Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 13 November 2021 (MN LATEX style file v2.2)

Exoplanet Transmission Spectroscopy using KMOS

Hannu Parviainen,[1]⋆ Suzanne Aigrain,[1] Niranjan Thatte,[1]
Joanna K. Barstow,[1] Thomas M. Evans,[2] and Neale Gibson[3]
[1]*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK*
[2]*Astrophysics Group, School of Physics, University of Exeter, Stocker Road, Exeter, EX4 4QL, UK*
[3]*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching bei München, Germany*

13 November 2021

ABSTRACT

KMOS (K-Band Multi Object Spectrograph) is a novel integral field spectrograph installed in the VLT's ANTU unit. The instrument offers an ability to observe 24 2.8″×2.8″ sub-fields positionable within a 7.2′ patrol field, each sub-field producing a spectrum with a 14×14-pixel spatial resolution. The main science drivers for KMOS are the study of galaxies, star formation, and molecular clouds, but its ability to simultaneously measure spectra of multiple stars makes KMOS an interesting instrument for exoplanet atmosphere characterization via transmission spectroscopy. We set to test whether transmission spectroscopy is practical with KMOS, and what are the conditions required to achieve the photometric precision needed, based on observations of a partial transit of WASP-19b, and full transits of GJ 1214b and HD 209458b. Our analysis uses the simultaneously observed comparison stars to reduce the effects from instrumental and atmospheric sources, and Gaussian processes to model the residual systematics. We show that KMOS can, in theory, deliver the photometric precision required for transmission spectroscopy. However, this is shown to require a) pre-imaging to ensure accurate centering and b) a very stable night with optimal observing conditions (seeing ∼0.8″). Combining these two factors with the need to observe several transits, each with a sufficient out-of-transit baseline (and with the fact that similar or better precision can be reached with telescopes and instruments with smaller pressure,) we conclude that transmission spectroscopy is not the optimal science case to take advantage of the abilities offered by KMOS and VLT.

Key words: Instrumentation: spectrographs–Techniques: photometric–Techniques: spectroscopic: Planets and satellites: atmospheres

1 INTRODUCTION

Transmission spectroscopy, the measurement of a transit depth as a function of wavelength, allows us to probe the existence and abundance of different atmospheric species–each with their wavelength-dependent extinction features–in planetary atmospheres (Brown 2001). However, the variations in the transit depth are minute, and high-precision spectroscopic time series are required in the characterization of the planetary transmission spectra.

Systematic trends from changing telluric and instrumental conditions impair the ground-based measurements, and the highest-quality transmission spectroscopy observations have been carried out using space-based HST until recently (Charbonneau et al. 2002; Berta et al. 2012; Sing et al. 2011, 2013; Knutson et al. 2014; Evans et al. 2013). However, the use of multi-object spectrographs in combination with modern data analysis methods has led to remarkable improvements in the precision that can be achieved from the ground. Simultaneous measurements of the target star and several comparison stars allow for the correction of common-mode systematics, in parallel to relative photometry (Bean et al. 2010; Jordán et al. 2013; Gibson et al. 2013a,b; Bean et al. 2013; Murgas et al. 2014). Further, the use of Gaussian processes has facilitated the correction of systematics by allowing for the robust modeling of correlated noise—including time correlation and correlations with auxiliary measurements such as seeing and humidity—in model-independent fashion (Gibson et al. 2012; Roberts et al. 2013; Rasmussen & Williams 2006).

The K-band Multi Object Spectrograph (KMOS, Sharples et al. 2013) is an integral field spectrograph installed in the Nasmyth-focus of the VLT's ANTU unit. KMOS consists of 24 arms that can be positioned (nearly) freely inside a 7.2′ diameter patrol field. The ability to observe multiple stars and sky fields simultaneously makes KMOS a potentially promising instrument for transmission spectroscopy. However, the small spatial extent of IFUs and the relative complexity of transforming the raw data to science datacubes may impair the photometric precision. The small

⋆ hannu.parviainen@physics.ox.ac.uk

per-IFU FOV can especially be expected to cause problems if the seeing varies significantly during the observations.

Here we document our experiences in assessing the KMOS' applicability to transmission spectroscopy based on the commissioning observations of a partial transit of WASP-19b, and our observations of the full transits of GJ 1214b and HD 209458b.[1]

2 OBSERVATIONS

2.1 Instrument

KMOS (Sharples et al. 2013) is installed in the Nasmyth-focus of the VLT's ANTU unit, and consists of 24 pickoff mirrors positionable inside a 7.2′ diameter patrol field. Each mirror covers a square 2.8″×2.8″ field, which is fed to an image slicer that divides the field into 14×14 spatial elements, each of which are further dispersed into spectra using a grating spectrometer. The 24 arms are divided into three groups, each with its own spectrograph. The spectra for each IFU and spatial element are arranged in the detector as one-dimensional columns, which can be constructed into three-dimensional image cubes through steps described by Davies et al. (2013).

The three spectrographs use substrate-removed 2kx2k 18μm-pixel Hawaii 2RG (HgCdTe) detectors (one per spectrograph) cooled to 40K. Detector readout is done using sample-up-the-ramp mode, that is, the detector is read out continuously without resetting it, and the counts for each pixel are computed by fitting the slope of the signal against time.[2]

2.2 WASP-19b

A partial transit of WASP-19b (H=10.6 Hebb et al. 2010) was observed on 29.03.2013 (03:11–05:10) as a part of the commissioning of the instrument. The observations were carried out in stare mode[3] using the HK grating covering the wavelength ranges from 1.48 to 2.44 μm with a spectral resolution of ~1800. The total exposure time was 60 s, with DIT = 10 s and NDIT = 6 (where DIT = detector on-chip integration time, NDIT = number of exposures averaged over a single frame.) Six reference stars and six sky fields were observed simultaneously with WASP-19. All stars were sufficiently well-centered, and the PSFs are well contained inside the IFUs (Fig. 1).

The airmass was at its minimum at the beginning of the observations, increasing from 1.10 to 1.35 by the end of the observations. The seeing, shown in Fig. 2, was slightly worse than the median Paranal seeing of (FWHM of 0.83″), varying from 0.74″to 1.23″with a median of 1.0″.

The observations cover only the last ~ 1/3 of the transit, but include a post-transit baseline of 1.2 h. The lack of sufficient in-transit coverage render the observations useless in transmission spectroscopy, but they can still be used to estimate the precision that could have been achieved if a full transit had been observed, as detailed later in Sec. 5.3.

[1] Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere, Chile. Proposals 60.A-9239(C), 60.A-9447(A), and 092.C-0812(A)

[2] A technical overview of KMOS can be found from `https://www.eso.org/sci/facilities/paranal/instruments/kmos/inst.html`

[3] Some dithered exposures were taken after the stare-mode observations, but these are excluded from our analysis.

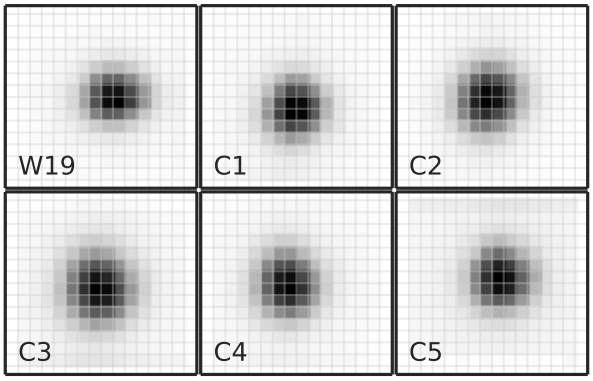

Figure 1. Log fluxes for WASP-19b and the five best comparison stars corresponding to a single exposure where the datacube has been flattened in the spectral direction.

