

A benchmark JWST near-infrared spectrum for the exoplanet WASP-39 b

Received: 7 March 2024

Accepted: 9 May 2024

Published online: 10 July 2024

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Observing exoplanets through transmission spectroscopy supplies detailed information about their atmospheric composition, physics and chemistry. Before the James Webb Space Telescope (JWST), these observations were limited to a narrow wavelength range across the near-ultraviolet to near-infrared, alongside broadband photometry at longer wavelengths. To understand more complex properties of exoplanet atmospheres, improved wavelength coverage and resolution are necessary to robustly quantify the influence of a broader range of absorbing molecular species. Here we present a combined analysis of JWST transmission spectroscopy across four different instrumental modes spanning 0.5–5.2 μm using Early Release Science observations of the Saturn-mass exoplanet WASP-39 b. Our uniform analysis constrains the orbital and stellar parameters within subpercentage precision, including matching the precision obtained by the most precise asteroseismology measurements of stellar density to date, and it further confirms the presence of Na, K, H_2O , CO, CO_2 and SO_2 as atmospheric absorbers. Through this process, we have improved the agreement between the transmission spectra of all modes, except for the NIRSpec PRISM, which is affected by partial saturation of the detector. This work provides strong evidence that uniform light curve analysis is an important aspect to ensuring reliability when comparing the high-precision transmission spectra provided by JWST.

WASP-39 b has a mass of $\sim 0.28 M_{\text{Jup}}$, an equilibrium temperature of $\sim 1,100$ K and a highly inflated radius of $\sim 1.27 R_{\text{Jup}}$ (ref. 1), making it an ideal target for transmission spectroscopy observations. Past optical and near-infrared observations of WASP-39 b with ground-based telescopes, the Hubble Space Telescope (HST) and Spitzer have found evidence for strong absorption features that are not severely affected by the muting effects of cloud extinction^{2–5}, which has been mirrored in the initial data releases for each of our James Webb Space Telescope (JWST) observations across the near-infrared^{6–10}. Furthermore, the host star WASP-39 is known to be relatively inactive^{1,7,11,12}. The atmosphere of WASP-39 b has a metal enrichment (metallicity) greater than that of its host star, although the range of metallicities that satisfy the observed JWST spectra differ between instrumental modes and can extend up to $\times 100$ solar metallicity^{7–10}. Similarly, the measured ratio of carbon and oxygen-bearing molecular species (C/O) in the atmosphere of

WASP-39 b seems to be substellar or stellar, depending on the instrumental mode. These differing compositional measurements are probably due to the preliminary modelling performed and the sensitivities of specific wavelength regions to relevant molecular tracers. Measurements of the metallicity and C/O ratio for an exoplanet are important indicators of its bulk atmospheric chemistry^{13–15} and formation history^{16–18}; therefore, these JWST observations must be analysed homogeneously so that the complementary constraining power of their different resolutions and wavelength ranges can be fully realized and the nature of WASP-39 b can be best understood.

We began our data analysis with the extracted spectral time series as reported in the initial data release publications for these observations^{7–10}. At this initial stage, the data have been corrected for both background and $1/f$ noise (additional correlated read noise due to, for example, biases in the detector readout electronics) when necessary.

✉ e-mail: aacarter@stsci.edu; erin.may@jhuapl.edu

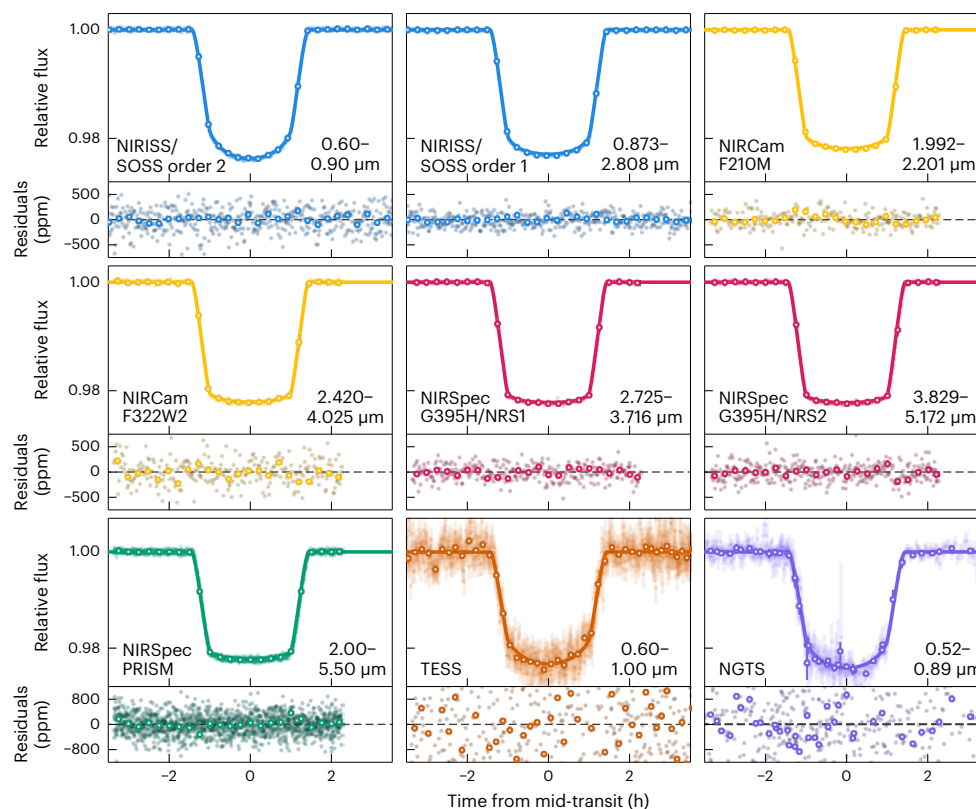


Fig. 1 | White light curves of WASP-39 b. All data are presented after correcting for systematics, including the G395H mirror tilt event⁸, with 1σ uncertainties. Solid lines indicate the best-fitting model for each of the datasets from our joint fitting analysis. Translucent points show the individual temporal measurements. Solid circles are the data after binning down to a lower temporal resolution (15 min for JWST datasets and 30 min for TESS and NGTS). The JWST light curve

data are very precise and primarily lie underneath the best-fitting model lines. For the TESS and NGTS data, all light curves are phase-folded and the displayed best-fitting model is an average across light curve fits. The residuals of the individual measurements compared to each best-fitting model are displayed underneath each light curve.

As several comparable reductions are presented in each of these publications, we selected the nominal case from each. Specifically, we chose the supreme-SPOON reduction^{9,19} for the single-object slitless spectroscopy (SOSS) mode of the near-infrared imager and slitless spectrograph (NIRISS)^{20,21}, the Eureka! reduction²² for the near-infrared camera (NIRCам) F210M + F322W2 (refs. 7,23), the ExoTIC-JEDI [V2] reduction²⁴ for near-infrared spectrograph (NIRSpect) G395H^{8,25–27} and the FIREFLY reduction²⁸ for NIRSpect PRISM^{6,10,25,26}. As WASP-39 b is the only exoplanet to date to have been observed with such a diversity of instrumental capability, this is the first opportunity for a one-to-one cross-comparison between these modes and a verification of their relative performance. We extracted seven separate white light curves from these data for our analysis using Eureka! (ref. 22): two from the separate NIRISS/SOSS orders 1 and 2 spectroscopy, one from the NIRCам F210M photometry, one from the NIRCам F322W2 spectroscopy, two from the NIRSpect G395H spectroscopy captured separately on the NRS1 and NRS2 detectors, and one from the NIRSpect PRISM spectroscopy (Methods, Extended Data Figs. 1–3 and Extended Data Table 1).

To constrain the parameters for WASP-39 b and its host star, we performed a joint fit using juliet (ref. 29) for the seven JWST white light curves, as well as a light curve from the Transiting Exoplanet Survey Satellite (TESS)³⁰, six separate light curves⁷ from the Next Generation Transit Survey (NGTS)³¹ and radial velocity measurements from CORALIE and the spectrograph for observing the phenomena of stellar interiors and exoplanets (SOPHIE)¹. All light curve fits are displayed in Fig. 1, and the best-fitting parameters are provided in Table 1. With the combined constraining power of these data, we were able to obtain exquisite constraints on the WASP-39 system. The period of WASP-39 b is constrained at subsecond precision (~ 0.3 s), with other physical and

orbital parameters constrained at subpercentage precision (~ 0.1 – 0.5%). Of particular interest is the constraint on stellar density provided by the fitting, as it is constrained to $\sim 0.3\%$, an equivalent precision to the most precise asteroseismology measurements made to date³². This is a direct consequence of the sampling of the transit events by the different JWST observations, which constrain the period and transit duration at unprecedented precision and, together with Kepler's third law, define the stellar density³³. If such precisions are common for JWST white light curves in general, then they could give rise to better constraints on orbital decay or transit timing variations. Improved stellar density measurements may improve the constraints on system ages.

Spectrophotometric light curves were extracted at the native spectral resolution from their corresponding spectral time series using the Eureka! package²², with orbital parameters fixed to the best-fitting values from fitting the white light curves. This resulted in 1,028 light curves for NIRISS/SOSS ($R \approx 350$ – $1,390$ and $\sigma_{\text{mean}} = 310$ ppm), 550 for NIRCам F322W2 ($R \approx 850$ – $1,360$ and $\sigma_{\text{mean}} = 294$ ppm), 1,163 for NIRSpect G395H ($R \approx 1,340$ – $2,630$ and $\sigma_{\text{mean}} = 496$ ppm) and 147 for NIRSpect PRISM ($R \approx 20$ – 290 and $\sigma_{\text{mean}} = 108$ ppm) for an overall total of 2,888 individual spectrophotometric light curves. As part of our analysis, we also investigated reductions at lower-resolution binning schemes and found that unless the underlying limb-darkening parameters were fixed during the light curve fitting, there were substantial wavelength-dependent variations in excess of 150 ppm between the native spectral resolution and $R = 100$ spectra (Methods and Extended Data Fig. 4). The combined transmission spectra of the JWST observations from this work, alongside those of the initial data release publications, are displayed in Fig. 2. The initial releases have different uncertainties at some locations, as they have a resolution different from the synthesized release.

Table 1 | Best-fitting orbital and instrumental parameters from fitting white light curves

Parameter	Value	Description
P	$4.0552842 \pm_{0.0000035}^{0.0}$	Orbital period (d)
a/R_s	11.390 ± 0.012	Scaled semimajor axis
b	0.4498 ± 0.0022	Impact parameter
i	87.7369 ± 0.0024	Inclination (deg)
e	0 (fixed)	Eccentricity
ρ_s	$1.6999 \pm_{0.0055}^{0.0054}$	Stellar density (g cm^{-3})
$T_{0,\text{NIRSpec/PRISM}}$	$2,459,771.335647 \pm_{0.000014}^{0.000013}$	NIRSpec PRISM mid-transit time ^a (days, BJD TDB)
$T_{0,\text{NIRCam}}$	$2,459,783.5015000 \pm_{0.0000069}^{0.0000068}$	NIRCam mid-transit time ^a (days, BJD TDB)
$T_{0,\text{NIRISS}}$	$2,459,787.5567843 \pm_{0.0000073}^{0.0000073}$	NIRISS mid-transit time ^a (days, BJD TDB)
$T_{0,\text{NIRSpec/G395H}}$	$2,459,791.6120684 \pm_{0.0000089}^{0.0000094}$	NIRSpec G395H mid-transit time ^a (days, BJD TDB)
$\sigma_{\text{NIRSpec/PRISM}}$	$218.9 \pm_{6.8}^{7.0}$	NIRSpec PRISM photometric jitter (ppm)
$\sigma_{\text{NIRCam/F322W2}}$	$235.6 \pm_{10.2}^{10.9}$	NIRCam F322W2 photometric jitter (ppm)
$\sigma_{\text{NIRCam/F210M}}$	$75.5 \pm_{11.1}^{11.3}$	NIRCam F210M photometric jitter (ppm)
$\sigma_{\text{NIRISS/order 1}}$	$110.3 \pm_{3.8}^{3.8}$	NIRISS/SOSS order 1 photometric jitter (ppm)
$\sigma_{\text{NIRISS/order 2}}$	$153.2 \pm_{7.9}^{8.1}$	NIRISS/SOSS order 2 photometric jitter (ppm)
$\sigma_{\text{NIRSpec/G395H/NRS1}}$	$138.6 \pm_{6.0}^{6.5}$	NIRSpec/G395H NRS1 photometric jitter (ppm)
$\sigma_{\text{NIRSpec/G395H/NRS2}}$	$147.5 \pm_{7.0}^{7.6}$	NIRSpec/G395H NRS2 photometric jitter (ppm)

^aA single time of transit was fitted for all datasets; we list the predicted time of transit for each instrument here based on that single fitted parameter.

