Astronomy & Astrophysics manuscript no. output
May 4, 2026

The atmospheric composition of the sub-Neptune K2-18 b and insights into its formation

Gareb Fernández-Rodríguez[1,2,3], Giuseppe Morello[4,5], Jonathan C. Tan[3,5], Enric Pallé[1,2], Mark R. Swain[7], Efthymios Poultourtzidis[3,8], Alfredo Biagini[5], Quentin Changeat[9], Chengzi Jiang[1,2], Francisco J. Pozuelos[4], Pedro J. Amado[4]

[1] Instituto de Astrofísica de Canarias (IAC), 38205 La Laguna, Tenerife, Spain
[2] Departamento de Astrofísica, Universidad de La Laguna (ULL), 38206 La Laguna, Tenerife, Spain
[3] Dept. of Space, Earth & Environment, Chalmers University of Technology, Gothenburg, Sweden
[4] Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain
[5] INAF- Palermo Astronomical Observatory, Piazza del Parlamento, 1, 90134 Palermo, Italy
[6] Dept. of Astronomy & Virginia Institute for Theoretical Astronomy, University of Virginia, Charlottesville, VA, USA
[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 October 16, 2025; accepted April 17, 2026

ABSTRACT

Unlocking the atmospheres of sub-Neptune planets is among 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 re-analysis of the JWST transmission spectrum of the temperate sub-Neptune K2-18 b, aimed at assessing the robustness of previously claimed atmospheric detections, exploring the planet's possible parameter space, and determining the implications behind its formation. The NIRISS/SOSS and NIRSpec/G395H observations were reduced using a combination of public and customised 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 performed atmospheric retrievals using `TauREx 3`, comparing models of varying complexity. We robustly detected CH_4 at a $>3\sigma$ significance across the majority of reduction and retrieval setups. Hints of CO_2 appear in most configurations, but the evidence is weak and strongly model-dependent. The tentative detection of dimethyl sulphide, 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 atmosphere characterised by a lower mean molecular weight. Furthermore, the assumed complexity of the retrieval model itself introduces significant biases; including more molecules systematically increases the retrieved CH_4 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 that is significantly greater than solar. We show that the physical properties of the K2-18 system planets (i.e. the 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. Furthermore, these properties predict a temperature of K2-18 b at the time of formation of $\gtrsim$ 500 K. This is much warmer than the current equilibrium temperature and just interior to the carbon 'soot' line, where an elevated C/O ratio and a correspondingly volatile-poor (dry) atmospheric composition is expected to be inherited from the protoplanetary disk, arguing against an Hycean world scenario.

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 $8.63\pm1.35\,M_\oplus$ (Cloutier et al. 2019) and a radius of $2.610\pm0.087\,R_\oplus$ (Benneke et al. 2019), the planet lies in a degenerate region of parameter space with respect to its interior composition. It is plausible that K2-18 b is host to a water-rich core with a modest H_2-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 stellar irradiation comparable to what is received by 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 *Hubble* Space Telescope (HST) Wide Field Camera 3 (WFC3), prior to the launch of JWST, revealed a prominent absorption feature at 1.4 μm. It was initially reported as a $\gtrsim3\sigma$ 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 μm 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 0.9–5.2 μm, combining NIRISS single Object Slit-

Article published by EDP Sciences, to be cited as https://doi.org/10.1051/0004-6361/202557718

less Spectroscopy (SOSS) and NIRSpec/G395H observations. The first analysis of these observations reported CH_4 ($\sim 5\sigma$), CO_2 ($\sim 3\sigma$), and tentative dimethyl sulphide ($(CH_3)_2S$ or DMS) detections (Madhusudhan et al. 2023). Along with the non-detection of H_2O, CO, and NH_3, 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 re-analysis of the same JWST data confirmed the CH_4 detection, but raised doubts about the CO_2 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 CO_2 at 3.7σ and non-detection of DMS.

The tentative detection of DMS in K2-18 b has sparked interest due to its potential as a biosignature gas (Seager et al. 2013), although an abiotic origin cannot be ruled out (Reed et al. 2024). A 6–12 μm transmission spectrum obtained with JWST's Mid-Infrared Instrument (MIRI) low-resolution spectroscopy (LRS) was initially reported as further evidence of 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. 2026).

In this work, we present an independent re-analysis 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 is also aimed at providing 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 the evidence points to a H_2-He dominated atmosphere with a high C/O ratio, which is consistent with the 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 (BOTS) mode, F290LP filter, SUB2048 sub-array, and the NRSRAPID readout pattern. The resulting spectrum was measured with the NRS1 and NRS2 detectors, spanning wavelengths of $2.8-3.72\,\mu$m and $3.83-5.17\,\mu$m 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 sub-array, and NISRAPID readout pattern, resulting in a wavelength coverage of $0.85-2.83\,\mu$m for the first spectral order and $0.6-1.1\,\mu$m 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 three 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 fittings and an extraction of the transmission spectrum. This approach allowed 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 `exoTEDRF` (Feinstein et al. 2023; Radica et al. 2023; Radica 2024). Our reduction followed 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 7σ for the time-domain outlier rejection algorithm (Radica et al. 2024), similarly to Ahrer et al. (2025), although we note that we used a more conservative value in this work. 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 the removal of the aforementioned noise, the background model was re-added and the ramp linearity is fitted. Background photons were 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; however, we did skip the flat-fielding step. In stage 3, we performed the 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 and, therefore, there is bound to be some spectral contamination. More complex extraction algorithms that are able to account for this effect are indeed accessible, such as the ATOCA algorithm; however, these effects can be neglected in relative flux measurements (Darveau-Bernier et al. 2022; Radica et al. 2022a). In this work, we only considered NIRISS SOSS order 1 data. The second-order spectroscopic light curves are more strongly affected by systematics due to lower overall flux levels,

making spectral extractions less reliable. Moreover, this selection makes a closer comparison with Madhusudhan et al. (2023) more feasible.

Table 1: Priors and posteriors for each instrument white light-curve fits.

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_{disk}/f_{spot}	$\mathcal{U}(0, 1)$	$0.816^{+0.026}_{-0.035}$
r_{disk}/r_{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	$\mathcal{U}(0, 1)$	$0.05402^{+0.00013}_{-0.00013}$

Notes. Orbital parameters in the case where NIRSpec white light-curve fits are shared. [(a)] Format is BJD TDB - 2400000.5

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 10σ, in line with the Early Release Science (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 the 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 $T_{\rm eff} = 3496$ K and $\log(g/[\mathrm{cm\,s^{-2}}]) = 4.858$ (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 0.85–2.83 μm 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 specialised 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_{spot}/f_{disk}) at the spot's centre, and the spot radius ratio (r_{spot}/r_{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.

From NIRSpec, the NRS1 and NRS2 light curves are integrate as 2.80 – 3.72 μm and 3.83 – 5.17 μm, respectively. In this case, we assumed a simpler transit model generated with `PYLIGHTCURVE` (Tsiaras et al. 2016). Our choice was motivated by the lack of statistical evidence of 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.

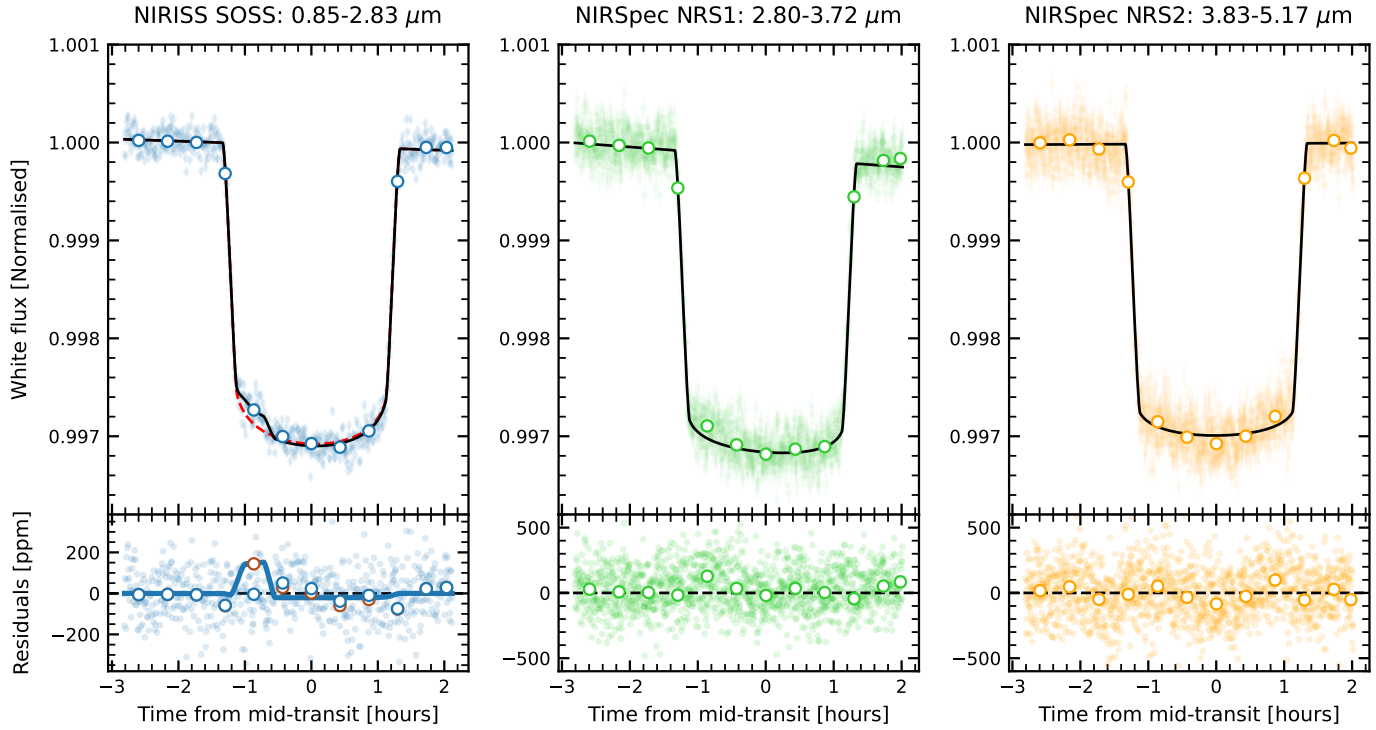

Fig. 1: White light curves for NIRISS, NIRSpec NRS1, and NIRSpec NRS2 with the best-fit models. The blue line in the left-most panel residuals (NIRISS) indicates the resulting profile from the subtraction of the spot and spotless models. The dashed red line is the spotless model. The coloured points are 25 minute binned light curves.

