Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **524,** 835–856 (2023)
Advance Access publication 2023 June 14

https://doi.org/10.1093/mnras/stad1762

# Awesome SOSS: transmission spectroscopy of WASP-96b with NIRISS/SOSS

Michael Radica [1]★ Luis Welbanks [2] Néstor Espinoza [3,4] Jake Taylor [1,5] Louis-Philippe Coulombe [1] Adina D. Feinstein [6] Jayesh Goyal [7] Nicholas Scarsdale [8] Loïc Albert [1] Priyanka Baghel [7] Jacob L. Bean [6] Jasmina Blecic [9,10] David Lafrenière [1] Ryan J. MacDonald [11] Maria Zamyatina [12] Romain Allart1 [1] Étienne Artigau [1,13] Natasha E. Batalha [14] Neil James Cook [1] Nicolas B. Cowan [15,16] Lisa Dang [1] René Doyon [1,13] Marylou Fournier-Tondreau [1] Doug Johnstone [17,18] Michael R. Line [2] Sarah E. Moran [19] Sagnick Mukherjee [8] Stefan Pelletier [1] Pierre-Alexis Roy [1] Geert Jan Talens [20] Joseph Filippazzo [3] Klaus Pontoppidan [3] and Kevin Volk[3]

*Affiliations are listed at the end of the paper*

Accepted 2023 June 9. Received 2023 May 25; in original form 2022 December 22

## ABSTRACT

The future is now – after its long-awaited launch in 2021 December, *JWST* began science operations in 2022 July and is already revolutionizing exoplanet astronomy. The Early Release Observations (ERO) programme was designed to provide the first images and spectra from *JWST*, covering a multitude of science cases and using multiple modes of each on-board instrument. Here, we present transmission spectroscopy observations of the hot-Saturn WASP-96 b with the Single Object Slitless Spectroscopy (SOSS) mode of the Near Infrared Imager and Slitless Spectrograph, observed as part of the ERO programme. As the SOSS mode presents some unique data reduction challenges, we provide an in-depth walk-through of the major steps necessary for the reduction of SOSS data: including background subtraction, correction of 1/*f* noise, and treatment of the trace order overlap. We furthermore offer potential routes to correct for field star contamination, which can occur due to the SOSS mode's slitless nature. By comparing our extracted transmission spectrum with grids of atmosphere models, we find an atmosphere metallicity between 1× and 5× solar, and a solar carbon-to-oxygen ratio. Moreover, our models indicate that no grey cloud deck is required to fit WASP-96 b's transmission spectrum, but find evidence for a slope shortward of 0.9 $\mu$m, which could either be caused by enhanced Rayleigh scattering or the red wing of a pressure-broadened Na feature. Our work demonstrates the unique capabilities of the SOSS mode for exoplanet transmission spectroscopy and presents a step-by-step reduction guide for this new and exciting instrument.

**Key words:** methods: data analysis – techniques: spectroscopic – planets and satellites: atmospheres – planets and satellites: gaseous planets – planets and satellites: individual: WASP-96 b.

## 1 INTRODUCTION

Transiting exoplanets provide astronomers with an ideal opportunity to study the atmospheres of distant worlds (Seager & Sasselov 2000). Spectroscopic observations during the transit or eclipse of an exoplanet have revealed the telltale signs of an abundance of molecular and atomic species in the atmospheres of giant exoplanets, both at high (e.g. Brogi et al. 2014; Hoeijmakers et al. 2020; Boucher et al. 2021, 2023) and low (e.g. Charbonneau et al. 2002; Kreidberg et al. 2015; Evans et al. 2016) spectral resolution, which have provided insights into the formation histories (Öberg, Murray-Clay & Bergin 2011; Madhusudhan, Amin & Kennedy 2014; Turrini et al. 2021) as well as the physical and chemical processes governing their atmospheres (Moses et al. 2011; Madhusudhan 2012; Parmentier, Showman & Lian 2013; Wakeford et al. 2017). Particularly, space-based observations using the Hubble Space Telescope (*HST*) and *Spitzer Space Telescope* have shed light on the population of giant exoplanet atmospheres, revealing near-ubiquitous detections of water as well as the presence of clouds and hazes (e.g. Sing et al. 2016; Barstow et al. 2017; Pinhas et al. 2019; Welbanks et al. 2019). The spectral signatures of alkali metals such as Na and K are also common, particularly in cloud-free atmospheres (e.g. Welbanks et al. 2019; Alam et al. 2021; Nikolov et al. 2022), as well as hints of carbon-bearing species such as CO and $CO_2$ (e.g. Kreidberg et al. 2018; Dragomir et al. 2020; Spake et al. 2020).

However, since neither observatory has observing modes specifically designed with exoplanet observations in mind, atmospheric studies with *HST* and *Spitzer* were far from ideal. Both presented astronomers with a number of technical challenges (e.g. Deming

★ E-mail: michael.radica@umontreal.ca

et al. 2013; Zhou et al. 2017), and moreover, atmospheric inferences from *HST* and *Spitzer* observations, even when used in conjunction, were often limited. The narrow bandwidth of the Wide Field Camera 3 (WFC3; 0.85–1.7 μm) and Space Telescope Imaging Spectrograph (STIS; 0.525–1.0 $\mu$m) instruments on board *HST* result in atmosphere spectra sensitive primarily to $H_2O$, alkalis, and clouds (Benneke & Seager 2012, 2013). *Spitzer* provided up to four photometric points at longer wavelengths (3.6, 4.5, 5.8, and 8 $\mu$m, although only the two bluest wavebands remained available after 2009), which could provide hints of the presence of carbon-bearing molecules. However, *Spitzer* photometry precluded any definitive detections, rendering constraints on bulk atmospheric metallicities or carbon-to-oxygen ratios (C/O) challenging (e.g. Kreidberg et al. 2015; Spake et al. 2020).

*JWST* is the natural successor to *HST* and *Spitzer*. Its four instruments each have observing modes specifically tailored to time series observations (TSOs) of transiting exoplanets, which greatly improve the wavelength range (0.6–14 μm) and spectral resolution (R∼50–3000) with which exoplanet atmospheres can be spectroscopically probed. Indeed, *JWST* observations of the giant planet WASP-39 b (Faedi et al. 2011) using the Near Infrared Spectrograph (NIRSpec) PRISM mode (Birkmann et al. 2022) have already yielded the first definitive detections of $CO_2$ (JWST Transiting Exoplanet Community Early Release Science Team 2023), and $SO_2$ (Alderson et al. 2023; Tsai et al. 2023) in an exoplanet atmosphere.

The Single Object Slitless Spectroscopy (SOSS) mode (Albert et al. 2023) of the Near Infrared Imager and Slitless Spectrograph (NIRISS) instrument (Doyon et al. 2023) on board *JWST* is already proving to be one of the workhorse modes for exoplanet atmosphere observations (Fu et al. 2022; Feinstein et al. 2023). NIRISS/SOSS provides medium resolution (R∼700) spectroscopy from 0.6–2.85 $\mu$m, yielding unprecedented coverage of the near-infrared (NIR) $H_2O$ bands, as well as coverage of signatures of Na and K at the bluest wavelengths, and potentially $CH_4$, CO, and $CO_2$ at the reddest wavelengths. The full SOSS wavelength range is divided between two spectral orders, with the wavelengths 0.85–2.85 $\mu$m and 0.6–1.0 $\mu$m covered by orders 1 and 2, respectively.

WASP-96 b (Hellier et al. 2014) is a highly inflated hot-Saturn with a mass of $0.48 \pm 0.03\,M_J$ and a radius of $1.20 \pm 0.06\,R_J$. It orbits its G8 host star in 3.42 d, which results in an equilibrium temperature of 1285 ± 40 K, making WASP-96 b an excellent candidate for transmission spectroscopy. Indeed WASP-96 b has already been observed in transmission by Nikolov et al. (2018) with the FOcal Reducer/low dispersion Spectrograph 2 (FORS2; 0.35–0.8 $\mu$m) on the *Very Large Telescope* (*VLT*), revealing the clear pressure-broadened profile of Na and suggesting that WASP-96 b hosts a mostly cloud-free atmosphere. These findings were later confirmed by the retrieval analyses of Welbanks et al. (2019) and Alam et al. (2021), the former of which reported a stellar-to-super-stellar Na abundance.

Yip et al. (2021) and Nikolov et al. (2022) also presented independent analyses of transmission observations of WASP-96 b with *HST*/WFC3 using both the G102 (0.8–1.15 $\mu$m) and G141 (1.08–1.7 $\mu$m) grisms, as well as with *Spitzer*/IRAC at 3.6 and 4.5-$\mu$m. Joint retrievals with the *VLT* observations point to solar-to-super-solar abundances of Na and oxygen. The oxygen abundance in particular was found to be consistent with the oxygen enrichment level of Jupiter in our own solar system – suggesting a bulk atmospheric metallicity for WASP-96 b consistent with the solar system mass-metallicity trend (e.g. Thorngren et al. 2016; Welbanks et al. 2019). Nikolov et al. (2022) also find a sub-solar C/O ratio, under the assumption of chemical equilibrium. Lastly, McGruder et al. (2022) presented Magellan/IMACS (0.475–0.825 $\mu$m) transmission spectra of WASP-96 b, independently confirming the presence of the broad Na feature. Their joint retrievals of all available transmission data also point to solar-to-super-solar atmospheric abundances of Na and $H_2O$.

Recently, Samra et al. (2023) explored the combined transmission spectrum of WASP-96 b, as collated by Nikolov et al. (2022), in the context of cloud formation. Although all previous analyses have concluded a clear atmosphere for WASP-96 b based on the clearly observable pressure-broadened wings of the Na feature, Samra et al. (2023) point out that the equilibrium temperature of the planet places it in a regime where asymmetric cloud cover of the terminator could be expected (Helling et al. 2023). Moreover, the general circulation models (GCMs) of Samra et al. (2023) point to homogeneous clouds, dominated by silicates, covering the terminator region of WASP-96 b. They perform two retrieval analyses on the Nikolov et al. (2022) transmission spectrum, with and without the inclusion of clouds, and find that cloudy solutions can accurately reproduce the observed transmission spectrum. Samra et al. (2023), furthermore, suggest avenues which could reconcile their GCM models with the previous cloud-free retrieval results. In particular, they indicate that reduced vertical mixing efficiency could cause an optically thick cloud layer to settle below the observable photosphere, or that increasing the porosity of cloud particles could lead to optically thin clouds, even if they remain within the observable atmosphere.

Here, we present transmission spectroscopy observations of WASP-96 b with NIRISS/SOSS taken as part of the JWST Early Release Observations (ERO) programme (Pontoppidan et al. 2022). As the SOSS mode presents several unique challenges, this paper undertakes a step-by-step overview of the data reduction procedures, such that the community can understand the critical steps necessary to extract atmosphere spectra from SOSS TSOs. The companion paper, Taylor et al. (2023), presents an in-depth exploration of the modelling and retrieval of this transmission spectrum, as well as the particular insights into atmospheric physics and chemistry which can be gained through NIRISS/SOSS observations. This work is organized as follows: Section 2 presents the observations, as well as the data reduction procedure. Section 3 outlines the light-curve fitting, and atmospheric grid modelling methods are explained in Section 4. We present our initial atmospheric inferences in Section 5, and summarize and discuss our results in Section 6.

## 2 OBSERVATIONS AND DATA REDUCTIONS

### 2.1 Outline of the observations

WASP-96 b was observed in transit using the SOSS mode of the NIRISS instrument as part of the JWST ERO programme (Pontoppidan et al. 2022). The TSO started on UTC 2022 June 21 and spanned 6.4 h, which covered the 2.4-h transit, as well as 2.5 h of baseline before the transit, and 1.5 h after. It used the standard GR700XD/CLEAR combination, along with the SUBSTRIP256 subarray which captures three diffraction orders of the target on the detector (Albert et al. 2023). In total, the TSO is composed of 280 individual integrations, each consisting of 14 groups – yielding an integration time of 76.9 s per integration, and a median signal-to-noise ratio of 125 per integration for order 1 at 1.5 $\mu$m. An optional second exposure, using the GR700XD grism in combination with the F277W filter was taken after the GR700XD/CLEAR exposure. The F277W filter limits the wavelength range of SOSS to $\lambda \gtrsim 2.6$ $\mu$m, and the exposure lasted only 0.25 h, using 11 integrations with the same exposure time per integration as the CLEAR.

### 2.2 Data reduction

The SOSS mode presents a number of particular challenges: the curved nature of the spectral trace, the unique background shape, the potential contamination by field sources, and the overlap between the first and second diffraction orders on the detector to name a few. Here, we outline the major challenges encountered during reduction of SOSS data, and present the `supreme-SPOON` pipeline (supreme-Steps to Process sOss ObservatioNs) for the reduction of SOSS data. `supreme-SPOON` is publicly available,[1] and has already been successfully applied to the SOSS TSOs of WASP-39 b (Feinstein et al. 2023) and WASP-18 b (Coulombe et al. 2023) taken as part of the JWST Transiting Exoplanet Community Early Release Science Programme (Stevenson et al. 2016; Bean et al. 2018). `supreme-SPOON` was briefly outlined in Feinstein et al. (2023), but here we provide a more in-depth discussion of the key steps, as well as introduce new methods to explicitly deal with field star contamination, which was not included in the reduction of the aforementioned TSOs. For additional verification of our results, three other pipelines (`NAMELESS`, `transitspectroscopy`, and `nirHiss`) were used to perform independent reductions. The particulars of these pipelines have also already been outlined in Feinstein et al. (2023), and a brief summary of each, as it pertains to the analysis of this TSO, is provided in Appendix A.

`supreme-SPOON` is composed of four stages, many steps of which are shared with the official `jwst` data reduction pipeline.[2] A summary of the major reduction steps are outlined below, and visualized in Fig. 1.

#### 2.2.1 *`supreme-SPOON` Stage 1 – detector level calibrations*

Like the official `jwst` pipeline, Stage 1 performs the 'detector-level' calibrations on the full four-dimensional (integrations, groups, spatial pixels, spectral pixels) TSO data cube. The Stage 1 calibrations of `supreme-SPOON` are separated into 10 main steps, a short description of which is provided below:

(i) `GroupScaleStep`: Rescale pixel values to account for on-board averaging of frames.
(ii) `DQInitStep`: Initialization of data quality flags.
(iii) `SaturationStep`: Flag pixels above the saturation limit.
(iv) `SuperBiasStep`: Subtraction of the detector bias level.
(v) `RefPixStep`: Perform initial 1/*f* noise and odd-even row corrections using reference pixels.
(vi) `BackgroundStep I`: Subtract the background level.
(vii) `OneOverFStep`: Perform further corrections for 1/*f* noise. Re-add background.
(viii) `LinearityStep`: Correct non-linearity effects.
(ix) `JumpStep`: Flag cosmic ray hits.
(x) `RampFitStep`: Calculate the mean count rates per pixel by fitting each pixel or each integration 'up-the-ramp' (that is, along the 'groups' axis).

The initial stages (i–v) are already well handled by the official `jwst` pipeline, and `supreme-SPOON` simply provides wrappers for these steps. For more in-depth information about each of these specific steps, please see the `jwst` documentation.[3] Fig. 1(a) shows a raw, uncalibrated data frame, and Fig. 1(b) shows the same data frame after the completion of these first five steps.

After the `RefPixStep`, `supreme-SPOON` begins to diverge from `jwst` by more comprehensively treating 1/*f* noise. 1/*f* noise is a unique noise source caused by the introduction of a small, arbitrary bias level when pixels are read off of the detector, and is present to varying degrees in all of the NIR *JWST* instruments (e.g. Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023). The bias level introduced varies over time. As the SOSS detector is read column-by-column, this manifests as column-correlated 'streaks' (e.g. Fig. 1c). The 1/*f* noise level technically also varies along a single column, as pixels at the bottom of the detector are read slightly after those at the top.

The ostensible purpose of the `RefPixStep` in the official `jwst` pipeline is to correct this 1/*f* noise using non-illuminated pixels on the top edge of the detector. Since these pixels are not light sensitive, whatever counts are measured must be introduced during readout. However, we have found this correction to be inadequate for the complete removal of the 1/*f* noise. This is likely due to the fact that there are only four rows of reference pixels, and since the 1/*f* noise also varies in a given column, the 1/*f* level determined in the top four reference pixels will not necessarily hold for the entire column.

We therefore develop an alternate strategy to handle the 1/*f* noise and implement it in `supreme-SPOON`. Since the 1/*f* noise is introduced during readout, it is one of the, if not the last, noise source injected into the data and should therefore be one of the first to be corrected. The background subtraction and correction of 1/*f* noise though are necessarily highly coupled, and for reasons that will be described more fully below, the background *must* be subtracted before considering the correction of 1/*f* noise.

The SOSS background has a unique structure due to the combination of the JWST 'pick-off' mirror and the GR700XD grism (Albert et al. 2023). The zodiacal background falling off the pick-off mirror creates a step in the background at around column ∼750 (e.g. Fig. 1e). Furthermore, the dispersal of this zodiacal light by the GR700XD grism causes low-frequency variations, which can be seen as a fading of the background level towards the right edge of the detector.

