Advances in ground-based characterization of transiting exoplanets using novel data analysis methods

Panwar, V.

Publication date
2023

Document Version
Final published version

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Advances in ground-based characterization of transiting exoplanets using novel data analysis methods

Vatsal Panwar
Advances in ground-based characterization of transiting exoplanets using novel data analysis methods
Thesis, Anton Pannekoek Institute, Universiteit van Amsterdam

Cover by Vatsal Panwar, Gemini image : Manuel Paredes, Sun from SDO, Sisi
Printed by Gildeprint

Advances in ground-based characterization of transiting exoplanets using novel data analysis methods

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel
op 12 januari 2023, te 13.00 uur

door

Vatsal Panwar

geboren te Bhilwara
Promotiecommissie:

Promotor(es): dr. J.M.L.B. Désert Universiteit van Amsterdam
Copromotor(es): prof. dr. Carsten Dominik Universiteit van Amsterdam

Overige leden: prof. dr. L. Kaper Universiteit van Amsterdam
prof. dr. R.A.D. Wijnands Universiteit van Amsterdam
dr. A. Oklopčić Universiteit van Amsterdam
prof. dr. H. Rauer DLR
dr. N. Gibson Trinity College Dublin

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
sitāroñ se aage jahāñ aur bhi haiñ
(There are yet more worlds beyond the stars)
Allama Iqbal
Contents

1 Introduction
  1.1 Three decades of exoplanet detection and characterization at a glance ... 1
  1.2 Towards studying Earth-like planets ........................................ 3
  1.3 Atmospheric characterization of transiting exoplanet atmospheres ....... 3
    1.3.1 Spectroscopy of transiting exoplanets at low and at high spectral res-
          olution .......................................................... 4
    1.3.2 Low-resolution spectroscopy ....................................... 4
    1.3.3 High-resolution spectroscopy ..................................... 8
  1.4 Challenges to ground-based low resolution spectroscopy of transiting ex-
      planets ............................................................. 9
    1.4.1 Earth’s atmosphere ................................................ 9
    1.4.2 Telescope and instrumental noise .................................. 12
    1.4.3 Stellar variability .................................................. 12
  1.5 Pathways to mitigate the challenges to ground-based low-resolution transit
      spectroscopy ......................................................... 14
    1.5.1 Multi-object spectroscopy: a solution to obtaining low-resolution spec-
          tra of transiting exoplanets ..................................... 14
    1.5.2 Data driven methods to model the systematics in ground-based transit
          light curves ...................................................... 19
  1.6 Strengths of applying Gaussian Processes regression to measuring exoplanet
      spectra .............................................................. 22
    1.6.1 Widening the scope of low-resolution transit spectroscopy to exoplan-
          ets orbiting bright stars ....................................... 23
    1.6.2 The bright star opportunity for ground-based follow-up of transiting
          rocky exoplanets ............................................... 24
  1.7 This thesis .............................................................. 24

2 Enhanced metallicity of an exoplanet atmosphere from precise alkali
      lines measurements:
Ground-based low-resolution detection of Na and K in the atmosphere of a
cloud-free gas-giant ......................................................... 27
  2.1 Introduction ........................................................... 28
    2.1.1 The XO-2 System ................................................. 29
2.2 Observations and Data Reduction ............................................. 30
  2.2.1 Observations ............................................................. 30
  2.2.2 Data reduction .......................................................... 31
2.3 Analysis ............................................................................. 33
2.4 Atmospheric modelling and interpretation of the transmission spectrum . 34
2.5 Results and Discussion ......................................................... 38
  2.5.1 Results .......................................................... 38
  2.5.2 Impact of stellar variability on the transmission spectrum ....... 40
  2.5.3 Implications of the high metallicity of XO-2Nb ................. 40
  2.5.4 Connection with the differential chemical abundances of XO-2N and XO-2S .............................................................. 42
  2.5.5 The large Na/K and the depletion of K ......................... 43
2.6 Conclusions ........................................................................ 46
2.7 Acknowledgements ............................................................. 47