The measured transmission spectra from the initial data releases display clear offsets relative to each other, whereas such offsets are reduced for the synthesized spectra after following the joint light curve fitting procedure described above (Fig. 2 and Methods). In the extreme case, the NIRSpec G395H and NIRISS/SOSS spectra have a mean offset in their overlapping region of 343 ± 16 ppm when using the initial spectra and 138 ± 16 ppm when using the synthesized spectra. This indicates that different assumptions and inferences during the light curve fitting process can substantially affect the final measured transmission spectrum, even with the constraining power of a single, highly precise, JWST white light curve. Despite these synthesizing efforts, offsets are still present between the different modes. The median offsets of the spectra from the higher-resolution modes relative to the NIRSpec PRISM spectrum are shown in Fig. 2c. Differences in the median offsets between the higher-resolution modes and NIRSpec PRISM are driven by the wavelength-dependent nature of NIRSpec PRISM systematics compounded with wavelength-independent offsets between the higher-resolution modes (Methods, Extended Data Fig. 5 and Extended Data Table 2).

Wavelength-independent offsets are commonly seen between different telescopes and instruments and are typically due to different orbital or stellar parameters assumed or different instrument sensitivities. However, an offset is present between NIRSpec PRISM and NIRSpec G395H, even though we use consistent orbital and stellar parameters and even though the NIRSpec PRISM detector is the same as the short wavelength NIRSpec G395H detector (NRS1). One potential

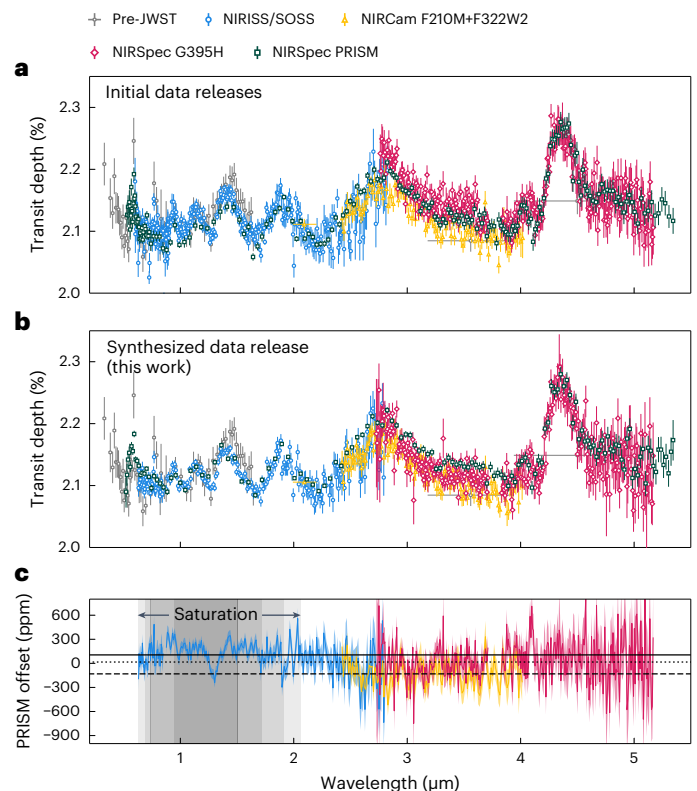


Fig. 2 | Measured transmission spectra of WASP-39. a, Measured transmission spectra for all four instrumental modes as reported in the initial data release publications^{7–10}. **b**, Measured transmission spectra at the native spectral resolution for NIRSpec PRISM and at one-fifth of the native spectral resolution for the other modes (Methods). **c**, Residuals of the synthesized data for each mode relative to a linear interpolation of the NIRSpec PRISM data, with coloured shading indicating the 1σ uncertainty bounds. Horizontal lines indicate the median difference relative to the SOSS (solid), F322W2 (dashed) and G395H (dotted) data. See Methods for a detailed quantitative discussion of the offsets between individual modes. Regions where the NIRSpec PRISM data experience detector saturation are marked in grey shading. These correspond to saturation after one (darkest) to four (lightest) groups. All displayed uncertainties correspond to 1σ .

explanation is that the low number of groups used in the NIRSpec PRISM observations may increase the influence of first-group effects and drive a more notable offset compared to the other observations, which utilize more groups. Such a shift has been observed in NIRCam transit observations³⁴ and may be present in this data also (Methods). Upon applying a shift to the NIRSpec PRISM data, it was possible to better match the other datasets across a broader wavelength range. However, a wavelength-dependent offset is still apparent from -0.6 to $2.0 \mu\text{m}$. It is across this wavelength range that the NIRSpec PRISM data are affected by detector saturation.

Despite an initial assessment that saturated data could be recovered for NIRSpec PRISM¹⁰, there is a clear discrepancy compared to the unsaturated NIRISS/SOSS data. In reality, the complex interaction of detector saturation, nonlinearity, first-group effects and pixel cross-talk modulate the measured transit depth as a function of wavelength. An investigation into determining a potential further correction to the saturated data is presented in Methods and Extended Data Figs. 6–8; however, for our model analysis, we rely solely on the NIRISS/SOSS measurements at the wavelengths impacted by saturation. For future analyses, we recommend exercising caution when inferring atmospheric properties from data that have been directly influenced by saturation. Such a conclusion is in agreement with past observations using earlier generations of infrared detectors, for which modelling

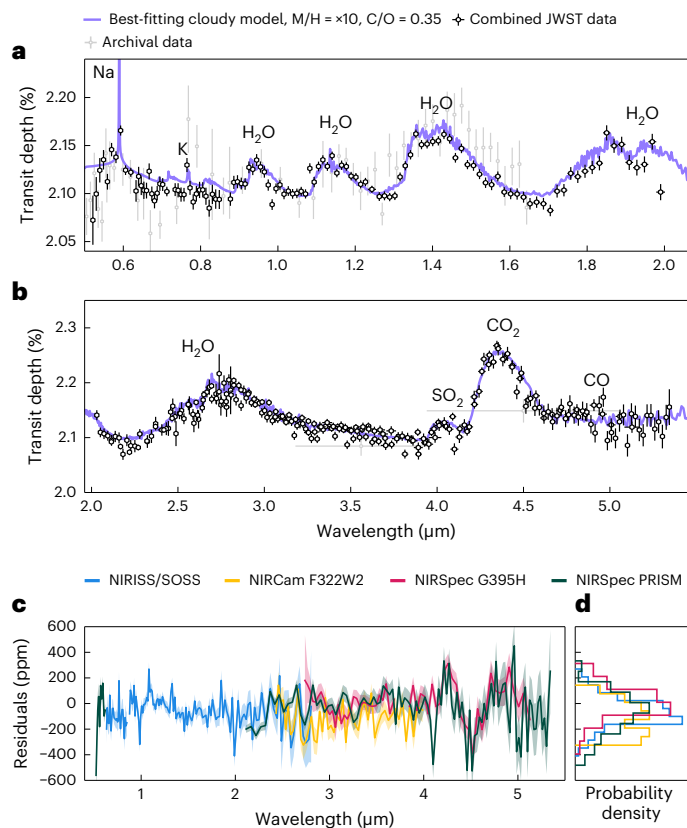


Fig. 3 | Data comparison against a 1D-RCPE model and residuals for the JWST spectrum of WASP-39 b. **a, b,** JWST transmission spectrum of WASP-39 b with 1σ uncertainties (black hexagons), split at $2.0\ \mu\text{m}$ for clarity alongside archival data (grey hexagons). **a,** $\lambda < 2.0\ \mu\text{m}$. **b,** $\lambda > 2.0\ \mu\text{m}$. For PRISM, the data are at the native spectral resolution. Data in the saturated region are excluded. For the other modes, the data are at $R = 100$. Results from the model with $\times 10$ solar metallicity and $\text{C/O} = 0.35$ are displayed in purple. **c,** Residuals of the data for each mode relative to the best-fitting model in addition to their 1σ uncertainties (shaded regions). **d,** Probability density of the residuals for each mode relative to the best-fitting model.

the detector response to saturation was difficult³⁵ so that data experiencing saturation were discarded³⁶. Even if absorption features are present across a region of partial saturation that qualitatively match the predictions of an atmospheric model, the precise structure and amplitudes of those features may not be reliable. If reliable inferences for the structure of an absorption feature are required, instead of obtaining saturated NIRSpec PRISM data, observers could consider using several higher-resolution modes. For example, NIRISS/SOSS and NIRSpec G395M/H cover a similar wavelength range as NIRSpec PRISM but saturate more slowly due to their higher resolving powers. Alternatively, the NIRSpec G140M/H mode could be utilized alongside NIRSpec PRISM, as its $0.97\text{--}1.89\ \mu\text{m}$ wavelength range spans the region of NIRSpec PRISM that has the highest throughput and is most prone to saturation but offers a ~ 2 mag improvement in brightness limit.

The combined transmission spectra exhibit a variety of spectroscopic features that can be attributed to absorption from elemental and molecular species. We investigated the origin of these spectroscopic features by comparing the observations to a self-consistent one-dimensional radiative-convective photochemical-equilibrium (1D-RCPE) model that assumes $\times 10$ solar metallicity and a subsolar carbon-to-oxygen ratio of 0.35 (Fig. 3), which is consistent with the inferred atmospheric properties from the initial data releases. We considered the possibility that inhomogeneous aerosols shaped the absorption features in our spectrum by post-processing the 1D-RCPE

model with clouds and hazes resulting from a fit to the data (Methods). Additionally, we allowed for a uniform offset to the NIRSpec PRISM data relative to the chosen model and found a median value of -177 ppm. The resulting atmospheric model confirms that the spectral features are best explained by absorption due to Na, K, H_2O , SO_2 and CO_2 . Also note the presence of a narrow absorption excess at $1.083\ \mu\text{m}$, which is evident across all explored binning schemes and may be indicative of absorption from metastable helium^{37,38}, but a further investigation is outside the scope of this paper. The spectral fit additionally confirms the need for cloud extinction, as expected from the relatively muted spectral features in the data. A more detailed analysis of these data that covers a broader variety of atmospheric modelling methodologies, explores offsets for all JWST modes and provides constraints on accessible atmospheric properties will be presented in a companion publication (Welbanks et al. in preparation, 2024).

Importantly, we found that a uniform analysis, namely joint white light curve fitting for consistent orbital parameters, resulted in an improved agreement over previously independent analyses between all JWST observing modes considered, with the exception of NIRSpec PRISM. Caution should be exercised when combining JWST spectra from different instruments without a uniform light curve analysis, particularly if those datasets have been analysed by independent teams. Until a more comprehensive limb-darkening investigation is performed, fixing limb-darkening parameters to models rather than fitting for these parameters at high resolution is necessary to improve consistency between spectra of the same dataset at different resolutions. Finally, although the included NIRSpec PRISM observations were impacted by several detector effects, particularly due to saturation (Methods), this mode remains a powerful tool for efficient characterization of planets around dimmer stars.

The continued effort to understand how best to combine data from several instruments is important for accurately characterizing exoplanet atmospheres. By combining their differing capabilities, these broad-wavelength, high-precision and high-resolution measurements will facilitate a wide range of model analyses, beginning with those presented in our companion publication (Welbanks et al. in preparation, 2024) and will greatly improve our understanding of the origins, histories and atmospheres of exoplanets. Finally, with future ultraviolet and mid-infrared transmission measurements of WASP-39 b also on the horizon (HST GO-17162 and JWST DDT-2783), we are poised to begin exploring the full potential of this new era of exoplanet characterization and the scientific advances that it can offer.

Methods

Data reduction

The data presented in this work were obtained from a selection of observations of WASP-39 b from the Panchromatic Transmission sub-programme within the JWST Transiting Exoplanet Community director's discretionary Early Release Science (ERS) programme^{39,40} (ERS 1366; PIs: N. M. Batalha, J. L. Bean and K. B. Stevenson). This includes primary transit observations with NIRISS/SOSS^{9,20,21} over 26–27 July 2022 (20:53–05:35 UT), NIRCam F210M + F322W2 (refs. 7,23) over 22–23 July 2022 (19:28–03:40 UT), NIRSpec G395H (refs. 8,25–27) over 30–31 July 2022 (21:45–06:21 UT) and NIRSpec PRISM^{6,10,25,26} on 10 July 2022 (15:05–23:39 UT). These observational modes span all three of JWST's near-infrared instruments with resolving powers of $R \approx 100\text{--}2,700$ depending on the mode and have overlapping wavelength ranges within a combined range of $0.518\text{--}5.348\ \mu\text{m}$.

The data reduction for this work began with the extracted spectral time series as presented in the initial ERS publications for these observations^{7–10}. At this stage, the data have undergone data processing steps such as detector-level corrections, ramp fitting, flat fielding, subtraction of background and $1/f$ noise, wavelength calibration and spectral extraction. For a detailed account of the precise analysis steps taken for each instrumental mode, we refer the reader to the initial ERS publications.