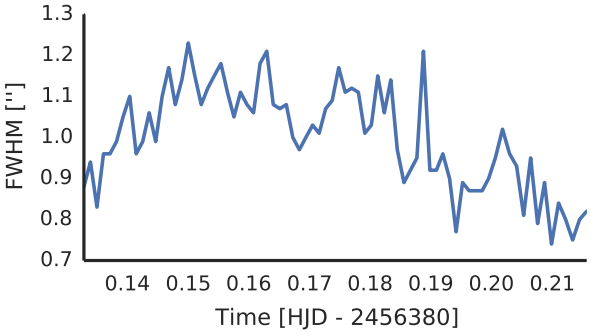

Figure 2. Seeing (as PSF FWHM) during the WASP-19b observations.

2.3 GJ 1214b

We observed a transit of GJ 1214b (H=9.1) on 29.06.2013 (2:34–5:15) using the HK grating. The total exposure time was 60 s, leading to average efficiency of 72% with DIT=20 s and NDIT=3 (average overheads were 0.39 min/frame). The observations covered the main target GJ 1214b (arm 2), four primary reference stars (arms 10,15,16,17), and seven secondary reference stars (arms 3,5,6,7,11,20,22). The observations were carried out in the stare mode, and nine arms were assigned to observe the sky. The average FWHM during the night was 0.8″, close to the median Paranal seeing, with a standard deviation of 0.1″. One of the primary reference stars (IFU 16) was not observed due to technical issues.

GJ 1214 has a high proper motion, and its position at the time of observations was calculated from the proper motion estimates that did not include uncertainty estimates. Unfortunately, the uncertainties were large enough to position the main target close to the upper-right corner of its IFU, as shown in Fig. 3, which led to severe systematics due to a varying amount of flux being lost outside the IFU with varying seeing.

2.4 HD 209458b

We observed a full transit of HD 209458b (H=6.4) on 28.10.2013 (23:49–03:53) in the stare mode using the IZ grating covering the wavelength range from 0.8 to 1.06 μm with a spectral resolution

[htp]

Target	Starting night	Time span	Grating	N_{sci}	N_{dark}	N_{flat}	N_{arc}	T_{exp} [s]	N_{ref}	N_{sky}
WASP-19b	29.03.2013	03:11–05:10	HK	64	5	4,1	1,1	10	6	6
GJ 1214b	29.06.2013	02:34–05:14	HK	116	5	18,3	6,1	20	11	9
HD 209458b	28.10.2013	23:49–03.53	IZ	226	5	18,3	6,1	60	5	18

Table 1. Observation details, where N_{sci} is the number of science exposures, N_{dark} is the number of dark exposures, N_{flat} is the number of flat field exposures (on,off), N_{arc} is the number of arc lamp exposures (on,off), T_{exp} is the science frame exposure time, N_{ref} is the number of reference stars, and N_{sky} is the number of sky fields.

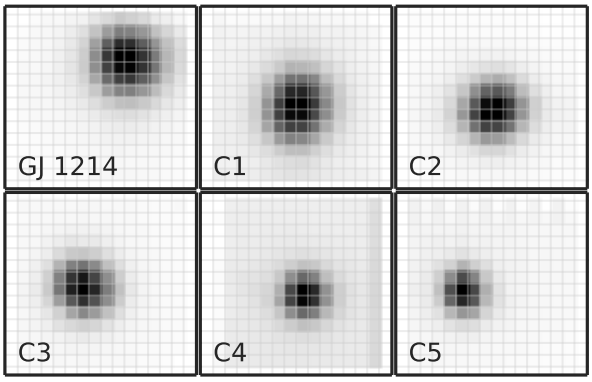

Figure 3. Log fluxes for GJ 1214b and the five best comparison stars.

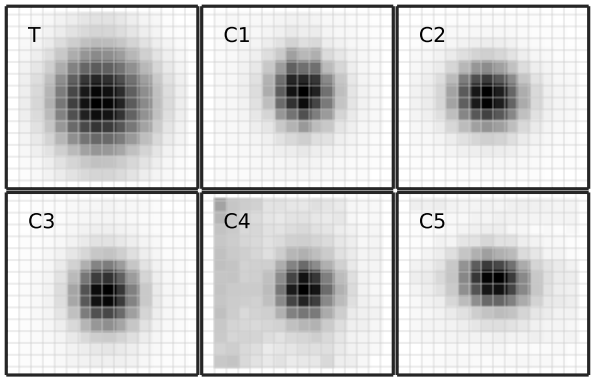

Figure 5. Log fluxes for HD 209458b (T) and five best comparison stars. The fourth comparison star (C4) was not used in the analysis.

of ∼3200. The total exposure time was 60 s with DIT = 10 s and NDIT = 6. Five reference stars and 18 sky fields were observed simultaneously with HD 209458. All the reference stars were significantly fainter than the target, the three brightest being 3.4%-5.6% of the HD 209458's flux each. The airmass was at its minimum in the beginning of the observations, increasing from 1.4 to 2.8 by the end of the observations. The seeing was non-optimal (median FWHM for the night was 1.06″, while the Paranal median FWHM is 0.83″), and varied from 0.75″ to 1.8″ during the night (Fig. 4).

All the stars were well-centered on their IFUs (Fig. 5), and we did not run into similar problems as with our observations of GJ 1214b. However, only a minimal pre-transit baseline was observed due to technical difficulties. Also, the post-transit baseline was shorter than optimal due to observing limitations. This lack of adequate out-of-transit baseline can be expected to have a strong impact on the precision of the transit depth estimates.

3 DATA REDUCTION

3.1 Overview

The official ESO KMOS-1.2.4 pipeline was used for the basic reductions (detailed in Davies et al. 2013), following the steps outlined in the "SPARK INstructional Guide for KMOS data" version 0.7 (Fairley 2012). The pipeline uses the calibration frames with the raw science frames to construct separate datacubes for each observed IFU, which were used in the following analysis. The basic calibrations consist of creation of master dark and flat frames, identification of hot and cold pixels, dark frame subtraction, and division by the master flat. The flat and arc sets consists of N_{on} exposures with the calibration lamp on, and N_{off} exposures with the calibration lamp off, for improved dark current removal. Hot pixels are identified from the dark frames, and cold pixels from the flat frames, and both are excluded from the rest of the analysis. A wavelength solution is computed using the arc frames, which is then used in the datacube creation. Cosmic ray detection and correction was not included into the basic reductions, and neither was the sky removal, which was carried out using proprietary methods. The number of calibration frames used in the analyses are given in Table 1.

We experimented with four approaches to photometry and two approaches to sky estimation. Since the observations were carried out in the stare mode, each target star was assigned with a nearby sky IFU (typically 0.4′–0.7′ from the object IFU) for sky subtraction. Thus, the photometry was carried out using a target and a sky cube for each target and exposure. The basic steps were similar for most approach combinations, PSF fitting excluded:

(i) Calculate the 1D target spectrum by collapsing the datacube in the spatial dimensions (sum the pixel values) after possible masking.
(ii) Calculate the 1D sky spectrum from either a simultaneously measured sky-cube or from target-cube pixels where the target flux does not significantly contribute to the total flux.
(iii) Subtract sky spectrum from the target spectrum.
(iv) Collapse the 1D sky-subtracted target spectrum after applying a bandpass filter to obtain a single photometric data point.

3.2 Sky estimation

We tried two approaches to sky estimation:

(i) **Using the sky IFUs.** A 1D sky spectrum was constructed by collapsing the target-specific sky-datacube in spatial dimensions with median operator (that is, calculating the median over the spatial pixels for each wavelength element). This spectrum was then subtracted from the 1D target spectrum.