Due to this unique structure, a constant background subtraction (i.e. estimating the background level based on one region of the detector, and subtracting this everywhere) will not be optimal. Indeed, as the strength of the background changes with wavelength (detector column), this will introduce a significant, wavelength-dependant dilution to transit (or eclipse) depths (e.g. Fig. 2). Instead, `supreme-SPOON` subtracts a model of the SOSS background, scaled to the flux level of each group. This model, shown in Fig. 1(e), was created during commissioning[4] and is available from jdox.[5] We note that other background models can be swapped into `supreme-SPOON` by the user if required. First, a median stack of each group is taken across the out-of-transit integrations. Then the background model is scaled to the flux level of each stack using a small region (x∈[250, 500], y∈[210, 250]) in the upper-left corner of the SUBSTRIP256 detector where the contribution from the target orders is minimal. This region should be adjusted as needed for each individual TSO as background sources may contaminate this

[1] https://github.com/radicamc/supreme-spoon
[2] https://jwst-pipeline.readthedocs.io/en/latest/jwst/pipeline/index.html. `jwst` v1.6.2 is used in this work.
[3] https://jwst-pipeline.readthedocs.io/en/latest/jwst/pipeline/calwebb_detector1.html
[4] Programme 1541
[5] https://jwst-docs.stsci.edu/jwst-calibration-pipeline-caveats/jwst-time-series-observations-pipeline-caveats/niriss-time-series-observation-pipeline-caveats#NIRISSTimeSeriesObservationPipelineCaveats-SOSSskybackground

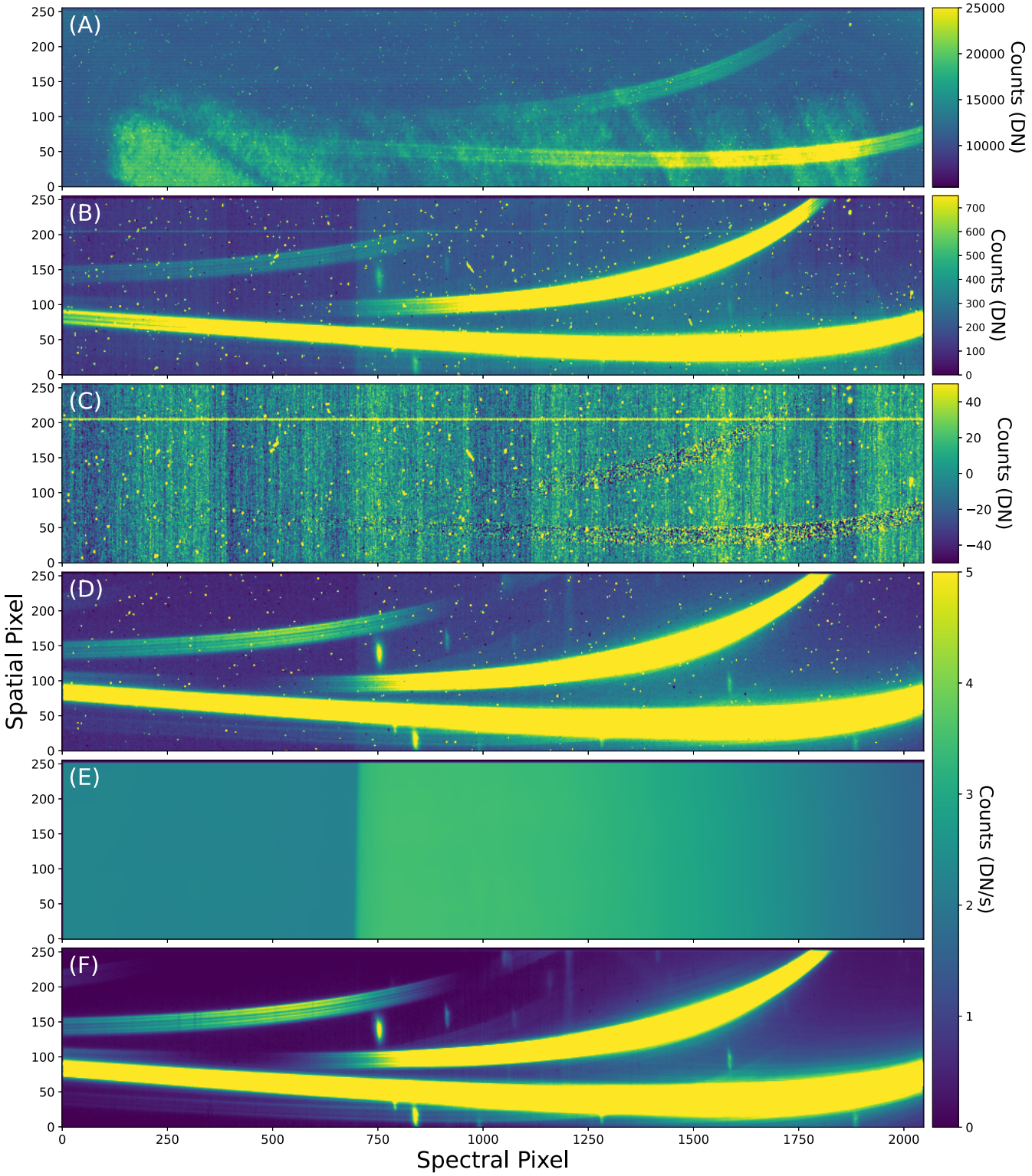


**Figure 1.** Visualization of the `supreme-SPOON` data products at several stages of the reduction process. **(a)** A raw, uncalibrated data frame in data numbers (DN). **(b)**: After superbias subtraction and reference pixel correction. **(c)** Frame (b) after the first background subtraction, and subtraction of the scaled median of all integrations to reveal the 1/*f* noise. **(d)** Data frame after ramp fitting and flat field correction. **(e)**: Background model scaled to the flux level of (d). **(f)**: Final calibrated data product. The horizontal stripe near row 200 in panels (b) and (c) is a known artifact resulting from FULL frame resets which occur before each subarray exposure.

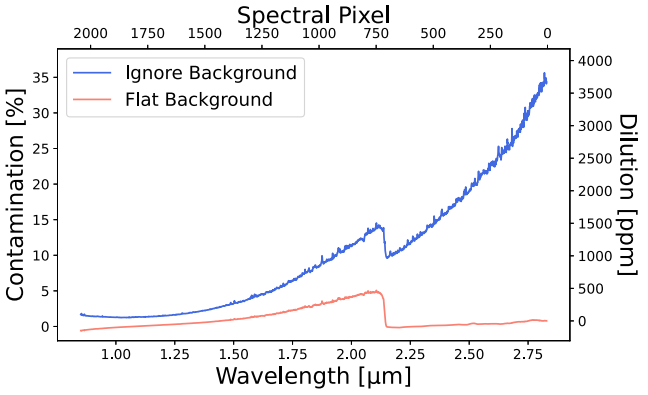


**Figure 2.** Dilution incurred through improperly treated background subtraction. The blue line shows the dilution to the WASP-96 b light curves if the background contributions are ignored entirely. In the red is the dilution which would occur if a constant background level is subtracted from the entire frame. Due to the structured nature of the SOSS background, this constant subtraction is ill-advised and can result in a dilution of several hundred ppm.

particular region. After the model is rescaled, it is then subtracted from each group in each integration.

We then proceed with the correction of the 1/*f* noise. The 1/*f* noise structures are generally quite faint (the streaks are barely visible in Fig. 1b). We therefore 'reveal' them by constructing difference images. To this end, we construct a group-wise median stack using all out-of-transit integrations; that is to say, we create one median stack for each group. For each integration, we then subtract each group frame from that group's median stack to remove the trace of the target star and reveal the 1/*f* noise structures (Fig. 1c). A column-wise median is then calculated on this difference image and subtracted from the original data frame. The core region of the target trace, any bad pixels, and bright background contaminants should also be masked during this procedure.

In order to accurately estimate the 1/*f* level, it is critical to have fully subtracted off the trace of the target star. The wings of the SOSS point spread function (PSF) are extremely broad, extending across virtually the entire detector. If residuals of the wings remain in the difference image when the 1/*f* noise level is calculated, this will again introduce a wavelength-dependent dilution. To subtract the target trace as fully as possible, it is critical to take into account the fact that it will dim due to the transit (or eclipse) of the planet. Therefore, we rescale the group-wise median to the flux level of each integration before subtracting it, using an estimate of the white light curve (e.g. a factor of 1.0 for out-of-transit frames, and 0.99 for in-transit frames of a 1 per cent transit). This procedure allows us to completely subtract the target trace, leaving only Gaussian noise in the trace core. In theory, wavelength dependant variations (due to the planet's atmosphere signal, for example) could also change the brightness of the trace within a given frame. However, these will generally be higher order effects and can safely be ignored in most cases. For planets with unusually strong atmospheric signals, or eclipses with large wavelength-dependent changes in brightness, this may become an important effect (e.g. Coulombe et al. 2023).

An additional consideration in this process is that when rescaling the group-wise median frame, one must be careful not to *also rescale the background*. The flux in the target trace will change due to a transit, however, the background level will naturally remain constant. This is why `supreme-SPOON` treats the background at the group level, and why the background is subtracted before correcting the 1/*f* noise. It is, furthermore, critical to adequately mask bright background sources (e.g. order 0 contaminants) during the 1/*f* correction. Since the flux level of these contaminants remains (to first order) constant over the course of the TSO, subtracting the scaled median frame from each integration will leave positive residuals at the locations of order 0 contaminants. These residuals bias the calculated 1/*f* level in affected columns to larger values – thereby resulting in the over-subtraction of 1/*f* noise and antidilution in the resulting transit depths.

The final level of complexity in the background-1/*f* coupling stems from the non-linearity correction. NIR integrations are taken in a non-destructive fashion, so-called 'up-the-ramp' reads, and final calibrated 2D images are composed of rate quantities (counts per second), obtained via fitting a linear trend to the pixel values of each group. For bright objects and large counts, the detector response begins to become non-linear, where the detector measures fewer counts that are received in reality. For SOSS, non-linearity effects become large around 35 000 counts (Albert et al. 2023); for this WASP-96 b TSO, the peak counts reached is only $\lesssim$20 000. This detector response correction is performed via the `LinearityStep`. It is important here to consider the order in which 'noise' is 'added' to the observations. The background flux arrives at the detector simultaneously to the photons from the target. The background is therefore subject to any non-linearity effects. However, the 1/*f* noise is introduced as the detector is read and *is not affected by non-linearity*. Therefore, for the most precise results, the 1/*f* noise should be treated before the non-linearity correction, and the background after. We therefore re-add the previously subtracted background to the data at the end of the `OneOverFStep`, and re-subtract it during Stage 2. We note here that performing the 1/*f* correction at the integration level (that is, after the non-linearity correction) does not lead to any systematic biases in the resulting transmission spectrum; however, it does result in overall less precise transit depths and more scatter (see Appendix B).

After the correction of the 1/*f* noise, the remaining steps for Stage 1 are well taken care of by the official `jwst` pipeline, and `supreme-SPOON` once again simply provides wrappers around the corresponding steps for steps viii–x. We note though that `supreme-SPOON` skips any dark current subtraction, as the dark reference file provided in the Calibration Reference Data System (CRDS) shows clear signs of superbias residuals and 1/*f* noise, which actually increases the noise level in the observations when subtracted. We find though that the dark level is in general $\lesssim$25 counts, and can therefore be safely ignored.

### 2.2.2 *supreme-SPOON Stage 2 – spectroscopic calibrations*

The second stage of `supreme-SPOON` performs further, high-level calibrations to SOSS data frames. A brief summary of the four major steps are outlined below:

(i) `FlatFieldStep`: Removal of flat field.
(ii) `BackgroundStep II`: Subtract the background level.
(iii) `BadPixStep`: Interpolation of hot, and other persistent bad pixels.
(iv) `TracingStep`: Determine the centroids and stability of the target trace.

The `FlatFieldStep` is a wrapper around the corresponding step of the official JWST pipeline, and performs the standard correction by dividing the data frames by a flat field reference image. This is the only step in this stage shared with the official pipeline. The `BackgroundStep` is once again performed, to

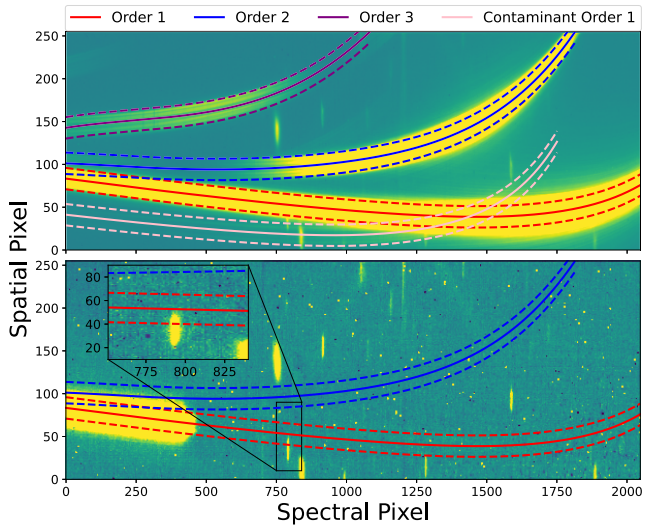


**Figure 3.** Visualization of field star contaminants in the WASP-96 b SOSS frame. *Top*: Median stack of the CLEAR exposure. The centroids of the target traces for the first, second, and third orders are indicated via the red, blue, and green solid lines, respectively. The 25-pixel extraction boxes are bounded by the dashed lines in each colour (Note that order 3 is not extracted, but included in this plot for completeness). The contaminating order 1 from a field star is indicated in pink and can be clearly seen to intersect the extract box of the target order 1 between pixels ∼1250–1600. *Bottom*: Median stack of the F277W exposure. The centroids and extraction boxes for orders 1 and 2 are again indicated identically to the above. Multiple field star order 0s are visible in this frame. One, centred at spectral pixel ∼790, clearly intersects the order 1 extraction box (see inset).

permanently remove the SOSS background. The process is identical to that described in Section 2.2.1, except only one median stack is constructed. The background model is then scaled to the flux level of this median stack, and subtracted from each integration. A single integration, before subtracting the background model is shown in Fig. 1(d), and the scaled background model itself in Fig. 1(e).

We skip both of the flux calibration steps, `pathloss` and `photom`, of the official `jwst` pipeline, as exoplanet atmospheric spectroscopy relies on relative measurements (e.g. in- vs out-of-transit), which renders an absolute flux calibration unnecessary.

`supreme-SPOON` then corrects any remaining hot pixels. This is accomplished by first constructing a median stack of all integrations. Pixels in the median stack which deviate by more than a given threshold from the surrounding pixels in time (for this work we use a threshold of $5\sigma$) are flagged and then corrected in each individual integration using the median of the neighbouring pixels. This is the final reduction stage, and an example of the resulting data frame is shown in Fig. 1(f).

The last step does not perform any additional reductions but is a utility step to aid in 1/*f* correction as well as eventual spectral extraction and light-curve fitting. The `TracingStep` extracts the centroids of the target traces, and optionally also determines the stability of the target trace over the course of the TSO. Centroids are extracted for each order from a final median stack of all integrations via the 'edgetrigger' algorithm (Radica et al. 2022). The centroids are overplotted on the final median stack for the WASP-96 b TSO in the top panel of Fig. 3. The centroids are necessary to define the extraction box (see Section 2.2.3), but also to define a trace mask which can be used to mask the trace core during 1/*f* correction.

Furthermore, the `TracingStep` can determine the stability of the target trace throughout the TSO. Pointing jitter, for example, may cause the target trace to shift positions on the detector over the course of a many-hour TSO. The stability is assessed through a cross-correlation function analysis. The median stack is shifted by a small fraction of a pixel individually in the vertical and horizontal direction and cross-correlated with the data frame from each integration. This enables the measurement of the position of the target trace to sub-pixel accuracy, relative to the position of the median trace. Furthermore, the full width at half-maximum (FWHM) of the trace is estimated by creating a 'super-profile' for each integration by stacking the PSFs of the first order near the peak in throughput (detector columns 1500–1750). A Gaussian profile is then fit to this super profile to estimate the FWHM. The change in FWHM, as well as *x* and *y* position, are shown in the bottom panel of Fig. C1. In general, the trace position is remarkably stable, with root mean square (RMS) deviations in *x* and *y* of ∼2 and ∼4 milli-pixels, respectively. The FWHM is also extremely stable, except during a mirror tilt event ∼1 hr after the transit midpoint (see Section 3.1).

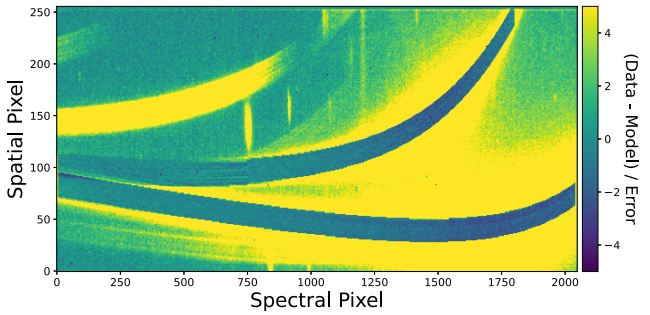


**Figure 4.** Example `ATOCA` decontamination results for integration 100. The first and second orders are well modelled across the entire detector, with residuals generally bounded within $1.5\sigma$.

### 2.2.3 *`supreme-SPOON` Stage 3 – 1D spectral extraction*

Stage 3 of the `supreme-SPOON` pipeline performs the 1D spectral extraction. The four major steps are listed below.

