)	$-0.96^{+0.45}_{-0.57}$ (4.16 σ)	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$
Blc-EQ-S Err $\times$ 1.08	CO ₂ +CH ₄	1.00	1450.17	$-4.39^{+1.97}_{-4.82}$ (1.86 σ)	$-1.77^{+0.66}_{-0.81}$ (4.24 σ)	131^{+49}_{-41}	$2.61^{+1.03}_{-0.26}$
	All	1.08	1448.90	$-4.47^{+2.42}_{-5.48}$ (N)	$-0.89^{+0.43}_{-0.57}$ (3.75 σ)	163^{+96}_{-57}	$4.59^{+3.26}_{-1.64}$
Bts-PhQ	CO ₂ +CH ₄ +DMS ^a	1.50	1443.52	$-5.54^{+4.45}_{-6.07}$ (3.89 σ)	$-1.57^{+0.67}_{-0.89}$ (5.94 σ)	134^{+80}_{-37}	$2.77^{+5.31}_{-0.41}$
	All	1.58	1444.07	$-1.18^{+0.32}_{-1.60}$ (2.33 σ)	$-0.64^{+0.33}_{-0.44}$ (3.83 σ)	328^{+184}_{-142}	$10.14^{+3.21}_{-3.39}$
Bts-PhQ Err $\times$ 1.225	CO ₂ +CH ₄ +DMS ^b	1.00	1449.56	$-5.80^{+4.28}_{-5.63}$ (2.32 σ)	$-1.83^{+0.79}_{-1.08}$ (4.54 σ)	129^{+71}_{-38}	$2.58^{+2.91}_{-0.25}$
	All	1.06	1449.81	$-1.62^{+0.70}_{-5.76}$ (N)	$-0.64^{+0.36}_{-0.58}$ (3.12 σ)	334^{+189}_{-154}	$9.71^{+3.98}_{-4.49}$

Notes. Testing the impact of inflated spectral error bars, potentially accounting for systematic errors or discrepancies with atmospheric models. ^(a) log (DMS) = $-4.60^{+0.95}_{-1.70}$ (2.13 σ); ^(b) log (DMS) = $-4.86^{+1.05}_{-2.91}$ (2.03 σ).

Appendix E: Comparison between retrieval configurations

Figure E.1 compares the posterior distributions of some selected parameters for three reduction configurations.

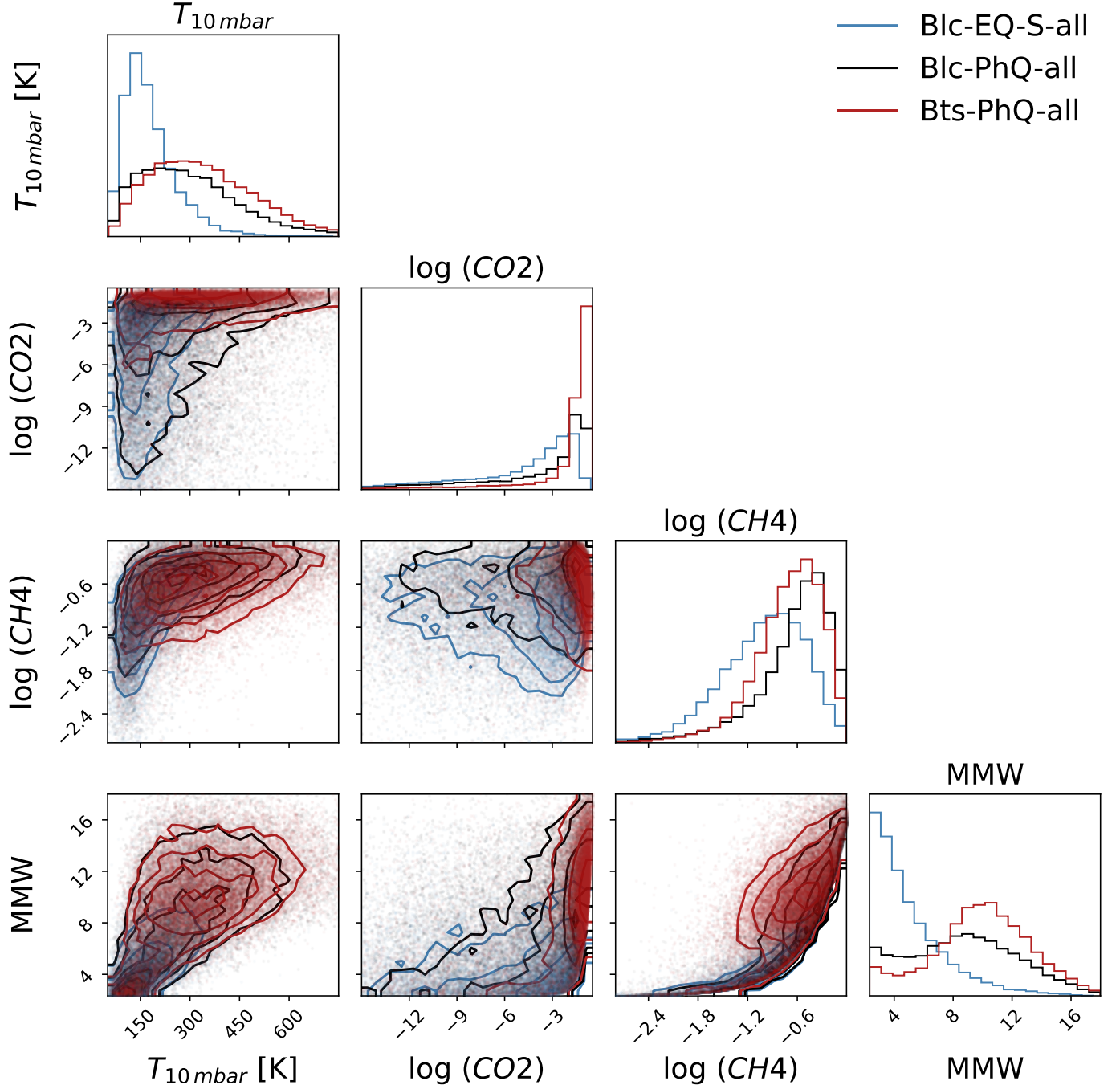


Fig. E.1: Posterior distributions of some selected parameters representative of the major reduction families with comprehensive retrievals. Blue shows our fiducial reduction corresponding to binned before fitting light curves - empirical limb darkening - with spot correction, black shows the binned before fitting light curves - PHOENIX model fixed quadratic limb darkening (no spot correction), and red the binned transmissions spectrum - PHOENIX model fixed quadratic limb darkening (no spot correction).