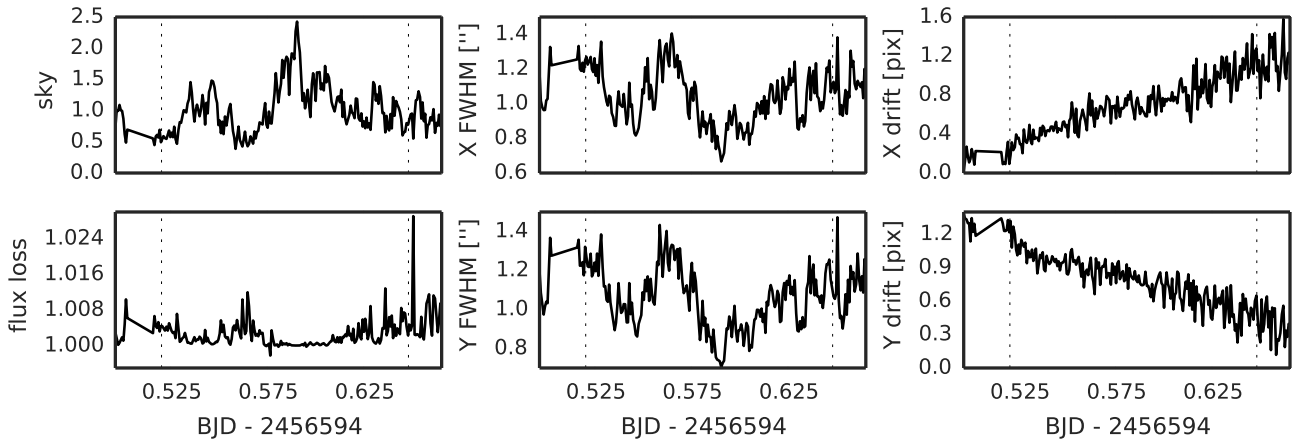

Figure 4. Auxiliary information from the instrument (FITS headers) and PSF fitting for the HD 209458b observations. The focus value is from the FITS headers, while the residual sky level, x and y FWHMs and x and y drifts are from the PSF fitting (averaged over the four brightest stars).

(ii) **Using the target IFUs.** A 1D sky spectrum was constructed from the target-cube pixels where the target star did not significantly contribute to the observed flux (based on 2D S/N maps created for optimal-mask photometry).

The first approach yielded the best sky removal for the bright stars, but was not optimal for the faint stars. The second approach yielded significantly better sky subtraction for the faint stars, but was not applicable for the bright stars for which all the pixels had a significant contribution from the stellar flux.

3.3 Photometry

We experimented with four different approaches on photometry. The major difference between the three first was the shape of the spatial mask applied to each datacube. The masks were constructed as 14×14 real arrays with values ranging from 0 to 1, and the masking was done by multiplying the datacube with the mask (along the wavelength axis). The approaches were:

(i) **No masking,** where the datacubes were collapsed first in the spatial dimensions without any masking. Next, the resulting 1D spectrum was multiplied with a smooth-edged window function (to select the desired passband), and collapsed into a photometric point.

(ii) **Soft aperture photometry**, where we used a soft-edged circular aperture mask centered on the star. The shape of the mask edge followed a smoothstep function

$$\alpha = \max\left(0, \min\left(3r_N^2 - 2r_N^3, 1\right)\right), \quad (1)$$

where α is the mask opacity, r_N is the distance from the aperture center divided by the aperture radius, and the min and max clamp the opacity to values between 0 and 1. The photometry was carried out as in the first approach, but the datacube pixels were first multiplied with the aperture mask (in the x,y-space).

(iii) **Optimal mask photometry**, where photometry masks were created for each target IFU based on the (x,y)-pixel values averaged over the wavelength axis and all individual exposures. We created masks that maximize the S/N ratio, but also experimented with somewhat larger masks. One mask per target was used for all exposures to minimize the scatter from varying mask shape.

(iv) **PSF fitting**, where an analytic PSF model was fitted to the observed data. Here we first collapsed the datacubes in the wavelength axis (after applying the passband mask), and fitted the PSF to the resulting 2D image. We tested two-dimensional (different FWHMs in x and y) Gaussian, Moffat, and two-component Gaussian PSF models to study whether one of the models would be superior to the others.

The first approach (no masking) yielded the lowest scatter in the resulting light curves for the bright stars, and masking improved the outcome for the fainter stars. The photometry calculated from PSF fitting had significantly higher scatter than any of the other approaches (for all the used PSF models). However, the PSF parameter estimates (centers, FWHMs, residual sky level) proved to be useful in the following Gaussian Process-based detrending, since they allow us to model the flux loss, which was the main cause of the strongest systematics in our data.

4 ANALYSIS

4.1 Overview

We detail the dataset-specific analyses for WASP-19b, GJ 1214b, and HD 209458b in Sects. 5, 6, and 7, respectively and review here the common parts.

The transmission spectroscopy is based on Bayesian parameter estimation, where we obtain a sample from the joint posterior distribution using Markov chain Monte Carlo (MCMC). The parameter estimates correspond to the marginal posterior medians, and their uncertainties to the 68% central posterior intervals. Most of the parameters are strongly constrained by informative priors based on previous studies in all the three cases covered, with the passband-specific planet-star radius ratios relative to the average broadband radius ratio (or planet-star area-ratios) being the most important unconstrained parameter of interest.

The noise in photometric time series is rarely white (Pont et al. 2006), and we use Gaussian processes (GPs, Gibson et al. 2012; Roberts et al. 2013; Rasmussen & Williams 2006) to model the correlated noise (together with a case-by-case selected set of simultaneously observed auxiliary parameters used as GP input parameters).

4.2 Priors

All of the planets in our study are well characterized by extensive amounts of previous studies using transit, occultation, and radial velocity observations. This allows us to set tight informative priors on the planet's orbital parameters and average radius ratios over wide passbands (indeed, given the quality of our data, we must set tight priors on the orbital parameters). We list the used priors in the analysis section for each planet.

4.3 Model

We model the broadband flux simultaneously with the narrow-band fluxes covering the wide passband, with the aim of estimating the narrow-band radius ratios (or area ratios) relative to the average broadband radius ratio. For WASP-19b and GJ 1214b we carry out the modeling for H and K bands separately, dividing the wide passbands into six sub-bands, and for HD 209458b we carry out the modeling for i and z bands (again with six narrow passbands).

The broadband light curve and the narrow-band light curves show OOT (out-of-transit) scatter on the same scale for all passbands. Thus, we can assume that the noise is dominated by a wavelength-independent systematic component and not by a photon noise. We model the systematic, wavelength-independent, component using the residuals from the simultaneous broadband modelling (similar to the often used divide-by-white approach), but in a way that also marginalizes over the uncertainties in the broadband modelling. The narrow-band light curves are divided by the broadband light curve scaled by a scaling factor that is a free parameter in the model (removing the truly constant systematics component) in order to obtain a time series of relative narrow-wide passband fluxes. From here, we have the option of using Gaussian processes (GPs) to model the residual systematics (that is, the systematics component that is not constant over the wide passband), using a set of simultaneously observed auxiliary parameters (or information derived from other processes, such as PSF fitting) as inputs to the GP, or, if no strong correlations between any of the inputs and the residual relative fluxes are found, we can assume the noise to be white and use the standard likelihood equation for normally distributed errors, described below.

We further assume that the stellar limb darkening is constant across the wide passband (which is a simplification), and use a quadratic limb darkening model with informative priors based on Claret et al. (2013) limb darkening tabulations.

