The Astronomical Journal, 159:13 (21pp), 2020 January https://doi.org/10.3847/1538-3881/ab55da

ACCESS: A Visual to Near-infrared Spectrum of the Hot Jupiter WASP-43b with Evidence of H_2O, but No Evidence of Na or K

Ian C. Weaver[1], Mercedes López-Morales[1], Néstor Espinoza[2,13,14], Benjamin V. Rackham[3,4,12], David J. Osip[5], Dániel Apai[3,4,6], Andrés Jordán[7,8], Alex Bixel[3,4], Nikole K. Lewis[9], Munazza K. Alam[1], James Kirk[1], Chima McGruder[1], Florian Rodler[10], and Jennifer Fienco[11]

[1] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA; iweaver@cfa.harvard.edu
[2] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[3] Department of Astronomy/Steward Observatory, The University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721, USA
[4] Earths in Other Solar Systems Team, NASA Nexus for Exoplanet System Science, USA
[5] Las Campanas Observatory, Carnegie Institution of Washington, Colina el Pino, Casilla 601 La Serena, Chile
[6] Lunar and Planetary Laboratory, The University of Arizona, 1640 E. Univ. Boulevard, Tucson, AZ 85721, USA
[7] Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diagonal las Torres 2640, Peñalolén, Santiago, Chile
[8] Millennium Institute for Astrophysics, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile
[9] Department of Astronomy and Carl Sagan Institute, Cornell University, 122 Sciences Drive, 14853, Ithaca, NY, USA
[10] European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago de Chile, Chile
[11] Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile
[12] Department of Earth, Atmospheric and Planetary Sciences, and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 2019 June 12; revised 2019 October 25; accepted 2019 November 7; published 2019 December 12

Abstract

We present a new ground-based visual transmission spectrum of the hot Jupiter WASP-43b, obtained as part of the ACCESS Survey. The spectrum was derived from four transits observed between 2015 and 2018, with combined wavelength coverage between 5300 and 9000 Å and an average photometric precision of 708 ppm in 230 Å bins. We perform an atmospheric retrieval of our transmission spectrum combined with literature *Hubble Space Telescope*/WFC3 observations to search for the presence of clouds/hazes as well as Na, K, Hα, and H_2O planetary absorption and stellar spot contamination over a combined spectral range of 5318–16420 Å. We do not detect a statistically significant presence of Na I or K I alkali lines, or Hα in the atmosphere of WASP-43b. We find that the observed transmission spectrum can be best explained by a combination of heterogeneities on the photosphere of the host star and a clear planetary atmosphere with H_2O. This model yields a log evidence of 8.26 ± 0.42 higher than a flat (featureless) spectrum. In particular, the observations marginally favor the presence of large, low-contrast spots over the four ACCESS transit epochs with an average covering fraction $f_{\rm het} = 0.27^{+0.42}_{-0.16}$ and temperature contrast $\Delta T = 132\,{\rm K} \pm 132\,{\rm K}$. Within the planet's atmosphere, we recover a log H_2O volume mixing ratio of $-2.78^{+1.38}_{-1.47}$, which is consistent with previous H_2O abundance determinations for this planet.

Unified Astronomy Thesaurus concepts: Exoplanet atmospheres (487); Observational astronomy (1145); Spectroscopy (1558); Exoplanets (498); Hot Jupiters (753); Stellar activity (1580)

1. Introduction

Observations of exoplanetary atmospheres offer the possibility of understanding the atmospheric physical properties and chemical composition of those worlds, as well as providing clues to their formation and evolution histories (e.g., Moses et al. 2013; Öberg et al. 2013; Mordasini et al. 2016; Espinoza et al. 2017). The first comparison studies of exoplanetary atmospheres using transmission spectra (see, e.g., Sing et al. 2016; Crossfield & Kreidberg 2017) found evidence of a gradual transition between clear and cloudy atmospheres, but no clear correlation of that transition with other system parameters, such as planetary mass, gravity, effective temperature, or stellar irradiation levels, and the chemical composition of the star. Recently, Pinhas et al. (2019) reanalyzed the Sing et al. (2016) sample and concluded that the majority of hot Jupiters have atmospheres consistent with subsolar H_2O abundances, with the log of those values ranging from $-5.04^{+0.46}_{-0.30}$ to $-3.16^{+0.66}_{-0.69}$.

High-altitude clouds/hazes have been inferred in the atmosphere of a number of exoplanets from scattering slopes in the visible (e.g., Pont et al. 2008, 2013; Sing et al. 2009, 2011, 2013; Gibson et al. 2013) and from the damping of pressure-broadened alkali Na I and K I lines originating from deeper within the atmosphere (e.g., Charbonneau et al. 2002; Wakeford & Sing 2015). The first detections of exoplanets with potentially clear atmospheres have only recently been made, e.g., WASP-96b (Nikolov et al. 2018).

This field is currently at the point where a number of efforts are underway to identify what system parameters, if any, correlate with the observed atmospheric properties of exoplanets. For example, the relationship between chemical abundance in an exoplanet's atmosphere and planet mass is actively being explored (Fraine et al. 2014; Kreidberg et al. 2014; Helling et al. 2016; Sing et al. 2016). Ground-based (e.g., ACCESS,[15] GPIES,[16] VLT FORS2,[17] LRG-BEASTS[18]) and

[13] Bernoulli Fellow.
[14] IAU-Gruber Fellow.

[15] Arizona–CfA–Católica–Carnegie Exoplanet Spectroscopy Survey (Rackham et al. 2017).
[16] Gemini Planet Imager Extra Solar Survey (Nielsen et al. 2019).
[17] Very Large Telescope FOcal Reducer and Spectrograph (Nikolov et al. 2018).
[18] Low Resolution Ground-Based Exoplanet Atmosphere Survey using Transmission Spectroscopy (Kirk et al. 2018).

space-based surveys (PanCET[19]) are working to provide homogeneous spectra of a statistically significant number of exoplanet atmospheres in the search for these correlations. In this paper, we present the ground-based visual to near-infrared (NIR) transmission spectrum of the hot Jupiter WASP-43b obtained as part of the ACCESS survey.

WASP-43b ($M_p = 2.052 \pm 0.053\ M_J$, $R_p = 1.036 \pm 0.012\ R_J$, $T_{eq} = 1440^{+40}_{-39}$ K; Gillon et al. 2012) is a hot Jupiter discovered by Hellier et al. (2011) transiting a $V = 12.4$ K7V-type dwarf star ($M_s = 0.717 \pm 0.025\ M_\odot$, $R_s = 0.667 \pm 0.010\ R_\odot$, $T_{eff} = 4520 \pm 120$ K; Gillon et al. 2012) every 0.81 days. WASP-43 is unusually active for its stellar type, as indicated by the presence of strong Ca H and K lines and perhaps due to star–planet interactions in this very short-period system (Staab et al. 2017). *Spitzer* secondary eclipse data in the 3.6 μm and 4.5 μm bands indicate brightness temperatures of 1670 ± 23 K and 1514 ± 25 K, respectively, which rule out a strong thermal inversion in the planet's dayside photosphere (Blecic et al. 2014). In addition, thermal emission observed in the K band (Chen et al. 2014) agrees with atmospheric models of WASP-43b, which predict poor day-to-night heat redistribution in an atmosphere with no thermal inversion present (Kataria et al. 2015).

The presence of water on the dayside of the planet was observed with *Hubble Space Telescope* (*HST*)/WFC3 emission measurements by Stevenson et al. (2014), with additional transmission observations by Kreidberg et al. (2014) finding water abundances comparable to solar values. In the visual regime, the Gran Telescopio Canarias's visual System for Imaging and low Resolution Integrated Spectroscopy (OSIRIS) instrument shows a tentative excess in R_p/R_s at Na I and a complete lack of one near the K I doublet (Murgas et al. 2014). That same study also notes an increasing trend of the planet-to-star radius ratio from 6200 to 7200 Å and a decreasing trend redward of 7200 Å. They attribute this pattern to the possible presence of VO and TiO.

In this work, we search further for the presence of Na I, K I, and Hα with new visual transit observations from the Magellan/Inamori-Magellan Areal Camera and Spectrograph (IMACS). In addition, we combine the *HST*/WFC3 transmission spectrum of WASP-43b (Kreidberg et al. 2014) with the new visual data to produce the full visual to NIR spectrum spanning 5317.90 Å–16420 Å that can further constrain the water absorption features present in the infrared spectrum and provide new information about water abundance. In our analysis, we find that the atmosphere of WASP-43b is best described by a clear atmosphere with water abundance consistent with solar. The planet's spectrum is also contaminated with stellar heterogeneity.

This paper is structured as follows: in Section 2, we present our Magellan/IMACS observations. In Section 3, we outline the data reduction process used in our observations and describe the selection of wavelength bins to optimize the signal-to-noise ratio (S/N), search for atomic features, and compare to other results. We present the detrended white and binned light curves for each data set. In Section 4, we give a qualitative analysis of the impact of observational stellar activity on the resulting combined transmission spectra from different visits. In Section 5, we present the final transmission spectrum and also compare to the results of Murgas et al. (2014). In Section 5.4, we present the results of a retrieval modeling analysis on the combined ACCESS and *HST* transmission spectrum to find the best-fit transmission model when the presence of a heterogeneous stellar photosphere is also taken into account. We summarize and conclude in Section 7.

[19] Panchromatic Comparative Exoplanetology Treasury (Wakeford et al. 2017).

Table 1
Target and Comparison Star Magnitudes and Coordinates from the UCAC4 Catalog (http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=I/322A&-to=3)

Star	R.A.	Decl.	B	V	J	K
WASP-43	10:19:38.0	−09:48:22.6	13.8	12.5	10.0	9.3
1	10:19:23.6	−09:36:24.9	13.1	12.5	11.4	11.0
2	10:19:30.7	−09:50:58.2	13.3	12.7	11.6	11.3
3	10:20:03.3	−09:34:16.3	13.3	12.8	11.6	11.2
4	10:18:55.5	−09:51:00.4	13.9	13.0	11.5	11.0
5	10:19:37.8	−09:32:22.0	14.0	13.2	11.6	11.2
6	10:19:33.5	−09:41:45.9	14.5	13.3	10.8	10.1

2. Observations

2.1. General Setup

We observed four transits of WASP-43b between 2015 and 2018 with the 6.5 m Magellan Baade Telescope and IMACS (Dressler et al. 2006) as part of ACCESS. For this study, we used the IMACS $f/2$ camera, which has a 27.′4 diameter field of view (FoV). With this large FoV, IMACS is able to observe several nearby comparison stars simultaneously to WASP-43 to effectively remove common instrumental and atmospheric systematics. We selected comparison stars less than 0.5 magnitude brighter and 1 magnitude fainter that WASP-43 and closest in $B - V/J - K$ color space, following Rackham et al. (2017). The selected comparison stars are shown in Table 1. We used a custom-designed multislit mask with $12'' \times 20''$ slits for the target and comparison stars. We used a similar calibration mask with $0.''5 \times 20''$ slits for arc lamp wavelength calibrations. We used a 300 line per millimeter grating with a 17.°5 blaze angle for all four data sets to achieve an average resolving power of $R \sim 1200$, or approximately 4.7 Å per resolution element, and access a full wavelength coverage of 4500–9260 Å. In practice, the S/N redward of 9000 Å and blueward of 5300 Å dropped to less than 25% of peak counts. For this reason, we omitted measurements outside of this range for the rest of the study. We omitted data taken at an air mass $Z > 2.0$ and/or during twilight as well.

2.2. Data Collection

We collected the two 2015 data sets on 2015 February 14 (UT 01:03–08:55, 433 science images) and 2015 March 9 (UT 04:35–08:22, 119 science images), collected a 2017 data set on 2017 April 10 (UT 00:35–03:35, 197 science images), and a final 2018 data set on 2018 June 3 (UT 23:36–02:29, 156 science images). During the 2017 and 2018 nights of observing, we introduced a blocking filter to reduce contamination from light at higher orders while also truncating the spectral range to 5300–9200 Å. We made observations in Multi-Object Spectroscopy mode with 2×2 binning in TURBO readout mode (30 s) for the 2015 data sets and in 2×2 binning in Fast readout mode (31 s) for the 2017 and 2018 data sets to take advantage of the reduced readout noise. During the observations, we adjusted the individual exposure times between 20 and 60 s to keep the number of counts per pixel roughly between 30,000

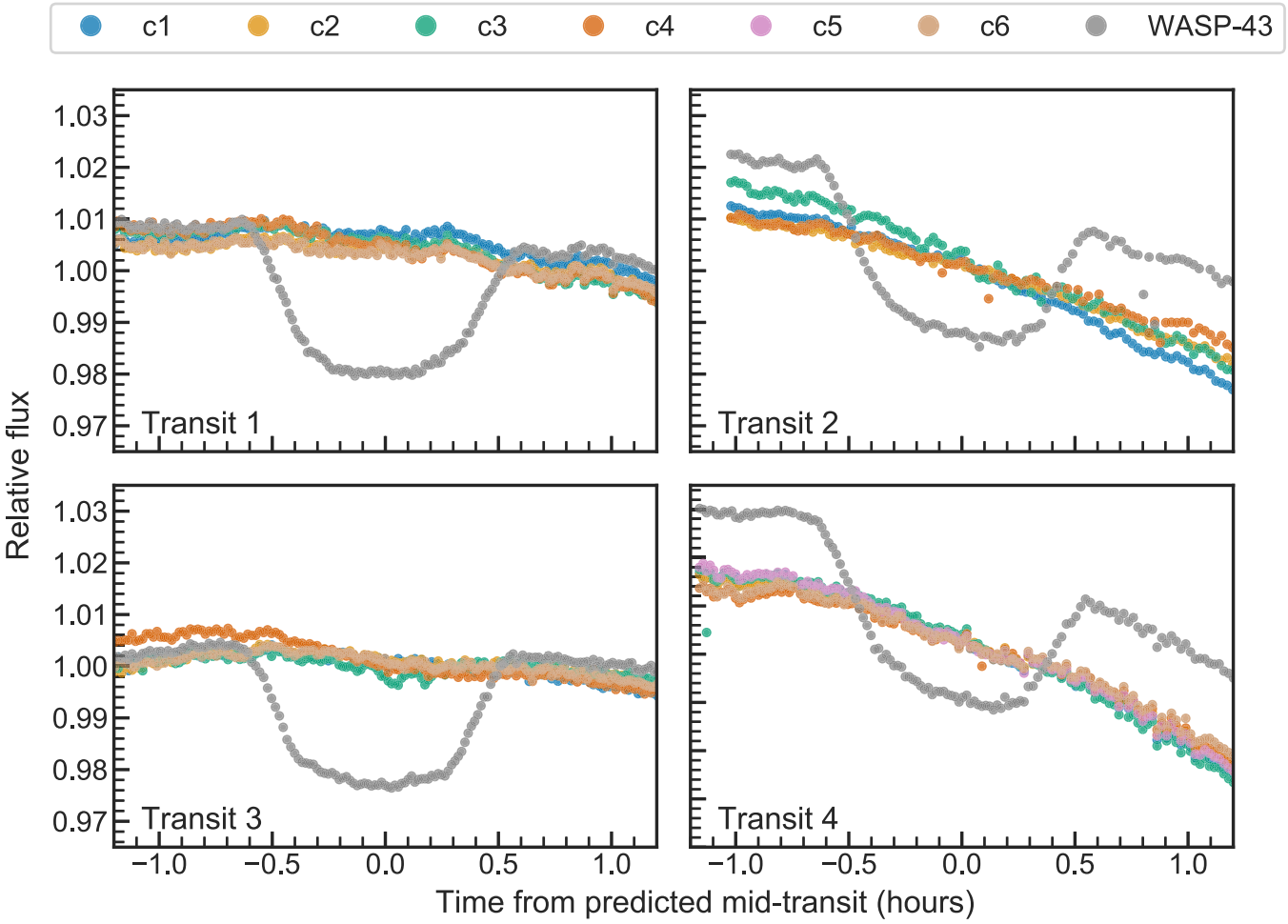

Figure 1. Raw integrated white-light-curve flux of WASP-43 (gray) and comparison stars (color) observed with IMACS, centered 1 hr around the predicted midtransit time. We calculated the predicted midtransit times with Swarthmore College's online transit finding tool (https://astro.swarthmore.edu/transits.cgi).

and 35,000 counts (ADU; gain = $1e^{-}$/ADU on $f/2$ camera), i.e., within the linearity limit of the CCD (Bixel et al. 2019).