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

4.2. Spectroscopic light curves

The 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 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; Carter et al. 2024; Taylor et al. 2023);
- Limb-darkening. We fit the binned light curves by 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.
- Occulted spot. We introduced a new method to account for the spot-crossing feature in spectroscopic light-curve fitting. The spot signal can be obtained as the difference between the spotted transit model from the white light-curve fit and the corresponding spotless model, created by setting $r_{\rm spot}/r_{\rm disk} = 0$. 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 any 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 minimises model

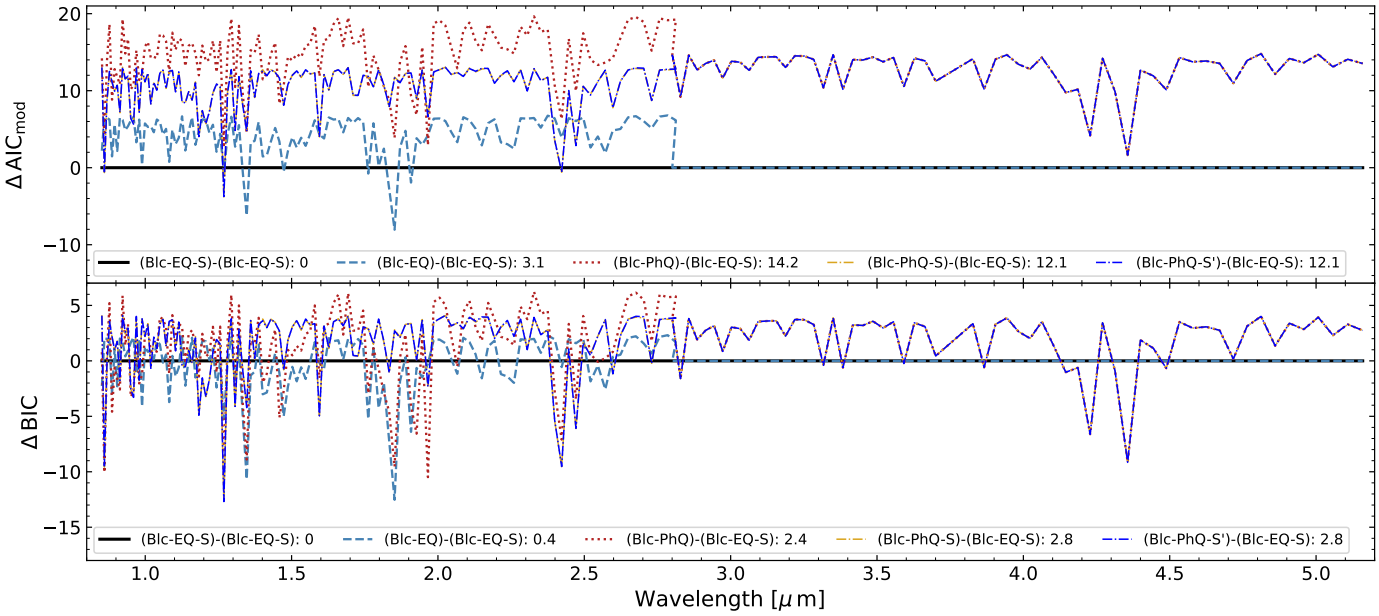

Fig. 2: Comparison between different reduction modified AICs in the top panel and BICs in the lower panel (all reduction names and their characteristics are described in Sect. 4.2) taking the Blc-EQ-S (binned before fitting light curves - empirical limb darkening - with spot correction) reduction as 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 $\Delta\,\mathrm{AIC_{mod}}$ and $\Delta\,\mathrm{BIC}$ lines.

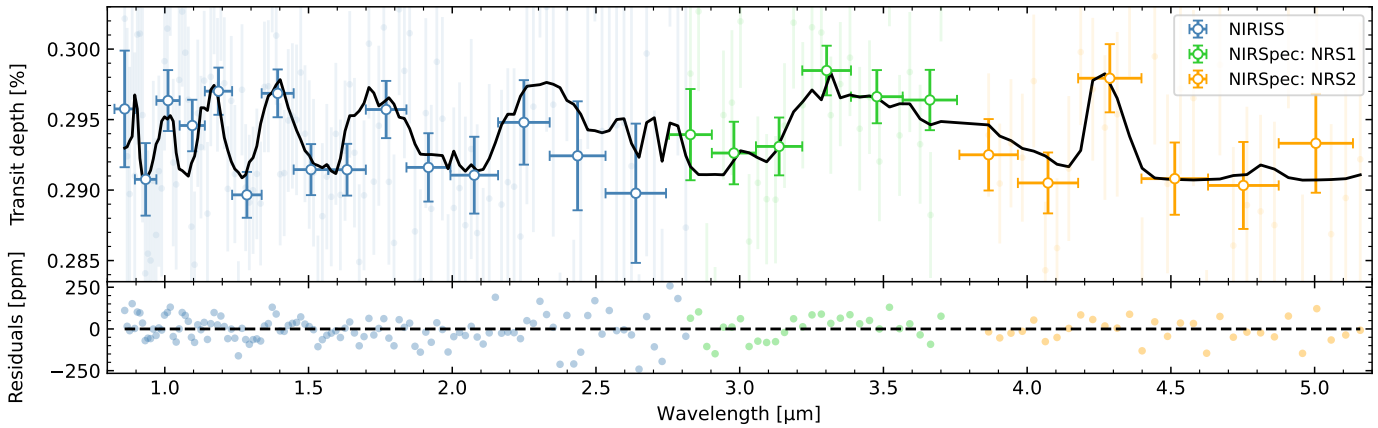

Fig. 3: Fiducial reduction and retrieval which includes: spot correction, two-coefficient empirical limb darkening law, light curves binned to a resolution of 100 before the fit (Blc-EQ-S). The data (lighter shaded points) are binned to a resolution of $R \sim 15$ (darker shaded points) to highlight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

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 Fig. 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. Hydrostatic equilibrium was assumed, with the planetary reference radius defined at the lower boundary of the model atmosphere (100 bar) and treated as a free parameter. The planetary mass and stellar radius were fixed to the literature values reported in Table 2.

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), a parametric Mie scattering prescription following the formalism of Lee et al. (2013), and an optically thick cloud deck defined by its top

Table 2: Parameters and priors used in the retrieval setups.

Parameter	Prior
Free	
R_p [$R_\oplus$]	$2.55514 \times \mathcal{U}(0.5, 1.5)$
T_{irr} [K]	$\mathcal{U}(10, 1000)$
$\log(k_{irr})$	$\log \mathcal{U}(10^{-4}, 0.5)$
$\log(k_1)$	$\log \mathcal{U}(10^{-4}, 0.5)$
$\log(k_2)$	$\log \mathcal{U}(10^{-4}, 0.5)$
α	$\mathcal{U}(0, 1)$
$\log(CH_4)$	$\log \mathcal{U}(10^{-15}, 0.99)$
⋮	...
$\log(P_{clouds})$ [bar]	$\log \mathcal{U}(10^{-7}, 100)$
R_{mie}^{lee} [μm]	$\log \mathcal{U}(10^{-5}, 10)$
Q_{ext}	$\mathcal{U}(1, 1000)$
$\log\left(P_{mie}^{lee}\right)$ [bar]	$\log \mathcal{U}(10^{-2}, 10^{7})$
$\log\left(\chi_{mie}^{lee}\right)$	$\log \mathcal{U}(10^{-7}, 100)$
Fixed	
M_p [$M_\oplus$]	8.63^a
He/H_2	0.17
$T_{int,Gui}$ [K]	60
$R_\star$ [$R_\odot$]	0.4445^a
T_{eff} [K]	3457^a
P_{min} [bar]	10^{-7}
P_{max} [bar]	100
$N_{P-layers}$	100

Notes. Only CH_4 is shown but the most comprehensive retrievals also include: H_2O, CO, CO_2, NH_3, H_2S, C_2H_6S, OCS, CH_3Cl, CS_2, N_2O, and DMS ($(CH_3)_2S$). All molecules have the same priors. [(a)] Benneke et al. (2019)

pressure. A fixed He/H_2 ratio of 0.17 was assumed throughout. In this framework, Rayleigh scattering corresponds to the small-particle limit of the Mie prescription; additional tests including explicit H_2 and He Rayleigh scattering produced indistinguishable results.

The temperature–pressure profile was parametrised 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). For all these molecules, we adopted cross-sections with a nominal spectral resolution of $R = 5 \times 10^4$ provided by the ExoMol team in a `TauREx`-compatible format. Additional retrievals were performed using alternative CH_4 opacity data to assess the robustness of our results (see Sect. 6.4). Furthermore, 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 (Karlovets et al. 2021), and $(CH_3)_2S$ or DMS. For DMS and CS_2, we used the absorption cross-sections provided by HITRAN (Sharpe et al. 2004; Kochanov et al. 2019; Gordon et al. 2022) at 1 bar and 298 K, assuming the same cross-sections at all other pressures and temperatures, as in 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 (see Table 2). 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; Tsiaras 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).

In this study, we converted the differences in log-evidence to detection significances following Schmidt et al. (2025), which adopts a more conservative mapping corresponding to one sigma lower than the commonly used conversion from Benneke & Seager (2013). For transparency and reproducibility, we additionally report Δ log ev and Bayes factors, allowing for a preferred significance convention in future studies (Kipping & Benneke 2025).

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 3 and C.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 $\log(CH_4) = -0.96^{+0.45}_{-0.57}$ and $\log(CO_2) = -3.63^{+1.75}_{-5.13}$. Notably, the posterior for CO_2 is significantly skewed, favouring higher abundances with a peak around $\log(CO_2) \sim -1.88$, 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 statistical significance of 3.2σ ($\Delta \log ev = 7.0$, $B = 1110$). 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 $\log(H_2O) < -2.29$ and $\log(NH_3) < -4.41$ (see Table 4 for the full list).

Table 3: Selected parameters from retrievals on the fiducial transmission spectrum (Blc-EQ-S).

Molecules	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	MMW	log (P_{cloud}/bar)
Flat[a]	1.26	1440.9	–	–	–	–	–
CO_2	1.34	−0.3	N	–			
CH_4	1.19	+7.7	–	$-1.78^{+0.64}_{-0.76}$	124^{+44}_{-38}	$2.53^{+0.75}_{-0.19}$	$-1.30^{+1.24}_{-0.60}$
				3.3σ, +7.7, 2122			
DMS	1.28	+0.6	–	–			
CO_2+CH_4[b]	1.17	+8.6	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$	$-1.14^{+1.35}_{-0.59}$
			1.0σ, +1.0, 2.6	3.6σ, +9.0, 7717			
CO_2+CH_4+DMS	1.18	+8.5	$-4.05^{+1.83}_{-4.86}$	$-1.65^{+0.65}_{-0.77}$	126^{+47}_{-38}	$2.71^{+1.31}_{-0.34}$	$-1.02^{+1.53}_{-0.75}$
All[c]	1.25	+7.5	$-3.63^{+1.75}_{-5.13}$	$-0.96^{+0.45}_{-0.57}$	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$	$-0.88^{+1.59}_{-1.10}$
			N, +0.6, 1.8	3.2σ, +7.0, 1110			

Notes. Retrieval setups including Guillot T-P profiles, an optically thick cloud deck, and Mie scattering. MMW denotes the mean molecular weight. Below the molecular abundance values, we report the detection significances (N denotes a non-detection), the corresponding $\Delta \log ev$ (relative to the analogous setup without that molecule) and the Bayes factors. [(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.