4.4 Posterior and likelihood

The unnormalized log posterior for our model is

$$\ln P(\boldsymbol{\theta}|D) = \ln P(\boldsymbol{\theta}) + \ln P(D|\boldsymbol{\theta}), \tag{2}$$

where $\ln P(\boldsymbol{\theta})$ is the log prior for the parameter vector $\boldsymbol{\theta}$. The second term, log likelihood for our data, is

$$\ln P(F|\boldsymbol{\theta}) = \ln P(F_{\mathrm{W}}|\boldsymbol{\theta}) + \sum_{i=1}^{N_{\mathrm{pb}}} \ln P(F_i|\boldsymbol{\theta}), \tag{3}$$

where the first term is the likelihood for the broadband data, and the second term is a sum over the narrow-band flux ratio log likelihoods. The likelihood for the broadband data follows either from the GP, or assumes normally distributed uncorrelated noise, which leads to the usual log likelihood equation

$$\ln P(W|\boldsymbol{\theta}) = -N \ln \sigma_{\mathrm{W}} - \frac{\ln 2\pi}{2} - \sum_{j=1}^{N} \frac{\left(W_{\mathrm{j}} - M_{\mathrm{W,j}}\right)^2}{2\sigma_{\mathrm{W}}^2}, \tag{4}$$

where W is the observed normalized broadband flux, M_{W} is the modeled broadband flux, N is the number of exposures, and σ_{H} the average broadband datapoint uncertainty. The likelihoods for the narrow-wide flux ratios also assume normally distributed white noise, yielding a log likelihood

$$\ln P(F_{\mathrm{i}}|\boldsymbol{\theta}) = -N \ln \sigma_{\mathrm{r}} - \frac{\ln 2\pi}{2} - \sum_{j=1}^{N} \frac{\left(\frac{\alpha F_{\mathrm{i,j}}}{1+\beta\left(W_{\mathrm{j}}-1\right)} - \frac{M_{\mathrm{i,j}}}{1+\beta\left(M_{\mathrm{W,j}}-1\right)}\right)^2}{2\sigma_{\mathrm{r}}^2}, \tag{5}$$

where F_{i} is the observed normalized flux for a narrow passband i, σ_{r} is the flux ratio scatter, α is the constant baseline level for the flux ratio, and β is a scaling factor applied to both observed and modeled broadband flux.

The approach is similar to the often-used method of first fitting the wide passband and subtracting the residuals from the narrow-band light curves, but slightly more robust, since we are marginalizing over the baseline and scale parameters α and β, and modeling the relative flux explicitly.

4.5 Numerical methods

The transit light curves were modeled using PyTransit,[4] a Python package implementing Mandel-Agol and Giménez transit models optimized for transmission spectroscopy. The MCMC sampling was carried out with emcee (Foreman-Mackey et al. 2013) a Python implementation of the affine invariant MCMC sampler by Goodman & Weare (2010). The sampler was initialized using a population of parameter vectors clumped around the local posterior maximum using PyDE,[5] a Python implementation of the Differential Evolution global optimization algorithm (Storn & Price 1997)

The analysis also uses the large set of tools build around SciPy and NumPy (van der Walt et al. 2011): IPython (Perez & Granger 2007), Pandas (Mckinney 2010), matplotlib (Hunter 2007), seaborn,[6] PyFITS,[7] and F2PY (Peterson 2009). The Gaussian Processes were computed using George.[8] (Ambikasaran et al. 2014)

5 WASP-19B: PARTIAL TRANSIT OBSERVED DURING THE KMOS COMMISSIONING

5.1 Overview

The partial WASP-19b transit observed during the KMOS commissioning cannot be considered for a serious transmission spectrum analysis due to a lack of in-transit coverage. Nevertheless, since the observation conditions were stable, it can be used to assess the precision of the transit depth estimates that can be achieved with the KMOS.

The observations covered only the last third of the transit, and alone cannot constrain any of the relevant properties (radius ratio, impact parameter, transit duration, etc.) However, since WASP-19b

[4] Available from `github.com/hpparvi/PyTransit`.
[5] Available from `github.com/hpparvi/PyDE`.
[6] `stanford.edu/~mwaskom/software/seaborn`
[7] PyFITS is a product of the Space Telescope Science Institute, which is operated by AURA for NASA
[8] `Availablefrom\url{dan.iel.fm/george}`

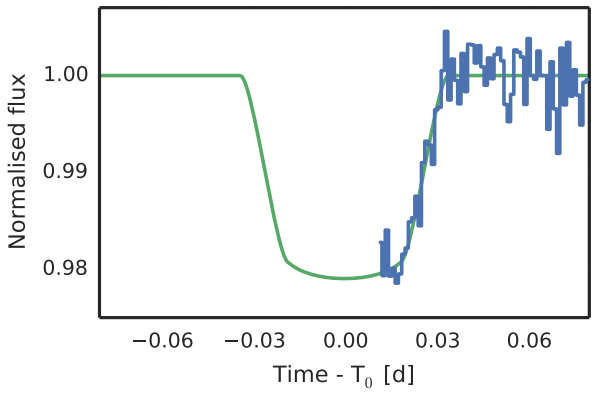

Figure 6. WASP-19b H-band transit light curve with the fitted model.

Parameter	Unit	Source	Prior
Orbital period	days	a	N(0.7888391, 1.1×10^{-7})
Average area ratio	$A_\star$	a	N(0.021, 0.001)
Impact parameter	$R_\star$	a	N(0.681, 0.008)
Semi-major axis	$R_\star$	b	N(3.552, 0.093)
Limb darkening a		c	N(0.138, 0.010)
Limb darkening b		c	N(0.252, 0.020)
Parameters per narrow passband with uninformative priors			
Relative radius ratio			
Flux ratio baseline			
Flux ratio scatter			

Table 2. Informative priors used in the analysis of WASP-19b partial transit. Sources: a) Bean et al. (2013), b) Hellier et al. (2011), c) Claret et al. (2013).

has been extensively studied in transit (Hebb et al. 2010; Bean et al. 2013; Mancini et al. 2013; Mandell et al. 2013; Huitson et al. 2013; Tregloan-Reed et al. 2013), occultation (Gibson et al. 2010; Mancini et al. 2013; Anderson et al. 2013), and RVs (Hellier et al. 2011), the planet's geometric and orbital parameters can be strongly constrained using informative priors, listed in Table 2.

5.2 Results from the partial transit analysis

We carry out the WASP-19b analysis using a broadband light curve covering most of the H-band (roughly from 1.5 μm to 1.8 μm, ~700 pixels), shown in Fig. 6, and six narrow-band light curves covering 43 nm (~100 pixels) each, as shown in Fig. 7. We detail only the H-band analysis, since the K-band results are qualitatively similar.

The final light curves were created by dividing the WASP-19 light curve with the sum of the six reference star light curves (this was found to yield the lowest final out-of-transit (OOT) ptp-scatter.) The OOT scatter of the final normalised broadband light curve was 2.3×10^{-3}, while the theoretical shot noise was 2.4×10^{-4}. The systematics were tested not to be correlated with any of the simultaneously measured auxiliary parameters, and we decided not to use Gaussian processes with this dataset.

We show our results for the six narrow bands in H in Fig. 7. The 68% posterior central interval widths are around 0.1%, which is promising considering observing full transits with sufficient pre-ingress and post-egress baselines. However, the variation in the estimates is greater than what expected from theory (or measured by others, such as Bean et al. 2013), and thus we must consider uncertainties significantly underestimated.

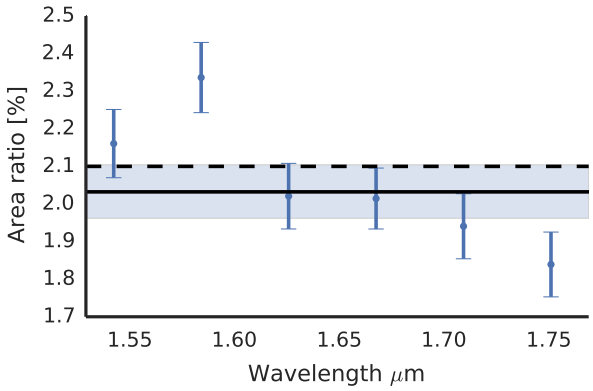

Figure 8. WASP-19b transmission spectrum estimated from the partial transit. The dots show the area ratio posterior medians, the errorbars correspond to the 68% posterior central intervals, the continuous black line shows the posterior median for the wide H-band area ratio, the shaded region its 68% posterior central interval, and the slashed black line the prior mean.