With the calibration mask in place, we took a series of wavelength calibration arcs using a HeNeAr lamp before each transit time-series observation. The narrower slit width of the calibration mask increased the spectral resolution of the wavelength calibration as well as avoided saturation of the CCD from the arc lamps. We took a sequence of high-S/N flats with a quartz lamp through the science mask to characterize the pixel-to-pixel variations in the CCD. We ended up not applying a flat-field correction to the science images after finding in all previous ACCESS studies (Rackham et al. 2017; Bixel et al. 2019; Espinoza et al. 2019) that flat fielding introduces additional noise in the data and does not improve the final results.

3. Data Reduction and Light-curve Analysis

3.1. Reduction Pipeline

We reduced the raw data using the ACCESS pipeline described previously (Rackham et al. 2017; Bixel et al. 2019; Espinoza et al. 2019). The detailed functions of the pipeline, including standard bias and flat calibration, bad pixel and cosmic-ray correction, sky subtraction, spectrum extraction, and wavelength calibration are described in detail in Jordán et al. (2013) and Rackham et al. (2017). We briefly summarize the data reduction here.

We applied the wavelength solution found with the arc lamps to the first science image, and the remaining science image spectra were then cross-correlated with the first's. We calibrated the spectra from all stars to the same reference frame by identifying shifts between the Hα absorption line minimum of the median spectra and air wavelength of Hα, and interpolated the spectra onto a common wavelength grid using b-splines. We aligned all spectra to within 2 Å, which is less than the average resolution element of 4.7 Å, assuming an average resolution of $R = 1200$ and shortest wavelength coverage of 5600 Å on IMACS. We also subtract scattered light within IMACS along each slit for every exposure, as described by Espinoza et al. (2019).

The final results are sets of wavelength-calibrated and extracted spectra for the target and each comparison star that can be used to produce integrated (white light; Figure 1) or spectroscopically binned light curves. The series of white-light curves (WLCs) produced in this fashion informed which comparison stars to omit in the rest of the analysis on a per data set basis. Based on the deviations of each comparison star's flux from the general trend of WASP-43b's flux in Figure 1, we omitted comparison star 5 from the first 2015 data set (ut150224, Transit 1), comparison stars 5 and 6 from the second 2015 data

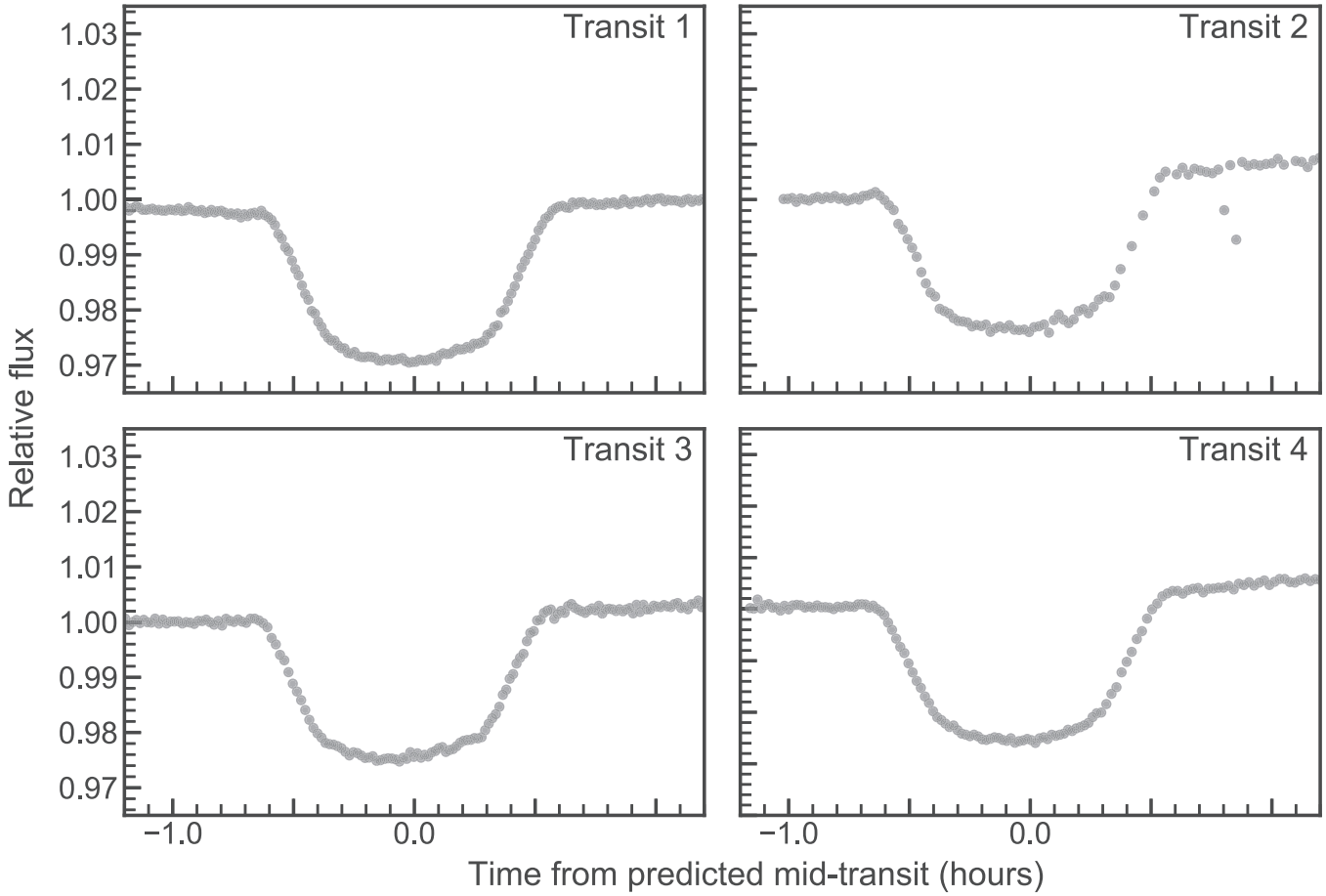

Figure 2. Raw white-light curve of WASP-43b divided by the sum of good comparison star fluxes each night. This informed which comparison stars and data points to omit from the study.

set (ut150309, Transit 2), comparison star 5 from the 2017 data set (ut170410, Transit 3), and no comparison stars from the 2018 data set (ut180603, Transit 4).

With this set of good comparison stars established, we made a normalized white-light curve ($F_{\rm div}$) for each observation by dividing the target WASP-43b flux $F_{\rm W43}$ by the sum of the N good comparison star fluxes (F_{C_n}),

$$F_{\rm div} = \frac{F_{\rm W43}}{\sum_n^N F_{C_n}}, \quad (1)$$

to produce the curves shown in Figure 2 for each data set. We omitted the remaining outliers (points with fluxes deviating at least 2σ from the median flux of 10 neighboring points) from the rest of our detrending procedures. Because the different wavelength bin schemes we explored (see Section 3.2.4) sometimes had different outlier points in their respective divided white-light curves, we omitted a superset of outlier points from all data sets to maintain uniformity and reduce systematics.

We detrended our original data, with the selected data points and comparison stars for each transit epoch selected, using three different methods described in Section 3.2.

3.2. Light-curve Analysis

We applied three detrending and transit-fitting methods to our white light and wavelength-binned data to test our results and verify that they were not dependent on the method used. The three detrending methods we used were polynomial wavelet (poly), polynomial wavelet followed by a common-mode correction (poly+CMC), and a Gaussian process combined with principal component analysis (GP+PCA).

3.2.1. Polynomial Wavelet Detrending

We performed a simultaneous transit model and systematics detrending fit on each of the transit white-light curves shown in Figure 2 using our Markov Chain Monte Carlo (MCMC) code, described in detail in Rackham et al. (2017). To briefly summarize, the divided light-curve model $F(t, \boldsymbol{\theta})$ can be written as

$$F(t, \boldsymbol{\theta}, \boldsymbol{\alpha}) = f(\boldsymbol{\theta}) P(t, \boldsymbol{\alpha}), \quad (2)$$

where t is the time, $f(\boldsymbol{\theta})$ is the analytic transit model described in Mandel & Agol (2002), $\boldsymbol{\theta}$ is the vector of the orbital and transit parameters (e, $a/R_{\rm s}$, i, ω, T, t_0, b, u_1, u_2, $R_{\rm p}/R_{\rm s}$), defined in Table 2, and $\boldsymbol{\alpha}$ is a vector of polynomial coefficients (α_0, α_1, ..., α_m), where

$$P(t, \boldsymbol{\alpha}) = \sum_{m=0}^{M} \alpha_m t^m \quad (3)$$

for an Mth-order polynomial that we fit to the out-of transit (OOT) flux of $F_{\rm div}$. We assumed a quadratic stellar limb-darkening profile and sampled the limb-darkening coefficients u_1

Table 2
Literature System Parameters

Bin Scheme	1	2	3
Eccentricity (e)	0	0	0
Semimajor axis/stellar radius (a/R_s)	4.867 ± 0.023	4.752 ± 0.066	4.872
Inclination (radians) (i)	1.426 ± 0.0056	1.433 ± 0.00175	1.433
Planet/stellar radius uncertainty (σ_{R_p/R_s})	0.0018	0.00145	0.00043[a]
Longitude of periastron (ω)	π	π	π
Period (days) (T)	$0.813473978 \pm 3.5 \times 10^{-8}$	$0.81347459 \pm 2.1 \times 10^{-7}$	0.81347436
Reference	Hoyer et al. (2016)	Murgas et al. (2014)	Kreidberg et al. (2014)

Notes. Literature system parameters corresponding to each bin scheme.
[a] We used σ_{R_p/R_s} from Hoyer et al. (2016) combined transit data because they are based on ground-based values while values from Kreidberg et al. (2014) are space based.

and u_2 according to Kipping (2013) to allow uninformative (uniform) priors to be placed on the transformed coefficients while avoiding the risk of sampling nonphysical values. We fit the divided light-curve model to determine the most likely values for $\boldsymbol{\alpha}$. We obtained the final detrended model ($F_{\rm det}$) by dividing through by the OOT best-fit flux model $P(t, \boldsymbol{\alpha})$, such that

$$F_{\rm det} = \frac{F_{\rm div}}{P(t, \boldsymbol{\alpha})}. \quad (4)$$

We performed a simultaneous MCMC fitting of transit parameters with `PyMC` (Salvatier et al. 2016). The likelihood was determined through the wavelet method described in Carter & Winn (2009). The fitted parameters include the midtransit time (t_0), planet-to-star radius ratio (R_p/R_s), three coefficients (a_i, $i = 0$, 1, 2) for the second-order polynomial used to fit the baseline OOT trend, two parameters for the transformed quadratic limb-darkening coefficients (q_1, q_2), one noise parameter (σ_w) for uncorrelated "white" noise, and one parameter (σ_r) for correlated "red" noise. We found that a second-order polynomial fit the OOT flux better than a lower order function.

We sampled t_0, σ_w, σ_r, q_1, and q_2 with uniform priors; α_i with a Gaussian prior with a width set by bootstrapping the uncertainty on α_i following Rackham et al. (2017); and R_p/R_s with a Gaussian prior with spread $5 \times \sigma_{R_p/R_s}$, where R_p/R_s and $\boldsymbol{\sigma}_{\boldsymbol{R}_p/\boldsymbol{R}_s}$ are taken from the literature. We used five chains, each composed of 100,000 steps, and started them at the estimated location of maximum a posteriori probability, using an additional 30,000 steps for burn in. We thinned the chains by sampling them at 10× their autocorrelation function half-life before combining them to produce the final posterior distributions. Each bin scheme has a corresponding source for the fixed system parameters, and we list them in Table 2. We fit the white-light curve while keeping the midtransit time t_0 and average transit depth R_p/R_s free.

We produced the binned light curves following the same procedure for producing the WLCs, with the only difference being that we kept the midtransit time t_0 found in the WLC analysis fixed when performing the fits for R_p/R_s in each wavelength bin for bin schemes 1, 2, and 3, described in Section 3.2.4.

3.2.2. Common-mode Correction

Following Sedaghati et al. (2016), we divided the detrended WLC obtained in the polynomial detrending method by its best-fit model to produce a common-mode correction (CMC) residual. We then divided that residual through the binned light curves to remove wavelength-independent variations. Next, we applied the CMC correction to each bin scheme.

3.2.3. Gaussian Process and Principal Component Analysis

Gaussian Process (GP) regression is a powerful tool for modeling data in the machine-learning community (Rasmussen & Williams 2005) that has started to gain more popularity in the exoplanets field (see, e.g., Aigrain et al. 2012; Gibson et al. 2012). Gibson et al. (2012) provide a good overview to this methodology applied to exoplanet transit light curves. Applying this methodology for a collection of N measurements ($\boldsymbol{f}$), such as the flux of a star measured over a time series, the log marginal likelihood of the data can be written as

$$\log \mathcal{L}(\boldsymbol{r}|\mathsf{X}, \boldsymbol{\theta}, \boldsymbol{\phi}) = -\frac{1}{2}\boldsymbol{r}^{\top}\Sigma^{-1}\boldsymbol{r} - \frac{1}{2}\log|\Sigma| - \frac{N}{2}\log(2\pi), \quad (5)$$

where $\boldsymbol{r} \equiv \boldsymbol{f} - T(\boldsymbol{t}, \boldsymbol{\phi})$ is the vector of residuals between the data and analytic transit function T; X is the $N \times K$ matrix for K additional parameters, where each row is the vector of measurements $\boldsymbol{x}_n = (x_{n,1}, \cdots x_{n,K})$ at a given time n; $\boldsymbol{\theta}$ are the hyperparameters of the GP; $\boldsymbol{\phi}$ are the transit model parameters; and Σ is the covariance of the joint probability distribution of the set of observations $\boldsymbol{f}$. In our analysis, we used six systematics parameters: time, FWHM of the spectra on the CCD, air mass, position of the pixel trace through each spectra on the chip of the CCD, sky flux, and shift in wavelength space of the trace. We used the Python package `batman` (Kreidberg 2015) to generate our analytic transit model. From here, the log posterior distribution $\log \mathcal{P}(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{f}, \mathsf{X})$ can be determined by placing explicit priors on the maximum covariance hyperparameters and the scale-length hyperparameters. From $\mathcal{P}$, the transit parameters can then be inferred by optimizing with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$.