(i) `SpecProfileStep`: Construct a model of the SOSS trace profile for all orders.
(ii) `SossSolverStep`: Determine the correct transform to match the CRDS reference files.
(iii) `Extract1dStep`: Perform the 1D extraction via either a box extraction or using the `ATOCA` algorithm.
(iv) `LightCurveStep`: Convert spectra into a convenient data format.

Within `supreme-SPOON`, the 1D spectral extraction can be performed via two different methodologies: either with a simple box aperture extraction or using the `ATOCA` algorithm (Darveau-Bernier et al. 2022). `ATOCA` was developed to explicitly treat the contamination resulting from the overlap of the first and second SOSS orders on the detector. Briefly, `ATOCA` constructs a linear model of every pixel on the detector, including contributions from both orders. It then 'decontaminates' the detector – that is, it produces a model of orders 1 and 2 which are then individually subtracted from the data to create decontaminated images of order 2 and 1, respectively. A box extraction can then be safely performed on these decontaminated frames completely free of contamination from the other order. An example of the decontamination for integration 100 of the WASP-96 b TSO is shown in Fig. 4.

Although this order contamination can introduce significant biases for absolute flux measurements, for relative flux measurements, like those used in exoplanet transmission and emission spectroscopy, the dilution caused by the contamination is negligible (Darveau-Bernier et al. 2022; Radica et al. 2022).

The `SpecProfileStep` is skipped for box extractions, but critical for extractions using `ATOCA`. This is because it constructs an estimate of the spatial PSF of the target trace upon which `ATOCA` relies using the `APPLESOSS` algorithm (Radica et al. 2022). A trace profile estimate, which was constructed with `APPLESOSS` on commissioning data of a quiet A0V star, is provided as part of the `specprofile` CRDS reference file. However, the fine structure of the SOSS PSF may change subtly over time due to, among other factors, changes in the wavefront caused by small temperature changes, or more drastically due to things such as tilt events. Therefore, it is recommended to construct a trace profile specifically for each TSO instead of relying on the reference profile provided. We additionally note that `APPLESOSS` has been significantly upgraded since the commissioning period to produce profiles which better reproduce the extended wing structure of the SOSS PSF (Fournier-Tondreau et al., in preparation).

Next, the `SossSolverStep` uses the extracted centroids to calculate the necessary parameters (rotation, $x$-shift, $y$-shift) to transform the reference centroids provided in the `spectrace` CRDS reference file to match the data. This is performed via a sequential least squares $\chi^2$ minimization implemented in the `scipy.optimize.minimize` routine. For WASP-96 b, we find that the centroids are already well-matched to those provided by the CRDS without applying any transform. The required transform is then passed to the `Extract1dStep` along with the calibrated data cube to do the 1D extraction. This is once again a wrapper around the official JWST pipeline step of the same name. We perform extractions using both the box and `ATOCA` methods, but since the results are identical (e.g. Fig. C3), we only consider the `ATOCA` results for the remainder of this work. In both cases, we use a box width of 25 pixels as this was found to maximize the signal-to-noise ratio (S/N) of the extracted stellar spectra. The extraction boxes as well as trace centroids for both orders are shown in the top panel of Fig. 3.

Lastly, the `LightCurveStep` interpolates any outliers in time above a user-specified threshold (in this work, we use $5\sigma$ as we find that sufficient to remove the handful of remaining outlier points) in the extracted stellar spectra, and packages the spectra into an easily portable Flexible Image Transport System (FITS) file format along with all the relevant extraction parameters to ensure reproducibility of the results.

### 2.2.4 *Spectrum post-processing*

The steps described above will be necessary for all SOSS TSOs, and can be performed, with slight fine-tuning, for any observation. However, the SOSS mode presents a further complexity, namely field star contamination, which must be treated on a case-by-case basis. As SOSS is a slitless mode, any background star lying within the field of view of NIRISS will also be dispersed onto the detector. The amount of field star contamination can be minimized by effectively choosing targets, as well as telescope roll angles. However, occasionally some level of contamination will be impossible to avoid.

The WASP-96 b TSO is an example of a case where background contamination was unavoidable. In Fig. 1(f), as well as the top panel of Fig. 3, two types of field star contaminants are clearly visible: order 0 contaminants, which appear as bright smudges at various locations, and an order 1 contaminant (whose trace is outlined in pink in the top panel of Fig. 3). Below we suggest a method for correcting each type of contaminant and apply them to the WASP-96 b TSO.

#### 2.2.4.1 *Order 0 contaminants*

Order 0 contaminants, due to their concentrated brightness, can potentially cause a large amount of dilution if they intersect the target trace. Although most order 0 contaminants are bright enough to be visible in a GR700XD/CLEAR exposure, we make use of the GR700XD/F277W exposure to more efficiently identify potential contaminants. The F277W filter only transmits light redder than $\sim$2.6 $\mu$m, effectively isolating the red end of the order 1 trace. This was initially posited as a key to understanding the extent to which the target orders self-contaminate, which in the end was proven negligible (see Section 2.2.3). Instead, this filter does allow for the effective identification of field star contaminants. We process the F277W filter exposures in a similar manner to that of the GR700XD/CLEAR science exposures. The bottom panel of Fig. 3 shows a median stack of the F277W filter exposure, as well as the extraction boxes for orders 1 and 2. Field star order 0s are clearly visible in the frame, even those which could potentially be hidden behind the target trace. We identify one order 0 (shown zoomed-in in the inset) centred at column $\sim$790 which intersects the target order 1 trace. There is a second contaminant, at pixel $\sim$1300 which appears to graze the order 1 extraction box; however, we were not able to ascertain that this contaminant causes any meaningful dilution, and we therefore ignore it.

Then, we proceed to estimate the dilution introduced by the order 0 contaminant at column $\sim$790. We roughly estimate the extent of the contaminant in the horizontal direction, and again construct a 'super-profile' by stacking the PSFs of the five columns on each side of the contaminant. We linearly interpolate the two super profiles over the pixels contaminated by the order 0 to construct a roughly 'uncontaminated' trace, which we subtract off of the detector to reveal the order 0 contaminant. Using the same extraction box as for the target trace, we then extract over the contaminant to estimate the amount of flux, and therefore the dilution factor, it introduces.

The dilution introduced by this order 0 contaminant is found to be quite significant, at a level of $\sim$750 ppm, but only over a very small wavelength region ($\sim$15 columns; e.g. Fig. C3).

#### 2.2.4.2 *Order 1 contaminant*

In the top panel of Fig. 3, a contaminant first order trace (outlined in pink) can be clearly seen to intersect the first order trace of the target. This will again introduce some dilution to the extracted light curves. In order to estimate the level of dilution from this contaminant, we first subtract off the order 1 trace of the target. We accomplish this by once again using the `ATOCA` algorithm to construct a model of the first order trace. Instead of using a width of 25 pixels as we did for our initial extraction, we instead select a width of 100 pixels in order to capture the extended wing structure in addition to the profile core.

After subtraction of the modelled target order 1, we then manually fit the extracted centroids of the target order 1, to the contaminant. We find that only order 1 is able to match the shape of the contaminant trace and that its shape is well represented by a shift of -510 pixels in the horizontal and -21 in the vertical direction from the target order 1. This also allows for the estimation of the wavelength solution for the contaminant via applying the same transform to the wavelength solution of the target. We then define an extraction box and extract the spectrum of the order 1 contaminant over pixel columns 0–500. It is critical that the order 1 of the target be subtracted off before extracting the contaminant spectrum, as the target order 1 wings introduce significant

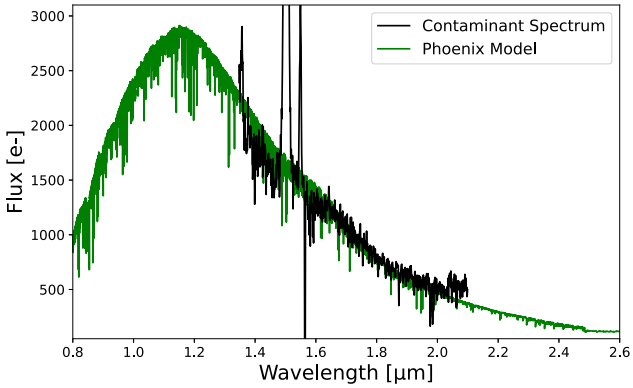


**Figure 5.** Comparison of the extracted spectrum of the contaminant order 1 field star and the best-fitting PHOENIX stellar model. The contaminant spectrum is well fit by a stellar model with parameters $T_{\rm eff} = 3900$ K, $\log g = 4.5$, and $Fe/H = 0.5$.

contamination to the spectrum of the contaminant; how the tables turn.

Using a custom SOSS contamination tool[6] we were able to identify the contaminant star in the Gaia DR3 catalogue[7] (Gaia Collaboration 2021). Cross-referencing with the *Transiting Exoplanet Survey Satellite* (Ricker et al. 2014) Input Catalogue, we retrieve the effective temperature and gravity of the source: $T_{\rm eff} = 3900$ K, $\log g = 4.65$. We then compared the extracted contaminant spectrum to a grid of PHOENIX stellar models (Husser et al. 2013), which have been corrected for the instrument throughput, and find the spectrum to be well matched by a model with $T_{\rm eff} = 3900$ K, $\log g = 4.5$, and $Fe/H = 0.5$ (Fig. 5) – which are consistent with the stellar parameters quoted above. We then smoothed the model stellar spectrum and used it to estimate the level of dilution introduced in the region where it intersects the target trace.

As the strength of any of the contaminants does not appear to vary in time, we then subtract this estimated contamination from the extracted stellar spectra at each integration to construct contamination-corrected spectra (but c.f., Fu et al. 2022 for a case of time-variable background star contamination). The transmission spectra before and after contamination correction are shown in Fig. C3. Compared to the order 0 contaminant, the order 1 contaminant introduces a much lower level of dilution ($\lesssim$100 ppm vs $\sim$750 ppm) but over a much larger wavelength range.

## 3 LIGHT-CURVE ANALYSES

### 3.1 White light-curve analysis

We sum the flux from all wavelengths to create white light curves for each order. We only consider wavelengths $<0.85$ $\mu$m for order 2 as at longer wavelengths the information is redundant with, and at a much lower signal-to-noise level than order 1. The white light curves are shown in Fig. 6.

We fit a `batman` transit model (Kreidberg 2015) to each white light curve using the flexible `juliet` package (Espinoza, Kossakowski & Brahm 2019). There appears to be a slight linear trend with time in the light curves for both orders; as well as a small ramp-like effect in the first $\sim$five integrations; but other than that they both appear free of systematic effects. We therefore cut these initial five integrations and also test a variety of different detrending techniques to find the optimal combination of transit and systematics models. We perform the `juliet` fits using its nested sampling capabilities, as implemented via `dynesty` (Speagle 2020), which allows for the estimation of the Bayesian evidence, $\ln \mathcal{Z}$, for each model. We test detrending against linear models with time ($\ln \mathcal{Z} = 2012.54$), trace $x$-position ($\ln \mathcal{Z} = 2015.99$), trace $y$-position ($\ln \mathcal{Z} = 2013.95$), and trace FWHM ($\ln \mathcal{Z} = 2015.12$). There is a slight preference ($\sim 2\sigma$ based on the scale from Benneke & Seager 2013) for detrending against the trace $x$-position. All systematics models are highly preferred ($>5\sigma$) over the case with no detrending ($\ln \mathcal{Z} = 1992.84$).

There was a so-called 'tilt-event' which occurred approximately 1.4 h after mid-transit (integration $\sim$200) during the WASP-96 b TSO. Tilt events are abrupt changes in the telescope wavefront potentially due to the thermal relaxation of the primary mirror segments (Rigby et al. 2023). As the SOSS PSF is so highly dispersed, it is incredibly sensitive to wavefront changes, and tilt events can sometimes manifest as large discontinuities in light curves as more (or fewer) photons fall within the extraction aperture (e.g. Coulombe et al. 2023). There is no clear evidence of a tilt-event-induced discontinuity in our white light curves. However, the tilt event is clearly visible in the FWHM of the trace (bottom panel of Fig. C1), as well as the 2D detector images themselves (Fig. C2). This is because the morphological change in the PSF which occurs during the tilt-event is contained entirely within our 25-pixel extraction box, and there is therefore no net change in flux within the aperture. However, using instead a 20-pixel aperture does result in a marked discontinuity in the light curve as the tilt-event causes additional flux to fall into the aperture (see Figs C1 and C2). We therefore do not explicitly model the tilt event in our light-curve fits (e.g. via detrending against an additional 'jump' parameter in the fits) as there is no evidence it affects the light curve at a level greater than our measurement precision, and a systematics model consisting of a jump term in addition to the trace $x$-position was disfavoured by $\sim 4\sigma$ over the case with just the trace $x$-position ($\Delta \ln \mathcal{Z} = 9.2$).

Our final white light transit model therefore has nine parameters: the transit mid-point time, $t_0$, the scaled planet radius, $R_p/R_*$, the transit impact parameter, $b$, the scaled orbital distance, $a/R_*$, the two parameters of the quadratic limb-darkening law, $u_1$ and $u_2$ (Claret 2000) sampled via the Kipping (2013) parameterization, a scalar jitter which is added in quadrature to the error bars, $\sigma$, as well as the two parameters of a linear systematics model with $x$-position, $\theta_0$ and $\theta_1$. Wide uninformative flat priors are used for each parameter. We additionally fix the period to 3.4252602 d (Nikolov et al. 2022) and assume a circular orbit. The best-fitting transit models are overplotted in black in the top panels of Fig. 6, and the residuals to the fits are shown in the middle and bottom panels. Additionally, the best-fitting transit parameters are shown in Table 1, along with those from all other reductions.

We assess the quality of the transit model fits to each order via the reduced chi-squared ($\chi^2_\nu$) metric. This value, along with the size of the average error bar, $\sigma$, and $e$; the error multiple necessary to reach a $\chi^2_\nu$ equal to unity are listed for each order in the top panel of Fig. 6. For both orders, we find a very favourable $\chi^2_\nu$ metric (1.15, and 1.14 for orders 1 and 2, respectively) indicating that the assumed model is an excellent fit to the data. The residuals are furthermore highly Gaussian, and bin down as would be expected from pure photon noise (e.g. Fig. C4).

[6] http://maestria.astro.umontreal.ca/niriss/SOSS_cont/SOSScontam.php
[7] Gaia Source ID 4990044874536787328

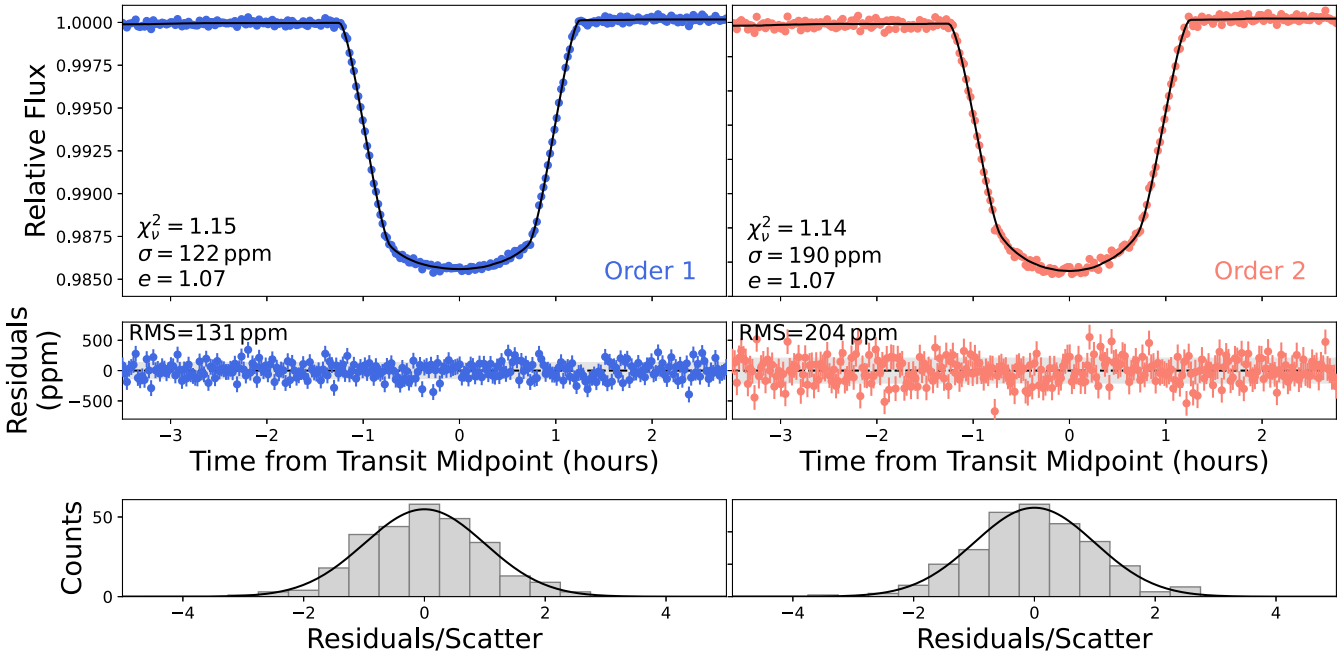


**Figure 6.** *Top*: White light curves for for order 1 (*left*) and order 2 (*right*). The best-fitting transit model is overplotted in black. The fit statistics are shown for each order: $\chi^2_\nu$, $\sigma$; the average error bar, and $e$; the error multiple to obtain a $\chi^2_\nu$ equal to unity. *Middle*: Residuals to the transit fits. The RMS scatter in the residuals is indicated on each panel. *Bottom*: Histogram of residuals. The histograms trace Gaussian distributions indicating that we have well handled systematic trends in the data.