5.3 Precision test with a mock dataset

While it was highly unlikely from the beginning that the analysis above would yield useful results, we can still use the observations to test the precision we could get if we would have observed a full transit with proper baseline. We use the model fitted to the wide H-band, the wide H-band residuals, and the narrow band residuals to create a mock dataset with similar time resolution as with the original observations. We model the broadband data by summing the broadband model and the broadband residuals (tiled periodically to cover the whole time-span), and show the resulting light curve in Fig. 9. Next, we create the narrow-band light curves by summing a narrow-band model and the narrow band residuals for each band. We set four of the narrow passbands to have exactly the same transit depth as the wide band (passbands 1,2,3, and 6), while the fourth narrow band is set to have the area ratio of 1.95% and the fifth an area ratio of 2.11%.

We show the transmission spectrum obtained using the mock data in Fig. 10. The area ratio estimates match well the input data, and the average 68% area ratio posterior interval width is ~0.05%, half of its value for the observed partial dataset.

Based on this test, we can be cautiously optimistic about using KMOS in transmission spectroscopy, given that we get sufficient pre-ingress and post-egress baseline. However, a note must be taken that the test above underestimates the effects from white noise, since repeating the residuals also repeats the white noise component.

6 GJ 1214B

6.1 Overview

The GJ 1214b observations suffered from inaccurate centering. The host star has a high proper motion, and it ended up being positioned close to the a corner of the IFU, as shown in Fig. 3. The poor

[t!]

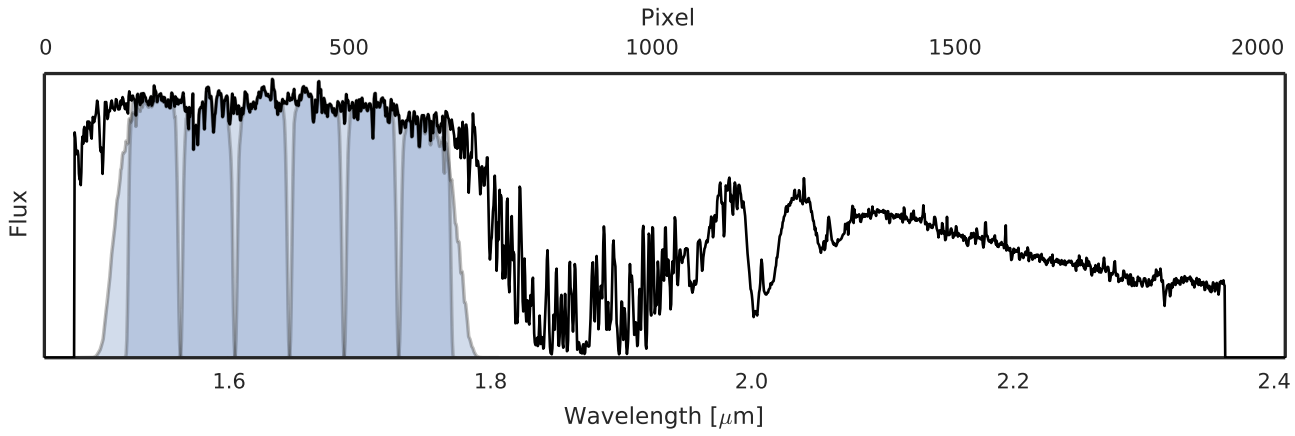

Figure 7. WASP-19 HK spectrum with the wide H-band marked as light blue and the narrow bands marked as darker blue.

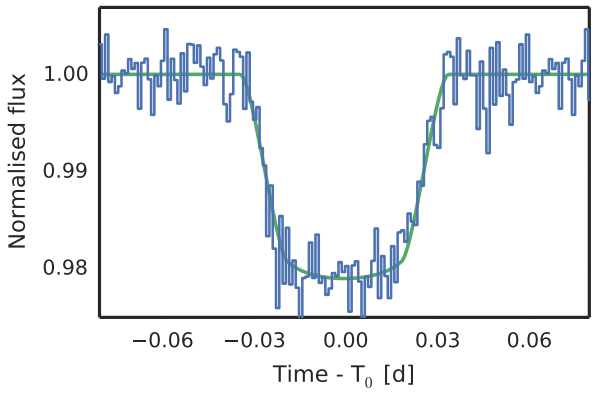

Figure 9. Mock WASP-19b H-band transit light curve with the fitted model.

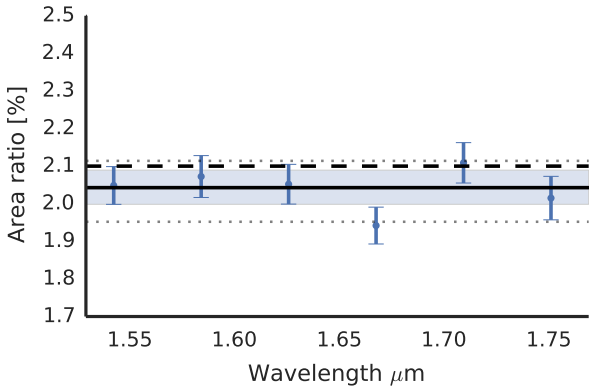

Figure 10. WASP-19b transmission spectrum from the mock dataset. The dots show the area ratio posterior medians, the errorbars correspond to the 68% posterior central intervals, the continuous black line shows the wide H-band area ratio posterior median, the shaded region its 68% posterior central interval, and the slashed black line the prior mean. The dotted lines show the area ratio values set to fourth and fifth passbands, while all other passbands have the same area ratio as the wide passband.

centering leads to strong seeing-related systematics due to varying amounts of flux being lost outside the IFU, as seen in Fig. 11. Dividing the target star light curve with the sum of the three best comparison star light curves mitigate the systematics (Fig. 12), but not sufficiently. One reason for this is that the flux loss is non-linear and depends on the location of the PSF center in the IFU (the elliptically symmetric PSF is bounded by a rectangular area). The systematics from flux loss can be mitigated by modeling the flux loss for each star based on PSF fitting, as carried out for the HD 209458 observations discussed later in Sect. 7, but the approach was not sufficient for this dataset.

Further, the light curves of comparison stars C1 and C2 in Fig. 11 show a sudden jump during the after-transit baseline observations. This jump relates to a sudden shift in the spectrum (in wavelength dimension) in some of the IFUs. The reason for the shift is unknown, and likely arises during the generation of datacubes from the raw data. The shift occurs at the same position for all IFUs where it is noticeable, and happens in detectorss 1 and 2. The photometric signal arises from the subtraction of a target IFU spectrum featuring a wavelength shift with a sky IFU spectrum without the shift.

6.2 Analysis

The strong systematics render the dataset useless in basic characterization, and more so in transmission spectroscopy. The transit is barely recognizable from the broadband relative photometry light curve (Fig. 12), and the systematics are not constant in wavelength. The measured broadband light curve rms scatter is $\sim 8{\times}10^{-3}$, while the theoretical shot noise level is $\sim 1.3 \times 10^{-4}$.

We nevertheless investigate whether Gaussian processes (Gibson et al. 2012; Roberts et al. 2013; Rasmussen & Williams 2006) can be used to model the systematics based on a set of auxiliary parameters (pressure, seeing, sky level), and the modeled flux loss based on PSF fitting. We condition a GP with an squared exponential kernel to the out-of-transit points (with a separate input-scale for each parameter) after first optimizing the GP likelihood as a function of kernel hyperparameters. The GP can explain some of

[t!]