We accomplished the above optimization problem with the Bayesian inference tool `PyMultiNest` (Buchner et al. 2014) and computed the log likelihoods from the GP with the `george` (Ambikasaran et al. 2014) package. We implemented this detrending scheme by simultaneously fitting the data with an exponential squared kernel for the GP under the assumption

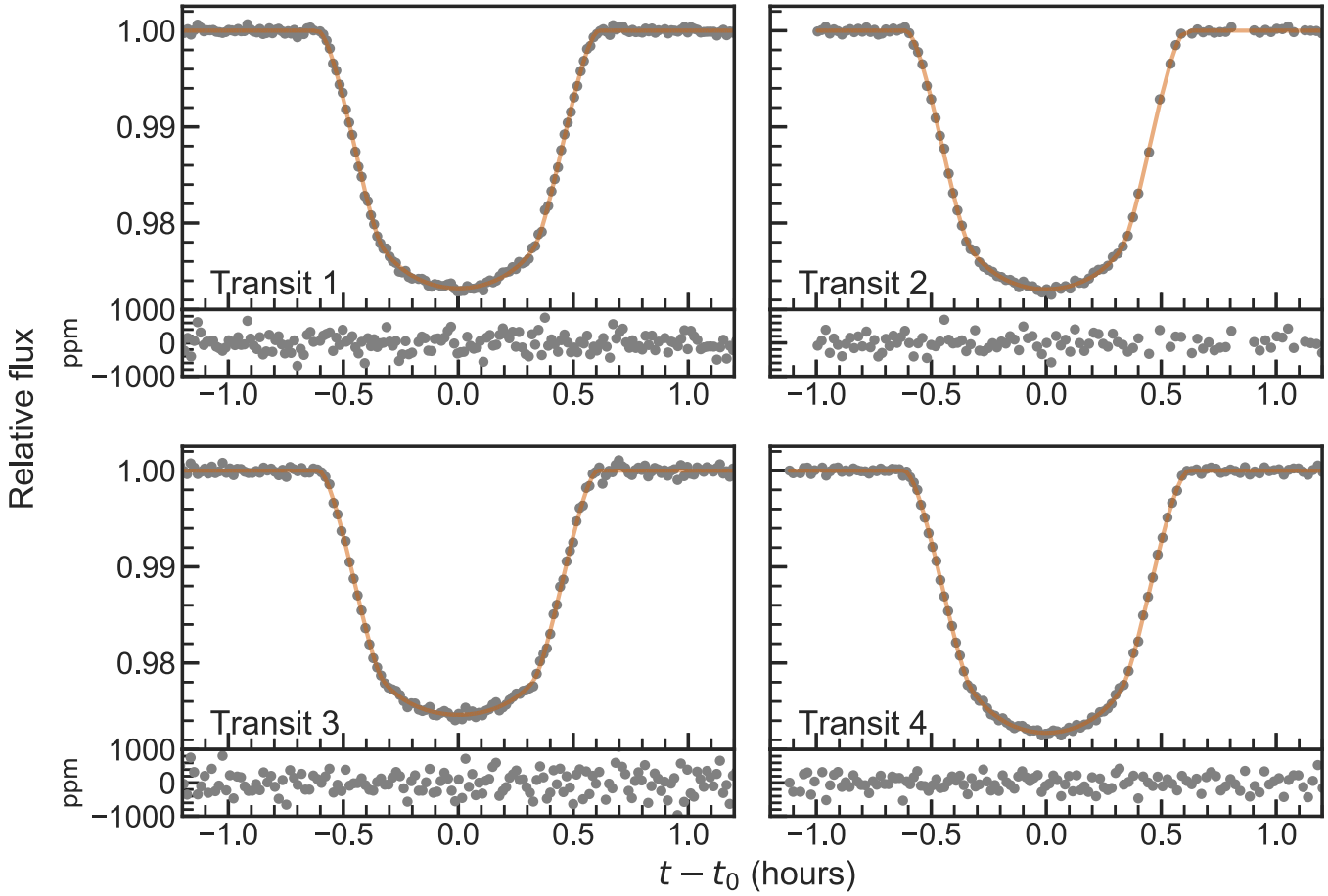

Figure 3. GP+PCA detrended white-light curves (gray) and associated models (solid line) for a representative wavelength coverage of 5300–9050 Å. We fit for the midtransit time t_0 each transit and used it to perform binned light-curve fitting later in the analysis. We center our data 1 hr around t_0. The parameters for each transit fit are given in Table 3. The associated corner plots are shown in the Appendix.

that points closer to each other are more correlated than points farther apart. The PCA methodology follows from Jordán et al. (2013) and Espinoza et al. (2019), where M signals, $S_i(t)$, can be extracted from M comparison stars and linearly reconstructed according to the eigenvalues, λ_i, of each signal. This allows for the optimal extraction of information from each comparison star to inform how the total flux of WASP-43 varies over the course of the night. We Bayesian model averaged the principal components together, which were determined by fitting with one, then two, up to M principal components, to create the final detrended WLC and model parameters of interest. We present the associated WLCs in Figure 3, the best-fit parameters in Table 3, and the associated corner plots in the Appendix. Based on the quality of the fits, discussed in Section 5.4, and to streamline our work, we show only the results of this method for the transmission spectra that informed our retrieval analysis.

We applied the same methodology on a wavelength bin by wavelength bin basis to produce the simultaneously fitted light curves for bin scheme 3 and create our final transmission spectrum, shown in Figure 8. We used the open source package `ld-exosim`[20] to determine that a square-root limb-darkening law was the most appropriate for WASP-43 and incorporated this into our GP analysis. We present the final transmission spectrum using this method in Figure 8.

[20] https://github.com/nespinoza/ld-exosim

3.2.4. White-light Curve and Binning Schemes

We applied the detrending methods described in Sections 3.2.1–3.2.3 to the following wavelength-binning schemes in our analysis:

1. Bin scheme 1: a set of uniform bins centered around the air wavelength values of key spectral features (Na I-D, Hα, K I, Na I-8200) to produce a transmission spectrum focused on these features.
2. Bin scheme 2: binning and system parameters identical to those in Murgas et al. (2014) to directly compare our transmission spectra with the ones presented in that study.
3. Bin scheme 3: similar 230 Å binning and system parameters to Kreidberg et al. (2014) to combine our visual measurement with their NIR measurements made with *HST*. We used this binning scheme to perform atmospheric retrievals described in Section 5.4. We centered 230 Å bins on the vacuum wavelength locations of Na I-D, Hα, K I, and Na I-8200 using smaller bins when necessary to have at least two wavelength bins between each feature.

We applied the poly+CMC detrending method described in Section 3.2.1 to all three bin schemes and applied GP+PCA described in Section 3.2.3 to bin scheme 2 (25 nm) and bin scheme 3 because of their similar wavelength binning and wavelength coverage in the visual (25 nm versus 23 nm). This allowed us to directly compare detrending methods between our study and Murgas et al. (2014).

Table 3
Fitted WLC Values from GP+PCA Detrending Method Shown in Figure 3

Parameter	Definition	Transit 1	Transit 2	Transit 3	Transit 4
R_p/R_s	Planet radius/star radius	$0.15854^{+0.00079}_{-0.00074}$	$0.15800^{+0.00218}_{-0.00278}$	$0.15436^{+0.00167}_{-0.00164}$	$0.16030^{+0.00127}_{-0.00118}$
d	Transit depth (ppm)	25134 ± 250	24963 ± 880	23828 ± 516	25695 ± 406
$t_0 - 2450000$	Midtransit (JD)	$7077.72325^{+0.00004}_{-0.00004}$	$7090.73888^{+0.00008}_{-0.00008}$	$7854.59100^{+0.00008}_{-0.00008}$	$8273.53019^{+0.00005}_{-0.00005}$
P	Period (days)	$0.81347^{+0.00000}_{-0.00000}$	$0.81347^{+0.00000}_{-0.00000}$	$0.81347^{+0.00000}_{-0.00000}$	$0.81347^{+0.00000}_{-0.00000}$
a/R_s	Semimajor axis/star radius	$4.92738^{+0.02856}_{-0.02850}$	$4.97737^{+0.05644}_{-0.06297}$	$4.90935^{+0.05881}_{-0.05815}$	$4.85836^{+0.03458}_{-0.03367}$
b	Impact parameter	$0.65644^{+0.00698}_{-0.00720}$	$0.65270^{+0.01433}_{-0.01372}$	$0.67153^{+0.01335}_{-0.01576}$	$0.66218^{+0.00822}_{-0.00818}$
i	Inclination	$82.34509^{+0.11731}_{-0.11874}$	$82.46517^{+0.23358}_{-0.25848}$	$82.13717^{+0.26874}_{-0.24363}$	$82.16642^{+0.14519}_{-0.14579}$
q_1	LD coeff 1	$0.70515^{+0.19778}_{-0.23016}$	$0.60216^{+0.23895}_{-0.17155}$	$0.50975^{+0.30047}_{-0.19881}$	$0.72406^{+0.18106}_{-0.19668}$
q_2	LD coeff 2	$0.38994^{+0.10395}_{-0.19375}$	$0.30680^{+0.17608}_{-0.18818}$	$0.41212^{+0.21415}_{-0.24787}$	$0.31602^{+0.10556}_{-0.16639}$

Notes. We share the associated corner plots in the Appendix. Note: we computed transit depths directly from R_p/R_s.

We discuss each binning scheme in more detail in Sections 5.2, 5.3, and 5.4. We also include the literature values for system parameters used in each bin scheme in Table 2. For bin scheme 1, where no associated literature values are being used for comparison, we adopt the most up-to-date values from Hoyer et al. (2016).

4. Stellar Activity

Before combining the transmission spectra from each night, we first considered the impact of stellar photospheric heterogeneity, which can have an observable effect on transmission spectra (Pont et al. 2008, 2013; Sing et al. 2011; Oshagh et al. 2014; Zhang et al. 2018), even if magnetically active regions are not occulted by the transiting exoplanet (McCullough et al. 2014; Rackham et al. 2018, 2019; Apai et al. 2018). Qualitatively, global variations in stellar activity could manifest themselves as an overall dimming or brightening of the star, which could lead to significant variations in transit depths. Changes in photometric activity roughly correlate with the covering fraction of starspots, which in turn can modulate the luminosity of the star and impact observed transit depths (Berta et al. 2011). Furthermore, those variations can be wavelength dependent, leading to slopes with spurious spectral features in the transmission spectrum.

The white-light transit depths we observed (Table 3) varied by as much as 1869 ppm between transits. To account for this offset between the data sets, we investigated the contribution due to stellar activity. Changes in photometric activity roughly correlate with the covering fraction of starspots, which in turn can modulate the luminosity of the star and impact observed transit depths (Berta et al. 2011). To assess the brightness variation of WASP-43 over the time frame of our observations, we used 2012 February 15–2018 May 21 activity data from 990 out-of-transit V-band images of WASP-43b taken by the Ohio State University's All-Sky Automated Survey for Supernovae[21] (ASAS-SN) program (Shappee et al. 2014; Kochanek et al. 2017).

The ASAS-SN photometric activity was sampled much more coarsely than our transit observations, so we used a regression routine (Alam et al. 2018) to fit the data and estimate the amplitude of the photometric variation induced by stellar activity during each of the four transit epochs of WASP-43b. Following Alam et al. (2018), we used a negative log-likelihood kernel ($\ln \boldsymbol{L}$) for the objective function given by

$$\ln(\boldsymbol{L}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln(\det \boldsymbol{K}) - \frac{1}{2}(\boldsymbol{y} - \mu)^{\top}\boldsymbol{K}^{-1}(\boldsymbol{y} - \mu), \quad (6)$$

where $\boldsymbol{y}$ is the data, μ is the model, n is the number of observations, and $\boldsymbol{K}$ is the covariance matrix. $\boldsymbol{K}$ describes the correlation weight between all possible pairs of photometric measurements and is populated with the GP kernel to quantify the correlation of pairs of observations. We used a gradient-based optimization routine to find the best-fit hyperparameters and used the 15.6 day stellar rotation period from Hellier et al. (2011). Figure 4 shows the GP regression model for the relevant ASAS-SN data.

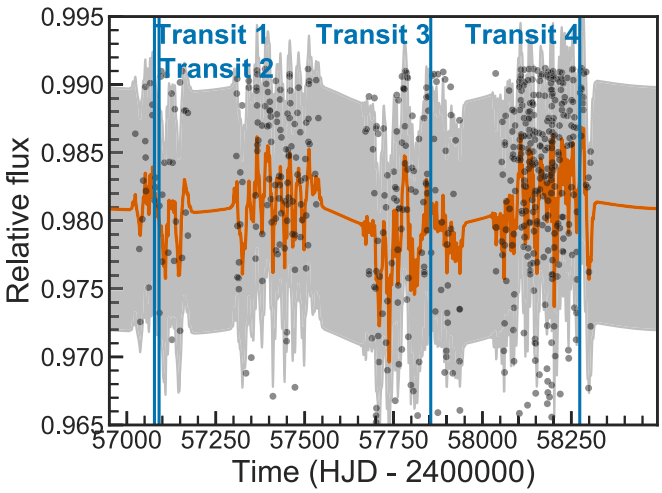

Figure 4. Ground-based photometric observations of WASP-43 from ASAS-SN (gray points) during the transit 1, transit 2, and transit 3 transit epochs (blue vertical lines). The data are flux relative to the average brightness of comparison stars. The Gaussian process regression model (red) and 1σ uncertainty (gray region) fit to ASAS-SN data are also overplotted.

Figure 4 shows the complete ASAS-SN light curve, and Figure 5 details the ASAS-SN photometry near each of our transit epochs. Overall, the relative flux from the photometric monitoring varies by as much as 3% from the median value obtained from the GP. Observations with ASAS-SN are too coarse to effectively sample the photometric activity during

[21] https://asas-sn.osu.edu

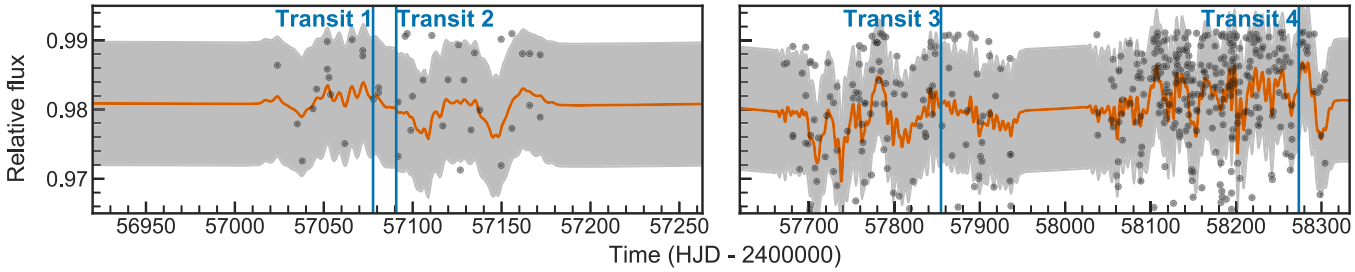

Figure 5. ASAS-SN photometric data from Figure 4 centered on all four transit epochs. Because the data are too sparse to cover even a single transit epoch, we do not use them to quantify the contribution of stellar activity to the transmission spectrum, but instead rely on the formalism discussed in Section 5.4.

times of transits. However, these data still give us a rough idea of differences in stellar flux between epochs.