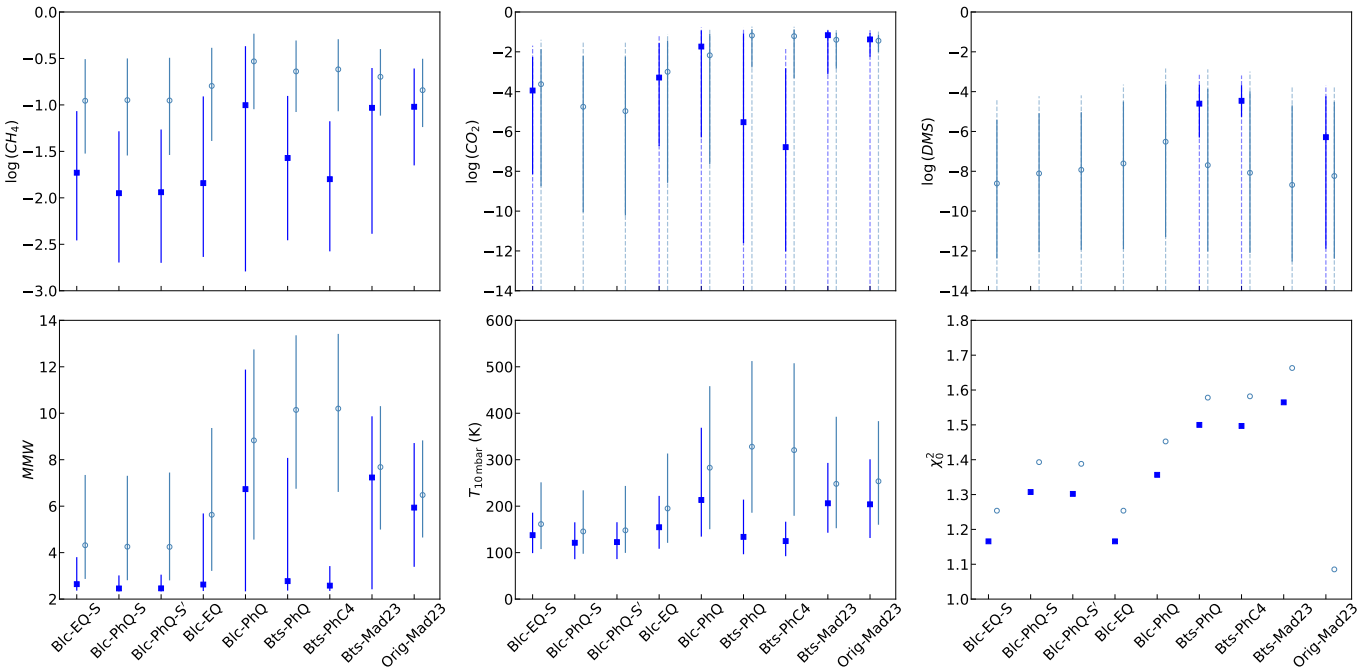

Fig. 4: Top: Main molecular abundances obtained from different spectral extractions, using simplified (blue) and comprehensive (light blue) retrieval setups. Note that CO_2 was not included in the Blc-PhQ-S and Blc-PhQ-S′ simplified retrievals. Similarly, DMS was not included in the majority of simplified retrievals. Dashed lines indicate 95% upper limits for CO_2 and DMS. Bottom: Retrieved MMWs, temperature at a pressure of 10 mbar, and reduced chi-squared values.

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 $\log(CH_4) = -1.73^{+0.66}_{-0.75}$ and $\log(CO_2) = -3.94^{+1.68}_{-4.20}$. This solution is only marginally favoured over a pure CH_4 scenario with $\log(CH_4) = -1.78^{+0.64}_{-0.76}$, corresponding to a $\sim 1\sigma$ ($\Delta \log ev = 1.0$, $B = 2.6$) detection of CO_2. Across retrieval setups, the inferred abundances are consistent within 1σ, though CH_4 in particular tends to shift toward higher values when including more molecules. This higher CH_4 abundance, together with the high upper limits for certain species, leads to a systematic increase of the mean molecular weight (MMW), ranging from $MMW = 2.53^{+0.75}_{-0.19}$ (CH4 only) up to $MMW = 4.32^{+3.02}_{-1.45}$ (all molecules).

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

Molecule	95^{th} Upper Limit
H_2O	<-2.29 / -2.83
CO	<-2.80 / -4.03
NH_3	<-4.41 / -4.96
HCN	<-2.54 / -3.22
H_2S	<-2.85 / -3.36
SO_2	<-3.20 / -3.90
DMS	<-4.39 / -4.98
OCS	<-4.61 / -5.77
CH_3Cl	<-3.72 / -4.84
N_2O	<-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 retrieved temperature-pressure profiles are nearly isothermal above the 0.01 bar level, with median posterior temperatures ranging from $T_{10\,\mathrm{mbar}} \sim 124 - 161$ K across all configurations, and 1σ credible intervals spanning $86 - 251$ K. In our reference setup including CO_2 and CH_4, we obtain $T_{10,\mathrm{mbar}} = 138^{+48}_{-39}$ K. All retrievals provide consistent hints of a deeper cloud layer, with the preferred setup indicating a cloud top pressure of $\log(P_{\mathrm{cloud}}/\mathrm{bar}) = -1.14^{+1.35}_{-0.59}$. 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 C.1). These simpler setups yield comparable Bayesian evidence, indicating that neither a thick cloud deck nor Mie scattering is statistically required by the data. The retrieved molecular 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 at least tentative piece of evidence for their presence ($\Delta \log \mathrm{ev} \gtrsim 1$), alongside the 'most comprehensive' model. Key results are summarised in Table B.1 and shown in Fig. 4.

In all cases, CH_4 is the most abundant species and is always detected at $> 2.6\sigma$ ($\Delta \log \mathrm{ev} > 4.9$, $B > 140$). However, its absolute abundance varies substantially across spectral extractions and retrieval setups, with 1σ intervals spanning $\log(CH_4) \in [-2.79, -0.23]$. The comprehensive retrievals yield systematically higher CH_4 abundances than the simplified chemical setups, though they remain consistent within 1σ. CO_2 is likely the second most abundant molecule, but it is not detected with statistical significance. Consequently, the corresponding 1σ intervals span an extremely broad range, $\log(CO_2) \in [-12, -0.93]$.

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 H_2–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 $\sim 3\sigma$ detection of CO_2 and showing hints of DMS in the optimal retrievals.

We found a general trend across our results in which higher MMWs correlate with higher retrieved temperatures (see Fig. 5) which is consistent with Benneke & Seager (2012) and Wakeford et al. (2017). 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.

6.3. Retrievals that include a free inter-instrument offset

Free offsets between spectral segments from different instruments or epochs are frequently included in atmospheric retrievals to account for potential instrumental systematics or unmodelled astrophysical effects. We performed retrievals that included a free offset between the NIRISS and NIRSpec segments. A model comparison based on the Bayesian evidence did not indicate a clear overall preference for the inclusion of this additional parameter in our spectral extractions; in fact, it is disfavoured in those that incorporate spot correction. Nevertheless, allowing a free offset increases the consistency of the retrieved atmospheric properties across the different spectral extractions, with parameters that are generally closer to those obtained from the Blc-EQ-S spectrum (with or without an offset). Quantitative comparisons are provided in Appendix D.

6.4. Sensitivity to methane opacity choices

The choice of molecular opacity data can impact atmospheric retrievals, potentially shifting abundances or species detection significances (Chubb et al. 2024). For our original grid of retrievals, we used CH_4 opacities in the form of cross-sections at $R = 50\,000$, based on the YT34to10 line list (Yurchenko et al. 2017) and provided by the ExoMol team for use with `TauREx`. We also tested more recent line lists, from HITEMP (Hargreaves et al. 2020), for which we computed cross sections at $R = 50\,000$, and the latest MM dataset from ExoMol (Yurchenko et al. 2024), for which `TauREx`-compatible cross-sections are available at $R = 15\,000$.

The retrieval results show slight variations, at the level of a fraction of a sigma, when only the CH_4 cross-sections are changed. These differences are accompanied by marginal or statistically insignificant improvements in the Bayesian log-evidence and reduced chi-squared. We observe a systematic trend toward slightly lower CH_4 mixing ratios, by a few tenths of a dex, when adopting the more recent cross-sections, particularly those from HITEMP. A detailed comparison of the retrieval outcomes is provided in Appendix E.

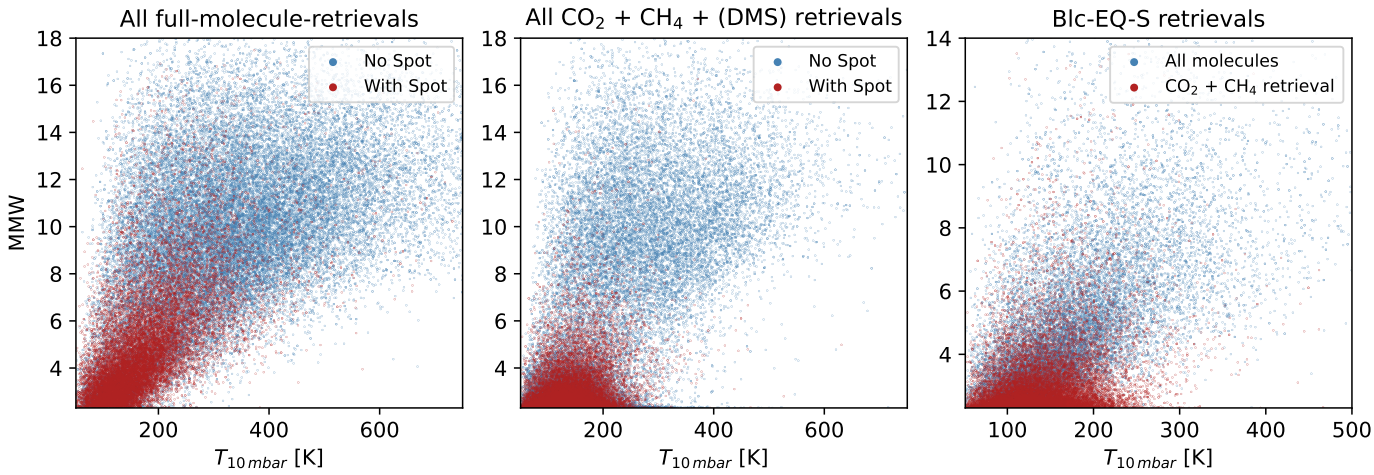

Fig. 5: Posterior distribution correlations between MMW and temperature at 10 mbar. Left-most panel: All retrievals that include the full suite of molecules. Middle panel: Same but only including CO_2, CH_4, and DMS. Blue and red indicate no spot and spot correction reductions respectively. Right-most panel: All retrievals for the fiducial spectrum (Blc-EQ-S, binned before fitting light curves - empirical limb darkening - with spot correction). Blue indicates those with all molecules and red those with only CO_2 and CH_4.