As the initial ERS publications provide a variety of different reductions to the data, spanning different pipelines and different methodologies, we selected just a single reduction from each for our analyses. For NIRCам F322W2, NIRSpec PRISM and NIRSpec G395H, we selected the reduction that matches that chosen in the initial publication, but for NIRISS/SOSS, we adopted a different reduction due to improvements in the out-of-transit baseline scatter. Specifically, this corresponds to the Eureka! reduction²² for NIRCам F322W2, the supreme-SPOON reduction^{9,19} for NIRISS/SOSS, the ExoTIC-JEDI[V2] reduction²⁴ for NIRSpec G395H and the FIREFLY reduction²⁸ for NIRSpec PRISM. Note that although a single reduction pipeline may be desirable, fundamentally different reduction procedures are required between instrumental modes, and a ‘jack of all trades’ pipeline will not necessarily produce the most consistent results. Further, analyses from the initial data releases^{7–10} demonstrated that different pipelines can reach good agreement to the resulting spectra. Median out-of-transit stellar spectra for each of the selected reductions are displayed in Extended Data Fig. 1. For the NIRCам F210M photometry, we did not repeat any data reduction procedures and adopted the existing extracted light curve⁷. Extended Data Table 1 gives an overview of the JWST observations included in this work.

White light curve analysis

Seven separate white light curves were obtained from these JWST datasets: one from NIRSpec PRISM, two from NIRSpec G395H (one from each detector), two from NIRISS/SOSS (one for each order), one from NIRCам F322W2 and one from NIRCам F210M photometry. For the spectroscopic observations, white light curves were constructed using Eureka!²² across similar wavelength ranges to those adopted in the initial ERS publications^{7–10}, except for NIRSpec PRISM, for which we excluded wavelengths below 2 μm due to saturation. This corresponds to wavelength ranges 0.873–2.808 μm for NIRISS/SOSS order 1, 0.6–0.9 μm for NIRISS/SOSS order 2, 2.420–4.025 μm for NIRCам F322W2, 2.725–3.716 μm for NIRSpec G395H NRS1, 3.829–5.172 μm for NIRSpec G395H NRS2 and 2.0–5.5 μm for NIRSpec PRISM.

We performed a joint fit to these light curves, in conjunction with six NGTS light curves, one TESS light curve (with three transits) and radial velocity measurements from CORALIE and SOPHIE¹; these auxiliary datasets were selected as they were readily available, did not show strong systematic effects and provided timestamps so that we were able to join them with our JWST measurements. The radial velocities were mean-subtracted before we performed any fitting. Dilution factors for each light curve were fixed to 1, implying that we assumed that there was no dilution from nearby contaminants on those light curves. High-contrast imaging observations of WASP-39 b reveal no nearby companions^{41,42}, and no contaminating sources are apparent in the NIRCам target acquisition image. In addition, the nominal joint fit presented and used in this work had eccentricity fixed to 0 (a fit leaving the eccentricity as a free parameter such that the priors described below constrain it to $e < 0.039$ with 99% credibility; the rest of the posterior parameters being consistent at 1σ with the ones here presented). All timestamps were converted to a barycentric Julian date (BJD) in Barycentric Dynamical Time (TDB).

The free parameters in this fit included:

- (1) The period, which had a normal prior distribution centred at the value reported in Maciejewski et al.⁴³, that is, 4.0552765 d but with a substantially larger standard deviation of 1 min to allow for possible timestamp mismatches between different BJD standards in the literature (for example, UTC or TDB⁴⁴)
- (2) The time of transit centre, which had a normal prior centred at 2,459,791.615201 BJD TDB (the time of the NIRSpec G395H observations) with a relatively large standard deviation of 0.1 d
- (3) The impact parameter, also centred at the value reported in Maciejewski et al.⁴³, that is, 0.45, but with a truncated normal

distribution between 0 and 1, and with a larger standard deviation of 0.1

- (4) The stellar density, whose prior was set to a log-uniform distribution between 0.1 and 10 g cm^{-3}
- (5) The radial velocity semi-amplitude, which had a uniform prior between 0 and 200 m s^{-1}
- (6) An individual radial velocity offset for each radial velocity instrument with data in Faedi et al.¹ (CORALIE and SOPHIE), with uniform priors between -100 and 100 m s^{-1} as well as jitter terms with log-uniform priors between 1 and 100 m s^{-1} for each
- (7) An individual planet-to-star radius ratio for each light curve, which had a uniform prior between 0 and 0.3 to account for possible wavelength-dependent planet-to-star radius ratio changes
- (8) The limb-darkening coefficients using a transformed quadratic law based on the uninformative sampling prescription of Kipping et al.⁴⁵, which implied two parameters per light curve, q_1 and q_2 , with uniform distributions between 0 and 1
- (9) A flux normalization term for each light curve, set with a normal prior centred at 0 and with a standard deviation of 100,000 ppm
- (10) A jitter term per light curve, set to a log-uniform prior between 0.1 and 10,000 ppm

To handle instrumental systematics in the light curves, based on analyses performed on the out-of-transit data, we decided to use the following models:

- (1) A Gaussian process on the NIRCам F322W2, NIRCам F210M photometry, NIRISS/SOSS orders 1 and 2, and TESS data. We chose a Matérn 3/2 kernel and used time as the only regressor. The prior on the amplitude of this Gaussian process was set with a log-uniform distribution from 0.01 to 100 ppm for the JWST light curves and from 0.001 to 100 ppm for the TESS light curves. The timescale also had a log-uniform prior distribution between 0.01 and 100 d. The Bayesian evidence suggests that adding a Gaussian process to the NIRSpec datasets did not provide an improvement to the fits.
- (2) A linear model for the NIRSpec G395H data with two regressors: a simple slope in time and a regressor that was 0 before the tilt event observed in the data⁸ and 1 after it.
- (3) A linear model for NIRSpec PRISM data with a simple slope in time.

Adding similar systematic models for the NGTS data did not change the results of our fit. In total, 84 free parameters were used to fit 12,206 data points, for which we used the dynamic nested sampling scheme as implemented in dynesty (ref. 46). Some of the resulting parameters from this joint fit are presented in Table 1. Posteriors for selected parameters are shown in Extended Data Fig. 2.

Wavelength binning scheme for spectrophotometric light curve extraction

To extract the spectrophotometric light curves it is necessary to define a wavelength binning scheme for each dataset. The largest number of spectrophotometric bins, and therefore the highest resolution, is reached by binning at the native pixel resolution. A spectrophotometric light curve is then extracted for each individual pixel column. However, for these JWST modes, the native pixel resolution was higher than the native spectral resolution, which defines the difference in wavelengths $\Delta\lambda$ that can be resolved at a given wavelength λ . As such, we adopted the native spectral resolution as a fundamental baseline when extracting spectrophotometric channels. Although higher-resolution schemes, including native pixel resolution, may theoretically offer access to narrower spectral features, understanding the potential and reliability of such an approach is beyond the scope of this work.

The native spectral resolving power, which defines the native spectral resolution, can be determined in units of pixels for each mode following $R_{\text{pix}} = \lambda/DR$, where λ is the wavelength, D is the dispersion of the instrumental mode and R is the spectral resolving power of the instrumental mode. For all modes, we took the dispersion and resolving power curves from the reference data files provided by the JWST exposure time calculator, Pandeia⁴⁷. Importantly, R_{pix} is a continuous function of wavelength, whereas individual pixel columns have discrete edges and cannot be meaningfully subdivided in wavelength. Therefore, we converted R_{pix} to integer pixel values using a ceiling function to ensure that pixel columns were not split across two separate wavelength bins. We defined the bin edges beginning with the lowest wavelength pixel column, where the wavelength at the lower edge of this column corresponds to the lower edge of the first spectrophotometric bin. Pixel columns were then added to this bin until the number of columns was equal to R_{pix} and the wavelength at the upper edge of the final column corresponds to the upper edge of the first spectrophotometric bin. This process was repeated using the previously determined upper edge as the starting lower edge for the next bin until all bin edges had been defined. In the event that there was a transition in the integer value of R_{pix} as pixel columns were added to a bin, the highest value of R_{pix} was used to define the number of pixel columns that must be included. If not enough pixels were available in the uppermost wavelength bin to satisfy this requirement, then those pixels were instead incorporated into the penultimate bin.

For each instrumental mode, we first extracted spectrophotometric light curves at the native spectral resolution following this binning scheme across a subset of the full wavelength range using the values adopted in the initial ERS publications for each of the instrumental modes^{7–10}. This corresponds to ranges of 0.873–2.808 μm for NIRISS/SOSS order 1, 0.630–0.853 μm for NIRISS/SOSS order 2, 2.420–4.025 μm for NIRCам F322W2, 2.725–3.716 μm for NIRSpec G395H NRS1, 3.829–5.172 μm for NIRSpec G395H NRS2 and 0.518–5.348 μm for NIRSpec PRISM. Additionally, for NIRISS/SOSS, we excluded 100 columns that were impacted by zeroth-order contamination from background sources⁹. The effect of resolution on the measured spectrophotometric transit depths is explored further below.

Spectrophotometric light curve fitting

Across all instrumental modes, we fitted the spectrophotometric light curves using the Eureka! package²², which jointly fits both a systematic and astrophysical model component to each of the light curves. The systematic model consists of a first-order polynomial in time, whereas the astrophysical transit models were computed using the batman package⁴⁸. We also fitted a step function to the NIRSpec G395H data to account for the flux drop close to mid-transit in the uncorrected light curves, which was driven by a mirror tilt event⁸. The orbital parameters were fixed during the fitting process using the values obtained from fitting the white light curve, as shown in Table 1. Limb-darkening was incorporated using a quadratic law, and the limb-darkening parameters were fixed in each of the light curve fits with initial values taken from the ExoTIC-LD package⁴⁹ using stellar parameters $[M/H] = 0.0$, $T_{\text{eff}} = 5,512$ and $\log(g) = 4.7$. Fitting for the limb-darkening parameters can produce wavelength-dependent biases as a function of wavelength binning resolution, which is investigated further below. The fitting itself was performed using a Markov chain Monte Carlo approach, as implemented by the emcee package⁵⁰, using 200 walkers and 1,100 steps and discarding the first 100 steps as a burn-in. The convergence was checked to ensure that the chains ran for at least $\times 50$ the auto-correlation time.

The transit-depth precision, native spectral resolving power and wavelength coverage resulting from the spectroscopic light curve fitting are displayed alongside similar properties for archival HST, Very Large Telescope and Spitzer data^{3,4} for WASP-39 b in Extended Data Fig. 4. Of the used JWST observational modes, NIRSpec PRISM provided

the best transit-depth precision at all wavelengths, at the expense of greatly reduced resolution. Where the wavelength ranges of the higher-resolution modes overlapped, NIRCам F322W2 provided the best precision from 2.4–2.9 μm at a slightly lower resolution, as well as providing unique access from 3.7 to 3.8 μm where NIRSpec G395H has no sensitivity due to the gap between the detectors that the spectrum falls across. The NIRISS/SOSS and NIRSpec G395H modes have similar resolutions in the narrow region where they overlap, with the NIRISS/SOSS providing superior precision below 2.75 μm . NIRSpec G395H has a similar precision to NIRCам F322W2 from 3.0 to 3.5 μm despite having a factor of 2 higher resolution due to it also having a factor of 2 higher throughput.

It is clear that JWST provides a dramatic improvement on previous capabilities for characterizing transiting exoplanet atmospheres, as it offers increased wavelength coverage, resolution and precision. The NIRISS/SOSS observations provide superior resolution to existing HST infrared data at 1–3 times higher transit-depth precision, and the NIRSpec PRISM offers a similar resolution at up to 8 times higher precision. At longer wavelengths, all four instruments provide unrivalled advantages, and the Spitzer photometry has been superseded by the spectroscopic capabilities of NIRSpec PRISM, NIRCам F322W2 and NIRSpec G395H. At native spectral resolution, the transit-depth precision of NIRCам F322W2 and NIRSpec G395H are a factor of 2–3 times lower than Spitzer, but offer over two magnitudes of improvement in resolving power compared to the Spitzer band-passes. Furthermore, NIRSpec PRISM, in addition to the $R = 100$ NIRCам F322W2 and NIRSpec G395H datasets, offers both 2–3 times higher transit-depth precision and 20–40 times improved resolving power. Nevertheless, ground-based telescopes and HST remain uniquely capable of accessing shorter wavelengths <0.5–0.6 μm , a wavelength range that is crucial for capturing and measuring the presence and strength of aerosol scattering and metal absorption lines⁵¹.