We conclude that the photometric activity data alone are not enough to constrain the contribution of unocculted heterogeneities on the surface of WASP-43b to the resulting transmission spectrum. Nonetheless, we argue that changes in disk coverage by unocculted heterogeneities likely drive the white-light light-curve depth variations that we observe between transit epochs. For this reason, we calculate and apply transit depth offset corrections as described in Section 5.1 before building the final transmission spectrum. In Section 5.4, we model the possible contribution of an unocculted heterogeneous photosphere to the resulting transmission spectrum without relying on photometric monitoring data.

5. Transmission Spectrum

5.1. Combining Nights

The uncertainties in transit depths from our individual transits range from 250 to 880 ppm, which are not enough to detect the atmosphere of WASP-43b (see Table 3). For example, a hydrogen-dominated composition for the atmosphere of WASP-43b would produce a signal ΔD of 435 ppm at 5 scale heights (based on Equation (11) of Miller-Ricci et al. 2009 and using the planet and star parameters given in Section 1). Therefore, we needed to combine the transmission spectra from the four transits to be sensitive to atmospheric features of the planet.

Transit depths between different epochs varied as much as 1867 ppm (see the difference in depths between transits 3 and 4 in Table 3), an effect that we attribute to stellar variability. We also expect some variability in the observed transit depths for two reasons: (i) stellar activity is stronger in the visual portion of the spectrum relative to in the IR and (ii) because of the large period of time over which we collected our data. For example, Kreidberg et al. (2014) observed six transits over one month, while our data span four transits over 3 yr. This leaves ample time for the star's intrinsic brightness to change, due to surface stellar heterogeneities and impact measured transit depths.

To be able to combine the data from each transit epoch, we needed to first consider the potential effects introduced by the stellar activity of WASP-43, as discussed in Section 4. Typically, studies have used photometric activity as a proxy for the presence of these stellar heterogeneities, but this (compounded with the fact that the photometric activity data we have are not well sampled enough in time to cover the given transits) has been shown to be insufficient to correct for these stellar contributions to transit depth variations (McCullough et al. 2014). Based on the lack of constraints on the contribution of occulted heterogeneities on the stellar photosphere to the transmission spectrum discussed in Section 4, we averaged the transmission spectrum from each night together, weighted by the wavelength-dependent uncertainty estimated from the wavelength-binned fitting. Before taking this weighted average, we first addressed the apparent offset visible in the resulting transmission spectra (colored points in Figure 8). We did this by subtracting the mean white-light transit depth of the four nights from the transmission spectrum of each night. After applying the offset, we combined the four transit epochs by averaging the transmission spectra from each night together, weighted by the uncertainties in the wavelength-dependent depths determined by the fitting. We took the maximum of this asymmetrical uncertainty to be conservative in our weighting. We applied this methodology to each bin scheme identified in the following section. Effectively, retrievals on the resulting transmission spectrum found probe for the average contribution from the stellar photosphere over all transit events.

5.2. Bin Scheme 1: Species-dependent Binning

In this binning scheme, we set the wavelength bin sizes based on the absorption bandwidths of features of interest, in particular Na I-D, Hα, K I, and Na I-8200. We set the minimum bin size for a given feature to be equal to the full width of its observed stellar absorption line, including the contributions from the wings of the line. This gives a set of four bin widths equal to 60, 10, 60, and 40 Å for the respective species listed above. For each species, we mapped a region covering five times its bin width above, centered on the air wavelength to resolve any potential peaks, and we used larger bins to cover the rest of the spectrum. We produce the combined spectra in Figure 6 following this procedure for the (poly+CMC) detrending scheme.

We observed an apparent peak near the Na I-8200 line, but the fact that the only point far from the baseline is also far from the air wavelength for this species indicates that this peak is most likely due to residuals from water tellurics. We also observed a potential K I peak 1σ above the median in bin two, 10 Å immediately redward of its air wavelength location, but believe that this is due to residual telluric absorption as well. Furthermore, we do not detect this peak at all in the GP+PCA detrended data. We also do not detect an absorption peak near Na I-D or Hα. To set this bin scheme apart from the other two bin schemes, we used more up-to-date system parameters from Hoyer et al. (2016).

5.3. Bin Scheme 2: Comparison to Ground-based Study

We applied the same detrending and combining methods using the system parameters reported by Murgas et al. (2014) to compare our transmission spectra with those from their similar study of WASP-43b. We adopted the following four

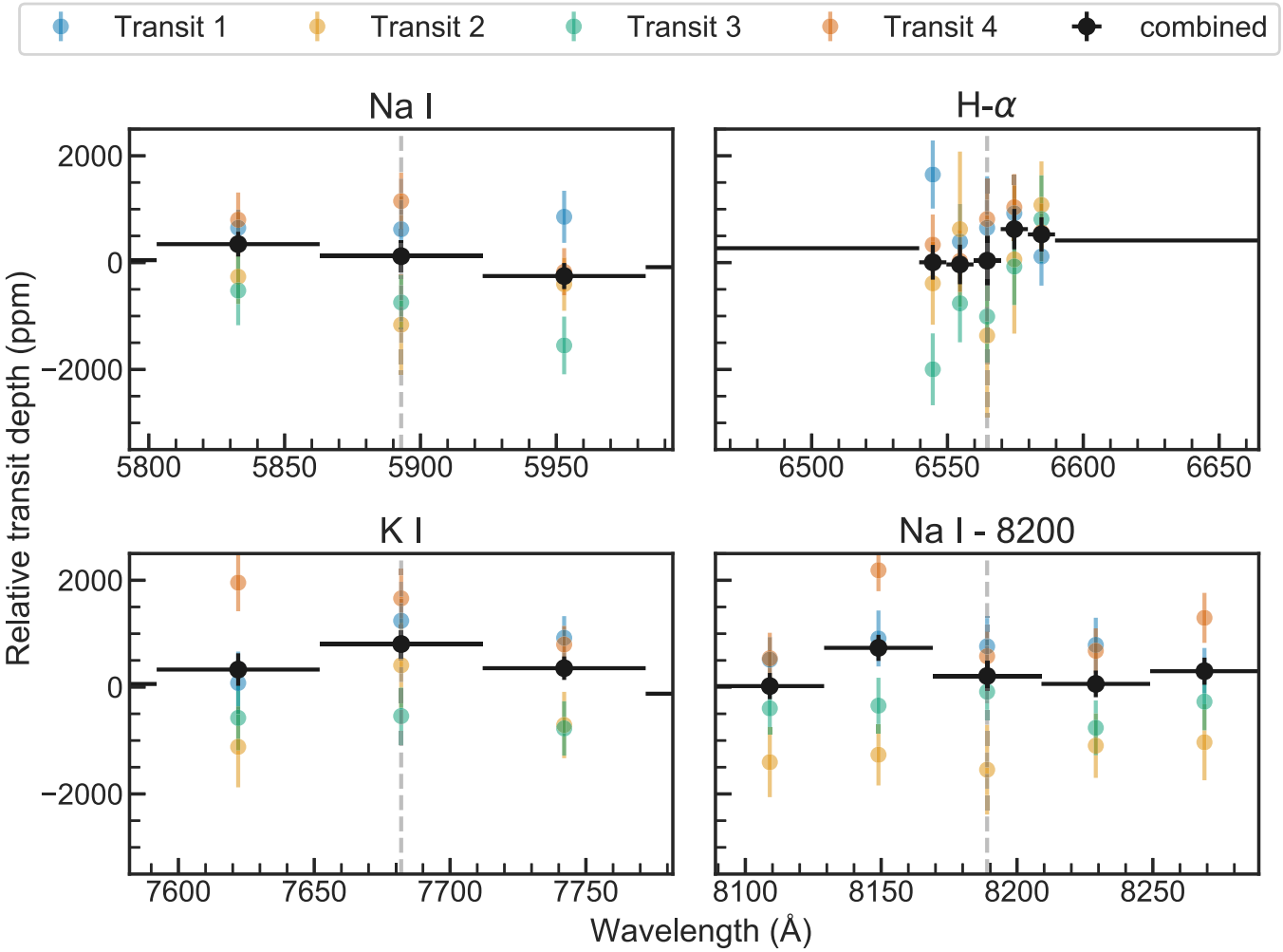

Figure 6. Portions of the transmission spectrum obtained using bin scheme 1, centered on the Na I, K I, and H_α features. Individual nights are shown as colored points; the combined spectrum is shown in black. We made the transmission spectrum sensitive to any potential features that may exist by decreasing the bin size as it approaches the air wavelength of each potential species we searched for. We do not detect an excess in transit depth for any of the species.

wavelength bin schemes identified in their work: (i) 100 Å bins ranging from 5445 to 8845 Å, (ii) 250 Å bins ranging from 5300 to 9050 Å, (iii) 750 Å bins ranging from 5300 to 9050 Å, and (iv) 180 Å bins centered near the K I 7665 and 7699 Å doublet to produce Figure 7.

Qualitatively, the shapes of the combined spectra in each binning scheme tend to follow the same slight upward curving slope near 7000 Å seen in Murgas et al. (2014). Unlike their finding, we do not observe an excess near 5892.9 Å that would indicate the presence of Na I in the atmosphere of WASP-43b.

5.4. Bin Scheme 3: Combining with NIR Study

We used the same bin width and system parameters (Table 2) from Kreidberg et al. (2014) to combine their NIR transmission spectrum with our visual spectrum. We found that although the polynomial+CMC detrending method tended to produce smaller error bars on average in the transmission spectrum (180 ppm versus 210 ppm), GP+PCA does a better job overall at fitting systematics in the light curves, resulting in an average standard deviation across the transit epochs of 396 ppm versus 274 ppm in the WLC model residuals. The average GP+PCA WLC residuals were also larger than the average white noise of 146 ppm estimated in the poly+CMC detrended WLCs, so the GP+PCA procedure was not overfitting the data. For these reasons, we adopted the GP +PCA detrended transmission spectrum, shown as black circles in Figure 8, as our final spectrum (reported in Table 7), which we combined with the NIR transmission spectrum from Kreidberg et al. (2014) for the atmospheric retrieval analysis.

6. Atmospheric Retrieval Analysis

With the final combined visual+NIR transmission spectrum from Section 5.4, we searched for signals in the atmosphere of WASP-43b. We used a Bayesian atmospheric retrieval code based on the same nested sampling Bayesian inference software we used to perform the GP+PCA detrending, `PyMultiNest`. The details of the retrieval code are given in our previous study of WASP-19b (Espinoza et al. 2019), and we only briefly summarize the methodology here.

Following the semianalytical formalism from Bétrémieux & Swain (2017) and Heng & Kitzmann (2017), we assume an isothermal and isobaric atmosphere, with an optically thick base region with radius $(R_p/R_s)_0$ and reference pressure P_0, which we interpret as the cloud-top pressure. Above this region is an optically thin planetary atmosphere with average temperature T that can have either (i) a set of atomic and molecular species and/or (ii) a scattering haze defined by

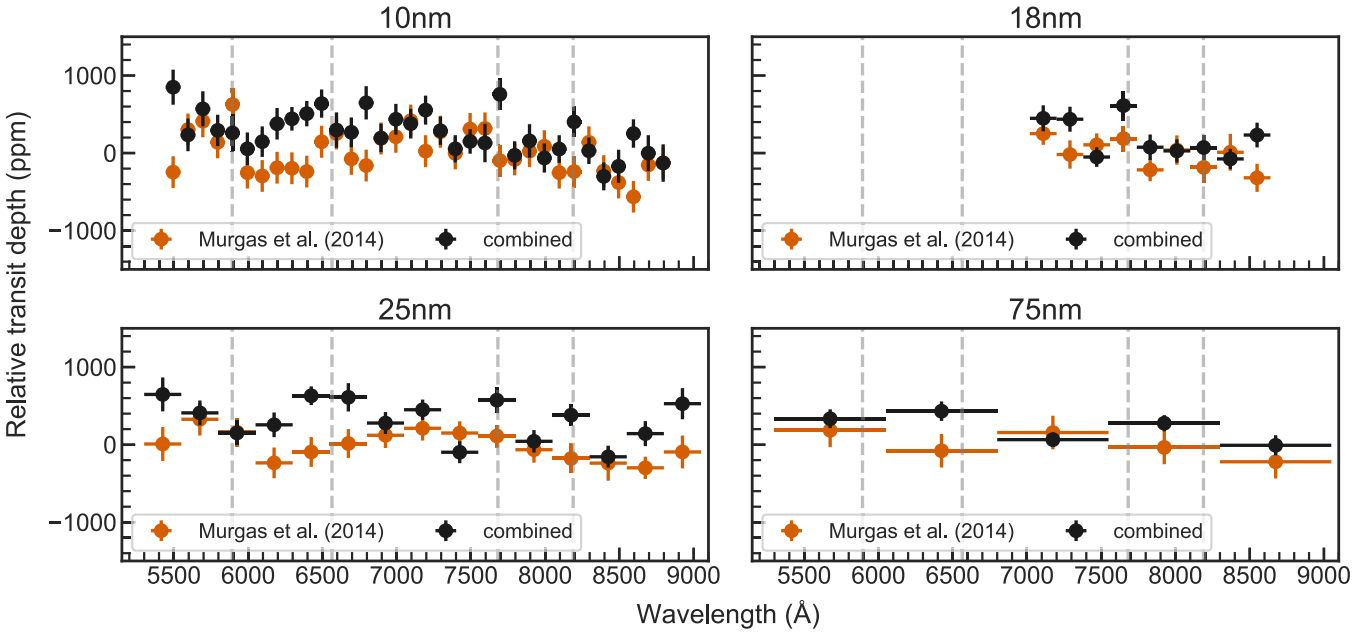

Figure 7. Combined poly+CMC detrended transmission spectrum comparing the four Murgas et al. (2014) binning schemes against our own spectrum (black). We were unable to find an associated data table for the 10 and 75 nm bins so we used a manual online digitizer instead (https://apps.automeris.io/wpd/). There appears to be a peak near K I, but we consider this to be a spurious detection. For example, we do not observe this peak in the GP+PCA detrended data. From left to right, the vertical dashed lines mark the air wavelength locations of Na I, Hα, K I, and K I-8200, respectively.