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. The detection significance of CH_4 decreases slightly but remains $> 2.1\sigma$ ($\Delta \log ev > 3.5$, $B > 33$) in all cases, while the significance of CO_2 drops well below 2σ even for the Bts-PhQ simplified retrieval (see Table F.1).

6.6. 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 re-binned 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. However, the CO_2 detection significances remain $< 3\sigma$, and $< 2\sigma$ for the comprehensive retrieval setups. Hints of DMS appears only for the Orig-Mad23 spectrum and simplified setup (1.3σ, $\Delta \log ev = 1.4$, $B = 4$).

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 adopted in this study. 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.

[1] Downloaded from: `https://osf.io/36djh/`

6.7. 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 evidence. Both indicators favour the simpler models from a statistical viewpoint. Slightly better metrics are also obtained with an isothermal T-P profile and/or without Mie scattering and clouds (see Table C.1), despite these setups being less physically realistic.

`TauREx` retrievals on the Orig-Mad23 spectrum yield χ^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.

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 Fig. A.1). For comparison, the mean spectral error bar at $R = 100$ is ~70 ppm.

In Appendix D we show that the mean offset between the NIRISS and NIRSpec spectra largely explains the different interpretations found when including spot corrections or not. Allow-

ing for a free offset between the datasets in the retrievals leads to significantly more consistent atmospheric parameters across the different spectral variants. Notably, the Blc-EQ-S, Blc-PhQ-S, and Blc-PhQ-S′ spectra require the smallest offsets, which are also not statistically required according to the Bayesian evidence (see Table D.2). Comparable consistency is obtained when fixed offsets are applied to equalise the mean separation between the NIRISS and NIRSpec segments. In short, spectral offsets appear to drive the variations in the retrieved CH_4 abundance and atmospheric MMW, while our spot-correction methods reduce the need for an offset between the two K2-18 b datasets. A comparison between multiple retrieval configurations can be seen in Fig. G.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 to firmly indicate 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 atmospheric inferences. In particular, at low S/N 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 when assessing the presence of certain absorbers. For instance, CO_2 and DMS are marginally supported only in the simplified setups (at $<3\sigma$), 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 spurious detections if the range of plausible models is not adequately explored. This might be the case for the initial claims of H_2O 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. 2026).

Conversely, the comprehensive retrievals yield systematically higher CH_4 and atmospheric MMW (see Table B.1). It may seem counterintuitive that molecules without notable spectral signatures can influence the retrieval results, when broad priors allow their abundances to remain minimal. However, a careful examination shows that most undetected species are only weakly constrained by the data, leading to loose upper bounds on their abundances.

This behaviour is analogous to the well-known issue of positively biased eccentricities in orbital fits of eclipsing binaries, where noisy data can lead to non-zero estimates even when the true eccentricity is zero (Lucy & Sweeney 1971). In our context, overly complex chemical setups may overfit the transmission spectra, with weakly constrained parameters interacting through correlations and degeneracies. For example, the inclusion of heavy molecules increases the atmospheric MMW, which reduces the scale height and consequently alters the strengths of spectral features, potentially biasing the inferred abundances of dominant absorbers such as CH_4.

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. Temperature constraints and model degeneracies

Our retrieved temperatures at a pressure depth of 10 mbar are generally lower than the zero-albedo, full-redistribution equilibrium value of ~275 K. From the Blc-EQ-S fiducial spectrum, using our simplified retrieval setup, we obtain $T_{10\,\mathrm{mbar}} = 138^{+48}_{-39}$. Changing the Guillot T–P profile to an isothermal atmosphere, varying the cloud prescription, and including free instrumental offsets all produce largely overlapping 1σ ranges spanning roughly 70–260 K. A similarly broad range is obtained when applying the comprehensive retrieval setup to the fiducial spectrum, yielding $T_{10\,\mathrm{mbar}} = 161^{+90}_{-54}$. We also note that using newer methane cross-sections leads to systematically lower temperatures by ~15–30 K. While the upper end of these temperature ranges is consistent with a Bond albedo of $A_B \sim 0.2$, most setups favour significantly higher albedos, with $A_B \gtrsim 0.7$ at 1σ.

Such high albedos would imply a high-altitude reflective cloud deck, but it would likely mute the absorption features to much smaller amplitudes than those observed. An alternative explanation is inefficient heat redistribution, with the terminator sampling temperatures closer to the colder nightside (MacDonald et al. 2020; Pluriel et al. 2022). Finally, we note that alternative spectral extractions, including Bts-PhQ, Bts-PhC4, and those from Madhusudhan et al. (2023), extend the retrieved temperature range up to ~500 K. This sensitivity to the choice of spectral extraction, combined with known parameter degeneracies and model limitations (Rocchetto et al. 2016; Jaziri et al. 2024), suggests that the atmospheric temperatures inferred from retrievals should be interpreted with caution.

7.4. Comparison with previous studies

Our reference analysis (Blc-EQ-S extraction + simplified retrieval) recovers a CH_4 abundance very similar to the three initial values reported by Madhusudhan et al. (2023), while our CO_2 posterior distribution has a median two orders of magnitude smaller and a long tail toward negligible values, though still consistent within 1σ. Considering all our retrieval results, the conservative 1σ range for CH_4 is slightly broader than that obtained by Schmidt et al. (2025), whereas the 1σ upper limit for CO_2 exceeds their 2σ upper limit. The latest study by Hu et al. (2025), incorporating new datasets, reports CH_4 and CO_2 abundances that fall in the upper halves of our ranges (most closely matching the CH_4 range from our comprehensive retrievals), while preserving a CO_2-to-CH_4 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).

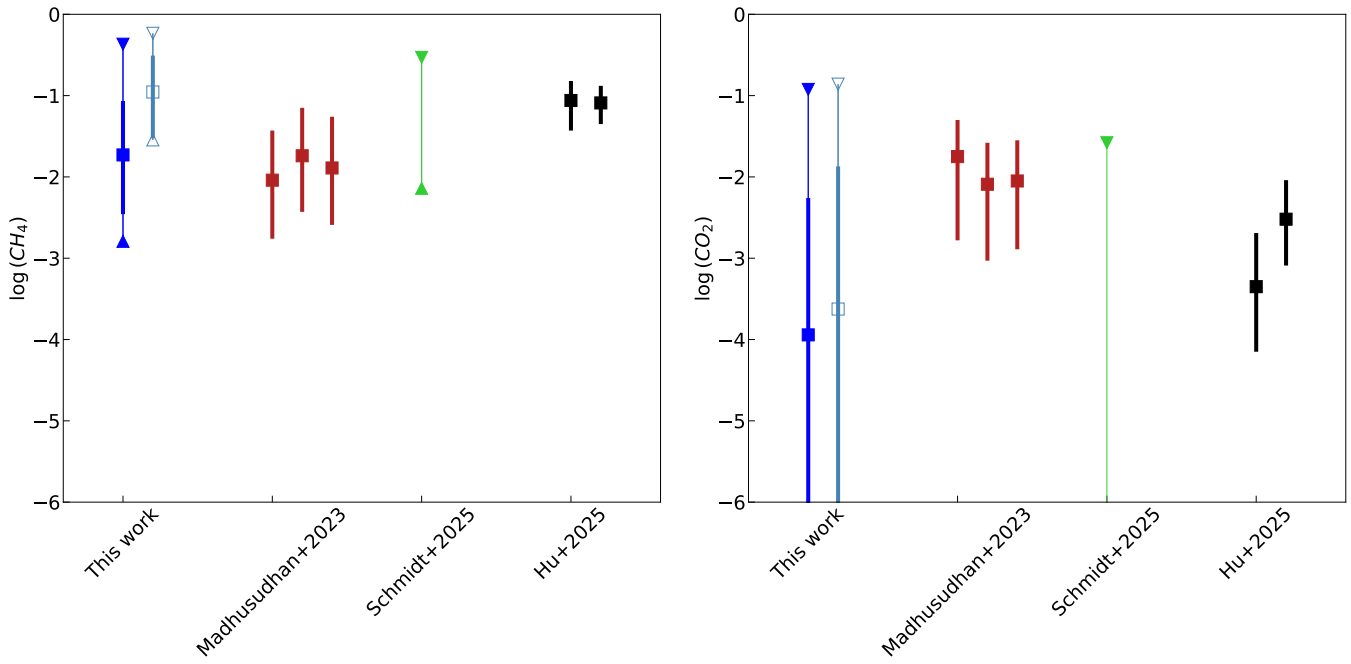

Fig. 6: Atmospheric abundances of CH_4 and CO_2 obtained in this work, compared with those reported by other papers. This work: blue denotes simplified retrieval, cyan is for comprehensive retrieval results; the thicker error lines are derived from the fiducial Blc-EQ-S spectrum, the thinner lines denote 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 Sect. 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 1-2σ.

8. Implications for planet formation

8.1. Overview of formation theories

The origin of the close-in super-Earth and sub-Neptune planets remains a matter of 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); and (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 $\sim$ 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 and silicates (e.g. Li et al. 2026). This soot is expected to sublimate at temperatures near $\sim$ 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 $\sim$ 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 H_2 and He with $n_{He} = 0.2 n_{H2}$. 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). 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 of pebble drift. However, in addition to gas accretion, the composition of the primordial atmosphere could be affected by continued pebble and/or planetesimal accretion, which would lead to enhanced abundances of refractory elements.

Once a primordial atmosphere has formed, its composition can be further altered by mass loss due to photoevaporation, especially if this tends to preferentially remove lighter species (i.e. H and He). Rogers & Owen (2021) modelled the super-Earth and sub-Neptune population as having common properties of a few Earth-mass rocky and iron-rich cores with initial primordial atmospheres of a few percent by mass, with the super-Earths (i.e. below the 'radius valley'), which is revealed by the photoevaporation of their atmospheres. Interactions with the planetary core could also alter the composition of a primordial atmosphere. For example, atmospheric material may dissolve in a global ocean. Furthermore, volatile materials might 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 subsection we outline such predictions for the in situ inside-out planet formation (IOPF) model (Chatterjee & Tan 2014).