Wavelength binning investigation and a dependence on limb-darkening

We also investigated and extracted transit spectra at coarser wavelength binning schemes following the procedure above at two to five times lower than native spectral resolution for all modes, as well as a $R = 100$ binning scheme for the NIRISS/SOSS, NIRCам F322W2 and NIRSpec G395H modes. We then binned the original native spectral resolution transmission spectra to an approximately similar resolution as each lower-resolution spectrum to explore the prevalence and extent of resolution-dependent offsets. This was performed by taking the weighted mean of the native spectral resolution transit depths within each wavelength bin of the lower-resolution spectrum. As each bin of the native spectral resolution spectrum constituted several pixel columns, the wavelength ranges of the binned native resolution transit spectrum could differ slightly from those binned to a lower resolution before fitting the light curve. However, the focus of this investigation was to identify broad deviations between different resolutions. A more detailed examination would require a future analysis at the native pixel resolution. Residuals from each comparison, both when fitting for or fixing the quadratic limb-darkening parameters, are shown in Extended Data Fig. 3.

When the limb-darkening parameters were free parameters in the fitting process, we saw notable differences between the binned native resolution spectra and those that were binned before fitting the light curve. Both NIRISS/SOSS and NIRSpec G395H exhibited broad wavelength-dependent offsets that became more pronounced towards lower resolutions. Specifically, at $R = 100$, NIRISS/SOSS had a mean difference of 32 ± 14 ppm (123 ± 58 ppm above 2.2 μm) and NIRSpec G395H had a mean difference of 110 ± 29 ppm (181 ± 51 ppm above 4.5 μm). NIRCам F322W2 exhibited a broad uniform offset at all wavelengths, with a mean difference of 58 ± 19 ppm at $R = 100$. Conversely, NIRSpec PRISM exhibited a nonsignificant mean difference of 3 ± 14 ppm at

one-fifth of the native spectral resolution. We re-emphasize that the wavelength ranges of the binned native resolution transit spectrum could differ slightly from those binned to a lower resolution before fitting the light curve. This is a probable driver of any observed narrow offset features, which are not explored in this work.

When the limb-darkening parameters were fixed during the fitting process, the agreements between the different resolution spectra were drastically improved. Nonsignificant mean transit-depth differences were exhibited by NIRISS/SOSS (3 ± 13 ppm but 8 ± 53 ppm above $2.2 \mu\text{m}$), NIRCам F322W2 (13 ± 17 ppm) and NIRSpec PRISM (-3 ± 12 ppm). In contrast, a wavelength-dependent offset remained for NIRSpec G395H, although its mean offset of 37 ± 27 ppm (87 ± 48 ppm above $4.5 \mu\text{m}$) was still reduced compared to when fitting for the limb-darkening parameters.

This stark difference in behaviour when fitting or fixing the limb-darkening parameters is indicative of underlying biases in the adopted limb-darkening model. These biases appear to be strongest at regions of lower received detector counts, as evidenced by the offsets at the ends of the NIRSpec G395H and NIRISS/SOSS data compared to the NIRCам F322W2 data (which had a relatively flat throughput and lower detector counts compared to other modes). Furthermore, these biases seem to be evident only for the higher-resolution modes and not the lower-resolution NIRSpec PRISM mode. As the native spectral resolution light curves have a lower signal-to-noise ratio than those that were binned to lower resolution before fitting the light curve, it is probable that they were more susceptible to biases introduced when fitting the limb-darkening parameters. As these biases are not necessarily Gaussian in nature (they may not be comparable to additional random noise), consistent results are not seen between the binned native resolution spectra and those that were binned before fitting the light curve.

For future model analyses of these data, we recommend using the native spectral resolution spectrum for NIRSpec PRISM, the lower resolution $R = 100$ spectra for the higher-resolution modes and, in all cases, the spectra that had their limb-darkening parameters fixed during the fitting process. We emphasize that this is not a global recommendation for all JWST datasets, but one that is specific to the current best understanding of these data. Fixing the limb-darkening parameters provided greater agreement across different resolution binning schemes. However, the underlying reality is that these datasets are now uniformly biased by our limb-darkening assumptions. The extent of such biases is difficult to estimate at present, and substantial future work will be required to explore the impact of different limb-darkening approaches on these data and those from other JWST observations.

Comparison of the wavelength overlap between instruments

Each of the four instrumental modes has an overlap between its wavelength coverage and the wavelength coverage of the other three modes, allowing for a comparison between their relative measurements of the transit depth. The broadest comparison comes from the NIRSpec PRISM mode, for which the wavelength coverage completely encompasses the coverage of the other modes but at substantially lower resolution. As already shown in Fig. 2, the NIRSpec PRISM data exhibit both a wavelength-independent offset across all wavelengths and a wavelength-dependent offset for data that experience saturation. Upon application of a -177 ppm uniform offset, as determined from the model analysis, the mean offset of NIRSpec PRISM was -124 ± 6 ppm relative to NIRISS/SOSS, 132 ± 13 ppm relative to NIRCам F322W2 and 17 ± 11 ppm relative to NIRSpec G395H. Although there was better agreement for NIRISS/SOSS before offsetting the NIRSpec PRISM spectra, its wavelength range overlaps heavily with the saturated region of NIRSpec PRISM, which is not completely reliable. When looking at wavelengths unaffected by saturation, the mean offset of NIRSpec PRISM was -10 ± 27 ppm relative to NIRISS/SOSS. There are more notable discrepancies at localized regions of the wavelength coverage.

In some cases (for example, the deviation at $2.6 \mu\text{m}$), the difference can be attributed to the lower resolution of PRISM acting to 'smooth over' atmospheric features that can be better captured at higher resolution.

Equivalent comparisons to Fig. 2c for the smaller wavelength overlaps between the higher-resolution NIRISS/SOSS, NIRCам F322W2 and NIRSpec G395H modes are displayed in Extended Data Fig. 5. We found excellent agreement between NIRISS/SOSS and NIRCам F322W2, with a nonsignificant mean offset of 11 ± 49 ppm, compared to 32 ± 46 ppm for the initial data release spectra. An offset was still present between NIRISS/SOSS and NIRSpec G395H of -372 ± 170 ppm, compared to -482 ± 132 ppm for the initial data release spectra. However, this wavelength range is at the edge of the NIRSpec G395H and may have been more substantially affected by systematic effects due to low throughput. An offset was also present between NIRCам F322W2 and NIRSpec G395H of -138 ± 16 ppm, but this is greatly diminished compared to -343 ± 16 ppm for the initial data release spectra. Furthermore, the distribution of residuals for NIRCам F322W2 versus NIRSpec G395H is close to the expected normal distribution but with a uniform offset (Extended Data Fig. 5). Given the agreement between NIRISS/SOSS and NIRCам F322W2, this probably suggests that a wavelength-independent bias remains in the NIRSpec G395H spectrum, even after performing a joint white light curve analysis.

All offsets as measured are presented in Extended Data Table 2 for ease of comparison. However, these values should not be interpreted as a generalizable property of the different detectors between the different instrumental modes. We predict that these offsets will be dependent on a currently unpredictable number of variables and will probably change between a given planet and observation. Instead, we emphasize that although offsets between JWST spectra have been identified, they can be mitigated through a uniform light curve analysis.

In totality, these comparisons demonstrate that the joint white light curve analysis has dramatically improved the agreement between these various JWST datasets. Nevertheless, this improvement is not perfect, and some offsets do remain between datasets. Notably, these offsets are only present relative to either the NIRSpec PRISM or NIRSpec G395H modes. This may be an early indication of an uncorrected systematic issue or bias specific to the NIRSpec instrument, especially considering the excellent agreement between NIRISS/SOSS and NIRCам F322W2. However, it is also possible that this agreement is a coincidence, and a firmer conclusion will require similar analyses across a wider range of JWST datasets. Where offsets are still present and saturation is not present, they appear to be close to normally distributed. Until an investigation even more detailed than that presented in this work is completed, the application of uniform offsets during model fitting and interpretation may be necessary.

NIRSpec PRISM saturation

With a 2MASS J magnitude of 10.66 (ref. 1), WASP-39 is above the brightness limit of the NIRSpec PRISM mode and produces detector saturation in the brightest pixel of the columns corresponding to $\sim 0.63\text{--}2.06 \mu\text{m}$. As the JWST detectors make use of nondestructive measurements to estimate the received flux (up-the-ramp sampling), if saturation occurs in a pixel after a large number of groups have been measured, then the flux of that pixel may still be reliably measured with a sufficient number of unsaturated groups. However, for these NIRSpec PRISM observations, there are only five groups per integration, with saturation occurring as early as the second group at the brightest part of the spectra. With so few measurements in each of these ramps, the ability to fit a slope and accurately estimate the flux for these pixels is diminished. When few groups are available, the linearity of the ramps is crucial for ensuring an accurate determination of the count rates.

Extended Data Fig. 6 demonstrates the differences between counts in neighbouring groups as a diagnostic of the true linearity of the ramps. We see that the regions of the detector with higher count rates demonstrate a pattern that is indicative of an unexpectedly low count

rate in the first group (2-1 is higher than 3-2, which suggests a similar effect to the first-group effect seen in NIRC*am* defocused imaging observations³⁴) or is potentially an insufficient nonlinearity correction (later group differences are lower than earlier group differences). Conversely, we see that rows 14 and 16 demonstrate a pattern that is indicative of a high count rate in the first group (2-1 is lower than 3-2) or is potentially due to pixel cross-talk or charge diffusion as the central row approaches saturation in later groups (higher count differences in later groups than earlier groups). Although we applied a more stringent saturation threshold in our updated PRISM analysis based on our analysis of the shapes of the ramps to avoid uncorrected nonlinearity (approximately 70–75% rather than the 80% full-well threshold used in the initial data release¹⁰), future work should more closely explore the accuracy of the NIRSpec nonlinearity correction with data that saturate more slowly and, therefore, contain many more groups to better determine the shape of the ramp and first-group impacts on partial saturation corrections. In particular, it would be useful to have several datasets of stars that saturate at different rates, in combination with a NIRISS/SOSS observation, to fully characterize the impact of nonlinearity, pixel cross-talk and first-group effects on partially saturated exoplanet transit data. Such an analysis will have important implications on the full-well threshold that is appropriate when attempting to recover a partially saturated region on the NIRSpec detector and could help determine how and when cross-talk occurs between neighbouring pixels. Note that regions of the detector that do not approach saturation in our data have flat group differences, which provides confidence in the extracted count rates in this region of the data.

The analysis of the differences between neighbouring groups makes it clear that the measured stellar flux rate is probably not representative of the true stellar flux rate in regions of the detector that rapidly approach saturation. The combination of the two regimes in Extended Data Fig. 6 together shape the measured count rates in the saturated region of the spectrum. As more groups are added, the impact of inaccurate group counts becomes less important; however, for the saturated region, we are limited in the number of groups available. In particular, where a low first group dominates and few groups are available, the measured ramps are steeper than reality, corresponding to a higher extracted flux and a diluted transit depth. As shown in Fig. 2, this is exactly what is seen in the saturated region of the offset NIRSpec PRISM transmission spectrum compared to the NIRISS/SOSS spectrum. To correct for this effect, we analysed how adding groups impacts the extracted spectrum to estimate the amount of excess flux measured within the saturated region. These excess flux measurements correspond directly to a dilution correction we can apply to the NIRSpec PRISM saturated region.

To explore how adding groups impacts the extracted spectrum, we reduced the data using the same number of groups across the entire wavelength range while masking the area that becomes saturated in that number of groups. This corresponds to a total of five median stellar spectra using one to five groups. Extended Data Fig. 7 shows these spectra relative to the five-group spectrum, which is equivalent to our standard extraction. We see a trend of increasing extracted flux within the saturated region (vertical regions shaded grey, where the darkest region saturates after one group and the lightest region saturates after four groups) when using a fewer number of groups, suggesting that the first-group effect is biasing the ramps when less than five groups are used. Extrapolating this trend by fitting a Gaussian, we could determine the median excess flux in regions of the spectrum that saturate in two, three or four groups. We do not report a correction for the region that saturates in one group because the one-group spectrum is very noisy and the first-group effect is not well-enough understood. For this reason, we suggest avoiding saturation after only one good group.

Extended Data Fig. 8 shows the residuals after subtracting the NIRISS/SOSS spectra from the NIRSpec PRISM spectra before and after applying the dilution corrections shown in Extended Data Fig. 7c.

We found that the median differences between SOSS and PRISM reduce from 0.4σ to 0.08σ . Although this is a marked improvement in agreement, it is important to note that the dilution corrections both assume that the shapes of the group differences are solely due to first-group effects and are extrapolated from a small number of groups, which therefore suggests that they may not be completely representative of the true effect. For this reason, and based on the extrapolated excess flux measurements, we suggest adopting a best practice of at least five groups before saturation. This is a safe regime and is not dominated by any of the effects that impact the group differences. Doing so will ensure an accurate measurement of the stellar flux in this region.