**Table 1.** Comparison of best-fitting white light-curve parameters.

| Reduction | $t_0$ [BJD − 2400000] | $R_p/R_*$ | b | $a/R_*$ |
|---|---|---|---|---|
| `supreme-SPOON` | $59751.82467 \pm 3 \times 10^{-5}$ | $0.1190 \pm 5 \times 10^{-4}$ | $0.7301 \pm 0.0037$ | $8.963 \pm 0.040$ |
| `nirHiss` | $59751.82471 \pm 5 \times 10^{-5}$ | $0.1197 \pm 9 \times 10^{-4}$ | $0.7276 \pm 0.0055$ | $8.988 \pm 0.058$ |
| `transitspectroscopy` | $59751.82464 \pm 5 \times 10^{-5}$ | $0.1191 \pm 5 \times 10^{-4}$ | $0.7243 \pm 0.0045$ | $9.011 \pm 0.044$ |
| `NAMELESS` | $59751.82469 \pm 4 \times 10^{-5}$ | $0.1193 \pm 9 \times 10^{-4}$ | $0.7287 \pm 0.0051$ | $8.978 \pm 0.051$ |
| Nikolov et al. (2018) | – | – | $0.749 \pm 0.020$ | $8.84 \pm 0.10$ |

Values listed above are for order 1 white light curves.

### 3.2 Spectrophotometric light-curve analysis

To obtain the final transmission spectrum, we proceed to fit the above transit model to light curves at the pixel level (that is one light curve per pixel column). Studies during commissioning have shown that binning detector columns after spectral extraction, but before light-curve fitting, results in higher than expected levels of scatter in the out-of-transit light-curve baseline. This has been attributed to unaccounted-for covariance between detector columns (see e.g. fig. 8 in Espinoza et al. 2023, for an example with NIRSpec BOTS). Espinoza et al. (2023) in particular, conclude that the optimal way to work with NIR spectra from *JWST* is to extract spectra and fit light curves at the pixel level, and bin observables, such as transit depths, at a later stage. We thus adopt this strategy, which results in 2040 bins for order 1 (2048 pixel columns minus eight reference pixel columns) and 567 bins for order 2, where we only consider wavelengths <0.85 $\mu$m. In each bin our `supreme-SPOON` reduction reached an average precision of 1.2 and 1.4× the photon noise for orders 1 and 2, respectively (e.g. Fig. 7).

Instead of fixing the orbital parameters of WASP-96 b to the best-fitting values from the white light curve, we instead used the $b$, and $a/R_*$ provided by Nikolov et al. (2022; last row of Table 1). This was chosen, in order to maintain a constant set of orbital parameters for the four different reductions. The transit mid-point time, however, was fixed to the best-fitting value from the order 1 white light curve. We note here that what follows pertains only to the `supreme-SPOON` reduction, and details on the light-curve fits for the other three reductions can be found in Appendix A. We placed Gaussian priors

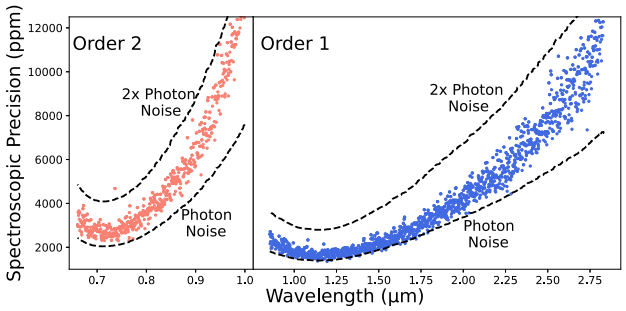


**Figure 7.** Spectroscopic precision per bin obtained in this NIRISS/SOSS TSO of WASP-96 b at the pixel level (bin width of one pixel) compared to the predicted photon noise limits (dashed black lines) for order 1 (right) and order 2 (left). The estimated read noise (∼21 ppm) is removed here from the precision calculated in each bin. Our `supreme-SPOON` reduction reaches an average precision of ∼1.2× and ∼1.4× the predicted photon noise for order 1 and order 2, respectively.

on both of the limb-darkening coefficients, centred on predictions from the `ExoTiC-LD` package (Sing 2010; Laginja & Wakeford 2020) using the 3D stellar models of Magic et al. (2015). In order to estimate both the position and width of the Gaussian prior, we calculated limb-darkening coefficients over a three-dimensional grid in the stellar effective temperature ($T_{eff}$), metallicity ([Fe/H]), and gravity ($\log g$) parameter space. The grid had five nodes, equally spaced in each dimension, and the extent of the grid was determined by the published uncertainties in each parameter. For all three stellar parameters, we used the following values and uncertainties from Hellier et al. (2014): $T_{eff} = 5400 \pm 140$ K, [Fe/H] $= 0.14 \pm 0.19$ dex, and $\log g = 4.42 \pm 0.02$ g cm$^{-3}$. The position of the Gaussian prior for the limb-darkening parameters at each wavelength was taken to be the mean values of the grid, and the width of the prior as the standard deviation across the grid. We therefore fit seven parameters to each bin. The light curves for several representative bins, as well as their best-fitting transit models are shown in the left-hand panel of Fig. 8, and the fit residuals on the right. The resulting transmission spectrum is shown in Fig. 9 at both the pixel resolution (faded points) and binned to a constant R = 125 for visual clarity.

Fig. 9 also includes the *VLT* and *HST*/WFC3 transmission spectra as published by Nikolov et al. (2022). We find offsets of 400 and 200 ppm for the *VLT* and *HST* transmission spectra relative to our SOSS spectrum. All of Yip et al. (2021), Nikolov et al. (2022), and McGruder et al. (2022) also find offsets between their individual reductions of the *VLT* and *HST* spectra, concluding that these offsets are linked to common-mode corrections applied to each data set in order to correct systematics in the binned light curves. Indeed, McGruder et al. (2022) explicitly explored the effects of common-mode corrections on retrieved transit depths, and find that although information regarding relative depths is preserved, absolute transit depths are poorly reproduced. Moving forward in the *JWST* era, we should therefore expect to find offsets between existing transmission spectra which require common-mode corrections (e.g. *HST*/WFC3), and those from *JWST* on which no common-mode corrections are applied and absolute transit depths are reliable.

## 4 GRID MODEL FITS

We interpret the NIRISS/SOSS observations by comparing them to a series of self-consistent radiative-convective equilibrium atmosphere models. We follow the procedure established by the JWST Transiting Exoplanet Community Early Release Science Team (Bean et al. 2018) and perform an initial atmospheric assessment using grids of models (e.g. Ahrer et al. 2023; Alderson et al. 2023; Feinstein et al. 2023; Rustamkulov et al. 2023). We defer detailed atmospheric retrieval analyses of these observations to the subsequent companion paper, Taylor et al. (2023). Below we describe the three grids of models used to interpret our WASP-96 b observations. Since all grids converge to broadly consistent results, we select the ScCHIMERA grid, which produces the overall best fit, as the reference for our discussion and conclusions.

We note that the model fits were performed on the observations at the different resolutions of R = 125, 250, 500, and pixel-level. The results for each grid are generally independent of the resolution of the observations, and we defer further analysis into any resolution-dependence of atmosphere inferences to Taylor et al. (2023). The best-fitting parameters for each model grid are included in Table 2.

### 4.1 PICASO

We use the modelling tool PICASO 3.0 (Batalha et al. 2019; Mukherjee et al. 2023) to create a grid of model atmospheres in 1D radiative-convective equilibrium for WASP-96 b. PICASO is based on the legacy 'Extrasolar Giant Planet (EGP)' code (Marley 1999; Fortney 2005; Marley et al. 2021) and uses the opacities listed in Marley et al. (2021). We generate two sets of models: one cloud-free and one with clouds. Both sets contain grids of models with atmospheric metallicities in the range (0.1, 0.3, 1, 3.1, 10, 31, 50, 100) times solar, and C/O ratios in the range (0.5, 1, 1.5, 2) times solar. We fix the energy redistribution between the day- and night-sides of the planet at 0.5, which represents full day-to-night heat redistribution in PICASO. Metallicity is computed by multiplying the abundances of elements heavier than hydrogen/helium by the appropriate factor. (C + O)/H ratio is held constant at a given metallicity so that changing the C/O ratio does not also change the metallicity. The chemistry grid for these metallicities and C/O ratios is computed using the thermocheical equilibrium models presented in Gordon & McBride (1994) and Visscher et al. (2010), and also presented in Marley et al. (2021). We use 0.458 for the solar C/O value (Lodders, Palme & Gail 2009). The cloudy grid additionally uses the Virga (Rooney et al. 2022) implementation of the Eddysed (Ackerman & Marley 2001) framework, with clouds parameterized by a vertical mixing coefficient $\log_{10}K_{zz}$ in the range (7, 9, 11; cgs units) and a sedimentation parameter $f_{sed}$ in the range (0.6, 1, 3, 6). This excludes grid edge points $\log_{10}K_{zz} = 5$ and $f_{sed} = 10$ used by the ERS collaboration previously in Feinstein et al. (2023). A $\log_{10}K_{zz} = 5$ is unphysically small at the expected temperatures of WASP-96 b (see fig. 2 in Moses et al. 2022, as well as references therein), while $f_{sed} = 10$ produces a compact, vertically thin cloud deck at pressures below the transmission contribution of our observations. Therefore, omitting these points does not affect the overall interpretation of WASP-96b. Both cloud parameters are taken to be constant with altitude. We account for cloud species MnS, $MgSiO_3$, Cr, $Mg_2SiO_4$, and Fe. Our grids use the orbital parameters described in Hellier et al. (2014) which yield an equilibrium temperature of 1291 K (for zero albedo). The best-fitting internal temperature is 200 K, yielding a planetary effective temperature of 1491 K.

### 4.2 ATMO

We also use a grid of self-consistent model atmospheres presented in Goyal et al. (2020) and previously applied to the WASP-96 b VLT and *HST* observations from Nikolov et al. (2022). This model grid has been generated using the 1D-2D planetary atmosphere model ATMO (Tremblin et al. 2015; Drummond et al. 2016; Goyal et al. 2018). In this grid, model transmission spectra are generated using radiative-convective equilibrium pressure–temperature (P–T) profiles consistent with equilibrium chemistry. These spectra are generated at R$\sim$1000 for all the opacity species listed in Table 1 of the supplementary material of Goyal et al. (2020) using the correlated-$k$ methodology with random overlap (Amundsen et al. 2017). For this work, model transmission spectra are generated for a range of heat re-distribution parameters (0.25, 0.50, 0.75, 1.0), metallicities (0.1, 1, 3, 5, 10, 50, 100, 200× solar), C/O ratios (0.35, 0.55, 0.7, 0.75, 0.90, 1.0), haze scattering parameters (1, 10, 150, 550, 1100× multigas Rayleigh scattering), and grey cloud parameters (0, 0.5, 1.0, 5.0× $H_2$ Rayleigh scattering cross-section at 350 nm). Here, a heat re-distribution parameter of 0.5 corresponds to efficient heat re-distribution. Radiative-convective equilibrium P–T profiles with four different re-distribution factors cover a range from hottest (1) to coolest (0.25) P–T profiles that could be encountered in the planet's atmosphere. The solar C/O ratio is 0.55, and a grey cloud parameter of 0 implies cloud free model spectra. The metallicity is perturbed in the model by multiplying the abundances of all the elements heavier

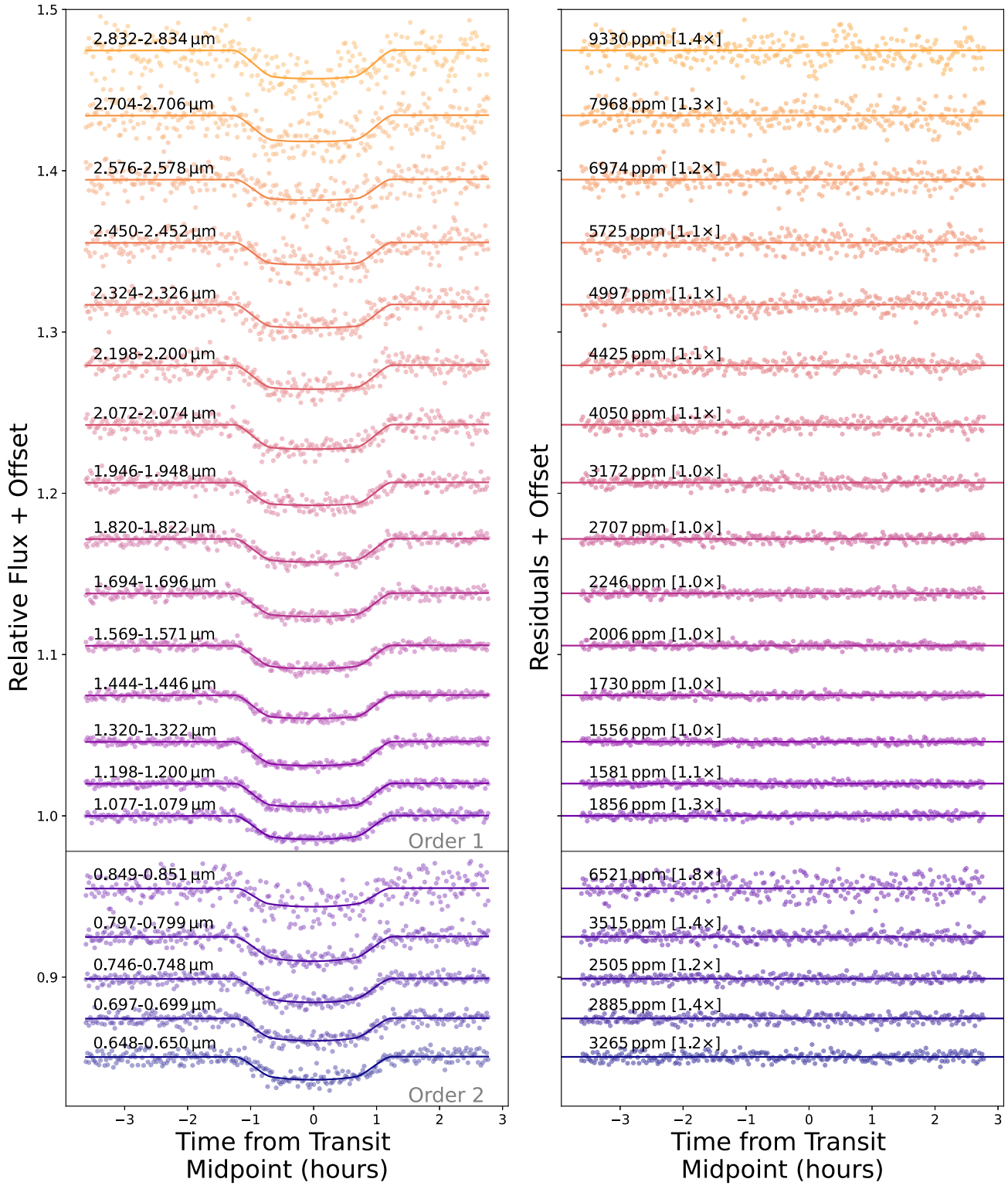


**Figure 8.** Spectrophotometric light curves and transit fits for the NIRISS/SOSS TSO of WASP-96 b. *Left*: Normalized spectrophotometric light curves at pixel resolution. The extent of each bin is indicated above each light curve. Overplotted on each is its best-fitting transit model. No systematic trends have been removed from the data. *Right*: Residuals for the light-curve fit in each bin. The RMS scatter of the residuals in each bin are indicated as well as the ratio to the predicted photon noise in brackets.

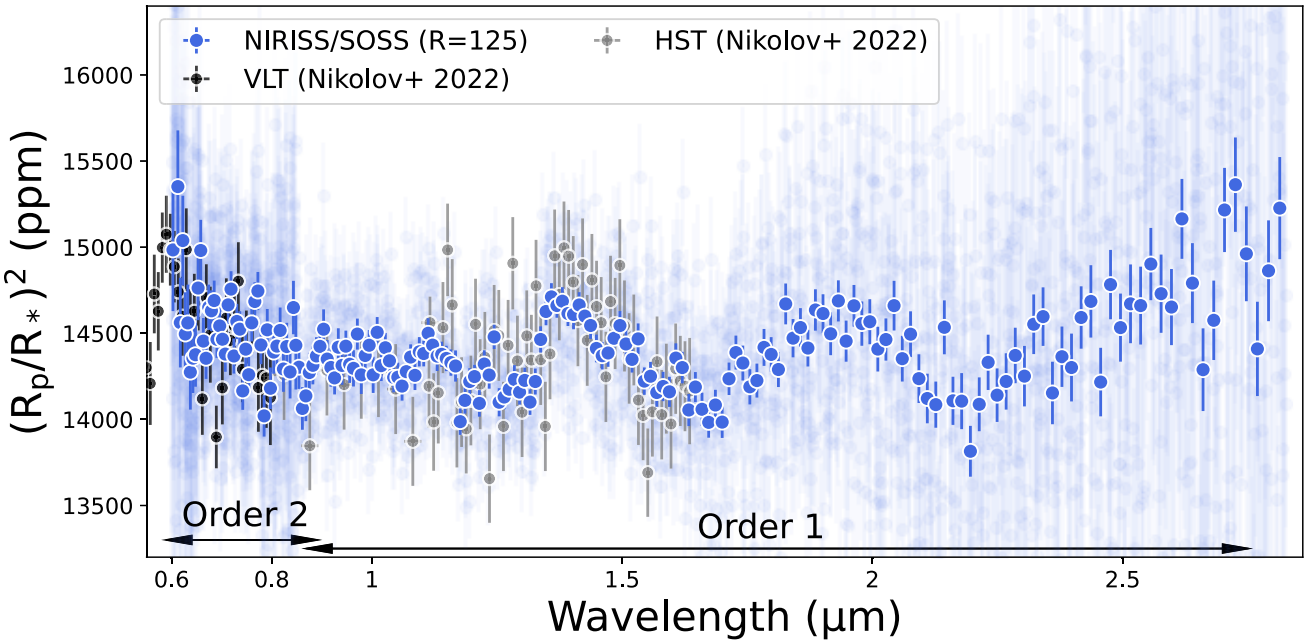


**Figure 9.** NIRISS/SOSS transmission spectrum of WASP-96 b at pixel resolution (faded blue points) and binned to a resolution of R∼125 (blue points). Also shown are the *VLT* (black points) and *HST*/WFC3 G102 and G141 (grey points) transmission spectra from Nikolov et al. (2022). Offsets of 400 and 200 ppm have been applied to the VLT and *HST* data, respectively (see text).