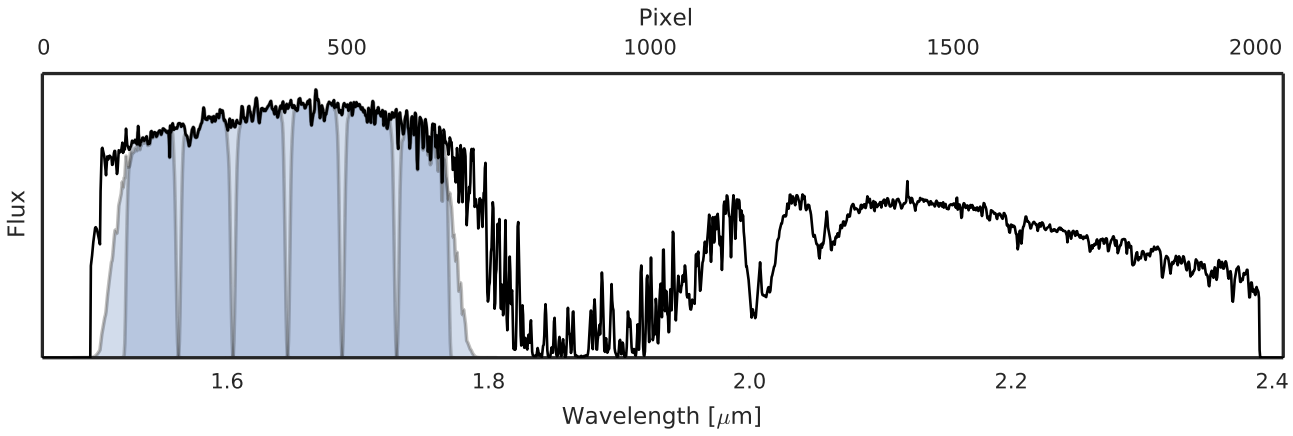

Figure 13. GJ 1214 HK spectrum with the wide H-band marked as light blue and the narrow bands marked as darker blue.

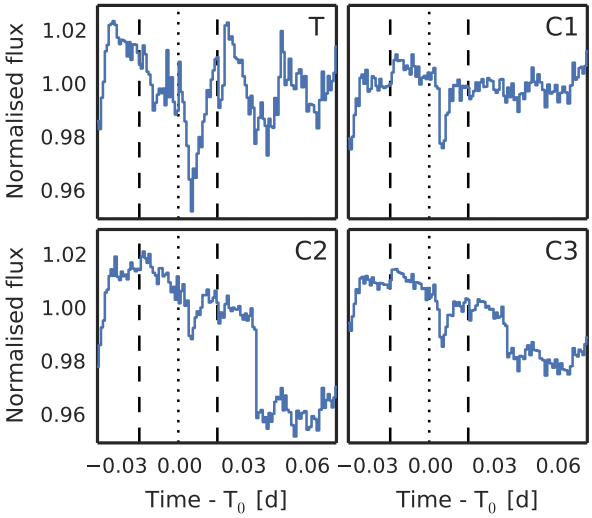

Figure 11. H-band light curves for GJ 1214 (T) and the three best comparison stars (C1-C3). The expected transit start, center, and end times are marked as vertical dotted lines. Both C2 and C3 feature a sharp jump during the second half of the observations.

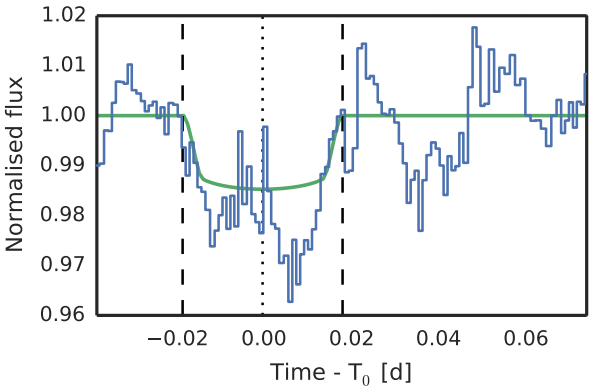

Figure 12. GJ 1214b H-band transit light curve with a model based on parameter estimates from Kreidberg et al. (2014). The dotted line shows the predicted transit center, and the dashed lines the beginning of ingress and the end of egress.

the variability, but not to a sufficient level for a meaningful analysis.

We further carry out an analysis similar to the WASP-19b transmission spectroscopy analysis. We set very narrow priors on the orbital parameter, limb darkening, and average radius ratio, based on the results by Kreidberg et al. (2014), and set to estimate the narrow-band transit depths over the H-band based on Eq. 5.

The results of the analysis are shown in Fig. 14. GJ 1214b is known to have an extremely flat transmission spectrum most likely dominated by clouds (Kreidberg et al. 2014), and we again conclude that the scatter seen in the transit depth is not a real feature of the planet's transmission spectrum, and that our uncertainties are significantly underestimated.

7 HD 209458B

7.1 Overview

The HD 209458 observations included pre-imaging to ensure that all the stars were well centered in the IFUs (Fig. 5). The centering is good, but the number of pre- and post-transit baseline exposures is limited due to technical issues and observational limitations. The flux loss due to the seeing variations (from 0.8″ to 1.4″) is again the main source of systematics, but this can be accounted for to a degree using a PSF-modeling based flux-loss model and Gaussian processes. The transit is not visible in the raw light curve (Fig. 15), and the quality of the data is lower than what could be expected based on the WASP-19 observations.

7.2 Analysis and Results

We begin by investigating how well the systematics can be modeled using modeled flux losses and Gaussian processes. First, we calculate the flux losses for each star used in the analysis based on

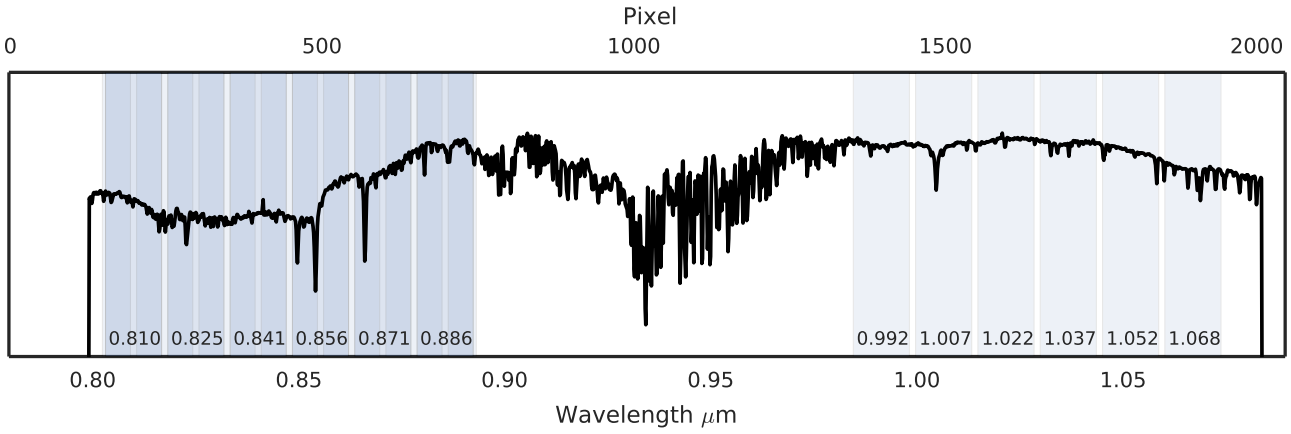

Figure 16. HD 209458b spectrum with the *i* and *z* bands marked with light blue, and 6 narrow passbands for each marked as darker blue. A separate analysis was carried out for the *i* band divided into 12 narrow passbands.