We further caution on relying on applying a similar correction to other NIRSpec PRISM data on a similarly bright star without NIRISS/SOSS data to compare to, particularly if the wavelength regime that saturates in the NIRSpec PRISM data is crucial to the science. Although the broad wavelength coverage of NIRSpec PRISM is unmatched, the results in this work demonstrate that (1) it can be offset from other modes, although in this work it may be partially due to the low total number of groups, and (2) the feature sizes in the saturated region are unreliable. Even with the application of a dilution correction, larger spectral differences are present relative to the NIRISS/SOSS data than are observed for all other modes relative to NIRISS/SOSS. Our recommendations are to avoid partial saturation of NIRSpec PRISM, particularly if the saturated wavelengths are important to the science case, unless future JWST calibration data better understands the first-group effect or improves the nonlinearity correction, therefore improving the ability to recover the saturated region. Although the strategy of using several higher-resolution modes would require the observation of a second transit, the higher spectral resolution may allow for additional science, such as enhanced stellar modelling⁵². NIRSpec PRISM remains a powerful tool for dimmer host stars as it obtains a wide wavelength coverage in an efficient single transit.

Modelling

To confirm the origin of the spectroscopic features present in the data from this synthesized release, we compared the observations against an atmospheric model of WASP-39 b.

We utilized all the spectroscopic JWST data presented in this work, except for data in the saturated region of NIRSpec PRISM. For NIRISS/SOSS, NIRC*am* F322W2 and NIRSpec G395H, we used the data binned to $R=100$, and for NIRSpec PRISM, we used the data at the native spectral resolution.

We were motivated by the atmospheric inferences from the initial data releases: NIRISS ($\times 10$ –30 solar metallicity and subsolar C/O)⁹; NIRSpec G395H ($\times 3$ –10 solar metallicity and subsolar C/O)⁸; NIRC*am* ($\times 1$ –100 solar metallicity and subsolar C/O)⁷ and NIRSpec PRISM (approximately $\times 10$ solar metallicity and subsolar C/O)¹⁰. We chose $\times 10$ solar metallicity, a subsolar carbon-to-oxygen ratio of 0.35 and atmospheric composition under the assumption of full day–night heat redistribution. The atmospheric model assumes 1D-RCPE. Calculating the RCPE models corresponds to coupling a thermochemical solver with a kinetics solver, as recently described in Bell et al.⁵³. First, the model was computed using the ScCHIMERA radiative-convective equilibrium solver⁵⁴, with recent updates and implementations to JWST data from the initial data releases^{9,10}. Then, the photochemical equilibrium corresponding to the atmospheric chemical state arising from the chemical kinetics due to photochemistry and vertical mixing was computed using the VULCAN tool^{55,56}. We iterated over the radiative-convective equilibrium and photochemical equilibrium to ensure that the temperature–pressure profile and gas mixing ratios did not change, resulting in a computation chain from ScCHIMERA to VULCAN to ScCHIMERA to VULCAN to ScCHIMERA.

We then considered the presence of inhomogeneous clouds and hazes by fitting a power-law and grey cloud-deck parametric model to the observations while assuming the resulting 1D-RCPE model. The

parametric cloud/haze model was fitted for a vertically uniform grey cloud opacity κ_{clid} and a power-law haze assuming a scaling law for its cross section, fitting for scattering slope γ and the scale a (ref. 57). Then, we allowed for the presence of inhomogeneous cloud cover by using a linear combination of the cloudy/hazy model and a cloud-free model⁵⁸, with a cloud fraction ϕ . When fitting the cloud parameters to the 1D-RCPE model, we allowed for an offset in the transit depth for the NIRSpec PRISM observations relative to all other instruments with a uniform prior of ± 500 ppm. Additionally, we allowed for a scaling to the planetary radius referenced to 1 bar pressure. The final transmission spectrum in Fig. 3 corresponds to the post-processed 1D-RCPE model with the median cloud parameters $\log(\kappa_{\text{clid}}) = -29.45$, $\gamma = 1.63$, $\log(a) = 1.95$ and $\phi = 0.84$, and a scaling of the planetary radius of 97%.

The fit for cloud and haze properties suggests that a non-negligible offset must be applied to the NIRSpec PRISM observations to match the 1D-RCPE model. For this specific atmospheric composition, we found that the best-fitting negative offset of 177 ppm was required to bring the data to the same transit-depth level as the model. Comparing this model to the remaining data from other instruments (for example, NIRISS/SOSS order 2) seems to suggest that additional offsets for each instrument and detector may be required. The presence of offsets in the data and their impact on the inferred atmospheric properties are explored in greater detail in the companion work of Welbanks et al.

Data availability

The data used in this paper are associated with JWST programme ERS 1366 (observations 1–4) and are available from the Mikulski Archive for Space Telescopes (<https://mast.stsci.edu>). Specific data products for the time series spectra, white and spectroscopic light curve fits, transmission spectra and model spectrum are available via Zenodo at <https://doi.org/10.5281/zenodo.10161743> (ref. 59).

Code availability

This publication made use of the following code software to analyse these data: NumPy (ref. 60), matplotlib (ref. 61), SciPy (ref. 62), pandas (ref. 63,64), Batman (ref. 48), emcee (ref. 50), Exotic-LD (ref. 49), dynesty (ref. 46,65), SpectRes (ref. 66), juliet (ref. 29) and Eureka! (ref. 22).

References

1. Faedi, F. et al. WASP-39b: a highly inflated Saturn-mass planet orbiting a late G-type star. *Astron. Astrophys.* **531**, A40 (2011).
2. Fischer, P. D. et al. HST hot-Jupiter transmission spectral survey: clear skies for cool Saturn WASP-39b. *Astrophys. J.* **827**, 19 (2016).
3. Nikolov, N. et al. VLT FORS2 comparative transmission spectroscopy: detection of Na in the atmosphere of WASP-39b from the ground. *Astrophys. J.* **832**, 191 (2016).
4. Wakeford, H. et al. The complete transmission spectrum of WASP-39b with a precise water constraint. *Astron. J.* **155**, 29 (2018).
5. Kirk, J. et al. LRG-BEASTS: transmission spectroscopy and retrieval analysis of the highly inflated Saturn-mass planet WASP-39b. *Astron. J.* **158**, 144 (2019).
6. Team, J. T. et al. Identification of carbon dioxide in an exoplanet atmosphere. *Nature* **614**, 649–652 (2023).
7. Ahler, E.-M. et al. Early Release Science of the exoplanet WASP-39b with JWST NIRCam. *Nature* **614**, 653–658 (2023).
8. Alderson, L. et al. Early Release Science of the exoplanet WASP-39b with JWST NIRSpec G395H. *Nature* **614**, 664–669 (2023).
9. Feinstein, A. D. et al. Early Release Science of the exoplanet WASP-39b with JWST NIRISS. *Nature* **614**, 670–675 (2023).
10. Rustamkulov, Z. et al. Early Release Science of the exoplanet WASP-39b with JWST NIRSpec PRISM. *Nature* **614**, 659–663 (2023).
11. Rackham, B. V., Apai, D. & Giampapa, M. S. The transit light source effect: false spectral features and incorrect densities for M-dwarf transiting planets. *Astron. J.* **853**, 122 (2018).
12. Rackham, B. V., Apai, D. & Giampapa, M. S. The transit light source effect. II. The impact of stellar heterogeneity on transmission spectra of planets orbiting broadly Sun-like stars. *Astron. J.* **157**, 96 (2019).
13. Madhusudhan, N. C/O ratio as a dimension for characterizing exoplanetary atmospheres. *Astrophys. J.* **758**, 36 (2012).
14. Moses, J., Madhusudhan, N., Visscher, C. & Freedman, R. Chemical consequences of the C/O ratio on hot Jupiters: examples from WASP-12b, CoRoT-2b, XO-1b, and HD 189733b. *Astrophys. J.* **763**, 25 (2013).
15. Mollière, P. et al. Model atmospheres of irradiated exoplanets: the influence of stellar parameters, metallicity, and the C/O ratio. *Astrophys. J.* **813**, 47 (2015).
16. Öberg, K. I., Murray-Clay, R. & Bergin, E. A. The effects of snowlines on C/O in planetary atmospheres. *Astrophys. J. Lett.* **743**, L16 (2011).
17. Espinoza, N. et al. Metal enrichment leads to low atmospheric C/O ratios in transiting giant exoplanets. *Astrophys. J. Lett.* **838**, L9 (2017).
18. Mollière, P. et al. Interpreting the atmospheric composition of exoplanets: sensitivity to planet formation assumptions. *Astrophys. J.* **934**, 74 (2022).
19. Radica, M. et al. Awesome SOSS: transmission spectroscopy of WASP-96b with NIRISS/SOSS. *Mon. Not. R. Astron. Soc.* **524**, 835–856 (2023).
20. Doyon, R. et al. The near infrared imager and slitless spectrograph for the James Webb Space Telescope. I. Instrument overview and in-flight performance. *Publ. Astron. Soc. Pac.* **135**, 098001 (2023).
21. Albert, L. et al. The near infrared imager and slitless spectrograph for the James Webb Space Telescope. III. Single object slitless spectroscopy. *Publ. Astron. Soc. Pac.* **135**, 075001 (2023).
22. Bell, T. et al. Eureka!: an end-to-end pipeline for JWST time-series observations. *J. Open Source Softw.* **7**, 4503 (2022).
23. Greene, T. P. et al. $\lambda = 2.4$ to $5 \mu\text{m}$ spectroscopy with the James Webb Space Telescope NIRCam instrument. *J. Astron. Telesc. Instrum. Syst.* **3**, 035001 (2017).
24. Alderson, L., Grant, D. & Wakeford, H. Exo-TiC/ExoTiC-JEDI: v0.1-beta-release. *Zenodo* <https://zenodo.org/records/7185855> (2022).
25. Jakobsen, P. et al. The near-infrared spectrograph (NIRSpec) on the James Webb Space Telescope. I. Overview of the instrument and its capabilities. *Astron. Astrophys.* **661**, A80 (2022).
26. Birkmann, S. et al. The near-infrared spectrograph (NIRSpec) on the James Webb Space Telescope. IV. Capabilities and predicted performance for exoplanet characterization. *Astron. Astrophys.* **661**, A83 (2022).
27. Espinoza, N. et al. Spectroscopic time-series performance of JWST/NIRSpec from commissioning observations. *Publ. Astron. Soc. Pac.* **135**, 018002 (2023).
28. Rustamkulov, Z., Sing, D. K., Liu, R. & Wang, A. Analysis of a JWST NIRSpec lab time series: characterizing systematics, recovering exoplanet transit spectroscopy, and constraining a noise floor. *Astrophys. J. Lett.* **928**, L7 (2022).
29. Espinoza, N., Kossakowski, D. & Brahm, R. juliet: a versatile modelling tool for transiting and non-transiting exoplanetary systems. *Mon. Not. R. Astron. Soc.* **490**, 2262–2283 (2019).
30. Ricker, G. R. et al. Transiting exoplanet survey satellite (TESS). *J. Astron. Telesc. Instrum. Syst.* **1**, 014003 (2015).
31. Wheatley, P. J. et al. The next generation transit survey (NGTS). *Mon. Not. R. Astron. Soc.* **475**, 4476–4493 (2018).
32. Silva Aguirre, V. et al. Standing on the shoulders of dwarfs: the Kepler asteroseismic LEGACY sample. II. Radii, masses, and ages. *Astrophys. J.* **835**, 173 (2017).