**Table 2.** Comparison of best-fitting model grid parameters.

| Model | Metallicity [×Solar] | C/O [×Solar] | Redistribution factor | Cloud parameters | Haze scattering parameters |
|---|---|---|---|---|---|
| PICASO | 10 | 1.5 | 0.5 (full redist.) | $f_{sed} = 1$, $\log_{10} K_{zz} = 7$ | — |
| ATMO | 5 | 1 | 0.5 (full redist.) | 0× grey cloud | 150× enhanced Rayleigh |
| ScCHIMERA | 1 | 1 | 1.319 (partial redist.) | $P_{cloud} = 1$ bar | $a = 1.78$, $\gamma = 4$ |

Results are from fits to R = 125 data. — indicates that this aspect was not considered in the grid. The ScCHIMERA results refer to model 2 described in the text.

than H and He by the appropriate metallicity factor. C/O ratio is varied in the model via the O/H ratio.

### 4.3 ScCHIMERA

Detailed descriptions of the ScCHIMERA framework are included in the previous works of Arcangeli et al. (2018), Piskorz et al. (2018), and Mansfield et al. (2021), and more recently in applications to *JWST* data in JWST Transiting Exoplanet Community Early Release Science Team (2023) and Feinstein et al. (2023). Generally, for a set of planetary parameters, this method pre-computes the P–T structure of the planet and the gas mixing ratio profiles under thermochemical equilibrium. The temperature pressure profile is computed under the assumption of radiative-convective equilibrium. The layer net fluxes are computed using the Toon et al. (1989) two stream source function technique and a Newton–Raphson iteration (McKay, Pollack & Courtin 1989) is used to march to convergence. Opacities are derived from correlated-K tables mixed on-the-fly using the random-overlap 'resort-rebin' procedure (e.g. Amundsen et al. 2016). The gas mixing ratios used to weigh the mixed correlated-K opacities are derived under the assumption of thermochemical equilibrium using the NASA CEA2 routine (Gordon & McBride 1994). While the calculation of the grid models fixes the internal temperature to 150K, we find that this value has no significant impact on the resulting P–T profile given the assumption of chemical equilibrium (e.g. Fortney et al. 2020). The computations are performed on a grid of atmospheric metallicity ([M/H], i.e. $\log_{10}$ enrichment relative to solar; Lodders, Palme & Gail 2009) spaced at 0.125 dex values between −1 and 2.5 (i.e. 0.1 to 316× solar) and C/O at values of 0.1, 0.2, 0.35, 0.45, 0.55, 0.65, 0.7, 0.725, 0.75, 0.775, 0.8, 0.85, 0.9, 0.95 (where 0.55 is solar). The grid also explores the energy redistribution ($f$) between the day and night sides of the planet (Fortney 2005), with values of 0.657, 0.721, 0.791, 0.865, 1.0, 1.03, 1.12, 1.217, 1.319 in our grid, where $f = 1.0$ corresponds to full day-to-night heat redistribution and $f = 2.0$ corresponds to dayside only redistribution.

The P–T structure and volume mixing ratio profiles of the different chemical species are then used to compute the transmission spectrum of the planet with CHIMERA (Line et al. 2013; Mai & Line 2019; Iyer & Line 2020) using correlated-$k$ tables at a resolution of R = 3000. The computed spectra are then compared to the observations using the Bayesian inference framework MultiNest (Feroz, Hobson & Bridges 2009) through its PYTHON implementation PyMultiNest (Buchner et al. 2014) to obtain an optimal (e.g. best-fit) solution for the atmospheric metallicity, C/O ratio, and heat redistribution. We also account for slight changes in the planetary radius by fitting for the planet radius at 1 bar, an arbitrary pressure with no direct impact on the inferred atmospheric properties (e.g. Welbanks & Madhusudhan 2019). Our atmospheric model considers the opacity

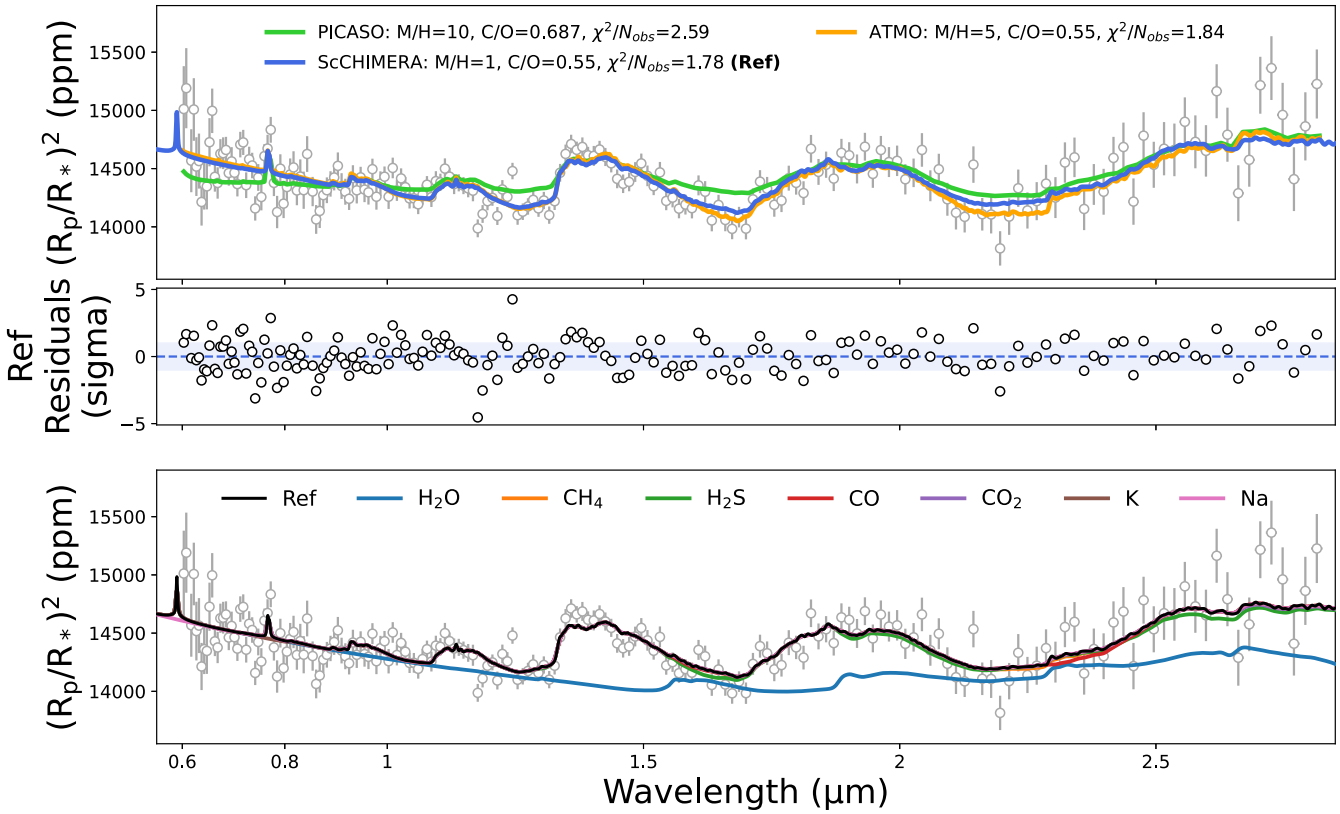


**Figure 10.** Grid modelling results of the WASP-96 b spectrum at R = 125. *Top*: Best-fitting models from the PICASO, ATMO, and ScCHIMERA grids overplotted with the R = 125 transmission spectrum. The ATMO and ScCHIMERA models prefer cloud-free solar-to-super-solar metallicity transmission spectra, with a solar C/O ratio and an enhanced Rayleigh scattering haze. The PICASO model prefers a super-solar metallicity and C/O ratio, as well as some grey cloud opacity – however, this is likely to be due to the fact that an enhanced scattering parameter was not included in the PICASO grids. *Middle*: Residuals [as (data – model)/error] to the reference ScCHIMERA model. *Bottom*: Impact on the transmission spectrum of removing the contributions of individual chemical species. $H_2O$ opacity dominates the spectrum over nearly the entire wavelength range, with a model preference (e.g. Bayesian evidence comparison, also known as 'detection significance'; Benneke & Seager 2013; Welbanks & Madhusudhan 2021) of $\sim 17\sigma$. The alkalis Na and K are marginally detected at $\sim 2\sigma$ and $\sim 3\sigma$, respectively. There are also hints of contributions from CO, $CO_2$ and $H_2S$ with marginal model preference of $\sim 2\sigma$.

sources expected to affect gas giant planets (e.g. Madhusudhan 2019), including $H_2$-$H_2$ and $H_2$-He CIA (Richard et al. 2012) alongside $H_2O$ (Freedman et al. 2014; Polyansky et al. 2018), $CO_2$ (Freedman et al. 2014), CO (Rothman et al. 2010), $CH_4$ (Rothman et al. 2010), $H_2S$ (Azzam et al. 2016), HCN (Barber et al. 2014), Na Kramida et al. (2018); Allard et al. (2019), and K (Kramida et al. 2018; Allard, Spiegelman & Kielkopf 2016). The opacities were computed following the methods described in Gharib-Nezhad et al. (2021) and Grimm et al. (2021).

Our atmospheric models also consider different cloud and haze treatments to explore the degree of cloudiness of WASP-96 b. We follow the model framework in Feinstein et al. (2023) and consider three base models: (1) a vertically distributed cloud opacity that is spatially uniform and grey (i.e. a single parameter $\kappa_{\rm cloud}$ which describes the opacity at all wavelengths). (2) A parameterization for scattering hazes following Lecavelier Des Etangs et al. (2008). That is, one parameter, $\gamma$, for the scattering slope and another, $a$, for the Rayleigh enhancement factor such that $\sigma_{\rm hazes} = a\sigma_0(\lambda/\lambda_0)^{\gamma}$. Here, $\sigma_0$ is the $H_2$ Rayleigh cross section at $\lambda_0$, given by $2.3 \times 10^{-27}$ cm$^2$ and 430 nm, respectively. This model also includes an optically thick cloud deck at a given atmospheric pressure, $P_{\rm cloud}$ (MacDonald & Madhusudhan 2017; Welbanks & Madhusudhan 2021). (3) A physically motivated droplet sedimentation model assuming enstatite grains with parameters to capture the behaviour of the eddy diffusion coefficient and the ratio of sedimentation velocity to the vertical mixing velocity (Ackerman & Marley 2001) following the prescription of Mai & Line (2019). This case has a total of four parameters: $P_{\rm base}$ the cloud-base pressure, $f_{\rm cond.}$ for the condensate mixing ratio at the cloud base, $K_{zz}$ for the Eddy diffusion coefficient, and $f_{\rm sed}$ for the ratio of sedimentation velocity to characteristic vertical mixing velocity. Furthermore, models 2 and 3 are considered as either fully cloudy or with inhomogeneous cloud cover by using the linear combination approach of Line & Parmentier (2016).

## 5 ATMOSPHERIC INFERENCE RESULTS

The best-fitting atmosphere models to the NIRISS/SOSS WASP-96 b transmission spectrum at R = 125 are shown in Fig. 10. The best-fitting solution found by PICASO ($\chi^2_\nu = 2.59$) suggests the atmosphere of WASP-96 b is best explained by a 10× solar metallicity atmosphere, a slightly super-solar C/O ratio of 0.687, and a preference for cloudy models in agreement with Samra et al. (2023). However, this preference for clouds over a cloud-free transmission spectrum may be driven by the lack of enhanced scattering opacity included in this particular grid of PICASO models. This grid only included self-consistent parameterized clouds from Virga with physically driven optical properties and particle sizes. As a result, the cloudy models within the PICASO grid that generated sufficient

scattering to match the short-wavelength slope from NIRISS/SOSS also overly suppressed the strength of molecular features at longer wavelengths. Therefore, the best-fitting model from the PICASO grid results in a model that compromises between the short and long wavelengths while fitting neither, resulting is a poorer fit, suggesting scattering particles not captured by the Virga clouds need to be introduced to explain the spectrum. On account of this, we do not consider the PICASO results in the remainder of our discussion.

The results from the ATMO grid provide a comparatively better fit ($\chi^2_\nu = 1.84$). The ATMO best-fitting suggests a metallicity of 5× solar and a solar C/O ratio, and also requires haze scattering as well as very low cloud opacity to explain the observations. The best fit ($\chi^2_\nu = 1.78$) across all grids was obtained by ScCHIMERA, which suggests a 1× solar metallicity, and solar C/O ratio. Across all resolutions, the ScCHIMERA grid finds that the best-fitting solutions are largely consistent with a solar composition, and correspond to the class of models with a power-law for the scattering hazes and an optically thick cloud deck (i.e. model 2 described above), which are strongly preferred ($\sim 12\sigma$) over the uniform vertically distributed cloud deck model (i.e. model 1). When using this model for clouds and hazes, only the scattering slope of the hazes has an impact on the spectrum and the optically thick deck cloud is placed deep in the atmosphere where it does not significantly affect the spectrum. That is, we find that the power law model for hazes best explains the spectrum, with no need for an optically thick cloud deck or inhomogeneous clouds. The median retrieved haze parameters are $\log_{10}(a) = 1.78$ and $\gamma = 4$, corresponding to enhanced Rayleigh-like scattering. The cloud deck pressure is constrained to pressures below the photosphere (e.g. $P_{\rm cloud} > 1$ bar). This will be referred to as our 'reference model' for some further tests described below. Overall, these results suggest that WASP-96 b's terminator region is mostly cloud-free at the pressures probed here with possible enhanced scattering due to small particle hazes. Furthermore, the consideration of inhomogeneous clouds and hazes does not significantly impact our results.

Assuming that both ScCHIMERA and ATMO provide an equally plausible scenario (given their similar $\chi^2_\nu$ values) for the atmospheric composition of WASP-96 b, our analysis suggests an atmospheric metallicity of 1–5× solar and a solar C/O ratio. Across both models, we find that there is a need for enhanced Rayleigh scattering to explain the blueward transit depths. However, we do not find strong evidence for optically thick clouds.

Our inferences on the metallicity and C/O ratio of WASP-96 b are enabled by the large spectral features present in the observations. The bottom panel of Fig. 10 shows the contribution to the spectrum by different molecules and atoms. The main absorber in our WASP-96 b transmission spectrum is $H_2O$, showing three clearly visible molecular bands. There are also signatures of absorption due to K near the known doublet peaks at ∼0.76 $\mu$m. The Na absorption feature seen by Nikolov et al. (2018) is not resolved in these observations, due to lack of wavelength coverage, but, as mentioned previously, we see a slope in the bluest wavelengths. Our models fit this slope with enhanced Rayleigh scattering, though we note that this could indeed be the red wing of a highly-broadened Na feature.

We therefore investigated the robustness of our inferences against this blueward slope. Given that these observations alone cannot robustly identify the nature of this slope, we aim to establish whether it has an impact on our inferred metallicity. We refit the R = 125 observations using the ScCHIMERA grid without the data shorter than 1 $\mu$m, finding a consistent inference of a 1× solar metallicity with a solar C/O ratio. We repeat this exercise for all other resolutions and find relatively consistent results, with an average best-fitting metallicity across all five model configurations with ScCHIMERA (i.e. three cloud/haze models + two inhomogeneous cloud/haze models) and the four resolutions tested (i.e. R = 125, 250, 500, pixel-level) of 2× solar with a standard deviation of 1× solar. The larger average is driven by the inferences at the pixel-level, which prefer metallicities of ∼4 × solar across most model configurations. For the same model/resolution combinations, when not considering the observations blue-wards of 1 $\mu$m, the average C/O is 0.46 with a standard deviation of 0.17; values consistent with solar expectations. This degeneracy will be further explored in the retrieval analyses of Taylor et al. (2023), but we note that jointly fitting NIRISS/SOSS transmission spectra with ground based transmission measurements, such as in this case the existing VLT/FORS2 observations, may be necessary to fully constrain Na and haze scattering properties.