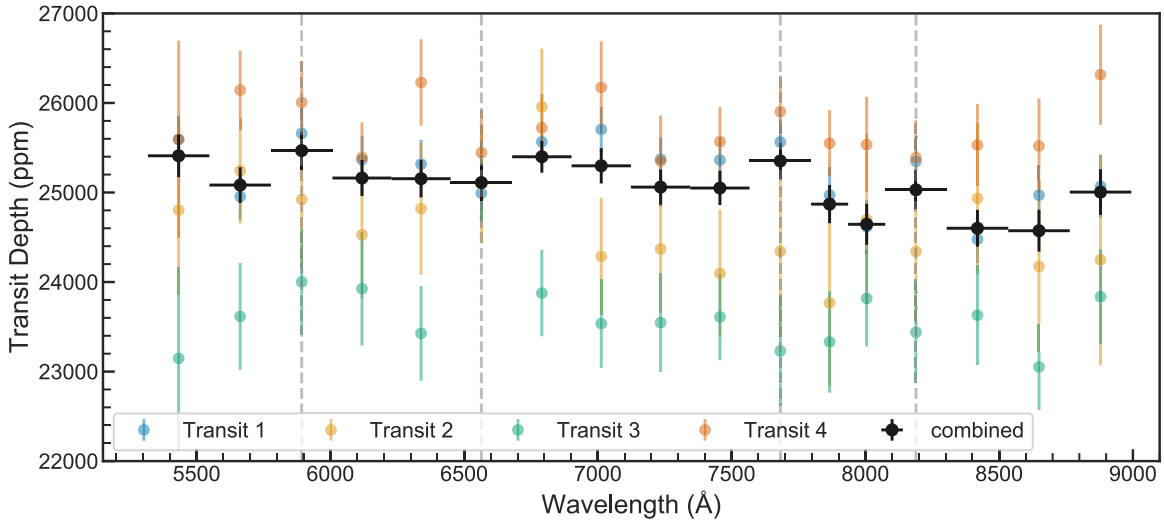

Figure 8. GP+PCA detrended transmission spectrum of bin scheme 3. The transit depths are all relative to the weighted mean of the white-light depths. From left to right, the vertical dashed lines mark the air wavelength locations of Na I, Hα, K I, and K I-8200, respectively. We share the table for the above data in Table 7 and associated binned light curves in the Appendix.

$\sigma_{\rm haze}(\lambda) = a\sigma_0(\lambda/\lambda_0)^{\gamma_{\rm haze}}$ (MacDonald & Madhusudhan 2017), where $\sigma_0 = 5.31 \times 10^{-27}$ cm^2 is the Rayleigh scattering cross section of H_2 at the reference wavelength $\lambda_0 = 350$ nm, and a and $\gamma_{\rm haze}$ are free parameters. We constrain $\gamma_{\rm haze}$ to be between 0 (uniform opacity) and -4 (Rayleigh scattering) to allow for a better constraint on a. Transmission spectra from separate studies can be combined by retrieving for an offset between the different data sets. A detailed overview of the retrieval framework is given in Appendix D of Espinoza et al. (2019).

Additionally, we explored the impact of a heterogeneous stellar photosphere on the observed transmission (Pinhas et al. 2018) by following the formalism described by Rackham et al. (2018, 2019). To summarize directly from the schematic in Figure 1 of Rackham et al. (2018): during a transit, exoplanet atmospheres are illuminated by the portion of a stellar photosphere immediately behind the exoplanet (the transit chord). Changes in transit depth must be measured relative to the spectrum of this light source. However, the light source is generally assumed to be the disk-integrated spectrum of the star. Any differences between the assumed and actual light sources will lead to apparent variations in transit depth.

In this framework, stellar contamination of the transmission spectrum from unocculted starspots and faculae is considered by placing constraints on the allowed spot- and faculae-covering fractions using a set of rotating photosphere models and then translating the covering fractions of potential stellar

Table 4
Priors Used in Retrieval Models

Model Component	Parameter	Units	Description	Prior Distribution
Offset	offset	ppm	Offset between Magellan/IMACS and *HST*/WFC3 data	Normal (Mean Depth, 1000 ppm)
Base	$(R_p/R_s)_0$	⋯	Radius corresponding to the top of the cloud layer or $\tau \gg 1$	Uniform (0.8, 1.2)
	P_0	bar	Reference pressure at $(R_p/R_s)_0$	Log uniform (10^{-4}, 1)
	T	K	Average temperature planet atmosphere	Uniform (0, 1500)
Atomic features	X	⋯	Mixing ratio of species X	Log uniform (10^{-14}, 1)
Haze	a	⋯	Amplitude of the haze cross-section power law	Log uniform (10^{-10}, 10^{20})
	γ	⋯	Index of the haze cross-section power law	Uniform (−10, 0)
Stellar photosphere	T_{occ}	K	Average temperature of the transit chord	Uniform (4000, 5000)
	T_{het}	K	Average temperature of the heterogeneous surface features	Uniform (4000, 5000)
	F_{het}	⋯	Fraction of the unocculted photosphere covered by spots	Uniform (0, 1)

contamination in the transmission spectrum. We incorporate this transit light-source effect into our retrieval framework with a three-parameter model for the stellar photosphere to fit this simultaneously with the planet's atmosphere. The three parameters are T_{chord}, the effective temperature of the transit chord; T_{het}, the mean effective temperature of the heterogeneous features not occulted by the transit chord; and f_{het}, the fraction of the projected stellar disk covered by these heterogeneous features. The impact of these heterogeneities on the transmission spectrum is expressed by the wavelength-dependent corrective factor ϵ_λ on the transit depth, where

$$\epsilon_\lambda \equiv \left[1 - f_{het}\left(1 - \frac{S_{\lambda,het}}{S_{\lambda,chord}}\right)\right]^{-1} = \left(\frac{R_{p,\lambda}}{R_{p,\lambda,0}}\right)^2. \quad (7)$$

Here, $R_{p,\lambda}$ and $R_{p,\lambda,0}$ are the apparent and actual planetary radius measured at wavelength λ, respectively; $S_{\lambda,het}$ is the spectrum of the unocculted photosphere determined by T_{het}; and $S_{\lambda,chord}$ is the spectrum of the portion of the photosphere inside of the transit chord. Following previous studies (e.g., McCullough et al. 2014; Rackham et al. 2017), we use PHOENIX stellar spectra (Husser et al. 2013) to model the emergent spectra of the photospheric components.

As Rackham et al. (2018) note, this formalism assumes that the transit chord can be described by a single emergent spectrum. Although this is not guaranteed for any one of our transits, they note that this formalism also holds for transits in which an occulted spot- or faculae-crossing event is present in the transit signal above the observational uncertainty and taken into account in the transit modeling. We also explored more complex models including multiple spot- and faculae-covering fractions, but we found that the data did not warrant the additional complexity of such a model.

Our combined retrieval approach uses the posterior Bayesian evidence $Z \equiv \mathbb{P}(D|H)$ computed by `PyMultiNest`, which is the probability of the data D given the hypothesis H to perform model comparisons. This property of the nested sampling algorithm allows us to study how complex our models have to be to explain the observed distortions to the light curve (such as number of spots) via the posterior odds, $\mathbb{P}(H_n|D)/\mathbb{P}(H_k|D)$, where the joint probability $\mathbb{P}(H_n|D) = P(D|H_n)\mathbb{P}(H_n)$, with $\mathbb{P}(H_n)$ being the prior probability on the hypothesis H_n. If we approximate model n and model k as having the same prior distribution on their respective hypothesis H_n, H_k, then the posterior odds simplify to just the ratio of the evidences,

$$\frac{\mathbb{P}(H_n|D)}{\mathbb{P}(H_k|D)} = \frac{\mathbb{P}(D|H_n)}{\mathbb{P}(D|H_k)} \equiv \frac{Z_n}{Z_k}. \quad (8)$$

In log space, this is the difference of the logs of each evidence and is denoted as $\Delta \ln Z \equiv \ln Z_n - \ln Z_k$. In the context of this study, $\Delta \ln Z$ is a measure of how statistically different a given model (Z_n) is from a flat atmospheric model ($Z_k \equiv Z_{flat}$). Trotta (2008) and Benneke & Seager (2013) review how these log odds translate to frequentist significance hypothesis testing. We note from that work that absolute log odds below 1 are usually considered inconclusive, near 2.5 can be interpreted as moderate evidence, and higher than 5 can be interpreted as highly significant. It is important to caution though that frequentist hypothesis testing has only one null hypothesis, whereas proper Bayesian model comparison considers a range

Table 5
$\Delta \ln Z$ for Various Models Relative to a Flat Spectrum for the ACCESS GP +PCA Detrended Data, Combined with the Kreidberg et al. (2014) Data

Model	H_2O	$Na+H_2O$	$Na+K+H_2O$
clear	4.98	6.42	6.08
haze	5.43	5.76	5.60
spot	8.26	7.90	7.74
spot+haze	6.92	6.70	6.59

Note. The largest value of 8.26 corresponds to the best retrieved transmission spectrum model shown in Figure 9.

Table 6
Retrieved Parameters for the Best-fit Retrieval Model Shown in Figure 9

Parameter	Description	Value
F_{het}	Spot-covering fraction	$0.27^{+0.42}_{-0.16}$
T	Planet temperature (K)	$352.91^{+206.14}_{-125.08}$
T_{het}	Spot temperature (K)	$4169.40^{+86.10}_{-92.33}$
T_{occ}	Occulted temperature (K)	$4300.91^{+94.04}_{-77.16}$
$\log H_2O$	Water volume mixing ratio	$-2.78^{+1.38}_{-1.47}$
$\log P_0$	Reference pressure (bar)	$-1.67^{+1.08}_{-1.16}$
$\text{offset}_{\text{Magellan/IMACS}}$	Offset between data sets (ppm)	$25749.40^{+101.54}_{-105.70}$
f	Planet radius normalization	$0.99^{+0.00}_{-0.00}$

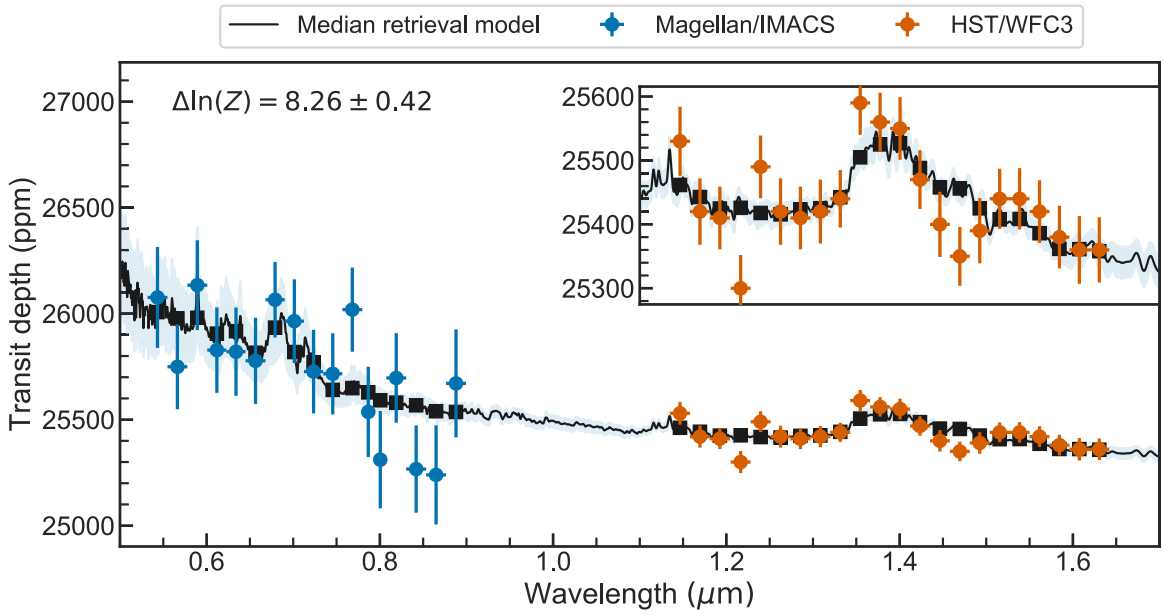

Figure 9. Best retrieved model and 1σ uncertainty (highlighted) for GP+PCA detrended transmission spectrum with data combined in the visual and NIR. The model includes H_2O in the atmosphere of WASP-43b and stellar heterogeneity. The black squares show the model binned to this study's Magellan/IMACS data and *HST*/WFC3 data (Kreidberg et al. 2014). The top-right panel is a zoom-in around the *HST*/WFC3 portion of the spectrum. We share the best retrieved parameters in Table 6 and in the associated corner plot in the Appendix.

Table 7
Associated Transit Depths of Figure 8, Shown Relative to the Mean Fitted GP+PCA WLC Depth of 25,071 ppm

Wavelength (Å)	Transit 1	Transit 2	Transit 3	Transit 4	Combined
5317.9–5547.9	523.0603 ± 260.4245	−267.3194 ± 938.3293	−1922.6752 ± 1020.5674	522.3093 ± 1102.5989	339.2619 ± 237.9392
5547.9–5777.9	−116.1220 ± 266.1869	172.6987 ± 589.2437	−1455.4281 ± 596.1788	1073.0221 ± 440.0050	12.2549 ± 200.1120
5777.9–6007.9	589.0840 ± 278.1240	−148.8073 ± 761.4993	−1069.2411 ± 595.0667	936.0414 ± 455.4287	397.0056 ± 211.7735
6007.9–6228.75	295.5726 ± 269.8283	−542.7774 ± 715.1426	−1145.5722 ± 635.7125	320.1664 ± 392.9886	90.7797 ± 201.4572
6228.75–6449.6	246.0972 ± 273.4700	−250.3903 ± 740.0324	−1642.9432 ± 528.5344	1157.7405 ± 482.7322	83.2450 ± 208.2044
6449.6–6679.6	−79.6483 ± 287.1381	24.9089 ± 660.5415	35.8304 ± 435.9645	373.9576 ± 469.4392	40.3179 ± 203.1956
6679.6–6901.6	495.3958 ± 236.3006	886.1126 ± 651.1368	−1194.4574 ± 482.4106	651.8302 ± 380.5127	328.2793 ± 178.2558
6901.6–7123.6	632.9288 ± 254.1061	−785.1288 ± 655.7699	−1535.3495 ± 493.4588	1103.0682 ± 514.3652	227.2646 ± 197.2615
7123.6–7345.6	300.9717 ± 245.2279	−700.5517 ± 718.5057	−1524.4228 ± 551.7447	281.0918 ± 508.2007	−10.5521 ± 197.1705
7345.6–7567.0	291.1467 ± 263.7409	−974.6394 ± 713.1503	−1460.5078 ± 485.1884	498.0201 ± 384.7823	−20.8910 ± 191.2337
7567.0–7797.0	493.7086 ± 254.0115	−726.8098 ± 960.7674	−1840.4984 ± 616.9723	832.4674 ± 403.5181	281.7274 ± 198.6125
7797.0–7935.5	−100.4397 ± 310.0690	−1304.6125 ± 934.2026	−1739.2184 ± 566.2540	479.8308 ± 368.1740	−200.6929 ± 212.9927
7935.5–8074.0	−453.8738 ± 302.4915	−372.3211 ± 954.3857	−1253.9094 ± 537.2197	463.4859 ± 534.4578	−426.0211 ± 229.4610
8074.0–8304.0	277.2830 ± 287.8497	−728.4844 ± 809.5412	−1631.6495 ± 572.5958	318.7489 ± 417.2814	−40.6945 ± 211.3458
8304.0–8534.0	−593.7829 ± 265.2407	−134.9286 ± 843.3740	−1441.1786 ± 556.1943	460.0621 ± 458.6014	−470.2350 ± 205.8146
8534.0–8764.0	−100.5637 ± 328.7061	−896.6887 ± 954.9673	−2019.0232 ± 479.6925	449.1528 ± 529.7768	−497.6864 ± 234.0147
8764.0–8994.0	−3.1396 ± 354.2233	−822.3481 ± 1178.9426	−1235.4496 ± 531.3372	1244.3340 ± 559.6876	−66.1594 ± 254.6280

of possible hypotheses, which limit the comparison with frequentist methods.