33. Seager, S. & Mallén-Ornelas, G. A unique solution of planet and star parameters from an extrasolar planet transit light curve. *Astrophys. J.* **585**, 1038–1055 (2003).
34. Schlawin, E. et al. JWST NIRCcam defocused imaging: photometric stability performance and how it can sense mirror tilts. *Publ. Astron. Soc. Pac.* **135**, 018001 (2023).
35. Zhou, Y., Apai, D., Lew, B. W. & Schneider, G. A physical model-based correction for charge traps in the Hubble Space Telescope wide field camera 3 near-IR detector and its applications to transiting exoplanets and brown dwarfs. *Astron. J.* **153**, 243 (2017).
36. Gibson, N. et al. Probing the haze in the atmosphere of HD 189733b with Hubble Space Telescope/WFC3 transmission spectroscopy. *Mon. Not. R. Astron. Soc.* **422**, 753–760 (2012).
37. Spake, J. et al. Helium in the eroding atmosphere of an exoplanet. *Nature* **557**, 68–70 (2018).
38. Oklopčić, A. & Hirata, C. M. A new window into escaping exoplanet atmospheres: 10830 Å line of helium. *Astrophys. J. Lett.* **855**, L11 (2018).
39. Stevenson, K. B. et al. Transiting exoplanet studies and community targets for JWST's Early Release Science program. *Publ. Astron. Soc. Pac.* **128**, 094401 (2016).
40. Bean, J. L. et al. The transiting exoplanet community Early Release Science program for JWST. *Publ. Astron. Soc. Pac.* **130**, 114402 (2018).
41. Ngo, H. et al. Friends of hot Jupiters. IV. Stellar companions beyond 50 au might facilitate giant planet formation, but most are unlikely to cause Kozai–Lidov migration. *Astrophys. J.* **827**, 8 (2016).
42. Evans, D. et al. High-resolution imaging of transiting extrasolar planetary systems (HITEP). II. Lucky imaging results from 2015 and 2016. *Astron. Astrophys.* **610**, A20 (2018).
43. Maciejewski, G. et al. New transit observations for HAT-P-30 b, HAT-P-37 b, TrES-5 b, WASP-28 b, WASP-36 b and WASP-39 b. *Acta Astron.* **66**, 55–74 (2016).
44. Eastman, J., Siverd, R. & Gaudi, B. S. Achieving better than 1 minute accuracy in the heliocentric and barycentric Julian dates. *Publ. Astron. Soc. Pac.* **122**, 935 (2010).
45. Kipping, D. M. Efficient uninformative sampling of limb darkening coefficients for two-parameter laws. *Mon. Not. R. Astron. Soc.* **435**, 2152–2160 (2013).
46. Speagle, J. S. DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Mon. Not. R. Astron. Soc.* **493**, 3132–3158 (2020).
47. Pontoppidan, K. M. et al. Pandeia: a multi-mission exposure time calculator for JWST and WFIRST. In *Proc. Observatory Operations: Strategies, Processes, and Systems VI* (eds Peck, A. B. et al.) 991016 (SPIE, 2016).
48. Kreidberg, L. batman: basic transit model calculation in Python. *Publ. Astron. Soc. Pac.* **127**, 1161 (2015).
49. Wakeford, D. G. Exo-TiC/ExoTiC-LD: ExoTiC-LD v3.0.0. *Zenodo* <https://zenodo.org/records/7437681> (2022).
50. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: the MCMC hammer. *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
51. Christiansen, J. et al. Understanding exoplanet atmospheres with UV observations. I. NUV and blue/optical. *Bull. Am. Astron. Soc.* **51**, 408 (2019).
52. Moran, S. E. et al. High tide or riptide on the cosmic shoreline? A water-rich atmosphere or stellar contamination for the warm super-Earth GJ 486b from JWST observations. *Astrophys. J. Lett.* **948**, L11 (2023).
53. Bell, T. J. et al. Methane throughout the atmosphere of the warm exoplanet WASP-80b. *Nature* <https://doi.org/10.1038/s41586-023-06687-0> (2023).
54. Tsai, S.-M., Lee, E. K.-H. & Pierrehumbert, R. A mini-chemical scheme with net reactions for 3D general circulation models. I. Thermochemical kinetics. *Astron. Astrophys.* **664**, A82 (2022).
55. Line, M. & Parmentier, V. The influence of nonuniform cloud cover on transit transmission spectra. *Astrophys. J.* **820**, 78 (2016).
56. Piskorz, D. et al. Ground- and space-based detection of the thermal emission spectrum of the transiting hot Jupiter KELT-2Ab. *Astrophys. J.* **156**, 133 (2018).
57. Tsai, S.-M. et al. VULCAN: an open-source, validated chemical kinetics Python code for exoplanetary atmospheres. *Astrophys. J. Suppl. Ser.* **228**, 20 (2017).
58. Lecavelier Des Etangs, A., Pont, F., Vidal-Madjar, A. & Sing, D. Rayleigh scattering in the transit spectrum of HD 189733b. *Astron. Astrophys.* **481**, L83–L86 (2008).
59. Carter, A. L. et al. Products and models for "Improving Consistency Between the JWST Near-Infrared Spectra of WASP-39b". *Zenodo* <https://doi.org/10.5281/zenodo.10161743> (2024).
60. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
61. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 905195 (2007).
62. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
63. Reback, J. et al. pandas-dev/pandas: Pandas. *Zenodo* <https://zenodo.org/records/3630805> (2020).
64. McKinney, W. Data structures for statistical computing in Python. In *Proc. 9th Python in Science Conference* (eds van der Walt, S. and Millman, J.) 56–61 (SciPy, 2010).
65. Koposov, S. et al. joshspeagle/dynesty: v2.1.0. *Zenodo* <https://zenodo.org/records/7600689> (2023).
66. Carnall, A. SpectRes: a fast spectral resampling tool in Python. Preprint at <https://arxiv.org/abs/1705.05165> (2017).

Acknowledgements

This work is based on observations made with 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. These observations are associated with programme JWST-ERS-01366. Support for programme JWST-ERS-01366 was provided by NASA through a grant from the Space Telescope Science Institute. This work is based in part on data collected under the NGTS project at the European Southern Observatory's La Silla Paranal Observatory. The NGTS facility is operated by a consortium of institutes with support from the UK Science and Technology Facilities Council (Projects ST/M001962/1, ST/S002642/1 and ST/W003163/1). L.W., J.D.T., M.M. and D.P. are or were NHFP Sagan Fellows. B.V.R. was a 51 Pegasi b Fellow. J.K. is an Imperial College Research Fellow. T.D. is an LSSTC Catalyst Fellow.

Author contributions

All authors played a substantial role in one or more of the following: developing the original proposal, managing the project, defining the target list and observation plan, analysing the data, theoretical modelling or preparing this paper. Some specific contributions are listed as follows. N.B., J.L.B., Z.K.B.-T., I.J.M.C., J.M.D., L.K., M.L., D.K.S., K.B.S. and H.R.W. provided overall programme leadership and management. D.K.S., E.M.-R.K., H.R.W., I.J.M.C., J.L.B., K.B.S., L.K., M.L.-M., M.L., N.M.B., V.P. and Z.K.B.-T. made substantial contributions to the design of the programme. K.B.S. generated the observing plan with input from the team. B.B., E.M.R.K., H.R.W., I.J.M.C., J.L.B., L.K., M.L.-M., M.L., N.M.B. and Z.K.B.-T. led or co-led working groups or contributed to substantial strategic planning efforts, such as the design and implementation of the prelaunch data challenges. A.L.C., D.K.S., E.S., N.E., N.P.G. and V.P. generated the simulated data for the prelaunch testing of the methods. A.L.C. and E.M.M. co-led the data synthesis effort, light curve generation, spectroscopic light curve fitting, limb-darkening investigation, saturation investigation and

production of the planetary spectrum. E.A., L.A., A.D.F., D.G., M.R., Z.R. and K.B.S. contributed substantially to the generation of, and provided, the spectral time series that were used in this paper. N.E., R.B., R.O., G.M. and J.T. performed the white light curve fitting. L.W. and M.L. generated the theoretical model for comparison with the data. A.L.C., E.M.M., N.E. and L.W. made substantial contributions to the writing of this paper. A.L.C. and E.M.M. generated the figures for this paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41550-024-02292-x>.

Correspondence and requests for materials should be addressed to A. L. Carter or E. M. May.

Peer review information *Nature Astronomy* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

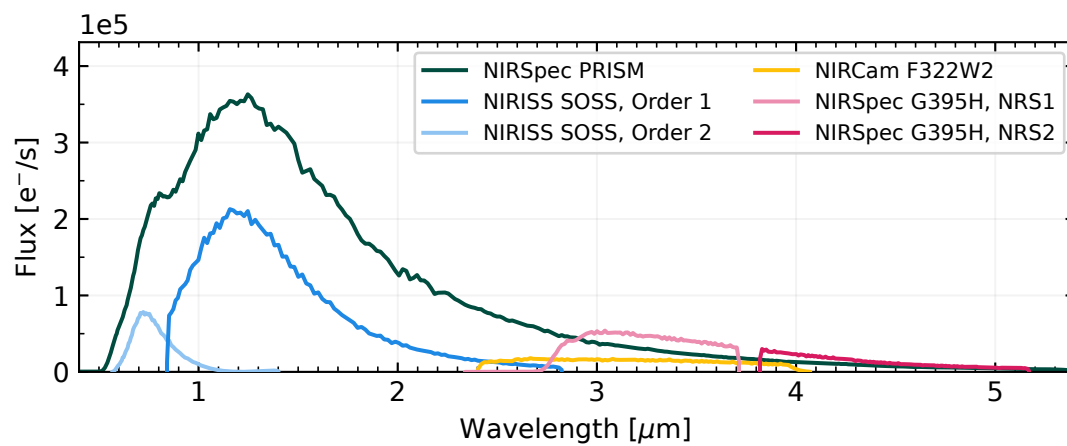
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

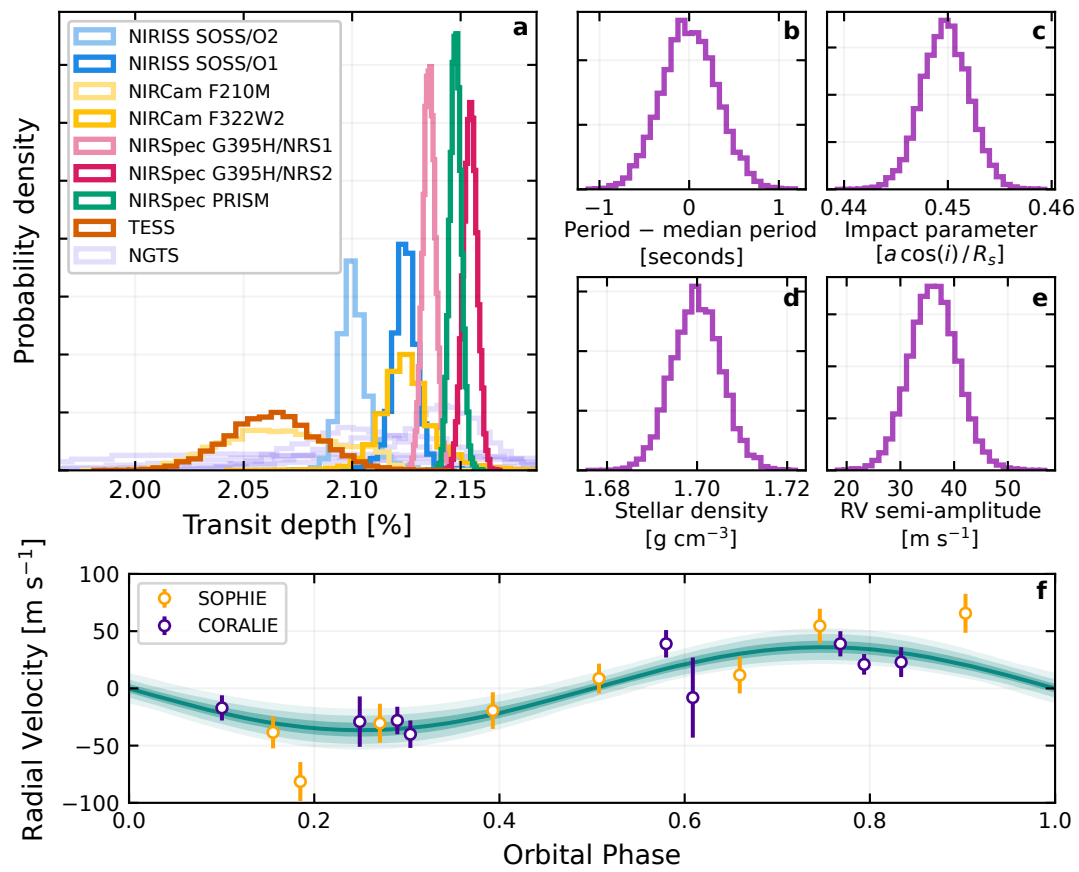
A. L. Carter^{1,2}✉, E. M. May³✉, N. Espinoza^{1,4}, L. Welbanks⁵, E. Ahrer^{6,7,8}, L. Alderson⁹, R. Brahm^{10,11,12}, A. D. Feinstein^{13,14}, D. Grant⁹, M. Line⁵, G. Morello^{15,16}, R. O'Steen¹, M. Radica^{17,18}, Z. Rustamkulov¹⁹, K. B. Stevenson³, J. D. Turner²⁰, M. K. Alam^{1,21}, D. R. Anderson^{6,7}, N. M. Batalha², M. P. Battley²², D. Bayliss⁷, J. L. Bean¹³, B. Benneke²³, Z. K. Berta-Thompson²⁴, J. Brande²⁵, E. M. Bryant²⁶, M. R. Burleigh²⁷, L. Coulombe²³, I. J. M. Crossfield²⁵, M. Damiano²⁸, J.-M. Désert²⁹, L. Flagg²⁰, S. Gill^{6,7}, J. Inglis³⁰, J. Kirk^{30,31}, H. Knutson³⁰, L. Kreidberg⁸, M. López Morales³², M. Mansfield³³, S. E. Moran³⁴, C. A. Murray²⁴, M. C. Nixon³⁵, D. J. M. Petit dit de la Roche²², B. V. Rackham^{36,37}, E. Schlawin³³, D. K. Sing^{4,19}, H. R. Wakeford⁹, N. L. Wallack²¹, P. J. Wheatley^{6,7}, S. Zieba^{8,38}, K. Aggarwal³⁹, J. K. Barstow⁴⁰, T. J. Bell⁴¹, J. Blečić^{42,43}, C. Cáceres^{44,45,46}, N. Crouzet³⁸, P. E. Cubillos^{47,48}, T. Daylan⁴⁹, M. de Val-Borro⁵⁰, L. Decin⁵¹, J. J. Fortney², N. P. Gibson⁵², K. Heng^{7,53}, R. Hu^{28,30}, E. M.-R. Kempton³⁵, P. Lagage⁵⁴, J. D. Lothringer^{1,55}, J. Lustig-Yaeger³, L. Mancini^{8,56,57}, N. J. Mayne⁵⁸, L. C. Mayorga³, K. Molaverdikhani^{53,59}, E. Nasedkin⁸, K. Ohno⁶⁰, V. Parmentier⁶¹, D. Powell³², S. Redfield⁶², P. Roy²³, J. Taylor^{23,63} & X. Zhang⁶⁴