Finally, some unresolved hints of carbon- (e.g. CO and $CO_2$ both with a model preference of $\sim 2\sigma$) and/or sulfur- (e.g. $H_2S$ with a model preference of $\sim 2\sigma$) bearing species may be present near the 2.5 $\mu$m feature in the spectrum (under the assumption of chemical equilibrium).

Additionally, we explored the reliability of our solar C/O ratio inference. We fixed the best-fitting parameters from ScCHIMERA to the R = 125 observations (with the exception of the C/O ratio) and investigated the impact of changing the best-fitting C/O ratio to sub-solar and super-solar values, while fixing the atmospheric metallicity and cloud/haze properties to their best-fitting values. We find that super-solar C/O ratios (i.e. 0.9) are incompatible with existing observations due to the expected signatures of carbon-bearing species, such as $CH_4$, which are not seen with NIRISS/SOSS. Sub-solar C/O ratios provide a worse fit by increasing the size of the observed $H_2O$ features. A more robust statistical inference on the atmospheric C/O ratio and metallicity will be possible with the more detailed retrieval study that follows our investigation (Taylor et al. 2023). Fig. 11 shows the results of our C/O ratio investigation and the reliability of our metallicity inferences for a partial fit to the observations explained above.

Our grid models find an atmospheric composition in good agreement with previous retrieval studies. The chemical equilibrium retrieval of McGruder et al. (2022), using the full ensemble of all pre-JWST data, found a metallicity of $Z/Z_\odot = 0.32^{+2.91}_{-0.20}$, which is consistent with our best-fitting metallicity at the $1\sigma$ level. The free- and equilibrium chemistry retrievals of Yip et al. (2021) and Nikolov et al. (2022), respectively, both point to solar-to-super-solar abundances of Na and O which, if extrapolated to a full atmosphere metallicity, are again consistent with our findings.

Moreover, previous observations of this target have pointed to an atmosphere free from optically thick clouds (e.g. Nikolov et al. 2018; Yip et al. 2021; McGruder et al. 2022; Nikolov et al. 2022) – in large part due to the pressure-broadened Na wings visible at optical wavelengths. The microphysical cloud models of Samra et al. (2023), however, predict a homogeneous cloud coverage of WASP-96 b's terminator. In Fig. 12, we show the P–T profile associated with the reference ScCHIMERA model, along with condensation curves for several prominent cloud condensates (Visscher et al. 2010; Morley et al. 2012). The shaded region shows the pressure range generally probed by the photosphere (e.g. Welbanks et al. 2019; Welbanks & Madhusudhan 2022). The best-fitting model has a P–T profile consistent with the planet's equilibrium temperature of ∼1285 K. Comparing the P–T profile to the condensation curves from (Visscher et al. 2010; Morley et al. 2012), our results do not preclude the formation of some cloud species. However, with the expanded wavelength range of

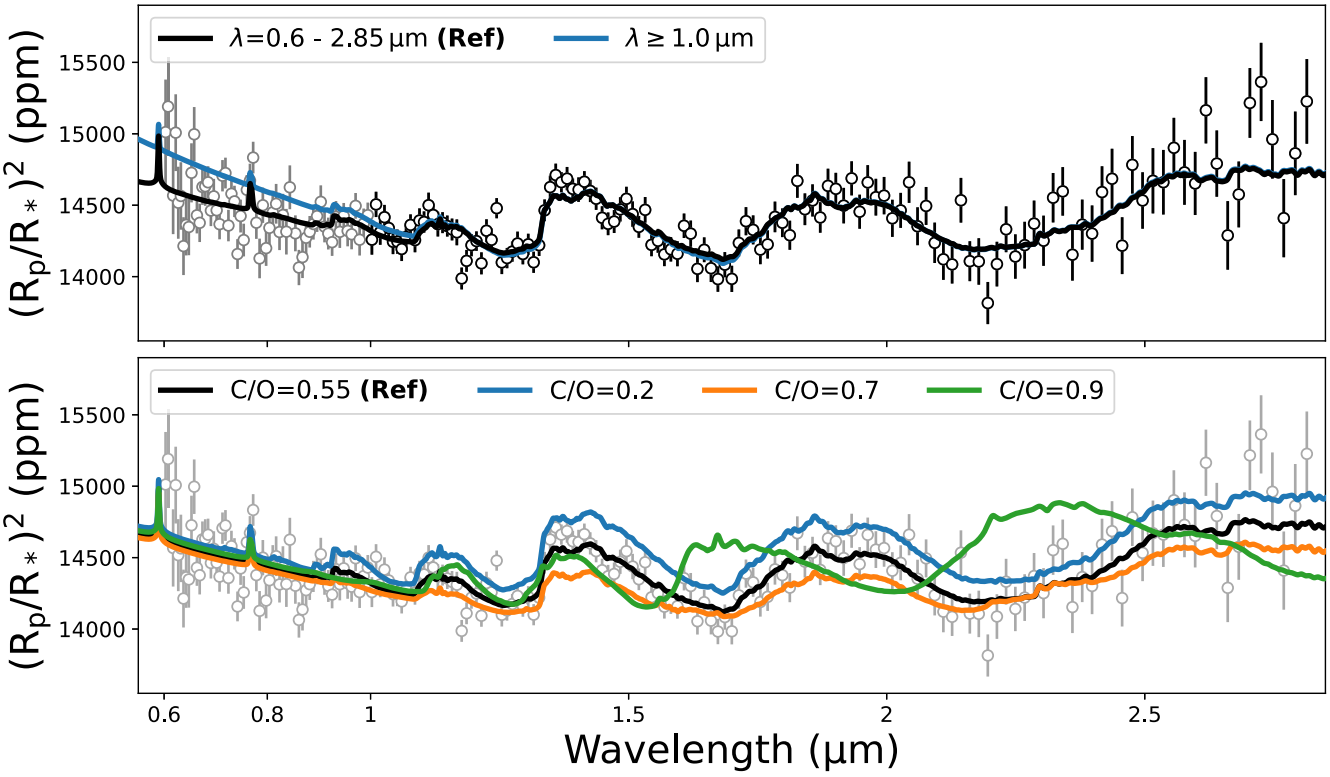


**Figure 11.** *Top*: Comparison of model fits to different wavelength ranges. A fit to the data for wavelengths $\lambda \gtrsim 1.0$ $\mu$m (blue) is compared to the reference ScCHIMERA model (black) fit to the entire wavelength range. At wavelengths $\lambda \gtrsim 1.0$ $\mu$m, the two fits are entirely consistent and yield identical C/O and metallicity, showing that the haze slope detected at shorter wavelengths is not biasing the other atmospheric inferences. *Bottom*: Investigation of the effects of C/O on the transmission spectrum. All model parameters are fixed to those from the reference ScCHIMERA model, except for the C/O ratio, which is allowed to vary. Four cases are compared here: 0.2, 0.55 (solar: best fit), 0.7, and 0.9. Super-solar C/O ratios result in large changes to the shape of the transmission spectrum (e.g. the emergence of $CH_4$, which we do not see in the data), allowing us to strongly rule out a super-solar C/O ratio for WASP-96 b.

NIRISS/SOSS, our transmission spectrum still shows no strong preference for optically thick clouds in the observable photosphere. We do, though, find a slope at the bluest wavelengths, which our models fit with an enhanced Rayleigh scattering haze, though we note that such a slope could also potentially be the red wing of the highly-broadened Na feature. Without sampling the wavelengths corresponding to the peak of the Na feature, these two scenarios are difficult to disentangle with SOSS observations alone, and jointly fitting the NIRISS/SOSS and VLT/FORS2 spectra will likely be a fruitful avenue to break this degeneracy. In this initial work, we also do not explicitly explore microphysical clouds models, which may potentially be able to help explain the blue-wavelength slope while having minimal opacity at redder wavelengths.

## 6 SUMMARY AND DISCUSSION

In this work, we have presented transmission spectroscopy observations of the hot-Saturn WASP-96 b, taken with NIRISS/SOSS as part of the JWST Early Release Observations Programme. As this is one of the very first SOSS data sets to be observed, we present a detailed walkthrough of the reductions for this new instrument, paying special attention to background subtraction and the correction of 1/*f* noise. We further suggest and implement strategies to mitigate contamination due to order 0 and order 1 traces of field stars.

We compare our transmission spectrum to grid models generated with the PICASO, ATMO, and ScCHIMERA codes. Overall, our grid fits suggest an atmosphere metallicity in the range of 1–5× solar, a solar C/O ratio, and a cloud-free upper atmosphere in the terminator region of WASP-96 b, which is in agreement with previous works using *HST* and ground-based observations. We also identify a strong slope towards bluer wavelengths, which may be either the red wing of a highly broadened Na feature or enhanced Rayleigh scattering from small particles in the atmosphere. This serves as a preliminary glimpse into the atmosphere of WASP-96 b, whose nature will be explored in more depth in Taylor et al. (2023).

The SOSS mode is slated to observe numerous exoplanet targets in Cycle 1, and will certainly continue to be a workhorse instrument, especially for observations of small, temperate worlds, throughout Cycle 2 and beyond. We therefore hope that this work will provide a useful reference to the community and aid in understanding this novel and awesome observing mode.

## SOFTWARE

(i) `astropy`; Astropy Collaboration (2013, 2018)
(ii) `batman`; Kreidberg (2015)
(iii) `dynesty`; Speagle (2020)
(iv) `ExoTiC-LD`; Laginja & Wakeford (2020)
(v) `ipython`; Pérez & Granger (2007)

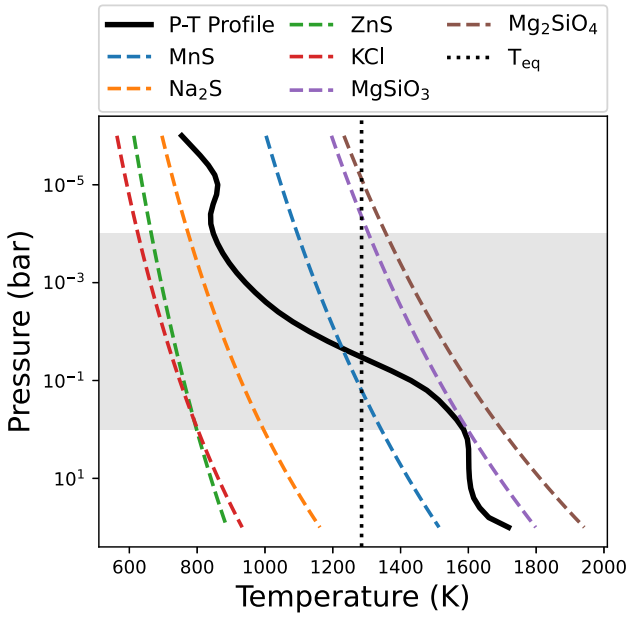


**Figure 12.** Pressure–temperature profile from the ScCHIMERA grid (black) along with condensation curves for prominent cloud species from Visscher, Lodders & Fegley (2010) and Morley et al. (2012). The shaded grey region shows the approximate pressure region of the photosphere (see e.g. Welbanks et al. 2019). Our best-fitting grid models find no evidence for an optically thick grey cloud deck in the observable photosphere of WASP-96 b despite the temperature structure not precluding conditions favourable for cloud condensation.

(vi) `juliet`; Espinoza et al. (2019)
(vii) `matplotlib`; Hunter (2007)
(viii) `numpy`; Harris et al. (2020)
(ix) `PyMultiNest`; Buchner et al. (2014)
(x) `scipy`; Virtanen et al. (2020)

## ACKNOWLEDGEMENTS

MR would like to acknowledge funding from the National Sciences and Research Council of Canada (NSERC), the Fonds de Recherche du Québec – Nature et Technologies (FRQNT), and the Institut Trottier de recherche sur les exoplanètes (iREx). He would also like to thank A, T, U, and Th for numerous helpful conversations and songs. DJ is supported by NRC Canada and by an NSERC Discovery Grant. LD acknowledges support from the Banting Postdoctoral Fellowship programme, administered by the Government of Canada. RA is a Trottier Postdoctoral Fellow and acknowledges support from the Trottier Family Foundation. LW and RJM acknowledge support provided by NASA through NASA Hubble Fellowships awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. This project was undertaken with the financial support of the Canadian Space Agency. This work is based on observations made with the NASA/ESA/CSA JWST. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for *JWST*.

LW is an NHFP Sagan Fellow, ADF is an NSF Graduate Research Fellow, RJM is an NHFP Sagan Fellow, RA is a Trottier Postdoctoral Fellow, and LD is a Banting Postdoctoral Fellow.

## DATA AVAILABILITY

All data used in this study is publicly available from the Barbara A. Mikulski Archive for Space Telescopes.[8]

## REFERENCES

Ackerman A. S., Marley M. S., 2001, ApJ, 556, 872
Ahrer E.-M. et al., 2023, Nature, 614, 653
Alam M. K. et al., 2021, ApJ, 906, L10
Albert L. et al., 2023, preprint (arXiv:2306.04572)
Alderson L. et al., 2023, Nature, 614, 664
Allard N. F., Spiegelman F., Kielkopf J. F., 2016, A&A, 589, A21
Allard N. F., Spiegelman F., Leininger T., Molliere P., 2019, A&A, 628, A120
Amundsen D. S. et al., 2016, A&A, 595, A36
Amundsen D. S., Tremblin P., Manners J., Baraffe I., Mayne N. J., 2017, A&A, 598, A97
Arcangeli J. et al., 2018, ApJ, 855, L30
Astropy Collaboration, 2013, A&A, 558, A33
Astropy Collaboration, 2018, AJ, 156, 123
Azzam A. A. A., Tennyson J., Yurchenko S. N., Naumenko O. V., 2016, MNRAS, 460, 4063
Barber R. J., Strange J. K., Hill C., Polyansky O. L., Mellau G. C., Yurchenko S. N., Tennyson J., 2014, MNRAS, 437, 1828
Barstow J. K., Aigrain S., Irwin P. G. J., Sing D. K., 2017, ApJ, 834, 50
Batalha N. E., Marley M. S., Lewis N. K., Fortney J. J., 2019, ApJ, 878, 70
Bean J. L. et al., 2018, PASP, 130, 114402
Bell T. J. et al., 2022, JOSS, 7, 4503
Benneke B., 2015, preprint (arXiv:1504.07655)
Benneke B., Seager S., 2012, ApJ, 753, 100
Benneke B., Seager S., 2013, ApJ, 778, 153
Benneke B. et al., 2019, ApJ, 887, L14
Birkmann S. M. et al., 2022, A&A, 661, A83
Boucher A. et al., 2021, AJ, 162, 233
Boucher A. et al., 2023, MNRAS, 522, 5062
Brogi M., de Kok R. J., Birkby J. L., Schwarz H., Snellen I. A. G., 2014, A&A, 565, A124
Buchner J. et al., 2014, A&A, 564, A125
Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377
Claret A., 2000, A&A, 363, 10
Coulombe L.-P. et al., 2023, preprint (arXiv:2301.08192)
Darveau-Bernier A. et al., 2022, PASP, 134, 094502
Deming D. et al., 2013, ApJ, 774, 17
Doyon R. et al., 2023, preprint (arXiv:2306.03277)
Dragomir D. et al., 2020, ApJ, 903, L6
Drummond B., Tremblin P., Baraffe I., Amundsen D. S., Mayne N. J., Venot O., Goyal J., 2016, A&A, 594, A69
Espinoza N., Kossakowski D., Brahm R., 2019, MNRAS, 490, 2262
Espinoza N. et al., 2023, PASP, 135, 018002
Evans T. M. et al., 2016, ApJ, 822, L4
Faedi F. et al., 2011, A&A, 531, A40
Feinstein A. D. et al., 2023, Nature, 614, 670
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Fortney J. J., 2005, MNRAS, 364, 649
Fortney J. J., Visscher C., Marley M. S., Hood C. E., Line M. R., Thorngren D. P., Freedman R. S., Lupu R., 2020, AJ, 160, 288
Freedman R. S., Lustig-Yaeger J., Fortney J. J., Lupu R. E., Marley M. S., Lodders K., 2014, ApJS, 214, 25
Fu G. et al., 2022, ApJL, 940, L35
Gaia Collaboration, 2021, A&A, 649, A1
Gharib-Nezhad E., Iyer A. R., Line M. R., Freedman R. S., Marley M. S., Batalha N. E., 2021, ApJS, 254, 34