6.1. Retrieval Results

We fit the combined GP+PCA ACCESS detrended spectrum and the *HST*/WFC3 NIR spectrum from Kreidberg et al. (2014) using the retrieval code described in the previous section. We fit for an offset between the two data sets, and a range of models. Those include combinations of clear and cloudy/hazy atmospheres with Na, K, and H_2O and contamination of the planet's transmission spectrum by stellar surface heterogeneity. Table 4 shows the prior distributions of the parameters used in the models. We used the prescription provided by Benneke & Seager (2013, Table 2) to interpret our relative log evidences (Table 5). Based on their values, we find that they all register as a strong detection, with no one model being statistically more likely than another, although the model with the largest log evidence relative to a flat atmosphere ($\Delta \ln Z = 8.26$) is the one including stellar heterogeneity, combined with a clear atmosphere with H_2O (but no Na or K). We adopt that model as the one that best fits the data and show it in Figure 9. From this model, we estimate an average spot contrast of $132\,\mathrm{K} \pm 132\,\mathrm{K}$ and covering fraction of $0.27^{+0.42}_{-0.16}$. The corner plot for that model solution is shown in the Appendix. The parameters of that model and their uncertainties are summarized in Table 6.

7. Summary and Conclusions

We have collected, extracted, and combined transmission spectra of WASP-43b from Magellan/IMACS over four transit epochs spanning the years 2015 to 2018. We combined this with IR data from *HST*/WFC3 to create a transmission spectrum with a total wavelength coverage of 5318–16420 Å. We analyzed the combined spectrum in a dynamic nested sampling framework with NIR data from Kreidberg et al. (2014), extending up to 16420 Å to search for the presence of different species. Assuming a water volume mixing ratio of 6.1×10^{-4} for a planetary atmosphere with solar abundances (Kreidberg et al. 2014), our retrieval yields a log H_2O volume mixing ratio of $-2.78^{+1.38}_{-1.47}$. ($2.72^{+65.26}_{-0.09}$ solar). Our retrieved water abundance is consistent with the 1σ range found by the joint transmission and emission spectrum analysis in Kreidberg et al. (2014, 0.4–3.5 solar) and phase-curve analysis in Stevenson et al. (2014). Our retrieved planetary temperature is also consistent with the ranges predicted from the more recent 2.5D theoretical phase-curve retrievals of the terminator region in Irwin et al. (2019, Figure 9). In the visual spectrum, we do not observe a statistically significant excess of K I given the data, as reported in Murgas et al. (2014). We also do not observe the presence of Na I or Hα in our combined spectra.

Our analysis also investigates the contribution of stellar heterogeneities to observed transmission spectra, given that WASP-43 is an active star. The best-fitting model to our transmission spectrum calls for the presence of spots in the surface of the star being a more favored model than a model with atmospheric hazes. However, we do note that the impact on the spectrum from both hazes and the contamination from surface stellar heterogeneities can be degenerate, and given the quality of the current data, it is not possible to fully distinguish between the two.

As the sample of available transmission spectra of different planets increases, it will become useful to compare their measured spectra and other system parameters (e.g., stellar irradiation levels, spectral type, metallicity, planet density) to look for possible correlations. Uniform data sets for a wide range of planets, such as the ones ACCESS is building, will be crucial in this analysis.

The results reported herein benefited from support, collaborations, and information exchange within NASA's Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by NASA's Science Mission Directorate. This paper includes data gathered with the 6.5 meter Magellan Telescopes located at Las Campanas Observatory, Chile. We thank the staff at the Magellan Telescopes and Las Campanas Observatory for their ongoing input and support to make the ACCESS observations presented in this work possible. A. J. acknowledges support from FONDECYT project 1171208 and by the Ministry for the Economy, Development, and Tourism's Programa Iniciativa Científica Milenio through grant IC 120009, awarded to the Millennium Institute of Astrophysics helped improve the quality of this work.

Facilities: Magellan: Baade, Smithsonian Institution High Performance Cluster (SI/HPC).

Software: Astropy (Astropy Collaboration et al. 2013), corner (Foreman-Mackey 2016), Matplotlib (Hunter 2007), NumPy (Oliphant 2006; van der Walt et al. 2011), PyMC (Salvatier et al. 2016), Multinest (Feroz et al. 2009), PyMultiNest (Buchner et al. 2014), SciPy (Virtanen et al. 2019), SPOTROD (Béky et al. 2014), batman (Kreidberg 2015), george (Ambikasaran et al. 2014).

Appendix

Below are the additional figures referenced in the text.

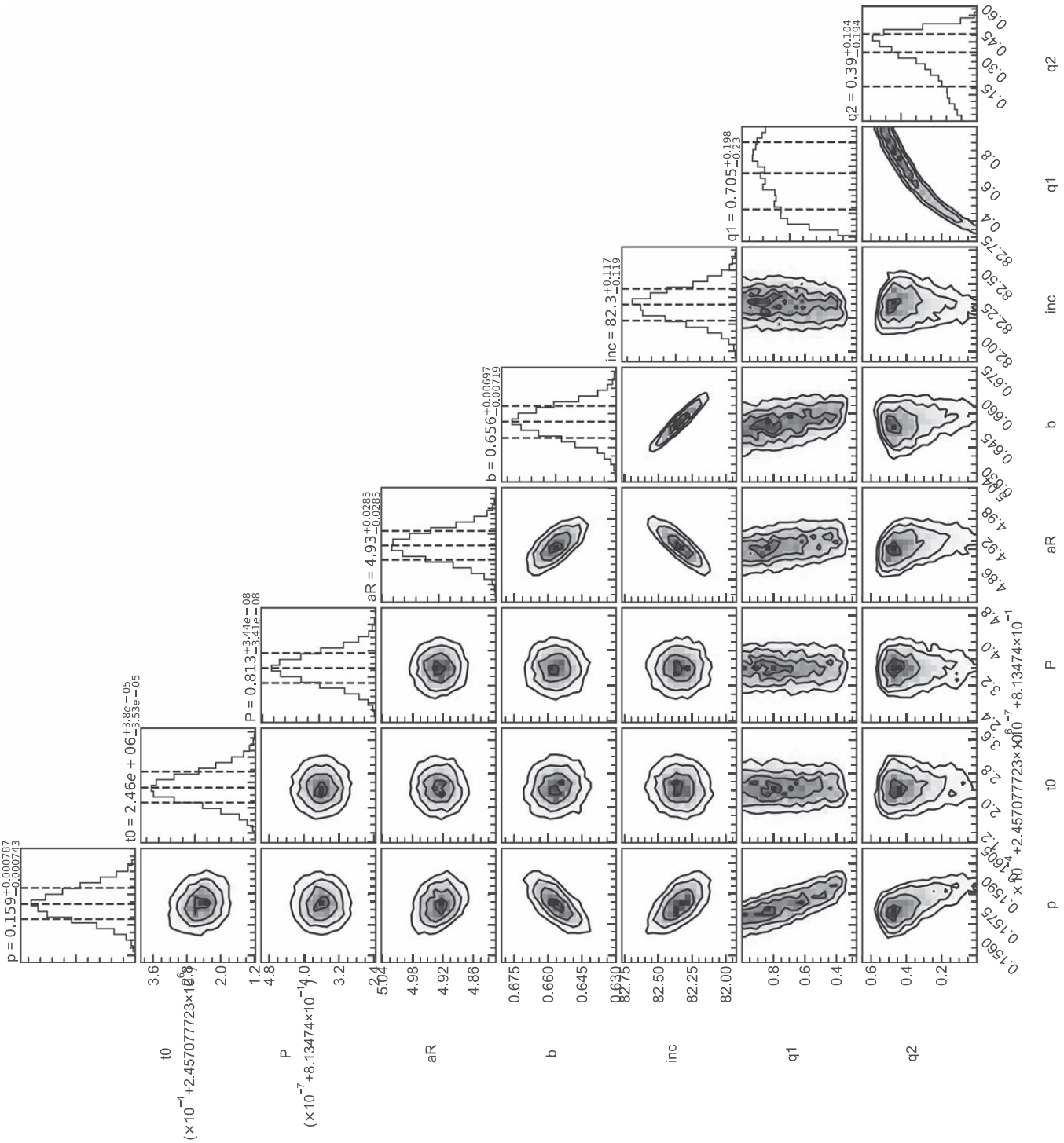

Figure 10. GP+PCA corner plot of fitted parameters for the transit 1 WLC in Figure 3. Vertical dashed lines mark the 16% and 84% quantiles. We share the best-fit values in Table 3.

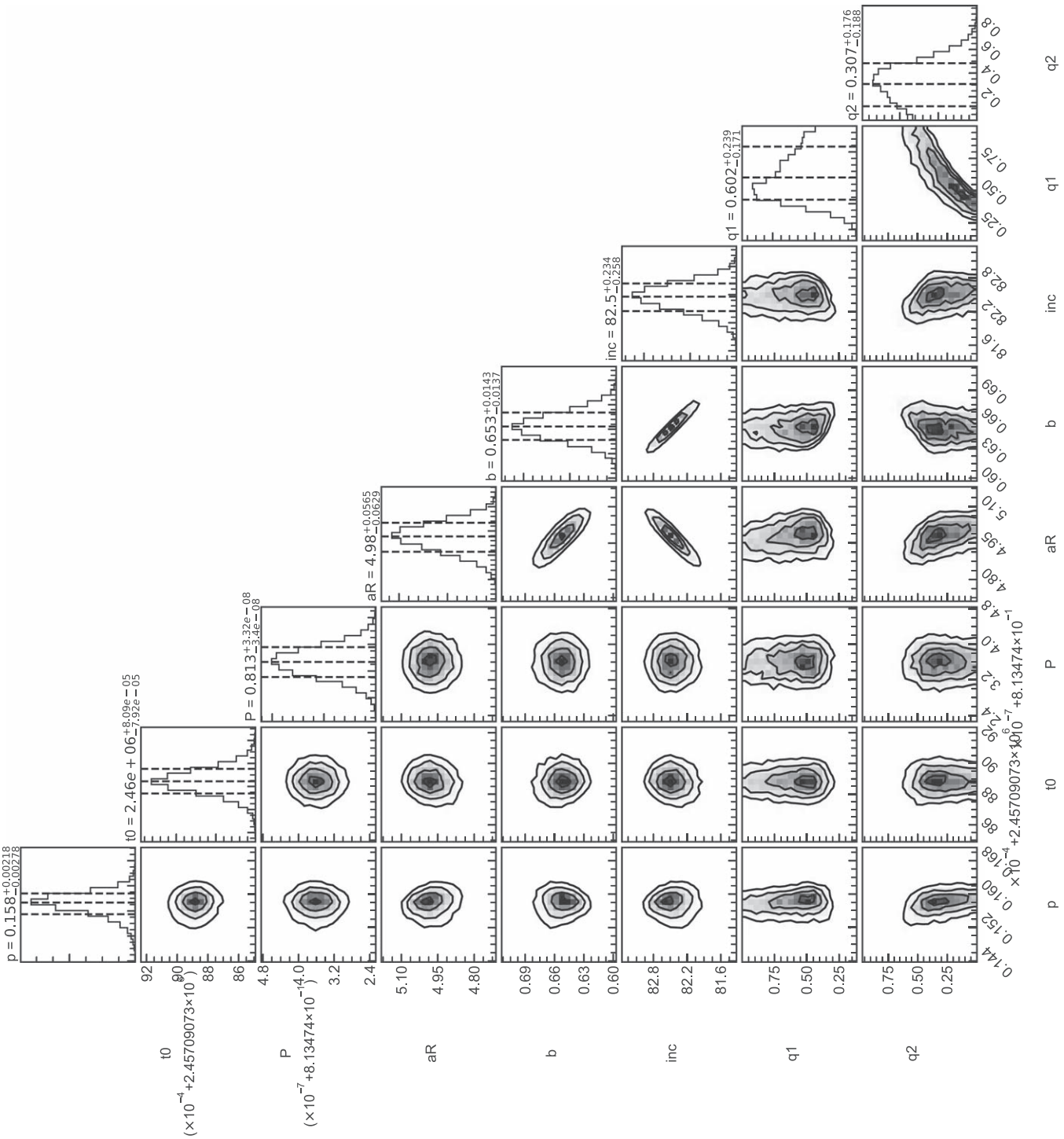

Figure 11. GP+PCA corner plot of fitted parameters for the transit 2 WLC in Figure 3. Vertical dashed lines mark the 16% and 84% quantiles. We share the best-fit values in Table 3.

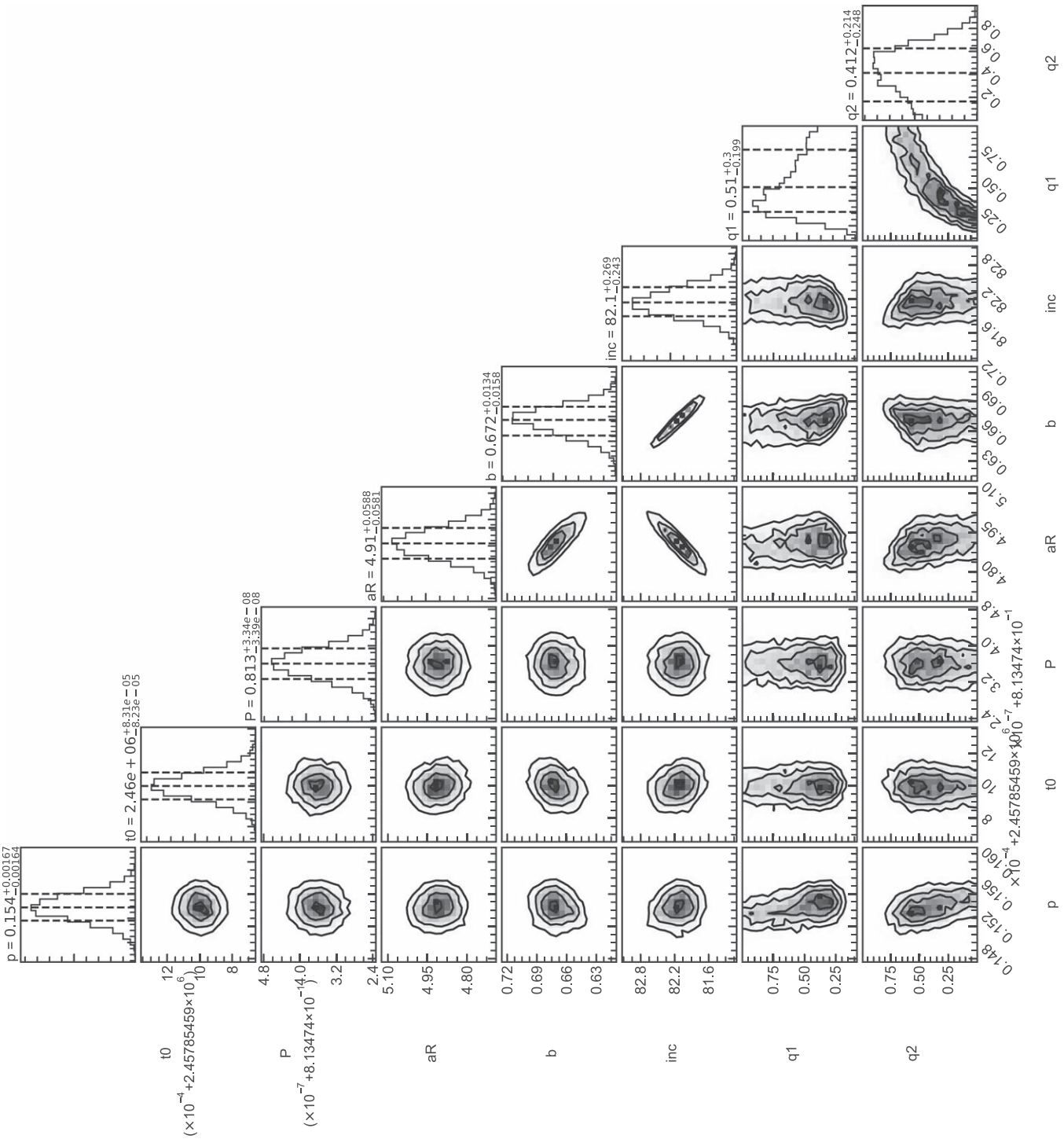

Figure 12. GP+PCA corner plot of fitted parameters for the transit 3 WLC in Figure 3. Vertical dashed lines mark the 16% and 84% quantiles. We share the best-fit values in Table 3.