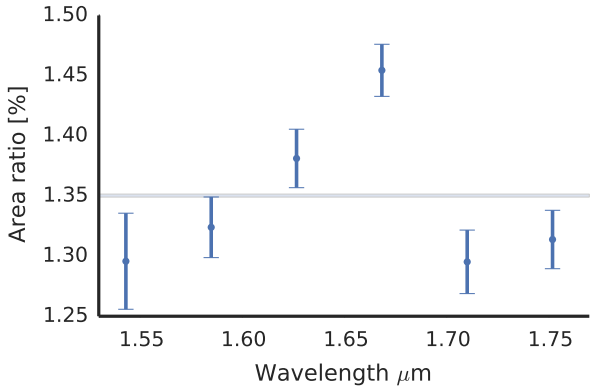

Figure 14. GJ 1214b H-band transmission spectrum estimated from the data. The light-blue shaded area covers the whole range of are ratio estimates by Kreidberg et al. (2014), assuming an average area ratio of 1.344%. The scatter can safely be assumed to be caused by unaccounted-for systematics, and the estimate uncertainties are severely underestimated.

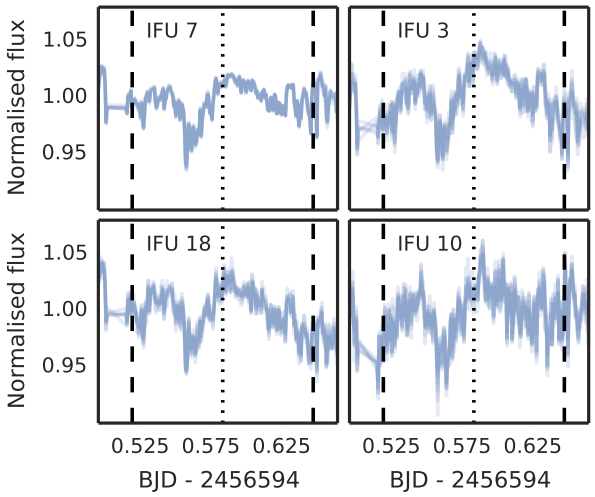

Figure 15. White light curves for HD 209458b (IFU 7) and the three brightest reference stars. The expected transit start, center, and end times are marked as vertical dotted lines.

PSF fitting. We continue by estimating the fractional flux loss not corrected by the relative photometry, shown in Fig. 4, and use this as an input parameter to a Gaussian process with a squared exponential kernel.

After accounting for the flux loss, the light curve contains still clear systematics from varying sky background. We include the sky level as a second GP input parameter, and end up with a GP characterized by two additive squared exponential kernels, each with an independent input scale parameter. We optimize the three GP hyperparameters (output scale, two input scales), and keep them fixed during the MCMC run.

The final *z* broadband light curve, the GP-based systematics, and the systematics-corrected light curve are shown in Fig. 17. Our broadband light curve rms scatter estimate of 3.5×10^{-3} is almost 20 times the theoretical shot noise estimate of $\sim 2\times10^{-4}$. The GP using the estimated sky level and the results from the flux loss modeling is capable of explaining most of the systematics. However, the average radius ratio estimate for the wide *z*-band disagrees with the previous studies, and the transmission spectrum for *i* and *z* bands, shown in Fig. 18, shows scatter with a significantly larger amplitude than what expected from theory (or observed previously, for example Sing et al. 2008; Evans et al. 2015). It is clear that our baseline observations are not sufficient to constrain the transit depth, and our parameter estimate uncertainties are again underestimated.

8 CONCLUSIONS

The test based on the observations of the partial transit of WASP-19b gives cautious support that KMOS can, in optimal conditions, deliver the photometric precision needed for transmission spectroscopy. However, the observations of GJ 1214b and HD 209458b teach us that the targets must be highly favorable (bright stars with several comparison stars of similar magnitude located within the patrol field), the IFU centering must be exact (which can be difficult

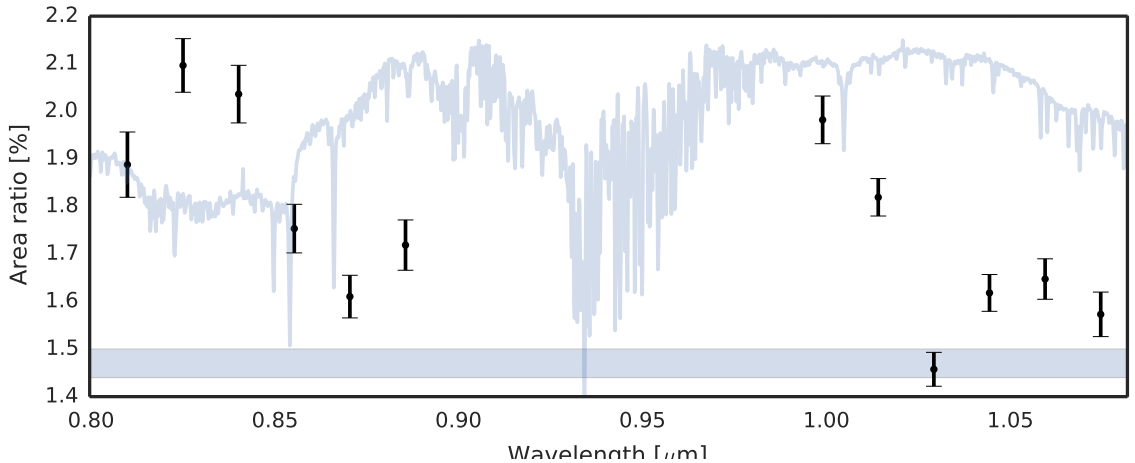

Figure 18. HD 209458b *i* and *z* transmission spectrum. The blue band shows the maximum expected variation in the are ratio.

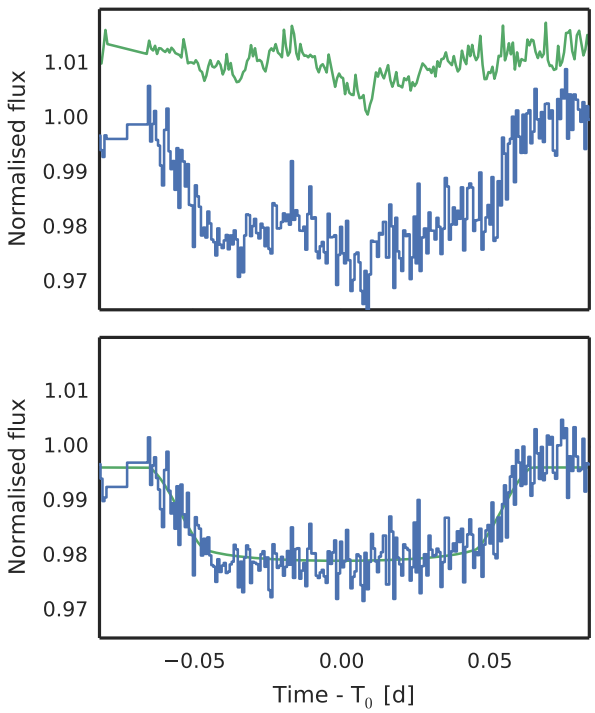

Figure 17. Top figure: HD 209458b z-band light curve divided by the sum of the fluxes from the three brightest reference stars (blue line) and the mean systematics explained by a Gaussian process model (green). Bottom figure: The relative photometry corrected with the mean systematics (blue) and the fitted transit model (green).

for targets with high proper motion) favorable targets, the observation conditions need to be stable due to the issues with flux loss, and sufficient out-of-transit baseline must be obtained. These requirements combined with time-critical nature of transit observations—and the need to repeat the observations several times—lead to rigid constraints on telescope scheduling, and the sensitivity on seeing variations means that the risk of not obtaining useful data is high.