8.2. K2-18 b and c in the context of inside-out planet formation

Inside-out planet formation (Chatterjee & Tan 2014) is a model for the in situ formation of close-in super-Earths and/or sub-Neptunes 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 active zone, with this location first set by thermal ionization of alkali metals at about 1,200 K. From Eq. 11 in Chatterjee & Tan (2014) the location of the DZIB, i.e. where the disk mid-plane temperature reaches about 1,200 K, is estimated to be

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

where $\phi_{\rm DZIB,0.5} \equiv \phi_{\rm DZIB}/0.5$ 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, $\kappa_{10} \equiv \kappa/10$ cm^2 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 normalisation guided by numerical simulations of dead zone regions of protoplanetary disks, $m_{*,1} \equiv m_*/M_\odot$ is the stellar mass with fiducial normalisation, where 1 $M_\odot$, $f_r \equiv 1 - \sqrt{r_*/r_{\rm DZIB}}$, r_* is the stellar radius and $\dot{m}_{-9} \equiv \dot{m}_*/10^{-9}\ M_\odot$ 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 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$_\oplus$[2], 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 $\dot{m}_{-9} = 1.80 \phi_{\rm DZIB,0.5}^{-9/4} \alpha_{-3}^{1/2}$, where f_r and other factors have been assumed to be unity. We note that this accretion rate (i.e. $\sim 10^{-9}\ M_\odot$ yr^{-1}) is typical of those observed in transition disk systems (Manara et al. 2014), namely, where dust 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; namely, pebble radius $a_p > 0.15$ cm (Hu & Tan 2026).

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

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

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 $M_{p,1} = 8.00 \phi_{\rm DZIB,0.5}^{-9/8} \alpha_{-3}^{1/2}\ M_\oplus$. Comparing to the observed mass of K2-18 c, i.e. $M_{p,1,\rm obs} \simeq 6.99^{0.96}_{0.99}$ M$_\oplus$ (note that we assume sin $i \simeq 1$; Radica et al. 2022b), 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 a gap-opening mass,

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

In other words, 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 considered 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 mid-plane 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 mid-plane 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 and/or 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 $\sim$ 1 to 10%. These conditions for the solid cores and initial primordial atmospheres have been inferred for the general super-Earth and 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 ~ 6 g cm^{-3}, then its radius would be $1.94\ R_\oplus$.

[2] The value obtained by Cloutier et al. (2019) is 5.62 ± 0.84 M$_\oplus$

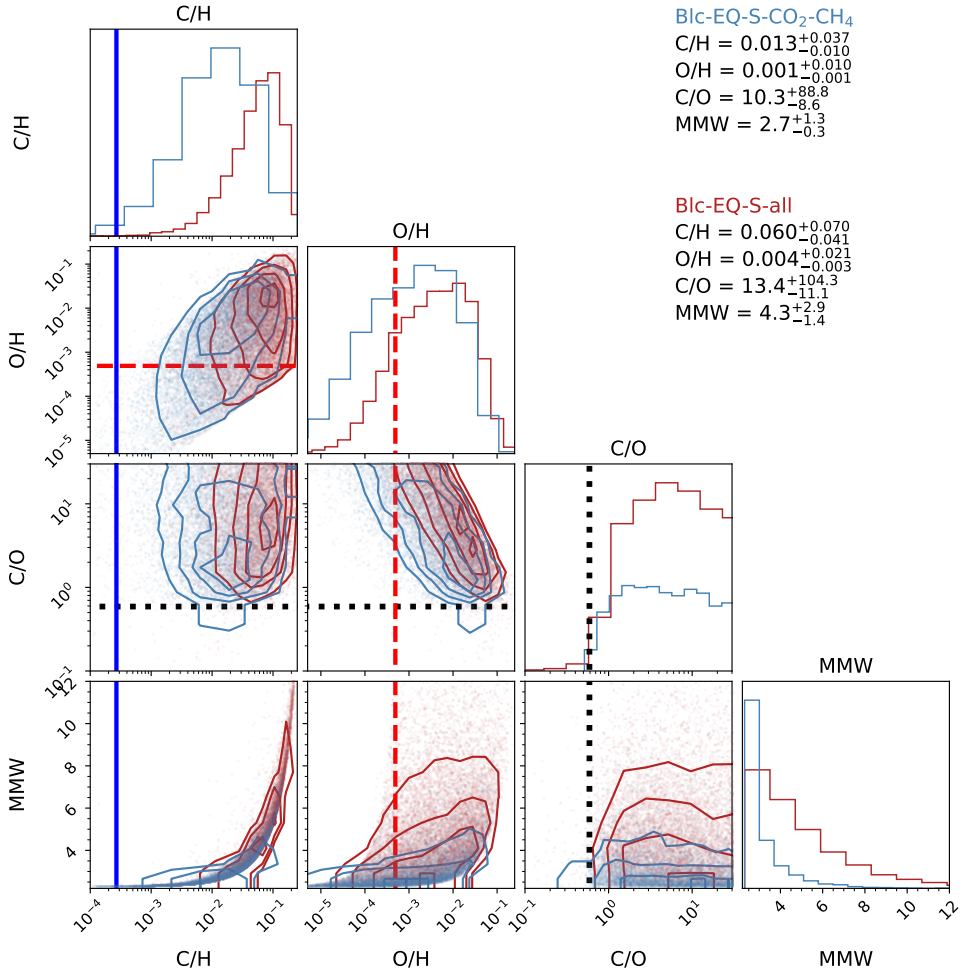

Fig. 7. Derived posterior distributions of carbon and oxygen abundances with respect to hydrogen, elemental C/O ratio and MMW for the fiducial reduction. Red shows the comprehensive retrieval and blue the simplified retrieval. The plots are zoomed in to highlight regions of interest such as C/O ~ 0.1 – 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.

Given the observed radius of K2-18 b of 2.61 ± 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 characterised by a low MMW atmosphere.

Detailed predictions for the atmospheric composition of the innermost planets forming via IOPF are not yet available. The 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 $r_{170K} = 0.587$ 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 $r_{500K} = 0.177$ 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 $[C_{soot}/H] \sim 1.6 \times 10^{-4}$, i.e. about 10× 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 Fig. 7 we highlight several metrics derived from our atmospheric retrievals of K2-18 b: C/H; O/H; C/O; and MMW. The results favour an atmosphere predominantly made of H_2 and He with significant metal enrichment. 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) recently presented the results of atmospheric retrievals of K2-18 b based on JWST NIRISS, NIRSpec and MIRI LRS data, while 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 made use of publicly available pipelines in combination with custom-made pipelines producing a total of 12 transmission spectra and exploring the effects of different treatments of limb darkening, spectral binning, error inflation, instrumental offsets, and spot correction. Notably, we have introduced a new semi-empirical method for correcting stellar spot crossing events. We also performed retrievals on the transmission spectrum based on Madhusudhan et al. (2023) and our binned version of it.

We recommend extracting a robust transmission spectrum by using binned light curves (R∼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 minimise 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 with respect to enhancing 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 this wavelength, depending on the atmospheric composition and S/N of the data; therefore, the optimal resolution or choice of bins might be case-dependent.

We also propose a suite of retrieval configurations to test atmospheric model dependencies, including different clouds, hazes, and temperature–pressure parametrisations, along with comprehensive and minimal sets of chemical species and criteria for selecting a simplified retrieval setup. We detected CH_4 with a significance of above 3–4σ in all configurations. The detection of CO_2 is configuration dependent and, in most cases, around 2σ. The absolute abundances of both species are broadly constrained and slightly dependent on the configuration, encompassing both scenarios with a primordial H_2-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 IOPF model, if K2-18 b did indeed form 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 the atmospheric chemical composition. Clearly, larger samples of planets with measured atmospheric compositions are needed for further study, ideally in multi-planet systems that span a range of disk conditions that straddle the soot and water ice lines.

Data availability

Due to the large volume of reductions and retrieval configurations presented in this work it is unfeasible to upload all files to an online platform. A selection of the most representative setup posterior distribution files can be found at the following repository. The remaining files are available on demand.

Acknowledgements. We thank the anonymous referee, whose comments and suggestions improved the quality of this manuscript. We thank Arturo Cevallos Soto, Kevin Heng, Greg Houlihan, Xiao Hu, Antonino Petralia and Shang-Min Tsai for helpful discussions. We acknowledge financial support from the Agencia Estatal de Investigación of the Ministerio de Ciencia e Innovación MCIN/AEI/10.13039/501100011033 and the ERDF "A way of making Europe" through projects PID2021-125627OB-C32 and PID2024-158486OB-C32. This work is supported by the European Union (ERC AdvG SPEAR, GA 101200674). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. 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. 2012, ApJ, 753, 100
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
Chubb, K. L., Robert, S., Sousa-Silva, C., et al. 2024, RAS Techniques and Instruments, 3, 636
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., Tsiaras, 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
Hargreaves, R. J., Gordon, I. E., Rey, M., et al. 2020, ApJS, 247, 55
Heng, K., Owen, J. E., & Tian, M. 2025, ApJ, 994, 28
Hu, R., Bello-Arufe, A., Tokadjian, A., et al. 2025, arXiv e-prints, arXiv:2507.12622. Submitted to AAS Journals
Hu, R., Damiano, M., Scheucher, M., et al. 2021, ApJ, 921, L8
Hu, X. & Tan, J. C. 2026, ApJ, 999, 220
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
Jaziri, A. Y., Pluriel, W., Bocchieri, A., et al. 2024, A&A, 684, A25
Karlovets, E. V., Gordon, I. E., Hashemi, R., et al. 2021, J. Quant. Spectr. Rad. Transf., 258, 107275
Kipping, D. & Benneke, B. 2025, arXiv e-prints, arXiv:2506.05392
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. 2026, ApJ, 997, L29
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. Submitted to AAS Journals
Lucy, L. B. & Sweeney, M. A. 1971, AJ, 76, 544
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
MacDonald, R. J., Goyal, J. M., & Lewis, N. K. 2020, ApJ, 893, L43
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
Pluriel, W., Leconte, J., Parmentier, V., et al. 2022, A&A, 658, A42
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., Lafreniére, 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
Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P.-O., & Tinetti, G. 2016, ApJ, 833, 120
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, AJ, 170, 298
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, AJ, 170, 257
Taylor, J. 2025, Research Notes of the American Astronomical Society, 9, 118
Taylor, J., Radica, M., Welbanks, L., et al. 2023, MNRAS, 524, 817
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
Wakeford, H. R., Sing, D. K., Kataria, T., et al. 2017, Science, 356, 628
Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107
Welbanks, L., Nixon, M. C., McGill, P., et al. 2026, Nature Astronomy, 10, 234
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 Sects. 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 ∼+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 2.5–4 μm 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 Sect. 6.2).