3 A new method to measure the spectra of transiting exoplanet atmospheres using multi-object spectroscopy 49
  3.1 Introduction ....................................................................... 50
  3.2 Observations .......................................................... 52
    3.2.1 The warm Neptune HAT-P-26b ................................... 52
    3.2.2 GMOS Transmission spectroscopy ............................... 53
  3.3 Data Reduction ............................................................. 55
    3.3.1 GMOS data .......................................................... 55
  3.4 Transit Light Curve Analysis ............................................. 58
    3.4.1 Modelling systematics in transit light curves ............... 60
    3.4.2 Analysis of White Transit Light Curves ....................... 65
    3.4.3 Analysis of Spectroscopic Light Curves ..................... 72
  3.5 Results and Discussion ..................................................... 77
    3.5.1 Comparison of the two Methods and Implications ........ 77
    3.5.2 Interpretation of the Optical to NIR Transmission Spectrum .... 95
    3.5.3 Transit Timing Variations ........................................ 98
  3.6 Conclusions .............................................................. 101
  3.7 Acknowledgements ......................................................... 103

Appendices 105
  3.A Testing robustness of using common-mode trend as a GP regressor for \lambda LC . 105
  3.B Model selection criteria for GP regressor combinations for fitting white transit light curves ............................................. 107

4 A new method to correct for host star variability in multi-epoch observations of exoplanet transmission spectra 115
  4.1 Introduction .............................................................. 116
  4.2 The case of WASP-19b ..................................................... 117
4.3 Multi-epoch transit observations of WASP-19b .............................. 120
  4.3.1 Gemini/GMOS transit observations of WASP-19b ....................... 120
4.4 Data Reduction .......................................................... 121
  4.4.1 Data reduction of the Gemini/GMOS observations ....................... 121
  4.4.2 Flat-Fielding ....................................................... 122
  4.4.3 Spectral extraction ............................................... 123
  4.4.4 Wavelength calibration ........................................... 123
  4.4.5 Extracting the light curves ...................................... 125
4.5 Analysis ................................................................. 126
  4.5.1 Modelling systematics in GMOS transit light curves .................... 126
  4.5.2 Analysis of White Transit Light Curves ................................ 127
  4.5.3 Analysis of Spectroscopic Light Curves ................................ 132
4.6 Discussion ............................................................... 138
  4.6.1 Effect of stellar activity on the transmission spectrum of WASP-19b . 138
  4.6.2 A new empirical approach to correct for stellar variability across multiple epochs ......................................................... 141
  4.6.3 Optical and near-infrared transmission spectrum of WASP-19 and comparison with forward atmospheric models .................. 145
  4.6.4 Comparison of the GMOS transmission spectrum of WASP-19b with previous studies ......................................................... 150
4.7 Conclusions .............................................................. 151
4.8 Acknowledgements ....................................................... 154

Appendices ................................................................. 155
  4.A Photometric monitoring of WASP-19 ..................................... 155
    4.A.1 TESS observations of WASP-19b .................................... 155
    4.A.2 Ground based photometric monitoring of WASP-19 by LCOGT ...... 158
  4.B Illustration of the transmission spectral slope vs offset analysis .... 160
  4.C Transmission spectra tables ........................................... 160
  4.D Supplementary material ................................................. 160

5 On the follow-up of small and rocky exoplanets orbiting bright stars using ground-based spectrophotometry ........................................ 191
  5.1 Introduction ............................................................ 192
  5.2 Observations ............................................................ 194
  5.3 Data reduction .......................................................... 194
  5.4 Transit light curve analysis ............................................ 195
    5.4.1 White light curve: 1D GP with and without comparison star as a GP regressor .............................................................. 197
    5.4.2 λLCs: 2D GP without using the comparison star as a GP regressor .... 199
  5.5 Results and Discussion ................................................ 201
    5.5.1 Transit parameters measured from the 1D and the 2D GP models .... 201
    5.5.2 Caveats of the GMOS measurements .................................. 206
6 Common trends in close-in gas giant exoplanet atmospheres from MOS surveys