[8] https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html

Gordon S., McBride B. J., 1994, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications. Part 1: Analysis Technical Report NASA-RP-1311,
Goyal J. M. et al., 2018, MNRAS, 474, 5158
Goyal J. M. et al., 2020, MNRAS, 498, 4680
Grimm S. L. et al., 2021, ApJS, 253, 30
Harris C. R. et al., 2020, Nature, 585, 357
Hellier C. et al., 2014, MNRAS, 440, 1982
Helling C. et al., 2023, A&A, 671, A122
Hoeijmakers H. J. et al., 2020, A&A, 641, A123
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Husser T.-O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, A&A, 553, A6
Iyer A. R., Line M. R., 2020, ApJ, 889, 78
JWST Transiting Exoplanet Community Early Release Science Team, 2023, Nature, 614, 649
Kipping D. M., 2013, MNRAS, 435, 2152
Kramida A., Ralchenko Y., Reader J., NIST ASD Team, 2018, NIST Atomic Spectra Database (ver. 5.6.1), National Institute of Standards and Technology, Gaithersburg, MD
Kreidberg L., 2015, PASP, 127, 1161
Kreidberg L. et al., 2015, ApJ, 814, 66
Kreidberg L. et al., 2018, AJ, 156, 17
Laginja I., Wakeford H., 2020, JOSS, 5, 2281
Lecavelier Des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008, A&A, 481, L83
Line M. R., Parmentier V., 2016, ApJ, 820, 78
Line M. R. et al., 2013, ApJ, 775, 137
Lodders K., Palme H., Gail H. P., 2009, Landolt Börnstein, 4B, 712
MacDonald R. J., Madhusudhan N., 2017, MNRAS, 469, 1979
McGruder C. D. et al., 2022, AJ, 164, 134
McKay C. P., Pollack J. B., Courtin R., 1989, Icarus, 80, 23
Madhusudhan N., 2012, ApJ, 758, 36
Madhusudhan N., 2019, ARA&A, 57, 617
Madhusudhan N., Amin M. A., Kennedy G. M., 2014, ApJ, 794, L12
Magic Z., Chiavassa A., Collet R., Asplund M., 2015, A&A, 573, A90
Mai C., Line M. R., 2019, ApJ, 883, 144
Mandel K., Agol E., 2002, ApJ, 580, L171
Mansfield M. et al., 2021, Nature Astron., 5, 1224
Marley M., 1999, PASP, 111, 1591
Marley M. S. et al., 2021, ApJ, 920, 85
Morley C. V., Fortney J. J., Marley M. S., Visscher C., Saumon D., Leggett S. K., 2012, ApJ, 756, 172
Moses J. I. et al., 2011, ApJ, 737, 15
Moses J. I., Tremblin P., Venot O., Miguel Y., 2022, Exp. Astron., 53, 279
Mukherjee S., Batalha N. E., Fortney J. J., Marley M. S., 2023, ApJ, 942, 71
Nikolov N. et al., 2018, Nature, 557, 526
Nikolov N. K. et al., 2022, MNRAS, 515, 3037
Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJ, 743, L16
Parmentier V., Showman A. P., Lian Y., 2013, A&A, 558, A91
Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Pinhas A., Madhusudhan N., Gandhi S., MacDonald R., 2019, MNRAS, 482, 1485
Piskorz D. et al., 2018, AJ, 156, 133
Polyansky O. L., Kyuberis A. A., Zobov N. F., Tennyson J., Yurchenko S. N., Lodi L., 2018, MNRAS, 480, 2597
Pontoppidan K. M. et al., 2022, ApJL, 936, L14
Radica M. et al., 2022, PASP, 134, 104502
Richard C. et al., 2012, J. Quant. Spec. Radiat. Transf., 113, 1276
Ricker G. R. et al., 2014, in Oschmann Jacobus M. J., Clampin M., Fazio G. G., MacEwen H. A.eds, Proc. SPIE Conf. Ser. Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 914320
Rigby J. et al., 2023, PASP, 135, 31
Rooney C. M., Batalha N. E., Gao P., Marley M. S., 2022, ApJ, 925, 33
Rothman L. S. et al., 2010, J. Quant. Spec. Radiat. Transf., 111, 2139
Rustamkulov Z. et al., 2023, Nature, 614, 659
Samra D., Helling C., Chubb K. L., Min M., Carone L., Schneider A. D., 2023, A&A, 669, A142
Seager S., Sasselov D. D., 2000, ApJ, 537, 916
Sing D. K., 2010, A&A, 510, A21
Sing D. K. et al., 2016, Nature, 529, 59
Spake J. J. et al., 2020, MNRAS, 500, 4042
Speagle J. S., 2020, MNRAS, 493, 3132
Stevenson K. B. et al., 2016, PASP, 128, 094401
Taylor J. et al., 2023, MNRAS, ,
Thorngren D. P., Fortney J. J., Murray-Clay R. A., Lopez E. D., 2016, ApJ, 831, 64
Toon O. B., McKay C. P., Ackerman T. P., Santhanam K., 1989, J. Geophys. Res., 94, 16287
Tremblin P., Amundsen D. S., Mourier P., Baraffe I., Chabrier G., Drummond B., Homeier D., Venot O., 2015, ApJ, 804, L17
Tsai S.-M. et al., 2023, Nature, 617, 483
Turrini D. et al., 2021, ApJ, 909, 40
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Visscher C., Lodders K., Fegley B. J., 2010, ApJ, 716, 1060
Wakeford H. R., Visscher C., Lewis N. K., Kataria T., Marley M. S., Fortney J. J., Mandell A. M., 2017, MNRAS, 464, 4247
Welbanks L., Madhusudhan N., 2019, AJ, 157, 206
Welbanks L., Madhusudhan N., 2021, ApJ, 913, 114
Welbanks L., Madhusudhan N., 2022, ApJ, 933, 79
Welbanks L., Madhusudhan N., Allard N. F., Hubeny I., Spiegelman F., Leininger T., 2019, ApJ, 887, L20
Yip K. H., Changeat Q., Edwards B., Morvan M., Chubb K. L., Tsiaras A., Waldmann I. P., Tinetti G., 2021, AJ, 161, 4
Zhou Y., Apai D., Lew B. W. P., Schneider G., 2017, AJ, 153, 243

## APPENDIX A: DETAILS OF ADDITIONAL REDUCTIONS

Here we provide details of the three independent reductions carried out on the WASP-96 b SOSS TSO using the `nirHiss` (Section A1), `transitspectroscopy` (Section A2), and `NAMELESS` (Section A3) pipelines. Although each pipeline was already described in Feinstein et al. (2023), below we provide a brief outline of each, especially noting any particular steps which differ from what was presented in that work. The final transmission spectrum for each case, along with that from the reference `supreme-SPOON` reduction described in the main text, are shown in Fig. C5.

### A1 `nirHiss`

As described in Feinstein et al. (2023), for the `nirHiss` reduction, we first process the TSOs through Stages 1 and 2 of the `Eureka` pipeline (Bell et al. 2022). From these outputs, `nirHiss` then follows three steps to remove additional background noise. First, we calculate the average scaling of the STScI JDox background model to a small region of the detector ($x \in [190, 250]$, $y \in [200, 500]$), and subtract this scaled background model from all integrations. For these data, we find the average scaling factor to be 0.448. Secondly, we use the F277W exposure, taken after the main TSO, to mitigate the effects of $0^{th}$ order contaminants which are present in the data. The F277W exposure consists of 11 integrations, and 14 groups per integration for a total exposure time of 846 s. We take the average in time of the F277W exposure and mask the trace. The background was then modelled as in Feinstein et al. (2023), and cosmic rays and bad pixels were identified as to not induce additional noise into the data. We then scale two 0th order contaminants to the TSO observations. These contaminants were located at $x_1 \in [700, 800]$, $y_1 \in [110, 160]$ and $x_2 \in [1850, 1950]$, $y_2 \in [220, 250]$. The scaling values

from each region were averaged and applied to all integrations; we find an average scaling of 2.04. Lastly, pixels with non-zero data quality flags are interpolated using the same method as Feinstein et al. (2023). Unlike Feinstein et al. (2023), after identification of the precise locations of all three diffraction orders, a simple box aperture extraction, as opposed to an optimal extraction routine, is performed on the first two orders using a width of 24 pixels.

A white light curve is constructed for both orders by summing the flux across all wavelengths (only wavelengths <0.85 $\mu$m are considered for order 2). The white light curves are then fit following the same procedure described in the main text for the `supreme-SPOON` reduction; the best-fitting parameters from the order 1 white light curve are listed in Table 1. The spectrophotometric light curves are then fit at the pixel level, again following the same procedure described in the main text.

### A2 `transitspectroscopy`

The `transitspectroscopy` reduction follows the same steps as those in Feinstein et al. (2023). We start with the `_rateints.fits` files produced by the official STScI pipeline, and use the STScI background model to subtract the zodiacal background from each SOSS integration. The background scaling for these data was found to be 0.466. 1/*f* noise is then corrected following the procedure outlined in Feinstein et al. (2023), and the stellar spectra for the first two orders are extracted using the `transitspectroscopy.spectroscopy.getSimpleSpectrum` routine and a box aperture of 30 pixels.

The white light curves for each order are fit with `juliet` using the same prior setup as described in Feinstein et al. (2023), except that the period is fixed to the 3.4252602 d from Nikolov et al. (2022). The best-fitting values from the white light-curve fit are shown in Table 1. For the spectrophotometric light-curve fits, the orbital parameters are fixed to the Nikolov et al. (2018) values and coefficients for the square-root limb-darkening law are calculated following the method described in Feinstein et al. (2023). The spectrophotometric fits are carried out at the pixel level.

### A3 `NAMELESS`

All steps of the `NAMELESS` are followed in an identical manner to those presented in Feinstein et al. (2023) except for the 1/*f* noise correction, for which we use a new method developed in Coulombe et al. (2023). This method is essentially similar in spirit to that described in 2.2.1, except it is applied at the integration level, and that instead of scaling the median image by an estimate of the white light curve to create the difference images, we allow each column to scale independently and simultaneously calculate these scaling factors with the 1/*f* noise. For a more in-depth description of the algorithm, see Coulombe et al. (2023). Stellar spectra are extracted from the corrected frames using the `transitspectroscopy.spectroscopy.getSimpleSpectrum` routine with a box width of 30 pixels.

We first fit for the white light-curves of both orders 1 and 2 separately using the ExoTEP framework (Benneke et al. 2019). We fit for the mid-transit time $t_0$, the planet-to-star radius ratio $R_p/R_*$, impact parameter $b$, semi-major axis $a/R_*$, and quadratic limb-darkening coefficients ($u_1$, $u_2$; Mandel & Agol 2002; Kreidberg 2015). We also fit for the scatter $\sigma$, as well as a linear systematics model with an offset $c$ and slope $v$. Uniform priors are considered for all parameters. For the spectrophotometric light curves, we follow the same process, but fix the impact parameter and semi-major axis to the Nikolov et al. (2018) values, and the time of mid-transit to its white light-curve best-fitting value. We fit 610 bins for order 1 and 161 for order 2.

## APPENDIX B: COMPARISON OF GROUP VS INTEGRATION LEVEL 1/F CORRECTIONS

In order to assess any potential biases introduced into a transmission (or emission) spectrum resulting from performing the 1/*f* noise correction at the group or integration level (i.e. before or after the non-linearity correction), we simulated TSO of WASP-96 b, analogous to those presented here, with the IDTSOSS simulator (Radica et al. 2022; Albert et al. 2023). The simulated TSO consisted of the same number of groups as integrations as the real TSO, and was seeded with a cloud-free, 10× solar metallicity, C/O = 0.25 atmosphere model generated with the SCARLET framework (Benneke 2015) under the assumption of chemical equilibrium. We then processed the simulated TSO through the `supreme-SPOON` pipeline in three different ways.

(i) Case 1: Group-level 1/*f* correction as described in the text.
(ii) Case 2: Correct 1/*f* at the group level. Subtract the background beforehand, but do not add it back after the `OneOverFStep`.
(iii) Case 3: Subtract the background and correct 1/*f* noise at the integration level.

These three cases allowed us to test the interplay between the background correction, 1/*f* correction, and the non-linearity. After extracting the stellar spectra for each case, we fit the spectroscopic light curves at the pixel level, fixing the orbital and limb-darkening parameters to the same values as were input to the simulation. The resulting transmission spectra, binned to a resolution of $R = 50$ are shown in Fig. C6. All three cases result in excellent agreement with the input spectrum ($\chi^2_\nu = 0.93$, 0.99, and 1.31 for Cases 1, 2, and 3, respectively), with no systematic biases resulting from either treating the background before, or the 1/*f* noise after the non-linearity correction. The integration level correction though does result in less precise transit depths (mean error bar of 144 ppm vs 193 ppm for Case 1), as well as a higher RMS residual scatter than the other two cases (139, 147, and 207 ppm for Cases 1, 2, and 3 respectively).

The importance of the non-linearity correction scales with the brightness of the target. Since these WASP-96 b TSOs remain well below the 35 000 counts threshold, it is possible that the excellent agreement of all three cases may stem from the relative unimportance of the non-linearity correction. We therefore simulated a second WASP-96 b TSO, but increased the brightness of the host star by 0.6 mag. This TSO has peak counts ∼25 000, so non-linearity effects will be more important. We processed this simulation following the three methodologies described above, and once again found little difference between the three cases. Case 1 again yielded the lowest RMS scatter, most precise transit depths, and best $\chi^2_\nu$, followed by Case 2 and then Case 3, with values for all metrics similar to those calculated for the normal brightness case. We therefore conclude that, even for significantly brighter targets, no biases result from a non-optimal treatment of the background-1/*f*-non-linearity coupling. This is likely due to the fact that, although non-linearity effects become more prominent for brighter targets, the relative importance of the background and 1/*f*-noise decrease correspondingly, and in the end the two effects effectively cancel out. We note here as well that the above discussion assumes that the SOSS non-linearity effects are perfectly characterized, which is not entirely true in reality.

## APPENDIX C: ADDITIONAL FIGURES AND TABLES

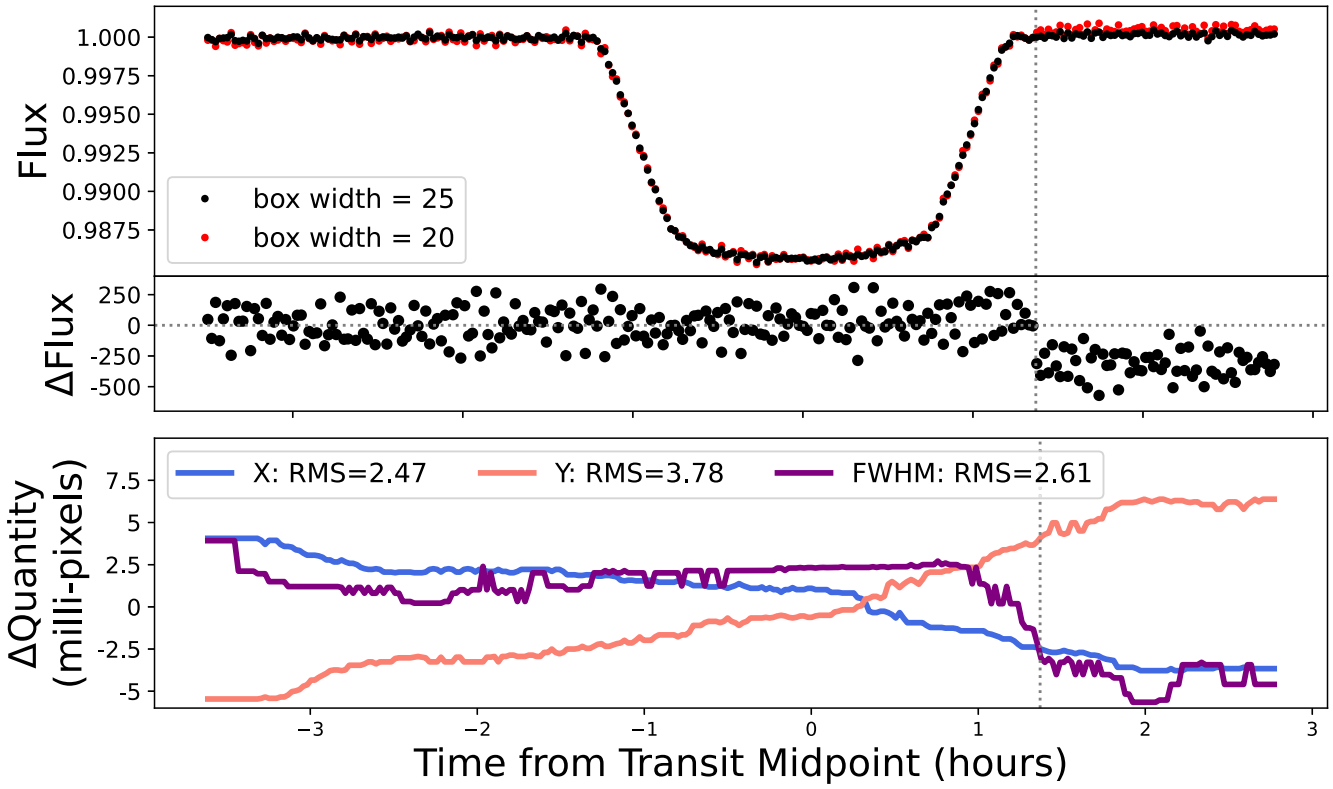


**Figure C1.** Detector level trends in the WASP-96 b SOSS TSO. *Top*: Order 1 white light curve extracted with a box aperture of 25 pixels (black) and 20 pixels (red). The light curves are identical until the tilt event ∼1.4 h after the transit midpoint (grey vertical line). *Middle*: Difference in white light flux between the 25 and 20 box aperture extractions. *Bottom*: Temporal trends in the X-position (blue), Y-position (red) and FWHM (purple) of the SOSS trace relative to the median stack through the TSO. The trace position is incredibly stable with RMS shifts in X and Y positions of <5 milli-pixels. The FWHM is also generally stable, except during the tilt event where there is an abrupt decrease of ∼6 milli-pixels.