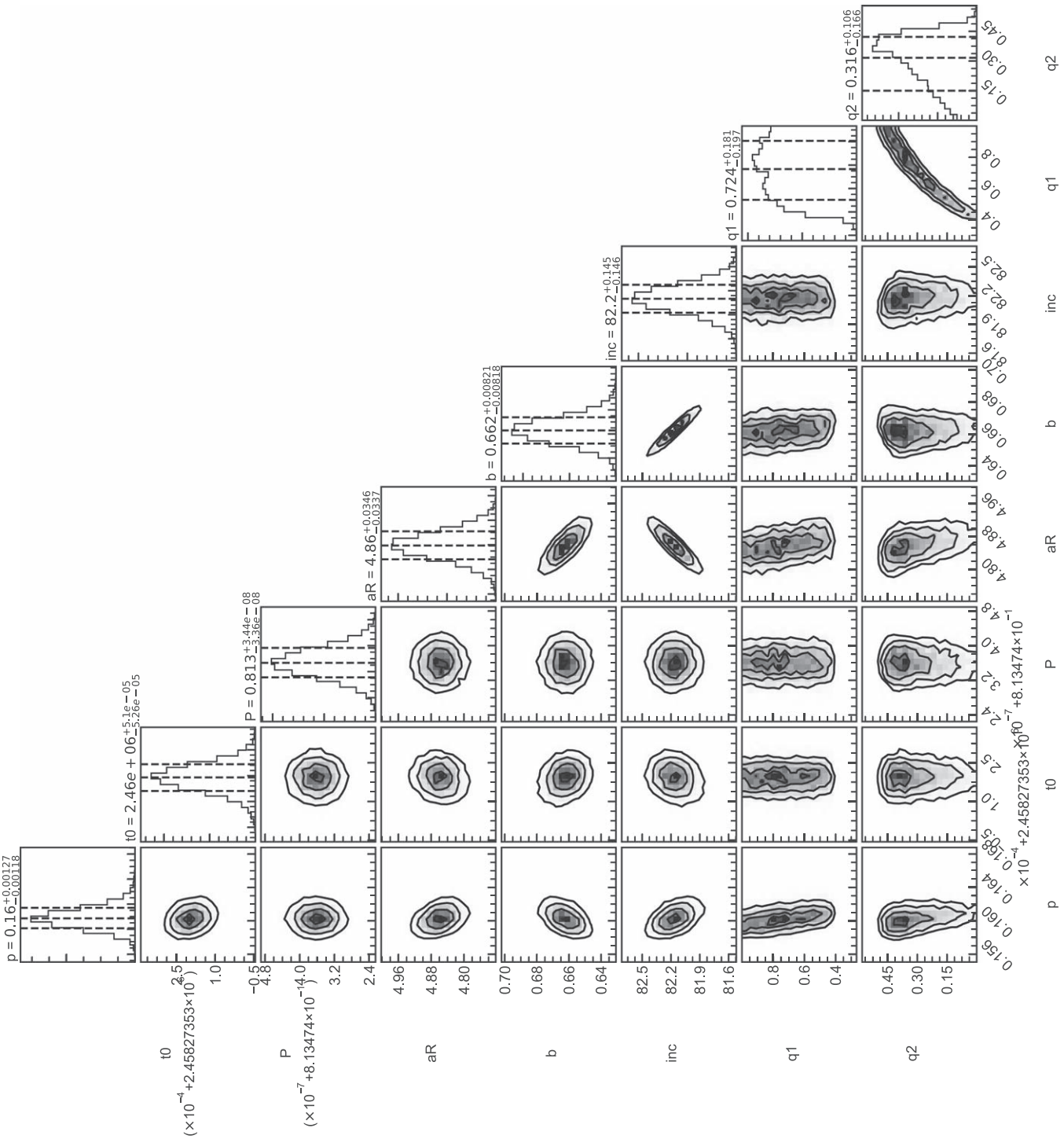

Figure 13. GP+PCA corner plot of fitted parameters for the transit 4 WLC in Figure 3. Vertical dashed lines mark the 16% and 84% quantiles. We share the best-fit values in Table 3.

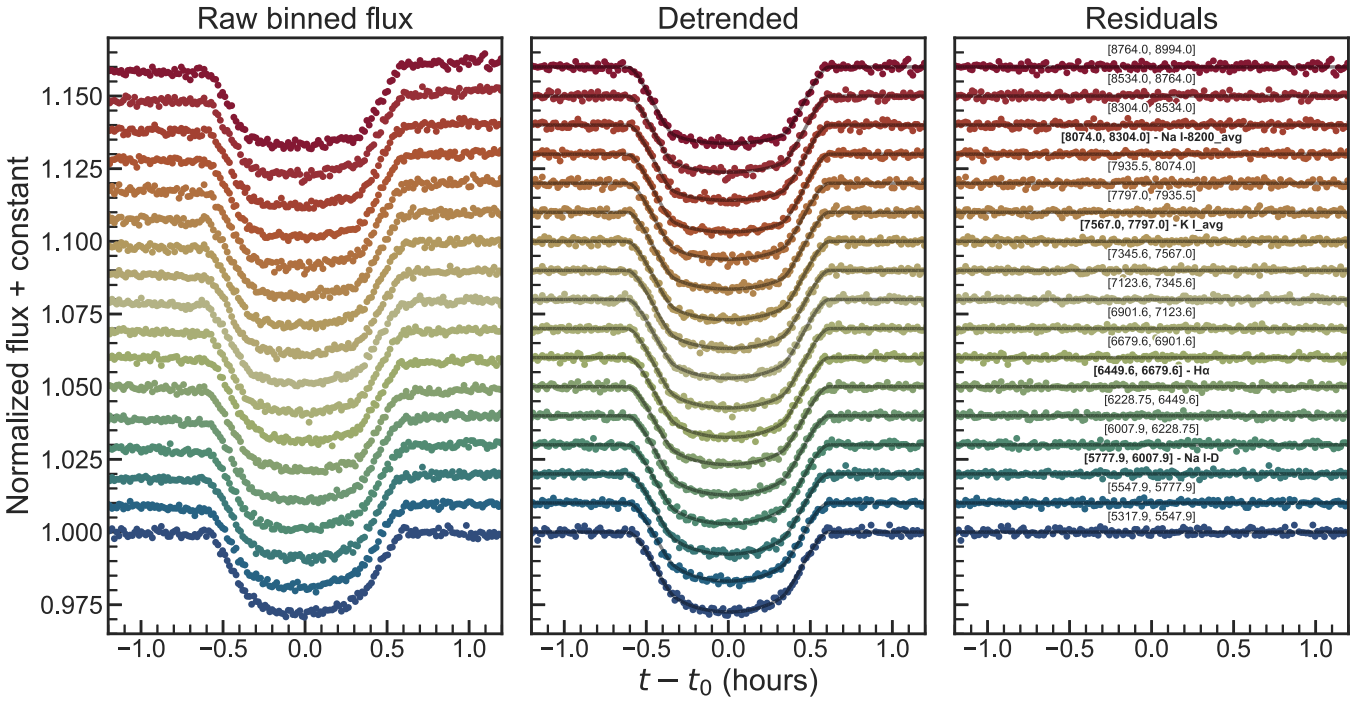

Figure 14. Binned light curves for transit 1 shown in Figure 8. Column 1 shows the raw observed flux, column 2 shows the GP+PCA detrended flux and model, and column 3 shows the residuals. We labeled the wavelength range of each bin in column 3 as well and marked the bins centered on the vacuum wavelength of the potential features of interest in bold. We centered all data 1 hr around the fitted midtransit time t_0 from the corresponding WLC in Figure 3. We share the binned transit data in Table 7.

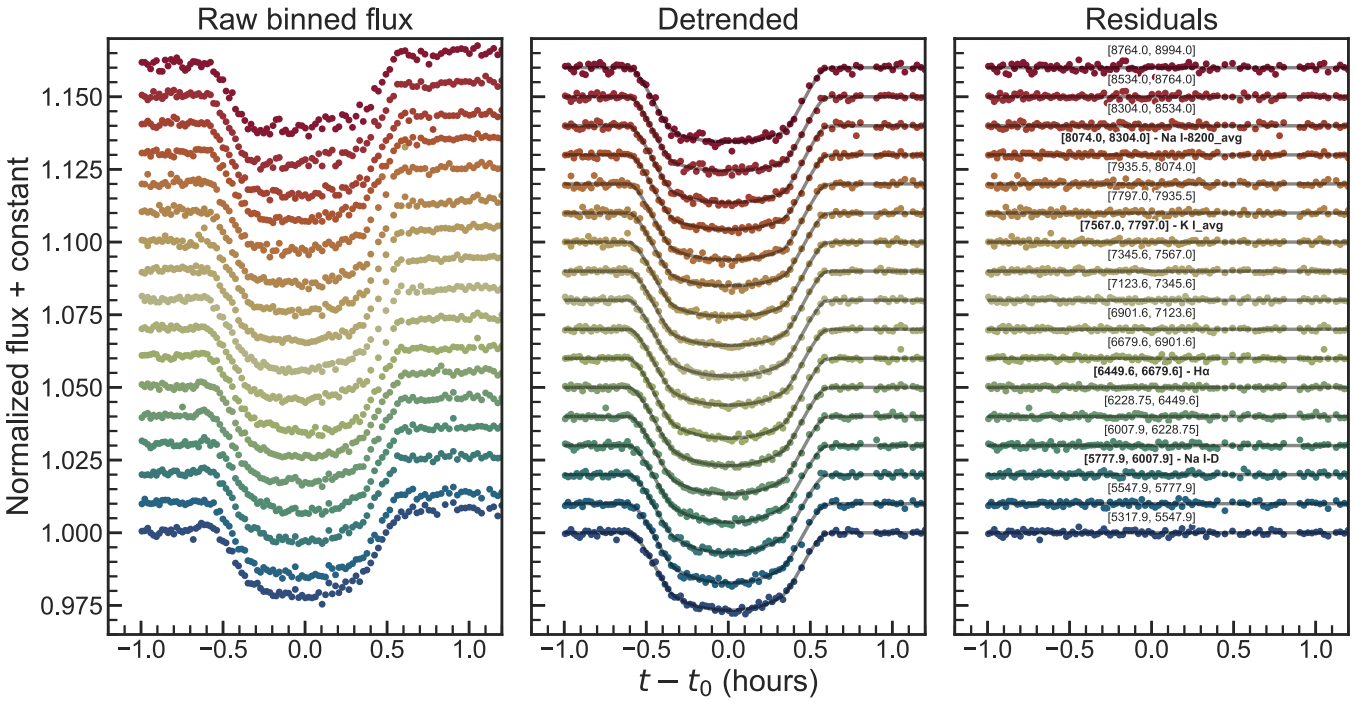

Figure 15. Binned light curves for transit 2 shown in Figure 8. Column 1 shows the raw observed flux, column 2 shows the GP+PCA detrended flux and model, and column 3 shows the residuals. We labeled the wavelength range of each bin in column 3 as well and marked the bins centered on the vacuum wavelength of the potential features of interest in bold. We centered all data 1 hr around the fitted midtransit time t_0 from the corresponding WLC in Figure 3. We share the binned transit data in Table 7.

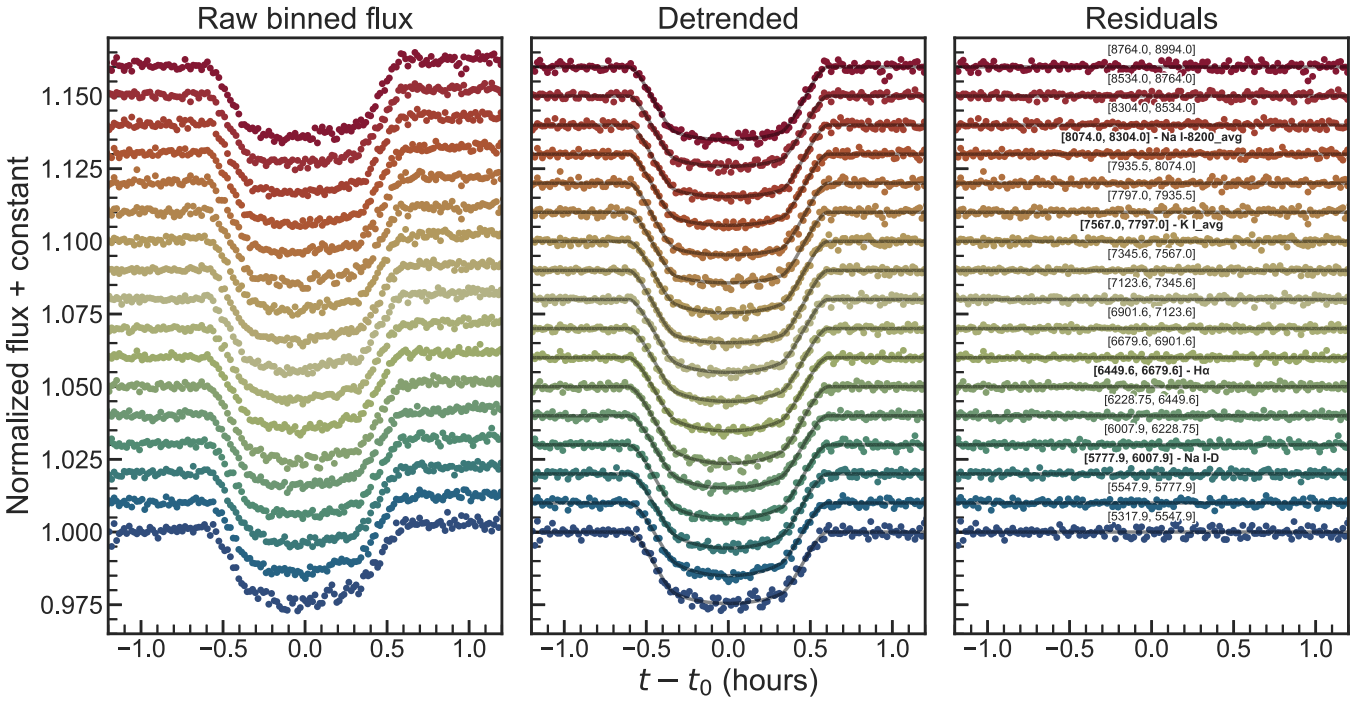

Figure 16. Binned light curves for transit 3 shown in Figure 8. Column 1 shows the raw observed flux, column 2 shows the GP+PCA detrended flux and model, and column 3 shows the residuals. We labeled the wavelength range of each bin in column 3 as well and marked the bins centered on the vacuum wavelength of the potential features of interest in bold. We centered all data 1 hr around the fitted midtransit time t_0 from the corresponding WLC in Figure 3. We share the binned transit data in Table 7.