Finally, the fact that similar or better quality can routinely be achieved with telescopes and instruments with smaller observation pressure (Bean et al. 2013, 2010, 2011; Croll et al. 2011; Crossfield et al. 2011), and the fact that transmission spectroscopy does not gain significantly from having spatially resolved spectra, we conclude that while transmission spectroscopy can be carried out with KMOS, it is not the optimal science case to take advantage of the abilities offered by the instrument and a 8.2 m telescope. If one nevertheless considers a case where the benefits can be larger than the risks, one should

- do pre-imaging to ensure that all the stars are well centered in their IFUs,
- ensure that all the target IFUs have good sky IFUs on the same detector as the targets,
- ensure that a proper OOT baseline is observed.

ACKNOWLEDGEMENTS

We warmly thank the anonymous referee for their useful comments and recommendations. HP has received support from the Leverhulme Research Project grant RPG-2012-661.

REFERENCES

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2014

Anderson D. R., Smith a. M. S., Madhusudhan N., Wheatley P. J., Collier Cameron A., Hellier C., Campo C., Gillon M., Harrington J., Maxted P. F. L., Pollacco D., Queloz D., Smalley B., Triaud a. H. M. J., West R. G., 2013, Mon. Not. R. Astron. Soc., 430, 3422

Bean J. L., Désert J.-M., Kabath P., Stalder B., Seager S., Miller-Ricci Kempton E., Berta Z. K., Homeier D., Walsh S., Seifahrt A., 2011, Astrophys. J., 743, 92

Bean J. L., Désert J.-M., Seifahrt A., Madhusudhan N., Chilingarian I., Homeier D., Szentgyorgyi A., 2013, Astrophys. J., 771, 108

Bean J. L., Kempton E. M.-R., Homeier D., 2010, Nature, 468, 669

Berta Z. K., Charbonneau D., Désert J.-M., Miller-Ricci Kempton E., McCullough P. R., Burke C. J., Fortney J. J., Irwin J., Nutzman P., Homeier D., 2012, Astrophys. J., 747, 35
Brown T. M., 2001, Astrophys. J., 553, 1006
Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, ApJ, 568, 377
Claret A., Hauschildt P. H., Witte S., 2013, Astron. Astrophys., 552, A16
Croll B., Albert L., Jayawardhana R., Kempton E. M.-R., Fortney J. J., Murray N., Neilson H., 2011, Spectrum, p. 12
Crossfield I. J. M., Barman T., Hansen B. M. S., 2011, Astrophys. J., 736, 132
Davies R. I., Agudo Berbel A., Wiezorrek E., Cirasuolo M., Förster Schreiber N. M., Jung Y., Muschielok B., Ott T., Ramsay S., Schlichter J., Sharples R., Wegner M., 2013, Astron. Astrophys., 558, A56
Evans T. M., Aigrain S., Gibson N., Barstow J. K., Amundsen D. S., Tremblin P., Mourier P., 2015, Mon. Not. R. Astron. Soc., 451, 680
Evans T. M., Pont F., Sing D. K., Aigrain S., Barstow J. K., Désert J.-M., Gibson N., Heng K., Knutson H. a., Lecavelier des Etangs A., 2013, Astrophys. J., 772, L16
Fairley A., 2012, Technical report, KMOS Data Reduction Library Design & User Manual
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publ. Astron. Soc. Pacific, 125, 306
Gibson N., Aigrain S., Pollacco D. L., Barros S. C. C., Hebb L., Hrudková M., Simpson E. K., Skillen I., West R., 2010, MNRAS, 404, 1
Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013a, Mon. Not. R. Astron. Soc., 428, 3680
Gibson N. P., Aigrain S., Barstow J. K., Evans T. M., Fletcher L. N., Irwin P. G. J., 2013b, Mon. Not. R. Astron. Soc., 436, 2974
Gibson N. P., Aigrain S., Roberts S., Evans T. M., Osborne M., Pont F., 2012, Mon. Not. R. Astron. Soc., 419, 2683
Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
Hebb L., Collier-Cameron A., Triaud A., Lister T., Smalley B., Maxted P., Hellier C., Anderson D., Pollacco D., Gillon M., Queloz D., West R., Bentley S., Enoch B., Haswell C., Horne K., Mayor M., Pepe F., Segransan D., Skillen I., Udry S., Wheatley P., 2010, Astrophys. J., 708, 224
Hellier C., Anderson D. R., Cameron a. C., Miller G. R. M., Queloz D., Smalley B., Southworth J., Triaud a. H. M. J., 2011, ApJ, 31, 7
Huitson C. M., Sing D. K., Pont F., Fortney J. J., Burrows A. S., Wilson P. a., Ballester G. E., Nikolov N., Gibson N. P., Deming D., Aigrain S., Evans T. M., Henry G. W., Lecavelier des Etangs A., Showman a. P., Vidal-Madjar A., Zahnle K., 2013, Mon. Not. R. Astron. Soc., 434, 3252
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jordán A., Espinoza N., Rabus M., Eyheramendy S., Sing D. K., Désert J.-M., Bakos G. A., Fortney J. J., López-Morales M., Maxted P. F. L., Triaud A. H. M. J., Szentgyorgyi A., 2013, Astrophys. J., 778, 184
Knutson H. A., Dragomir D., Kreidberg L., Kempton E. M.-R., McCullough P. R., Fortney J. J., Bean J. L., Gillon M., Homeier D., Howard A. W., 2014, Astrophys. J., 794, 155
Kreidberg L., Bean J. L., Désert J.-M., Benneke B., Deming D., Stevenson K. B., Seager S., Berta-Thompson Z., Seifahrt A., Homeier D., 2014, Nature, 505, 69
Mckinney W., 2010, in 9th Python Sci. Conf. Vol. 1697900, Data Structures for Statistical Computing in Python. pp 51–56
Mancini L., Ciceri S., Chen G. e. a., 2013, Mon. Not. R. Astron. Soc., 436, 2
Mandell A. M., Haynes K., Sinukoff E., Madhusudhan N., Burrows A. S., Deming D., 2013, Astrophys. J., 779, 128
Murgas F., Pallé E., Zapatero Osorio M. R., Nortmann L., Hoyer S., Cabrera-Lavers A., 2014, Astron. Astrophys., 563, A41
Perez F., Granger B., 2007, Comput. Sci. Eng., pp 21–29
Peterson P., 2009, Int. J. Comput. Sci. Eng., 4, 296
Pont F., Zucker S., Queloz D., 2006, MNRAS, 373, 231
Rasmussen C. E., Williams C., 2006, Gaussian processes for machine learning. The MIT Press
Roberts S., Osborne M., Ebden M., Reece S., Gibson N., Aigrain S., 2013, Philos. Trans. A. Math. Phys. Eng. Sci., 371, 20110550
Sharples R., Bender R., Berbel A., 2013, The Messenger, pp 2012–2014
Sing D. K., Lecavelier des Etangs A., Fortney J. J., Burrows A. S., Pont F., Wakeford H. R., Ballester G. E., Nikolov N., Henry G. W., Aigrain S., Deming D., Evans T. M., Gibson N. P., Huitson C. M., Knutson H., Showman A. P., Vidal-Madjar A., Wilson P. A., Williamson M. H., Zahnle K., 2013, Mon. Not. R. Astron. Soc., 436, 2956
Sing D. K., Pont F., Aigrain S., Charbonneau D., Désert J.-M., Gibson N., Gilliland R., Hayek W., Henry G., Knutson H., des Etangs A. L., Mazeh T., Shporer A., 2011, Mon. Not. R. Astron. Soc., 416, 1443
Sing D. K., VidalMadjar A., Désert J., Lecavelier des Etangs A., Ballester G., 2008, Astrophys. J., 686, 658
Storn R., Price K., 1997, J. Glob. Optim., 11, 341
Tregloan-Reed J., Southworth J., Tappert C., 2013, Mon. Not. R. Astron. Soc., 428, 3671
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22