We also compare our transmission spectra with the first reduction published by Madhusudhan et al. (2023), re-binned 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.

Appendix B: Retrievals on different reduction configurations

Table B.1 displays the results from comprehensive and simplified retrievals on all standard reduction configurations presented in this work.

Appendix C: Retrievals with alternative T–P profiles, clouds, and haze treatments

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

Appendix D: Retrievals with instrumental offsets

Table D.1 summarises the retrieval results from different spectral extractions using the simplified chemical setups with or without a free offset parameter. Table D.2 reports retrievals results using fixed offsets to resemble the mean separation corresponding to the Blc-PhQ configuration (or Bts in the last rows). The mean separation is the difference between the average transit depth of the NIRISS and NIRSpec spectrum. Fixing these offsets allows to test whether the spot correction behaves as a mostly uniform offset or if it accounts for other effects.

Appendix E: Retrievals with alternative opacity sources

Table E.1 summarises the results from different retrievals on the fiducial spectrum using different opacity data for CH_4, as described in Sect. 6.4.

Appendix F: Retrieval tests with inflated error bars

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

Appendix G: Comparison between retrieval configurations

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

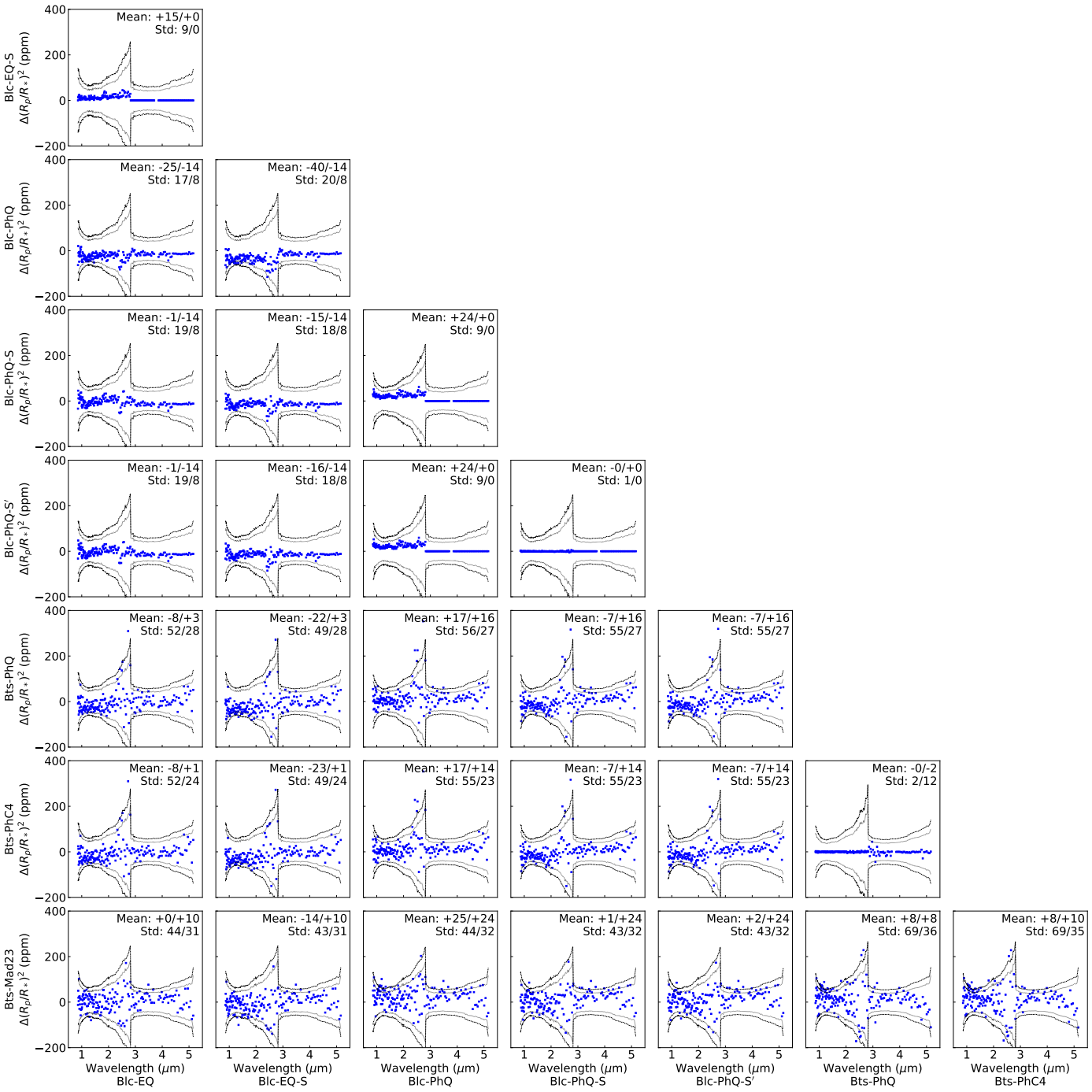

Fig. A.1: Corner plot of the difference between transmission spectra extracted with various methods (see Sect. 4.2).

Table B.1: 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)	$T_{10\,\mathrm{mbar}}$ (K)	MMW
Blc-EQ-S	CO_2+CH_4	1.17	1449.4	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
				1.0σ, +1.0, 2.6	3.6σ, +9.0, 7717		
	All	1.25	1448.3	$-3.63^{+1.75}_{-5.13}$	$-0.96^{+0.45}_{-0.57}$	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$
				N, +0.6, 1.8	3.2σ, +7.0, 1110		
Blc-PhQ-S	CH_4	1.31	1450.5	–	$-1.95^{+0.67}_{-0.75}$	121^{+44}_{-35}	$2.46^{+0.56}_{-0.13}$
					3.6σ, +8.7, 5892		
	All	1.39	1449.8	$-4.76^{+2.56}_{-5.30}$	$-0.95^{+0.45}_{-0.60}$	146^{+89}_{-48}	$4.26^{+3.05}_{-1.44}$
				N, +0.4, 1.5	3.1σ, +6.6, 755		
Blc-PhQ-S′	CH_4	1.30	1450.9	–	$-1.94^{+0.68}_{-0.76}$	123^{+43}_{-37}	$2.46^{+0.59}_{-0.13}$
					3.6σ, +8.7, 5833		
	All	1.39	1450.2	$-4.98^{+2.74}_{-5.24}$	$-0.95^{+0.46}_{-0.59}$	148^{+96}_{-49}	$4.25^{+3.20}_{-1.44}$
				N, +0.3, 1.4	3.0σ, +6.4, 629		
Blc-EQ	CO_2+CH_4	1.19	1446.6	$-3.29^{+1.73}_{-3.44}$	$-1.84^{+0.93}_{-0.79}$	155^{+67}_{-46}	$2.63^{+3.06}_{-0.28}$
				1.0σ, +1.0, 2.6	3.5σ, +8.5, 4943		
	All	1.27	1446.1	$-3.00^{+1.56}_{-5.58}$	$-0.80^{+0.41}_{-0.59}$	195^{+118}_{-74}	$5.63^{+3.74}_{-2.42}$
				N, +0.7, 2.0	2.8σ, +5.5, 244		
Blc-PhQ	CO_2+CH_4	1.36	1446.6	$-1.74^{+0.81}_{-4.56}$	$-1.00^{+0.64}_{-1.79}$	213^{+155}_{-79}	$6.73^{+5.15}_{-4.40}$
				1.3σ, +1.5, 4.4	4.0σ, +10.5, 37246		
	All	1.45	1447.2	$-2.17^{+1.07}_{-5.46}$	$-0.53^{+0.30}_{-0.51}$	283^{+175}_{-132}	$8.83^{+3.92}_{-4.27}$
				N, +0.6, 1.8	2.6σ, +4.9, 140		
Bts-PhQ	CO_2+CH_4+DMS[a]	1.50	1443.5	$-5.54^{+4.45}_{-6.07}$	$-1.57^{+0.67}_{-0.89}$	134^{+80}_{-37}	$2.77^{+5.31}_{-0.41}$
				2.9σ, +3.6, 35	4.9σ, +15.7, 6.4×10^6		
	All	1.58	1444.1	$-1.18^{+0.32}_{-1.60}$	$-0.64^{+0.33}_{-0.44}$	328^{+184}_{-142}	$10.14^{+3.21}_{-3.39}$
				1.3σ, +1.6, 4.8	2.8σ, +5.8, 321		
Bts-PhC4	CO_2+CH_4+DMS[b]	1.50	1443.6	$-6.79^{+3.96}_{-5.23}$	$-1.80^{+0.62}_{-0.78}$	125^{+42}_{-32}	$2.58^{+0.84}_{-0.23}$
				2.0σ, +3.1, 23	4.8σ, +15.1, 3.5×10^6		
	All	1.58	1444.0	$-1.21^{+0.34}_{-2.11}$	$-0.62^{+0.32}_{-0.45}$	321^{+187}_{-141}	$10.20^{+3.21}_{-3.59}$
				1.1σ, +1.3, 3.5	2.9σ, +5.8, 341		
Bts-Mad23	CO_2+CH_4	1.56	1423.5	$-1.17^{+0.24}_{-1.96}$	$-1.03^{+0.43}_{-1.36}$	206^{+87}_{-63}	$7.23^{+2.63}_{-4.81}$
				2.1σ, +3.5, 34	5.1σ, +16.8, 2.0×10^7		
	All	1.66	1423.7	$-1.40^{+0.35}_{-1.44}$	$-0.70^{+0.30}_{-0.42}$	248^{+144}_{-95}	$7.68^{+2.62}_{-2.69}$
				1.1σ, +1.2, 3.2	2.8σ, +5.7, 288		
Orig-Mad23	CO_2+CH_4+DMS[c]	1.08	28820.9	$-1.38^{+0.32}_{-0.87}$	$-1.02^{+0.41}_{-0.63}$	204^{+97}_{-73}	$5.94^{+2.78}_{-2.55}$
				2.6σ, +5.1, 164	5.2σ, +17.5, 3.8×10^7		
	All	1.09	28820.7	$-1.44^{+0.31}_{-0.57}$	$-0.84^{+0.34}_{-0.40}$	254^{+129}_{-94}	$6.48^{+2.35}_{-1.83}$
				1.9σ, +2.9, 18	3.2σ, +7.1, 1267		

Notes. [a] log (DMS) = $-4.60^{+0.95}_{-1.70}$ (1.1σ) ; [b] log (DMS) = $-4.46^{+0.78}_{-0.81}$ (1.3σ) ; [c] log (DMS) = $-6.29^{+2.06}_{-5.60}$ (1.3σ) . 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. MMW denotes the mean molecular weight. We report the detection significances, $\Delta \log ev$ and the Bayes factors (as in Table 3).