¹Space Telescope Science Institute, Baltimore, MD, USA. ²Department of Astronomy & Astrophysics, University of California, Santa Cruz, Santa Cruz, CA, USA. ³Johns Hopkins APL, Laurel, MD, USA. ⁴Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD, USA. ⁵School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA. ⁶Centre for Exoplanets and Habitability, University of Warwick, Coventry, UK. ⁷Department of Physics, University of Warwick, Coventry, UK. ⁸Max Planck Institute for Astronomy, Heidelberg, Germany. ⁹School of Physics, University of Bristol, Bristol, UK. ¹⁰Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago, Chile. ¹¹Millennium Institute for Astrophysics, Santiago, Chile. ¹²Data Observatory Foundation, Santiago, Chile. ¹³Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL, USA. ¹⁴Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, CO, USA. ¹⁵Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg, Sweden. ¹⁶Instituto de Astrofísica de Canarias (IAC), Tenerife, Spain. ¹⁷Department of Physics, Université de Montréal, Montreal, Quebec, Canada. ¹⁸Trottier Institute for Research on Exoplanets, Université de Montréal, Montreal, Quebec, Canada. ¹⁹Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA. ²⁰Department of Astronomy and Carl Sagan Institute, Cornell University, Ithaca, NY, USA. ²¹Earth and Planets Laboratory, Carnegie Institution for Science, Washington DC, USA. ²²Département d'Astronomie, Université de Genève, Sauverny, Switzerland. ²³Department of Physics and Institute for Research on Exoplanets, Université de Montréal, Montreal, Quebec, Canada. ²⁴Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO, USA. ²⁵Department of Physics & Astronomy, University of Kansas, Lawrence, KS, USA. ²⁶Mullard Space Science Laboratory, University College London, Dorking, UK. ²⁷School of Physics and Astronomy, University of Leicester, Leicester, UK. ²⁸Astrophysics Section, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. ²⁹Anton Pannekoek Institute for Astronomy, University of Amsterdam, Amsterdam, The Netherlands. ³⁰Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA. ³¹Department of Physics, Imperial College London, London, UK. ³²Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA. ³³Steward Observatory, University of Arizona, Tucson, AZ, USA. ³⁴Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ, USA. ³⁵Department of Astronomy, University of Maryland, College Park, MD, USA. ³⁶Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ³⁷Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA, USA. ³⁸Leiden Observatory, University of Leiden, Leiden, The Netherlands. ³⁹Indian Institute of Technology, Indore, India. ⁴⁰School of Physical Sciences, The Open University, Milton Keynes, UK. ⁴¹BAER Institute, NASA Ames Research Center, Moffett Field, CA, USA. ⁴²Department of Physics, New York University Abu Dhabi, Abu Dhabi, UAE. ⁴³Center for Astrophysics and Space Science (CASS), New York University Abu Dhabi, Abu Dhabi, UAE. ⁴⁴Instituto de Astrofísica, Universidad Andrés Bello, Santiago, Chile. ⁴⁵Centro de Astrofísica y Tecnologías Afines (CATA), Santiago, Chile. ⁴⁶Núcleo Milenio de Formación Planetaria (NPF), Santiago, Chile. ⁴⁷INAF – Osservatorio Astrofisico di Torino, Pino Torinese, Italy. ⁴⁸Space Research Institute, Austrian Academy of Sciences, Graz, Austria.

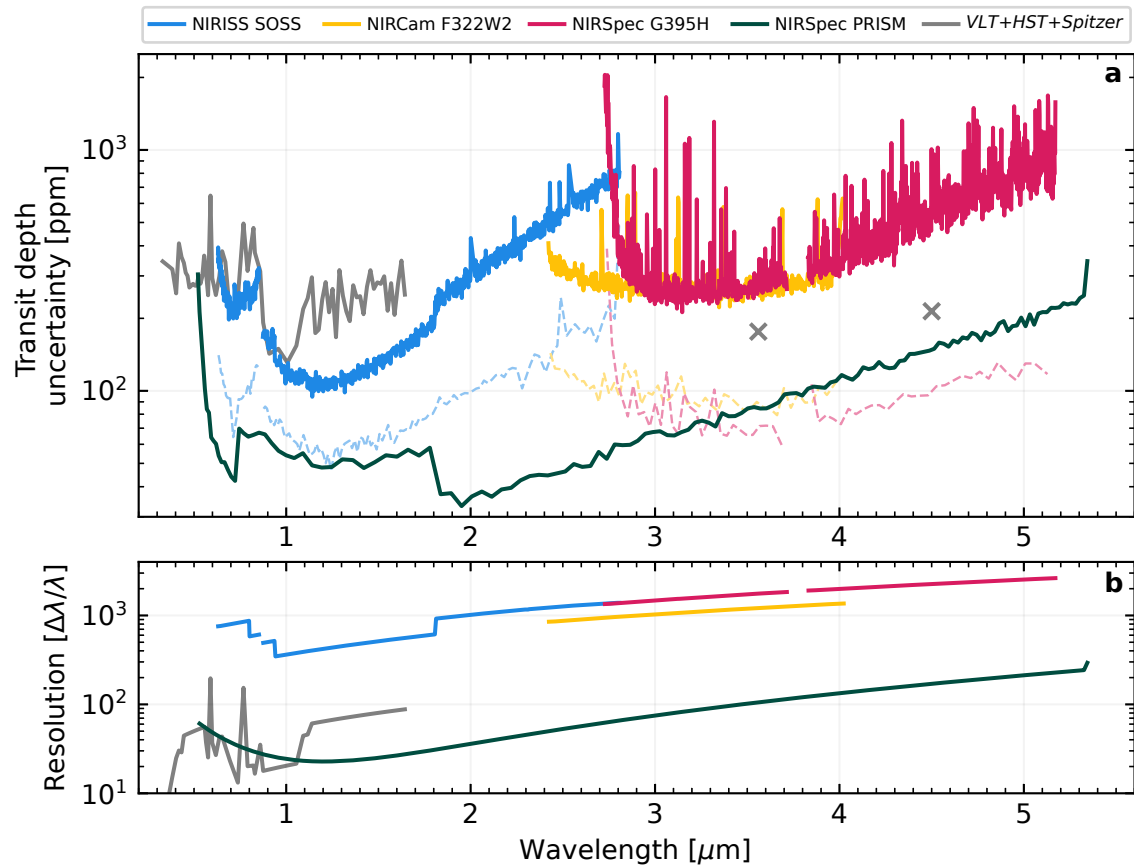
⁴⁹Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA. ⁵⁰Planetary Science Institute, Tucson, AZ, USA. ⁵¹Institute of Astronomy, Department of Physics and Astronomy, KU Leuven, Leuven, Belgium. ⁵²School of Physics, Trinity College Dublin, Dublin, Ireland. ⁵³Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Munich, Germany. ⁵⁴Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, Gif-sur-Yvette, France. ⁵⁵Department of Physics, Utah Valley University, Orem, UT, USA. ⁵⁶Department of Physics, University of Rome ‘Tor Vergata’, Rome, Italy. ⁵⁷INAF - Turin Astrophysical Observatory, Pino Torinese, Italy. ⁵⁸Department of Physics and Astronomy, University of Exeter, Exeter, UK. ⁵⁹Exzellenzcluster Origins, Garching, Germany. ⁶⁰Division of Science, National Astronomical Observatory of Japan, Mitaka-shi, Tokyo, Japan. ⁶¹Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS Laboratoire Lagrange, Nice, France. ⁶²Astronomy Department and Van Vleck Observatory, Wesleyan University, Middletown, CT, USA. ⁶³Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK. ⁶⁴Department of Earth and Planetary Sciences, University of California Santa Cruz, Santa Cruz, CA, USA. ✉e-mail: aacarter@stsci.edu; erin.may@jhuapl.edu



Extended Data Fig. 1 | Median out-of-transit stellar spectra for WASP-39 in units of photoelectrons recorded at the detector. All datasets have been binned in wavelength to match the NIRSpec PRISM data. Residual differences between datasets are a result of the different instrumental throughputs.

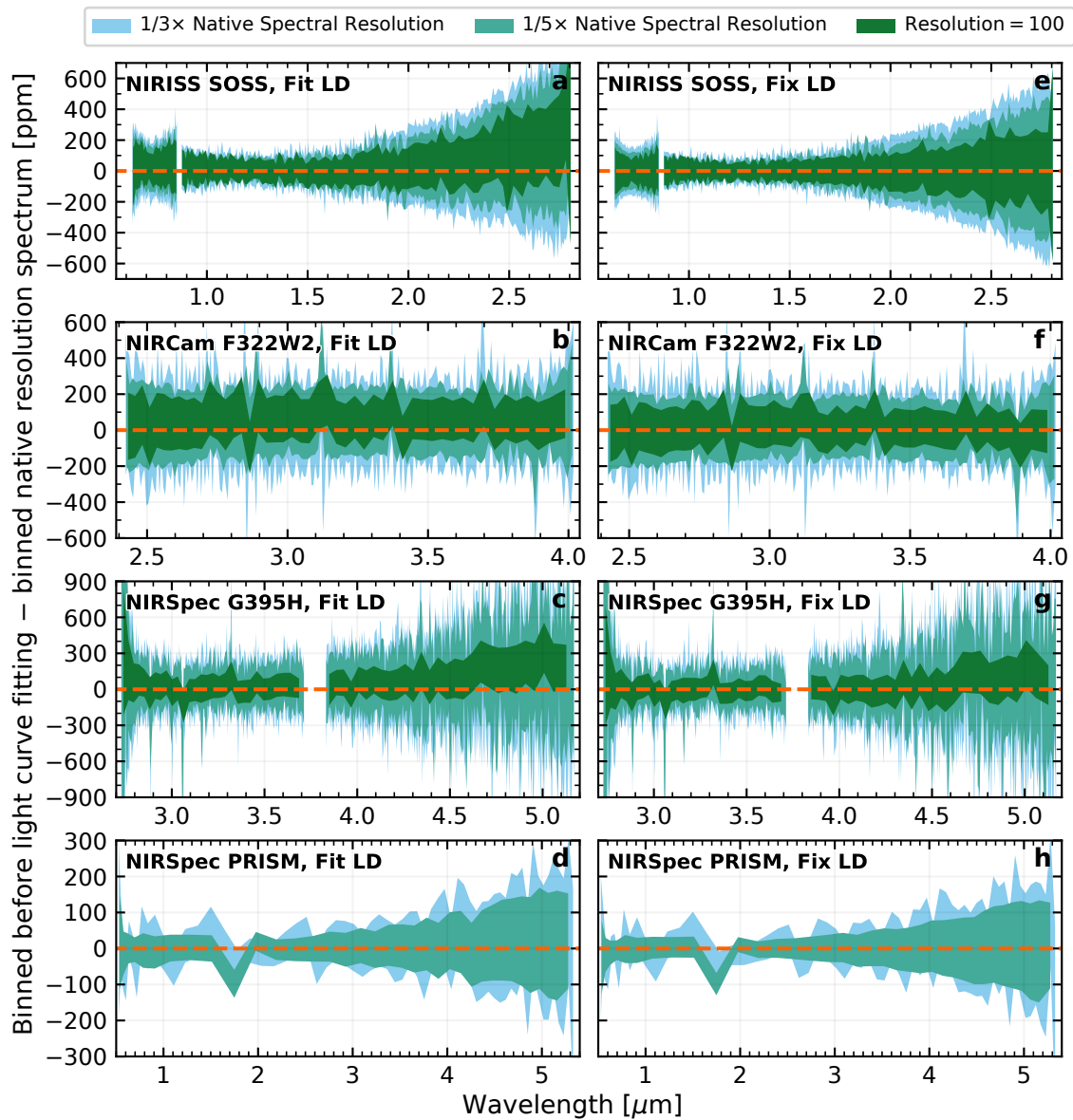


Extended Data Fig. 2 | Best-fit orbital parameters for WASP-39b (a-e) and the radial velocity measurements (f). The measured transit depth varies between instrumental modes as they each span a different wavelength range. The best-fitting model to the radial velocity data (solid line) alongside its 1, 2, and 3 σ contours (shaded regions) are also indicated.