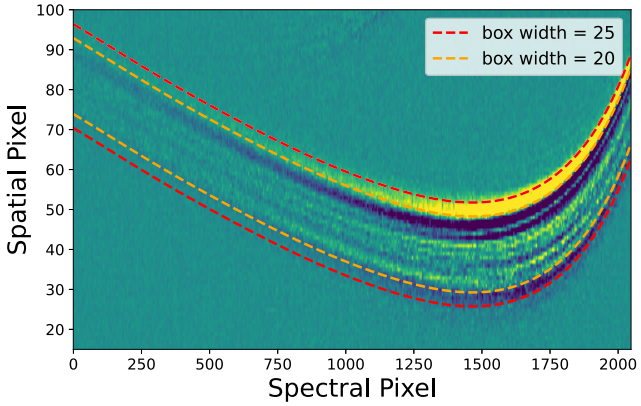


**Figure C2.** Visualization of the tilt event in 2D. Shown is the difference between a median stack of the first 50, and median stack of the last 50 integrations of the TSO, zoomed-in on the order 1 trace. The morphology change during the tilt event can be clearly seen at the upper edge of the trace. The width of a 25, as well as 20 pixel extraction box are shown in red and orange, respectively. The morphological change is entirely contained within the 25 pixel box aperture, explaining why we do not see any evidence for the tilt event in our analysis. However, when using a 20 pixel-wide box, some additional flux falls into the aperture after the tilt event, resulting in a discontinuity in the light curve (e.g. Fig. C1).

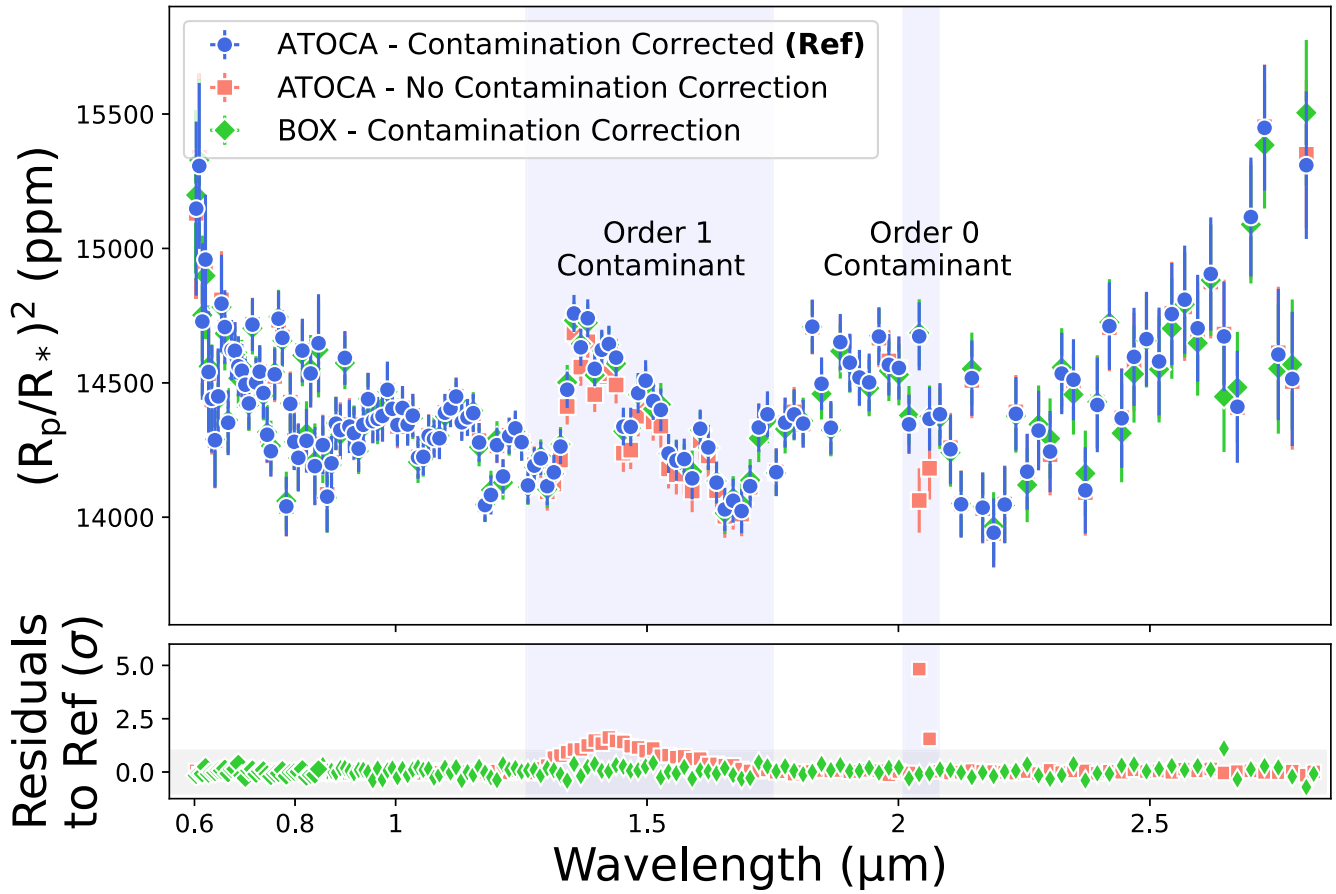


**Figure C3.** Comparison of WASP-96 b transmission spectra obtained with different methodologies. *Top*: The reference transmission spectrum extracted with ATOCA and corrected for dilution from background sources (blue) compared to an ATOCA spectrum without dilution correction (red), and a box-extracted spectrum with dilution correction (green). All spectra have been binned to R = 150 here for visual clarity. The wavelength regimes affected by each background contaminant are denoted with faded blue boxes. *Bottom*: Residuals between each spectrum shown above and the reference spectrum, normalized by the reference spectrum error bars. The $\pm 1\sigma$ range is shaded in grey.

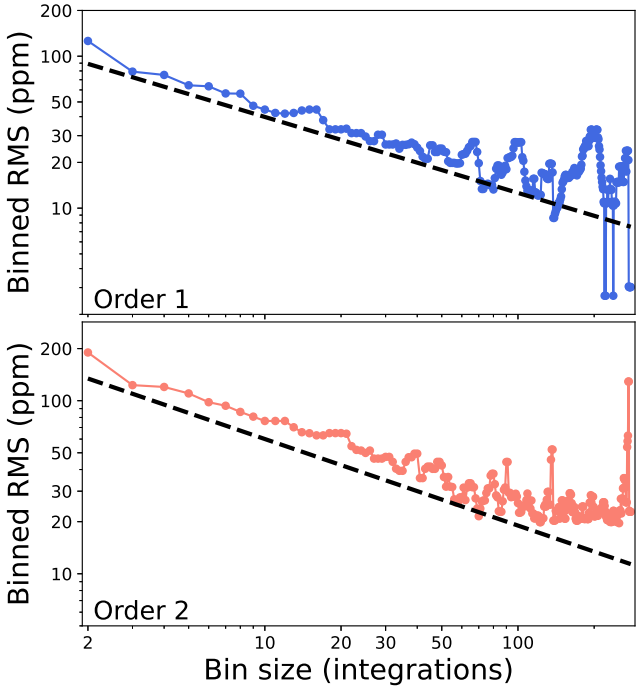


**Figure C4.** Allan variance plots for order 1 (*top*) and order 2 (*bottom*). The coloured lines in each panel are the white light-curve residuals binned to different bin widths. The black dashed lines represent the trends for pure photon noise. In general, the binned residuals trace well the pure photon noise trend.

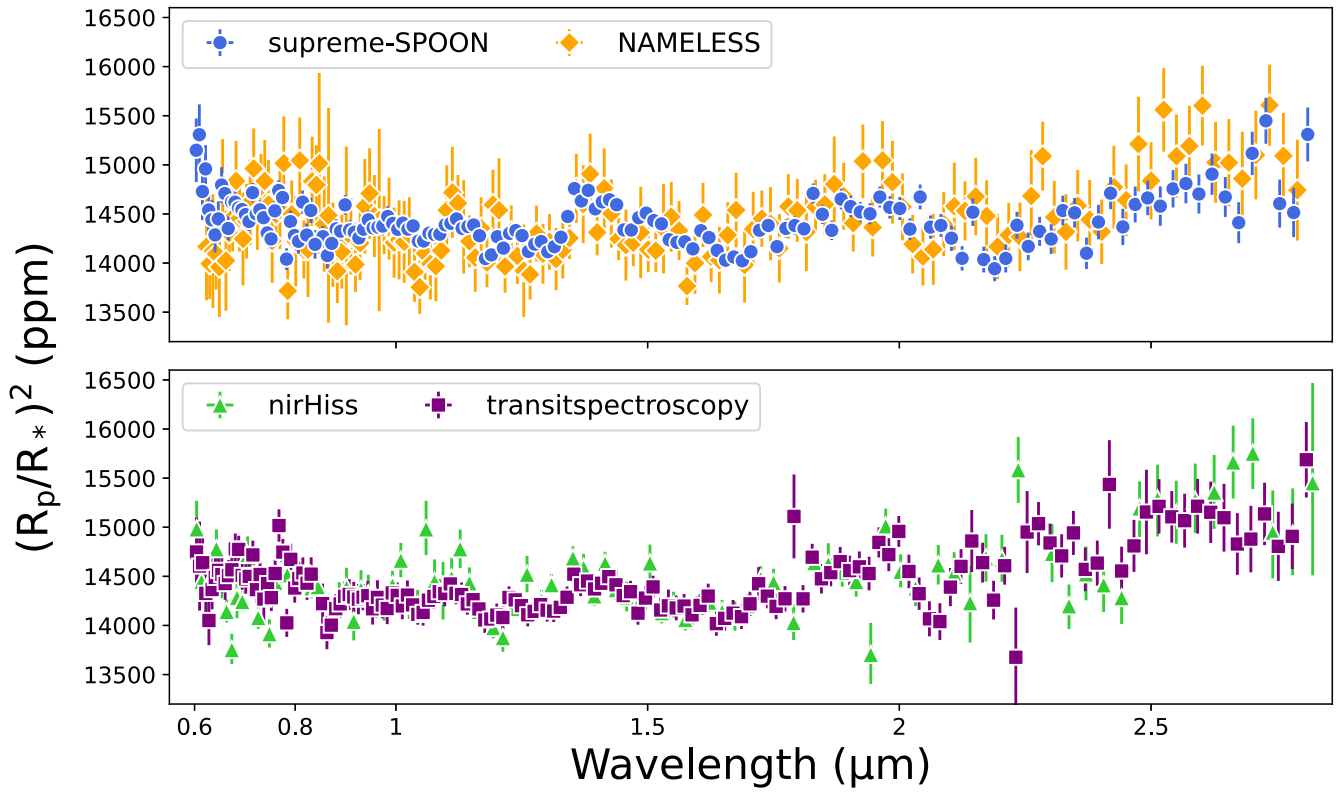


**Figure C5.** Comparison of transmission spectra for WASP-96 b obtained through four different pipelines: `supreme-SPOON` (blue), `nirHiss` (green), `transitspectroscopy` (purple), and `NAMELESS` (orange). All transmission spectra here are binned to R = 100. The four independent spectra are in good agreement, showing consistent transit depths and features across the full wavelength range of SOSS. Note that only the `supreme-SPOON` reduction is completely corrected for contamination from the background order 1 and order 0 contaminants, and thus shows a slightly larger 1.4-$\mu$m water feature.

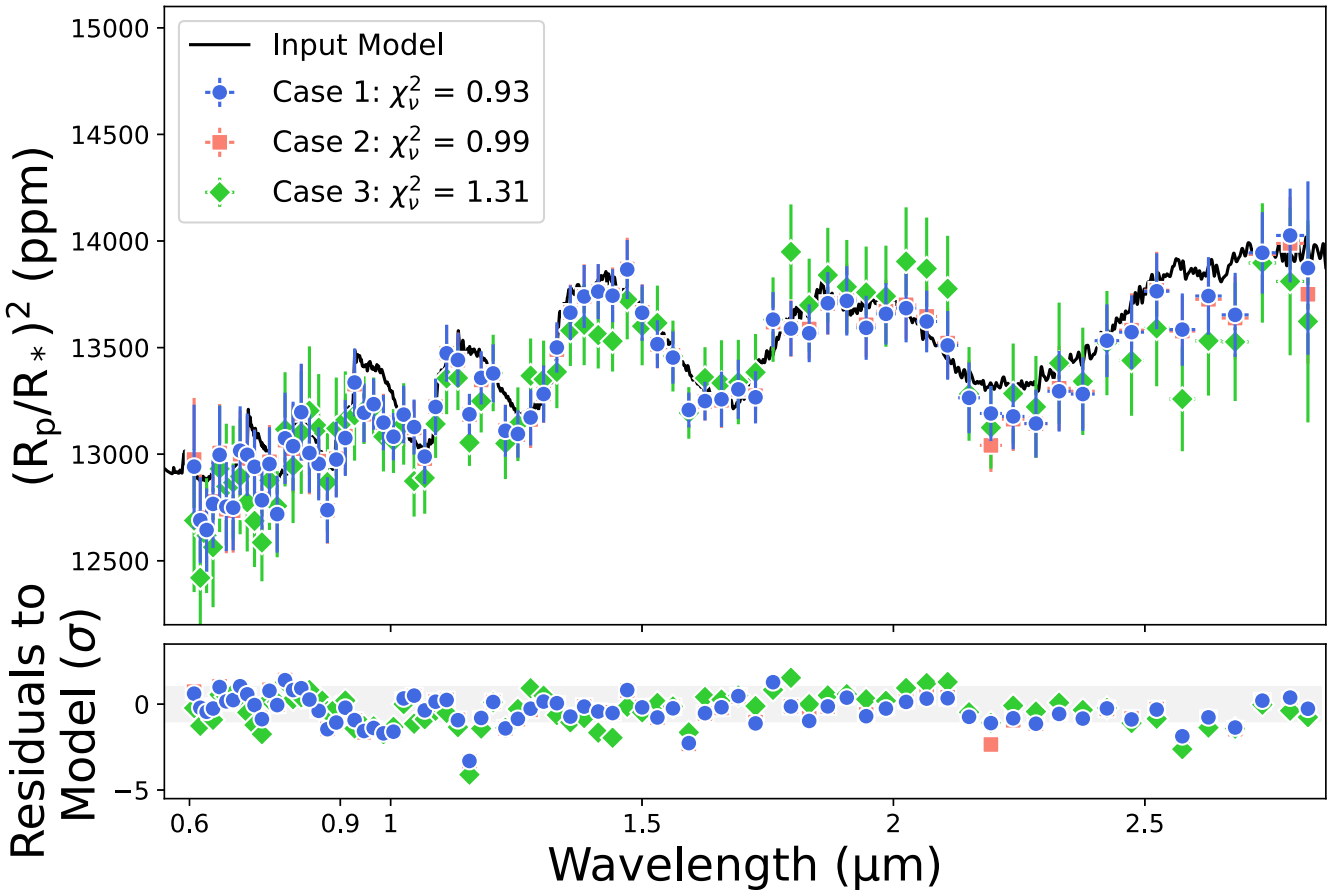


**Figure C6.** *Top*: Transmission spectra resulting from three different reductions of a simulated WASP-96 b SOSS TSO to test different 1/*f* noise correction methodologies. Case 1 represents the reduction described in the main text body, which is a group-level 1/*f* correction where the background is re-added after the 1/*f* correction is performed. Case 2 is the reduction where the background was not re-added after the 1/*f* correction, and Case 3 represents an integration-level background and 1/*f* noise correction. The input atmosphere model is shown in black. Case 1 results in the best $\chi^2_\nu$, as well as the lowest residual RMS and most precise transit depths. However, no systematic biases result from the other two cases. *Bottom*: Residuals divided by the error bar on each point between each transmission spectrum and the input atmosphere model.

[1]*Institut Trottier de recherche sur les exoplanètes and Département de Physique, Université de Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal, QC, H2V 0B3, Canada*
[2]*School of Earth and Space Exploration, Arizona State University, 781 Terrace Mall, Tempe, AZ, 85287, USA*
[3]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[4]*Department of Physics and Astronomy, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218, USA*
[5]*Department of Physics (Atmospheric, Oceanic and Planetary Physics), University of Oxford, Parks Rd, Oxford OX1 3PU, UK*
[6]*Department of Astronomy & Astrophysics, University of Chicago, 5640 S Ellis Ave, Chicago, IL 60637, USA*
[7]*School of Earth and Planetary Sciences (SEPS), National Institute of Science Education and Research (NISER), HBNI, Jatani, Odisha 752050, India*
[8]*Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95060, USA*
[9]*Department of Physics, New York University Abu Dhabi, PO Box 129188 Abu Dhabi, UAE*
[10]*Center for Astro, Particle, and Planetary Physics (CAP3), New York University Abu Dhabi, PO Box 129188 Abu Dhabi, UAE*
[11]*Department of Astronomy, University of Michigan, 1085 S. University Ave., Ann Arbor, MI 48109, USA*
[12]*Department of Physics and Astronomy, Faculty of Environment, Science and Economy, University of Exeter, Exeter, EX4 4QL, UK*
[13]*Observatoire du Mont-Mégantic, Université de Montréal, Montréal, QC, H3C 3J7, Canada*
[14]*NASA Ames Research Center, Moffett Field, CA 94035, USA*
[15]*Department of Physics, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada*
[16]*Department of Earth and Planetary Sciences, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada*
[17]*NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Rd, Victoria, BC, V9E 2E7, Canada*
[18]*Department of Physics and Astronomy, University of Victoria, Victoria, BC, V8P 5C2, Canada*
[19]*Department of Planetary Sciences and Lunar and Planetary Laboratory, University of Arizona, Tuscon, AZ 85721, USA*
[20]*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA*

This paper has been typeset from a TEX/LATEX file prepared by the author.