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

Setup	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	log (P_{cloud}/bar)
Gui-Cl-Mie	1.17	1448.3	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$-1.14^{+1.35}_{-0.59}$
Gui-Cl-0	1.14	1449.3	$-4.30^{+1.84}_{-4.90}$	$-1.86^{+0.80}_{-0.82}$	131^{+50}_{-41}	$-1.12^{+1.42}_{-0.60}$
Gui-0-Mie	1.16	1449.0	$-3.92^{+1.91}_{-4.13}$	$-1.89^{+0.79}_{-0.97}$	108^{+42}_{-32}	–
Gui-0-0	1.14	1448.4	$-3.56^{+1.84}_{-3.92}$	$-1.99^{+1.01}_{-1.07}$	103^{+52}_{-30}	–
Iso-Cl-Mie	1.15	1449.6	$-4.18^{+1.75}_{-4.97}$	$-1.84^{+0.68}_{-0.78}$	141^{+39}_{-31}	$-1.04^{+1.68}_{-0.69}$
Iso-Cl-0	1.12	1449.4	$-3.81^{+2.04}_{-5.53}$	$-1.48^{+0.83}_{-0.96}$	162^{+95}_{-47}	$-1.11^{+1.79}_{-0.82}$
Iso-0-Mie	1.14	1449.4	$-4.10^{+1.87}_{-4.10}$	$-2.05^{+0.89}_{-0.90}$	122^{+43}_{-25}	–
Iso-0-0	1.17	1448.9	$-2.84^{+1.46}_{-4.25}$	$-1.39^{+0.93}_{-1.33}$	142^{+140}_{-41}	–

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

Table D.1: Selected parameters from retrievals on different spectral extractions with or without a free offset parameter.

Spec./Mol.	Offset (ppm)	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	MMW
Blc-EQ-S	–	1.17	1449.4	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
(CO_2+CH_4)	+13 ± 8	1.17	1444.9	$-4.15^{+1.91}_{-5.18}$	$-1.31^{+0.47}_{-0.69}$	139^{+58}_{-44}	$3.11^{+1.43}_{-0.65}$
Blc-PhQ-S	–	1.31	1450.5	–	$-1.95^{+0.67}_{-0.75}$	121^{+44}_{-35}	$2.46^{+0.56}_{-0.13}$
(CH_4)	+16 ± 8	1.30	1446.2	–	$-1.46^{+0.52}_{-0.73}$	115^{+50}_{-35}	$2.78^{+1.08}_{-0.39}$
Blc-PhQ-S′	–	1.30	1450.9	–	$-1.94^{+0.68}_{-0.76}$	123^{+43}_{-37}	$2.46^{+0.59}_{-0.13}$
(CH_4)	+16 ± 8	1.29	1446.8	–	$-1.44^{+0.51}_{-0.72}$	117^{+51}_{-36}	$2.80^{+1.13}_{-0.40}$
Blc-EQ	–	1.19	1446.6	$-3.29^{+1.73}_{-3.44}$	$-1.84^{+0.93}_{-0.79}$	155^{+67}_{-46}	$2.63^{+3.06}_{-0.28}$
(CO_2+CH_4)	+26 ± 8	1.16	1444.4	$-4.24^{+1.94}_{-5.26}$	$-1.24^{+0.43}_{-0.62}$	143^{+59}_{-47}	$3.22^{+1.44}_{-0.70}$
Blc-PhQ	–	1.36	1446.6	$-1.74^{+0.81}_{-4.56}$	$-1.00^{+0.64}_{-1.79}$	213^{+155}_{-79}	$6.73^{+5.15}_{-4.40}$
(CO_2+CH_4)	+36 ± 8	1.30	1447.9	$-6.15^{+3.20}_{-5.30}$	$-1.30^{+0.45}_{-0.62}$	125^{+54}_{-38}	$3.05^{+1.30}_{-0.57}$
Bts-PhQ	–	1.50	1443.5	$-5.54^{+4.45}_{-6.07}$	$-1.57^{+0.67}_{-0.89}$	134^{+80}_{-37}	$2.77^{+5.31}_{-0.41}$
(CO_2+CH_4+DMS)	+40 ± 8	1.44	1444.2	$-5.12^{+2.52}_{-5.61}$	$-1.60^{+0.51}_{-0.65}$	109^{+42}_{-30}	$2.71^{+0.85}_{-0.32}$
Bts-PhC4	–	1.50	1443.6	$-6.79^{+3.96}_{-5.23}$	$-1.80^{+0.62}_{-0.78}$	125^{+42}_{-32}	$2.58^{+0.84}_{-0.23}$
(CO_2+CH_4+DMS)	+40 ± 8	1.45	1444.1	$-5.37^{+2.75}_{-5.41}$	$-1.58^{+0.51}_{-0.66}$	111^{+44}_{-31}	$2.73^{+0.93}_{-0.34}$
Bts-Mad23	–	1.56	1423.5	$-1.17^{+0.24}_{-1.96}$	$-1.03^{+0.43}_{-1.36}$	206^{+87}_{-63}	$7.23^{+2.63}_{-4.81}$
(CO_2+CH_4)	+41 ± 8	1.49	1425.8	$-3.65^{+1.30}_{-3.10}$	$-1.28^{+0.37}_{-0.53}$	148^{+46}_{-40}	$3.13^{+1.02}_{-0.58}$

Table D.2: Selected parameters from retrievals on additional spectral extractions with a fixed offset.

Spectrum	Molecules	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	MMW
Blc-PhQ-S	CH_4	1.31	1450.6	–	$-1.95^{+0.67}_{-0.75}$	121^{+44}_{-35}	$2.46^{+0.56}_{-0.13}$
	All	1.39	1449.9	$-4.76^{+2.56}_{-5.30}$	$-0.95^{+0.45}_{-0.60}$	146^{+89}_{-48}	$4.26^{+3.05}_{-1.44}$
Blc-PhQ-S -24 / 0 [a]	CO_2+CH_4	1.36	1444.7	$-1.11^{+0.33}_{-3.89}$	$-0.91^{+0.55}_{-1.82}$	243^{+137}_{-98}	$9.19^{+4.06}_{-6.85}$
	All	1.45	1445.7	$-1.88^{+0.85}_{-4.95}$	$-0.47^{+0.27}_{-0.44}$	359^{+180}_{-162}	$9.99^{+3.69}_{-4.09}$
Blc-PhQ	CO_2+CH_4	1.36	1446.6	$-1.74^{+0.81}_{-4.56}$	$-1.00^{+0.64}_{-1.79}$	213^{+155}_{-79}	$6.73^{+5.15}_{-4.40}$
	All	1.45	1447.1	$-2.17^{+1.07}_{-5.46}$	$-0.53^{+0.30}_{-0.51}$	283^{+175}_{-132}	$8.83^{+3.92}_{-4.27}$
Blc-EQ-S	CO_2+CH_4	1.17	1449.5	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
	All	1.25	1448.4	$-3.63^{+1.75}_{-5.13}$	$-0.96^{+0.45}_{-0.57}$	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$
Blc-EQ-S -26 / 0 [a]	CO_2+CH_4	1.23	1443.6	$-1.36^{+0.47}_{-2.94}$	$-1.17^{+0.68}_{-1.65}$	211^{+120}_{-74}	$6.76^{+4.49}_{-4.42}$
	All	1.31	1444.5	$-1.71^{+0.66}_{-4.24}$	$-0.51^{+0.29}_{-0.45}$	331^{+172}_{-141}	$9.38^{+3.64}_{-3.62}$
Bts-PhQ	CO_2+CH_4+DMS	1.50	1443.5	$-5.54^{+4.45}_{-6.07}$	$-1.57^{+0.67}_{-0.89}$	134^{+80}_{-37}	$2.77^{+5.31}_{-0.41}$
	All	1.58	1444.1	$-1.18^{+0.32}_{-1.60}$	$-0.64^{+0.33}_{-0.44}$	328^{+184}_{-142}	$10.14^{+3.21}_{-3.39}$
Bts-PhQ +26 / 0 [a]	CO_2+CH_4	1.44	1449.2	$-4.20^{+1.81}_{-5.12}$	$-2.05^{+0.62}_{-0.65}$	132^{+41}_{-34}	$2.48^{+0.55}_{-0.14}$
	All	1.53	1448.2	$-4.26^{+2.26}_{-5.51}$	$-1.19^{+0.49}_{-0.55}$	143^{+93}_{-46}	$3.34^{+3.17}_{-0.89}$

Notes. [(a)]: The first number indicates the offset (in ppm) added to the corresponding NIRISS spectrum and the second number the NIRSpec offset. The offsets are chosen such that the Blc-PhQ-S (-24/0) and Blc-EQ-S (-26/0) have mean separation between NIRISS and NIRSpec identical to that of Blc-PhQ. For Bts-PhQ (+26/0) the mean separation is intermediate between those of Blc-PhQ-S and Blc-EQ-S.

Table E.1: Selected parameters from retrievals adopting different CH_4 opacity data.