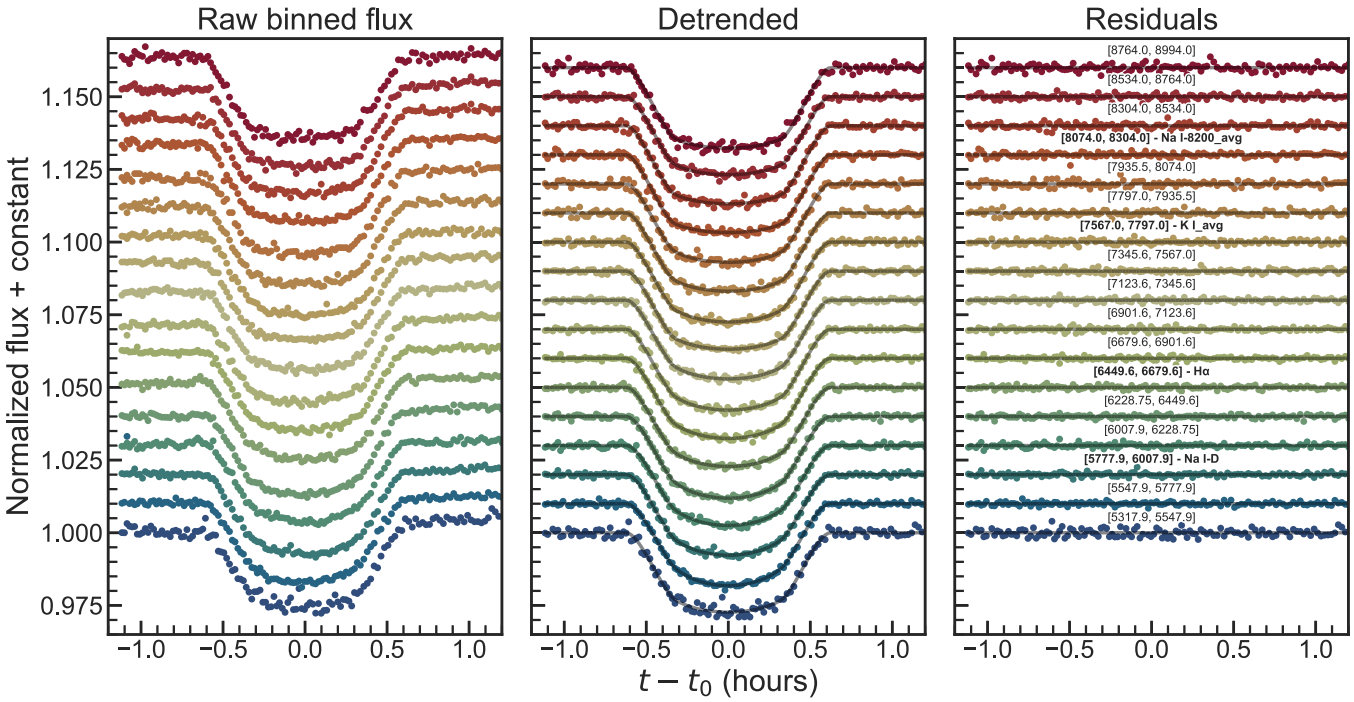

Figure 17. Binned light curves for transit 4 shown in Figure 8. Column 1 shows the raw observed flux, column 2 shows the GP+PCA detrended flux and model, and column 3 shows the residuals. We labeled the wavelength range of each bin in column 3 as well and marked the bins centered on the vacuum wavelength of the potential features of interest in bold. We centered all data 1 hr around the fitted midtransit time t_0 from the corresponding WLC in Figure 3. We share the binned transit data in Table 7.

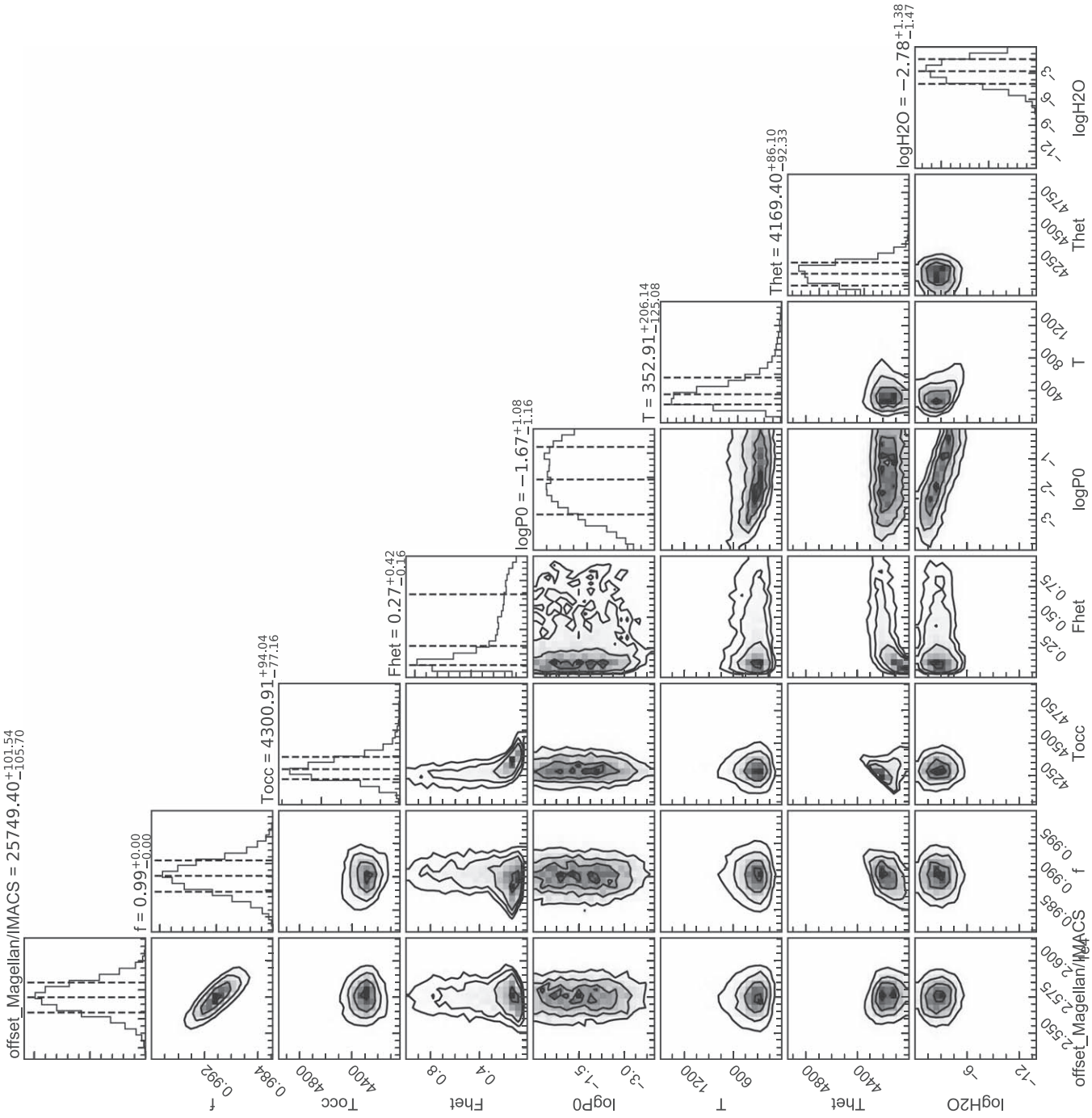

Figure 18. Corner plot for the best-fit transmission spectrum retrieved in Figure 9.

ORCID iDs

Ian C. Weaver https://orcid.org/0000-0001-6205-6315
Mercedes López-Morales https://orcid.org/0000-0003-3204-8183
Néstor Espinoza https://orcid.org/0000-0001-9513-1449
Benjamin V. Rackham https://orcid.org/0000-0002-3627-1676
Dániel Apai https://orcid.org/0000-0003-3714-5855
Andrés Jordán https://orcid.org/0000-0002-5389-3944
Alex Bixel https://orcid.org/0000-0003-2831-1890
Nikole K. Lewis https://orcid.org/0000-0002-8507-1304
Munazza K. Alam https://orcid.org/0000-0003-4157-832X
James Kirk https://orcid.org/0000-0002-4207-6615
Florian Rodler https://orcid.org/0000-0003-0650-5723

References

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Alam, M. K., Nikolov, N., López-Morales, M., et al. 2018, AJ, 156, 298
Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O'Neil, M. 2014, ITPAM, 38, 252
Apai, D., Rackham, B. V., Giampapa, M. S., et al. 2018, arXiv:1803.08708
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Béky, B., Kipping, D. M., & Holman, M. J. 2014, MNRAS, 442, 3686
Benneke, B., & Seager, S. 2013, ApJ, 778, 153

Berta, Z. K., Charbonneau, D., Bean, J., et al. 2011, ApJ, 736, 12
Bétrémieux, Y., & Swain, M. R. 2017, MNRAS, 467, 2834
Bixel, A., Rackham, B. V., Apai, D., et al. 2019, AJ, 157, 68
Blecic, J., Harrington, J., Madhusudhan, N., et al. 2014, ApJ, 781, 116
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Carter, J. A., & Winn, J. N. 2009, ApJ, 704, 51
Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, ApJ, 568, 377
Chen, G., van Boekel, R., Wang, H., et al. 2014, A&A, 563, A40
Crossfield, I. J. M., & Kreidberg, L. 2017, AJ, 154, 261
Dressler, A., Hare, T., Bigelow, B. C., & Osip, D. J. 2006, Proc. SPIE, 6269, 62690F
Espinoza, N., Fortney, J. J., Miguel, Y., Thorngren, D., & Murray-Clay, R. 2017, ApJL, 838, L9
Espinoza, N., Rackham, B. V., Jordán, A., et al. 2019, MNRAS, 482, 2065
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foreman-Mackey, D. 2016, JOSS, 1, 24
Fraine, J., Deming, D., Benneke, B., et al. 2014, Natur, 513, 526
Gibson, N. P., Aigrain, S., Barstow, J. K., et al. 2013, MNRAS, 436, 2974
Gibson, N. P., Aigrain, S., Roberts, S., et al. 2012, MNRAS, 419, 2683
Gillon, M., Triaud, A. H. M. J., Fortney, J. J., et al. 2012, A&A, 542, A4
Hellier, C., Anderson, D. R., Collier Cameron, A., et al. 2011, A&A, 535, L7
Helling, C., Lee, G., Dobbs-Dixon, I., et al. 2016, MNRAS, 460, 855
Heng, K., & Kitzmann, D. 2017, MNRAS, 470, 2972
Hoyer, S., Pallé, E., Dragomir, D., & Murgas, F. 2016, AJ, 151, 137
Hunter, J. D. 2007, CSE, 9, 90
Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6
Irwin, P. G. J., Parmentier, V., Taylor, J., et al. 2019, arXiv:1909.03233
Jordán, A., Espinoza, N., Rabus, M., et al. 2013, ApJ, 778, 184
Kataria, T., Showman, A. P., Fortney, J. J., et al. 2015, ApJ, 801, 86
Kipping, D. M. 2013, MNRAS, 435, 2152
Kirk, J., Wheatley, P. J., Louden, T., et al. 2018, MNRAS, 474, 876
Kochanek, C. S., Shappee, B. J., Stanek, K. Z., et al. 2017, PASP, 129, 104502
Kreidberg, L. 2015, PASP, 127, 1161
Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, ApJL, 793, L27
MacDonald, R. J., & Madhusudhan, N. 2017, MNRAS, 469, 1979
Mandel, K., & Agol, E. 2002, ApJL, 580, L171
McCullough, P. R., Crouzet, N., Deming, D., & Madhusudhan, N. 2014, ApJ, 791, 55
Miller-Ricci, E., Seager, S., & Sasselov, D. 2009, ApJ, 690, 1056
Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, ApJ, 832, 41
Moses, J. I., Line, M. R., Visscher, C., et al. 2013, ApJ, 777, 34
Murgas, F., Pallé, E., Zapatero Osorio, M. R., et al. 2014, A&A, 563, A41
Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, AJ, 158, 13
Nikolov, N., Sing, D. K., Fortney, J. J., et al. 2018, Natur, 557, 526
Öberg, K. I., Boamah, M. D., Fayolle, E. C., et al. 2013, ApJ, 771, 95
Oliphant, T. E. 2006, Guide to NumPy (Provo, UT: Trelgol Publishing), http://www.numpy.org/
Oshagh, M., Santos, N. C., Ehrenreich, D., et al. 2014, A&A, 568, A99
Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, MNRAS, 482, 1485
Pinhas, A., Rackham, B. V., Madhusudhan, N., & Apai, D. 2018, MNRAS, 480, 5314
Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, MNRAS, 385, 109
Pont, F., Sing, D. K., Gibson, N. P., et al. 2013, MNRAS, 432, 2917
Rackham, B., Espinoza, N., Apai, D., et al. 2017, ApJ, 834, 151
Rackham, B. V., Apai, D., & Giampapa, M. S. 2018, ApJ, 853, 122
Rackham, B. V., Apai, D., & Giampapa, M. S. 2019, AJ, 157, 96
Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, PeerJ Computer Science, 2, e55
Sedaghati, E., Boffin, H. M. J., Jeřabková, T., et al. 2016, A&A, 596, A47
Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48
Sing, D. K., Désert, J.-M., Lecavelier Des Etangs, A., et al. 2009, A&A, 505, 891
Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Natur, 529, 59
Sing, D. K., Lecavelier des Etangs, A., Fortney, J. J., et al. 2013, MNRAS, 436, 2956
Sing, D. K., Pont, F., Aigrain, S., et al. 2011, MNRAS, 416, 1443
Staab, D., Haswell, C. A., Smith, G. D., et al. 2017, MNRAS, 466, 738
Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014, Sci, 346, 838
Trotta, R. 2008, ConPh, 49, 71
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Virtanen, P., Gommers, R., & Oliphant, T. E. 2019, arXiv:1907.10121
Wakeford, H. R., & Sing, D. K. 2015, A&A, 573, A122
Wakeford, H. R., Stevenson, K. B., Lewis, N. K., et al. 2017, ApJL, 835, L12
Zhang, Z., Zhou, Y., Rackham, B. V., & Apai, D. 2018, AJ, 156, 178