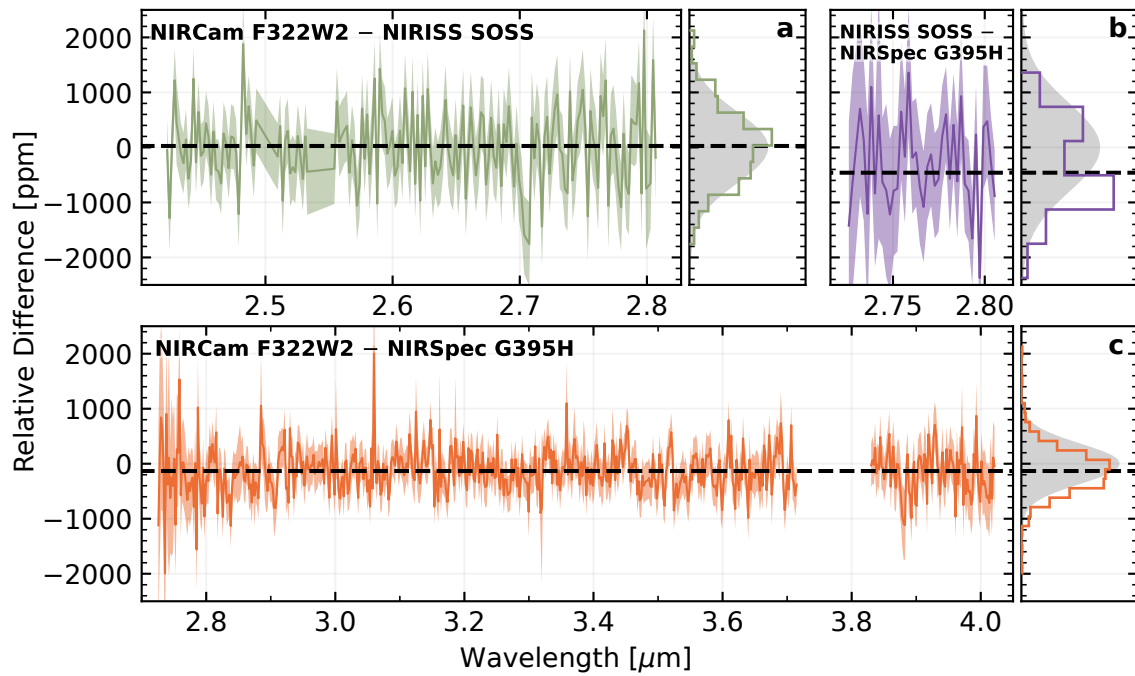
Extended Data Fig. 3 | The achieved transit depth precision (a) and resolving power (b) across all datasets. For the transit depth uncertainty, solid lines correspond to the native spectral resolution datasets, and dashed lines correspond to the $R = 100$ datasets. Archival *HST*, *VLT*, and *Spitzer* data are also displayed (gray lines and crosses). As the reported archival transit depths are an

averaged combination of multiple transits, we inflate the reported transit depth uncertainties by a factor of $\sqrt{N}$, where N is the number of transits for a given bandpass, to more accurately compare the signal-to-noise provided by a single transit across different instruments.



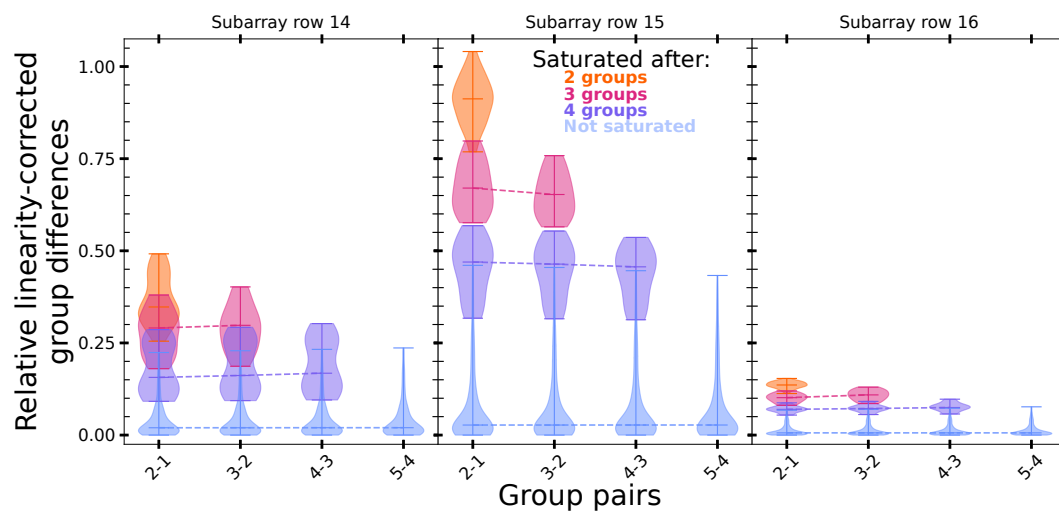
Extended Data Fig. 4 | Comparison between the transmission spectra under different binning schemes and limb-darkening approaches for each instrumental mode. Each panel shows the residual when comparing the spectroscopic transit depths as determined by binning before light curve fitting is performed versus those determined by directly binning the measured

transmission spectrum at the native spectral resolution. Panels **a–d** indicate the residuals for each marked filter, where limb darkening parameters are fit for, and panels **e–h** indicate the residual for each marked filter, with limb darkening parameters fixed during the fit. The width of each shaded region corresponds to the 1σ uncertainty in the measured residual. Dashed lines indicate zero deviation.



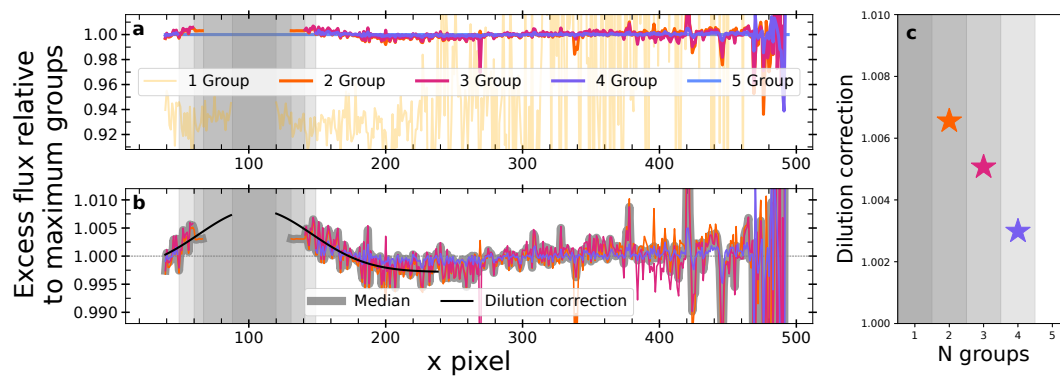
Extended Data Fig. 5 | Comparison of *JWST* instrumental modes across their overlapping wavelength regions. Each panel shows the residual when comparing the spectroscopic transit depths as determined by binning before light curve fitting is performed versus those determined by directly binning the measured transmission spectrum at the native spectral resolution. Panels **a–d**

indicate the residuals for each marked filter, where limb darkening parameters are fit for, and panels **e–h** indicate the residual for each marked filter, with limb darkening parameters fixed during the fit. The width of each shaded region corresponds to the 1σ uncertainty in the measured residual. Dashed lines indicate zero deviation.



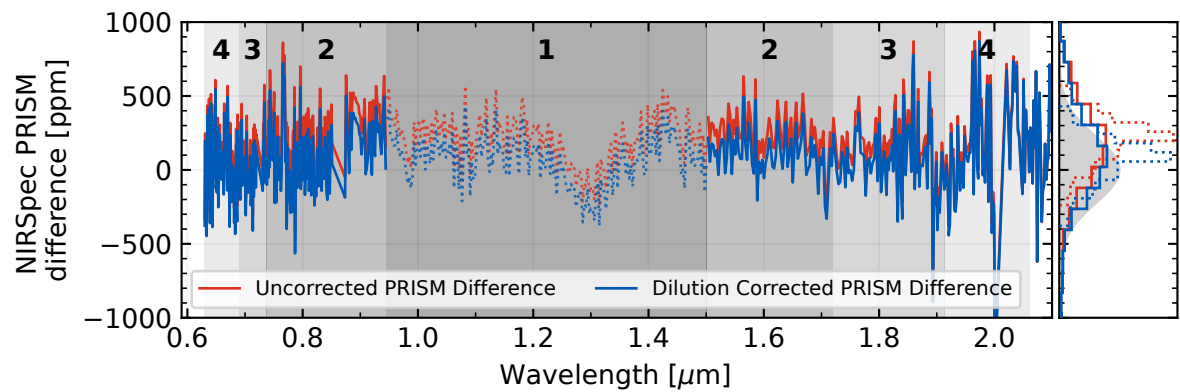
Extended Data Fig. 6 | Differences in counts between neighboring groups after the non-linearity correction. A linear ramp would show a flat line difference between neighboring groups. Residual slopes in differences between neighboring groups are indicative of detector effects that are not fully corrected (denoted by the dashed lines connecting the median of group pairs). Above we show the distributions in differences between neighboring groups for three

detector rows where row 15 corresponds to the central location of the trace, and rows 14 and 16 are the neighboring detector rows. Distributions are colored by the number of groups that are usable before reaching our saturation threshold. Note that we follow the same methodology as¹¹ in flagging an entire column as saturated when the central row saturates.



Extended Data Fig. 7 | Analysis of excess flux relative to the nominal 5-group spectrum. The 5-group spectrum uses a varying number of groups within the saturated region, as is done in the standard PRISM extraction. **(a)** Extracted spectra using different numbers of groups relative to the 5-group spectrum. Data are not shown in regions that are already saturated, or where the 5-group spectrum uses the same number of groups. **(b)** Zoomed-in view of the Top Left

panel, specifically the region in panel a that is enclosed in a box, excluding the noisy 1-group spectrum. A Gaussian is fit to the excess flux spectrum. **(c)** The dilution corrections derived from the median excess flux fit. In all panels the vertical-gray shaded regions denote the level of saturation, with the darkest corresponding to saturation after 1 group and the lightest corresponding to saturation after 4 groups.



Extended Data Fig. 8 | Comparison of NIRISS SOSS to NIRSpec PRISM before and after the estimated dilution correction. Shaded regions indicate the wavelength ranges that experience saturation after the number of groups indicated in bold. A dilution correction cannot be computed for the region that experiences saturation after one group (dotted lines); however, we show an example correction using the correction value for the regions experiencing saturation after two groups. The probability density of the residuals is displayed in the right panel, alongside a Gaussian distribution with zero mean and a

standard deviation equal to the median error on the residual (gray shading). We note that the downwards peak around 1.3 micron is likely due to this spectral region having the lowest resolution across the instrument, resulting in the highest counts in pixels at these wavelengths such that these detector columns get closer to saturation within the first group. Due to insufficient non-linearity correction within the *JWST* pipeline and the first-group effect, lower-than-expected flux is measured in this region, resulting in deeper transit events.

Extended Data Table 1 | Overview of *JWST* observations

Instrument	Observation Date	Wavelength Range (μm)	Resolving Power ($\lambda/\Delta\lambda$)	Number of Integrations	Groups per Integration	Integration Time (seconds)
NIRSpec PRISM	July 10th 2022	0.5 - 5.5	~ 100	21,500	5	1.13
NIRCam F210M + F322W2	July 22nd 2022	2.42 - 4.025	~ 1600	366	12	79.45
NIRISS SOSS	July 26th 2022	0.6 - 2.8	$\sim 700 / \sim 1400$	537	9	49.45
NIRSpec G395H	July 30th 2022	2.725 - 3.716 3.829 - 5.172	$\sim 2,700$	465	70	63.14

Observations are listed in the order they were acquired. For more detail on observational set up for these four modes, see^{5,7,11}; and¹², respectively.

Extended Data Table 2 | Offsets between spectra

Instrument	NIRSpec PRISM	NIRcam F322W2	NIRISS SOSS	NIRSpec G395H
NIRSpec PRISM	-	132 ± 13 ppm	-124 ± 6 ppm	17 ± 11 ppm
NIRCam F322W2	-132 ± 13 ppm	-	-11 ± 49 ppm	-138 ± 16 ppm
NIRISS SOSS	124 ± 6 ppm	11 ± 49 ppm	-	-372 ± 170 ppm
NIRSpec G395H	-17 ± 11 ppm	138 ± 16 ppm	372 ± 170 ppm	-

All offsets should be read as the selected spectrum in the left hand column relative to a spectrum of a different column. It is critical to emphasize that these offsets are unique to the spectra presented in this paper, and should not be interpreted as inherent offsets between JWST's instrumental modes.