Spec./Mol.	Offset (ppm)	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	MMW
Blc-EQ-S	CH_4: ExoMol YT34to10, R = 50 000						
CO_2+CH_4	–	1.17	1449.4	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
”	+13 ± 8	1.17	1444.9	$-4.15^{+1.91}_{-5.18}$	$-1.31^{+0.47}_{-0.69}$	139^{+58}_{-44}	$3.11^{+1.43}_{-0.65}$
All	–	1.25	1448.3	$-3.63^{+1.75}_{-5.13}$	$-0.96^{+0.45}_{-0.57}$	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$
”	+10 ± 9	1.25	1442.9	$-4.18^{+2.06}_{-5.15}$	$-0.95^{+0.40}_{-0.50}$	156^{+78}_{-51}	$4.16^{+2.51}_{-1.24}$
Blc-EQ-S	CH_4: ExoMol MM, R = 15 000						
CO_2+CH_4	–	1.15	1450.5	$-3.94^{+1.59}_{-3.68}$	$-1.84^{+0.59}_{-0.67}$	122^{+44}_{-33}	$2.58^{+0.74}_{-0.22}$
”	+14 ± 8	1.15	1445.8	$-4.41^{+2.01}_{-5.14}$	$-1.49^{+0.49}_{-0.63}$	124^{+51}_{-36}	$2.85^{+1.01}_{-0.42}$
All	–	1.23	1449.4	$-4.18^{+1.86}_{-5.04}$	$-1.32^{+0.50}_{-0.54}$	130^{+57}_{-35}	$3.16^{+1.60}_{-0.61}$
”	+10 ± 9	1.23	1444.0	$-4.58^{+2.13}_{-5.06}$	$-1.22^{+0.42}_{-0.51}$	132^{+58}_{-37}	$3.30^{+1.50}_{-0.68}$
Blc-EQ-S	CH_4: HITEMP, R = 50 000						
CO_2+CH_4	–	1.16	1449.7	$-4.48^{+1.87}_{-4.69}$	$-2.00^{+0.64}_{-0.74}$	115^{+40}_{-34}	$2.49^{+0.58}_{-0.15}$
”	+13 ± 8	1.16	1445.0	$-4.89^{+2.25}_{-5.24}$	$-1.58^{+0.53}_{-0.71}$	114^{+43}_{-34}	$2.74^{+0.93}_{-0.35}$
All	–	1.24	1448.6	$-4.24^{+1.97}_{-5.04}$	$-1.25^{+0.57}_{-0.58}$	134^{+62}_{-39}	$3.33^{+2.41}_{-0.75}$
”	+10 ± 8	1.24	1443.0	$-4.89^{+2.40}_{-5.40}$	$-1.15^{+0.46}_{-0.54}$	129^{+57}_{-38}	$3.46^{+1.95}_{-0.81}$

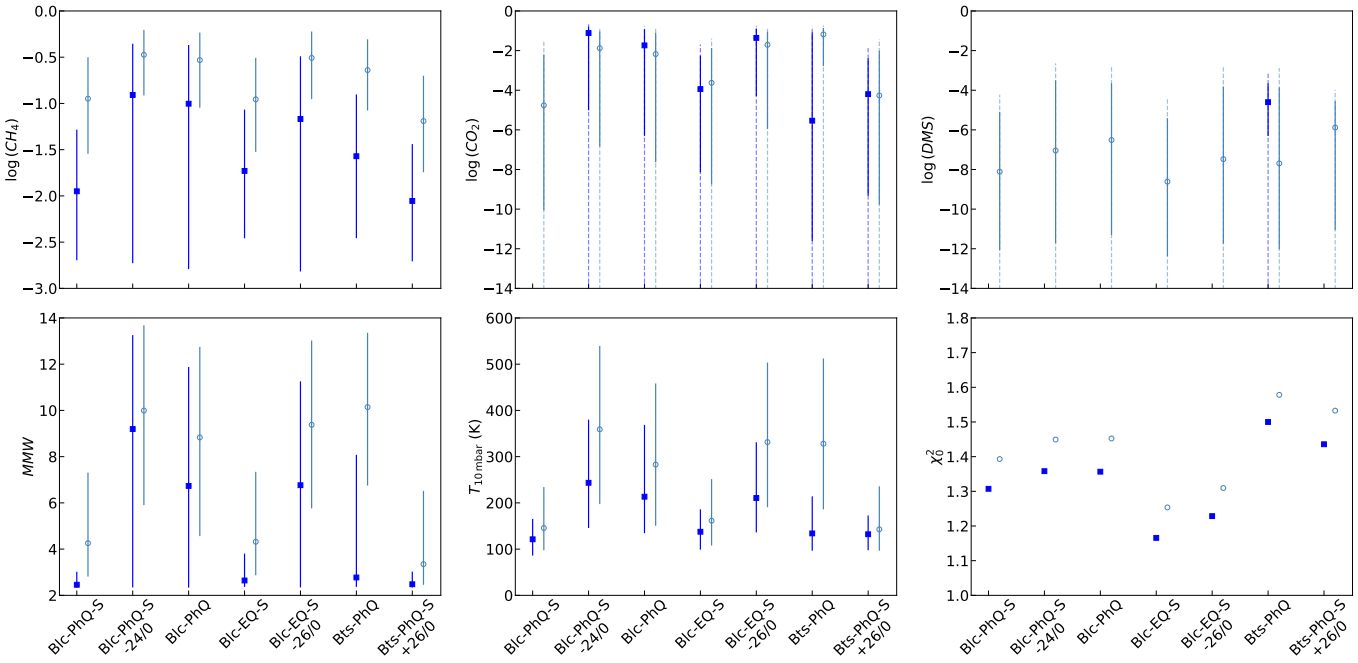

Fig. F.1: Analogous to Fig. 4 but comparing spectra with and without offsets (see also Table D.2).

Table F.1: Selected parameters from retrievals on additional spectral extractions, analogous to Table B.1.

Spectrum	Molecules	χ^2_0	log ev	log (CO_2)	log (CH_4)	$T_{10\,\mathrm{mbar}}$ (K)	MMW
Blc-EQ-S	CO_2+CH_4	1.17	1449.5	$-3.94^{+1.68}_{-4.20}$	$-1.73^{+0.66}_{-0.73}$	138^{+48}_{-39}	$2.64^{+1.16}_{-0.28}$
	All	1.25	1448.4	$-3.63^{+1.75}_{-5.13}$	$-0.96^{+0.45}_{-0.57}$	161^{+90}_{-54}	$4.32^{+3.02}_{-1.45}$
Blc-EQ-S Err × 1.08	CO_2+CH_4	1.00	1450.2	$-4.39^{+1.97}_{-4.82}$	$-1.77^{+0.66}_{-0.81}$	131^{+49}_{-41}	$2.61^{+1.03}_{-0.26}$
	All	1.08	1448.9	$-4.47^{+2.42}_{-5.48}$	$-0.89^{+0.43}_{-0.57}$	163^{+96}_{-57}	$4.59^{+3.26}_{-1.64}$
Bts-PhQ	CO_2+CH_4+DMS[a]	1.50	1443.5	$-5.54^{+4.45}_{-6.07}$	$-1.57^{+0.67}_{-0.89}$	134^{+80}_{-37}	$2.77^{+5.31}_{-0.41}$
	All	1.58	1444.1	$-1.18^{+0.32}_{-1.60}$	$-0.64^{+0.33}_{-0.44}$	328^{+184}_{-142}	$10.14^{+3.21}_{-3.39}$
Bts-PhQ Err × 1.225	CO_2+CH_4+DMS[b]	1.00	1449.6	$-5.80^{+4.28}_{-5.63}$	$-1.83^{+0.79}_{-1.08}$	129^{+71}_{-38}	$2.58^{+2.91}_{-0.25}$
	All	1.06	1449.8	$-1.62^{+0.70}_{-5.76}$	$-0.64^{+0.36}_{-0.58}$	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) = $-4.60^{+0.95}_{-1.70}$; [(b)] log (DMS) = $-4.86^{+1.05}_{-2.91}$.

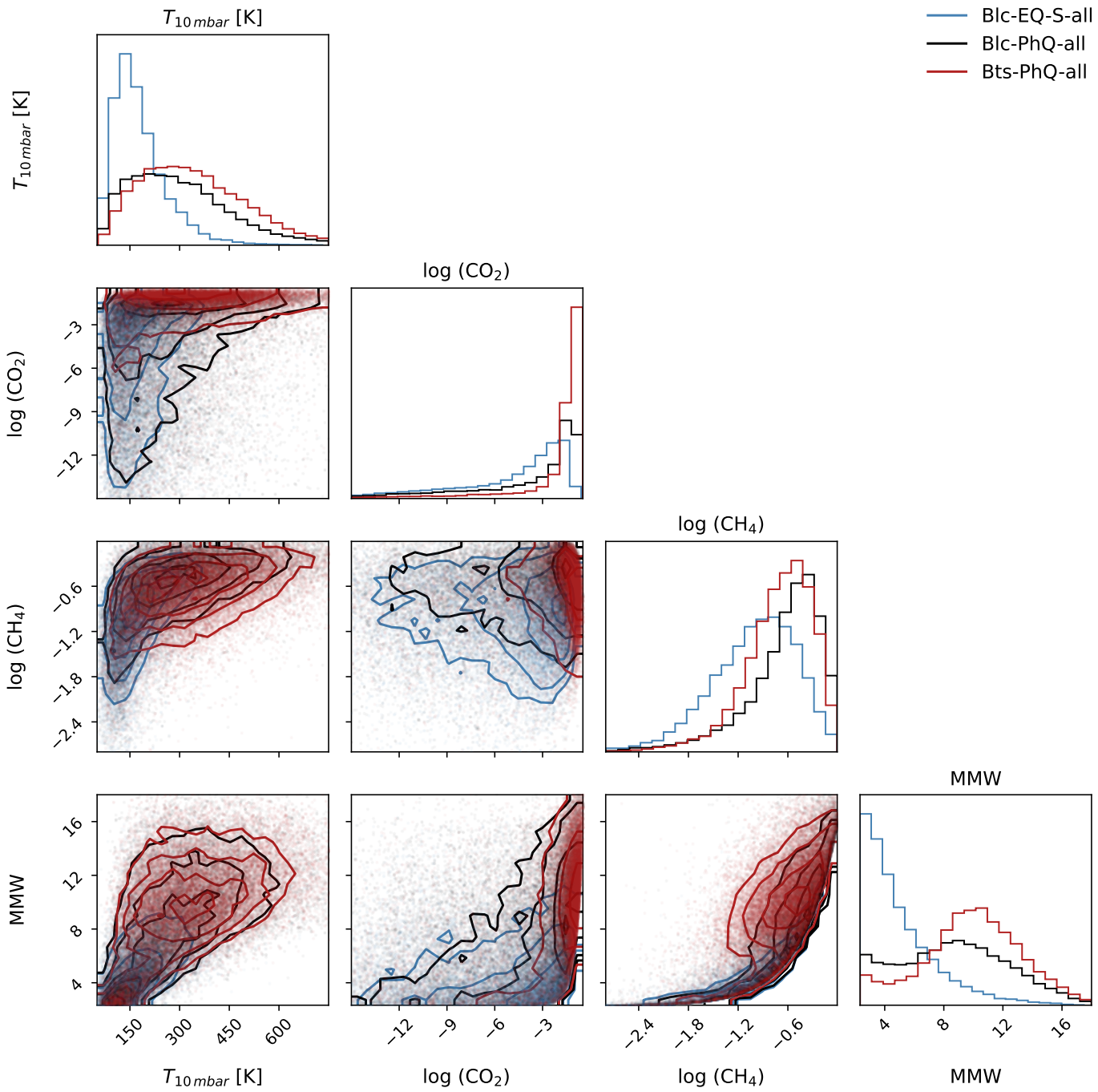

Fig. G.1: Posterior distributions of some selected parameters representative of the major reduction families with comprehensive retrievals. Blue shows our fiducial reduction corresponding to binned before fitting light curves - empirical limb darkening - with spot correction, black shows the binned before 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).