*Using a library of tools for modelling spectrophotometric transit light curves with Gaussian Processes (GHALIB)*

6.1 Introduction ......................................................... 216
6.2 Overview of ghalib .................................................. 218
6.3 Observations from a ground-based survey of transiting gas-giants ................. 220
  6.3.1 Gemini/GMOS observations .................................. 220
  6.3.2 Keck/MOSFIRE observations .................................. 223
6.4 Data reduction ......................................................... 223
6.5 Application of ghalib to GMOS and MOSFIRE observations ...................... 225
6.6 Results and Discussion ................................................ 226
  6.6.1 Transit parameters from the white light curves ..................... 226
  6.6.2 Transmission and emission spectra across the sample ................. 226
  6.6.3 Conclusions and future perspective ............................... 227
6.7 Acknowledgements .................................................. 230

Appendices ....................................................................... 231
  6.A ghalib fits to white transit light curves from GMOS and MOSFIRE .......... 231
  6.B Measured transit properties from ghalib .................................. 231

Appendices ....................................................................... 231

Bibliography ................................................................. 243

Contribution from co-authors ............................................. 259

Summary .......................................................................... 261

Samenvatting .................................................................... 265

Acknowledgements ............................................................ 269
Chapter 1

Introduction

1.1 Three decades of exoplanet detection and characterization at a glance

The study of extrasolar planets (exoplanets) forms the backbone of the contemporary quest towards understanding the origins of our solar system, the earth, and the life thereupon. A quick glance at the milestones leading up to and since the detection of the first exoplanet around a solar type star (Mayor et al. (1995), Mayor & Queloz (1995)) shows that the science of exoplanets has been observationally driven. 51 Peg b, a Jupiter mass planet in an unexpectedly compact orbit around its host star, was detected by Mayor & Queloz (1995) using the echelle spectrograph ELODIE (Baranne et al. (1996)). ELODIE employed the then state-of-the-art high-resolution spectroscopy instrumentation to push the technique of cross-correlation spectroscopy to the radial velocity precision needed for detecting Jupiter mass objects around sun like stars.

Soon after its detection, the first attempts to obtain time series spectroscopy of 51 Peg b were made to search for atmospheric signatures during the transit of its extended exosphere by Coustenis et al. (1997) and Rauer et al. (2000). Eventually the hunt for exoplanets and their atmospheres transiting their host stars got a boost with the discovery of the first transit of the hot Jupiter HD 209458b (Charbonneau et al. (2001), Henry et al. (2000)). Transit- ing exoplanets opened the doors to measuring an exoplanet’s radius and also removing the degeneracy between the planetary mass and planetary orbital inclination inherent in radial velocity measurements. Simply using the mass and radius of a planet, its interior bulk composition can be constrained by comparing its position in the mass-radius space (e.g. Fortney et al. (2007), Seager et al. (2007)) with the theoretically predicted mass-radius relationship for various compositions (e.g., see Figure 1.1).

The number of exoplanets, most of them gas-giants, continued to grow as a result of a number of ground-based radial velocity and transit photometry surveys. The next breakthrough for exoplanet detection came with the advent of space-based transit photometry surveys like CoRoT (Barge et al. (2008)) and most notably from Kepler (Borucki et al. (2010)). Transit photometry has in fact been the most prolific method of exoplanet detection and characterization. In terms of simply the total number, the transit method has been used to discover
2 Introduction

Figure 1.1: Mass radius relationship between a selected range of confirmed exoplanets discovered using transits (green), radial velocity (red), timing variations (yellow, including transit, eclipse, pulsar, and pulsation timing variations), and orbital brightness modulation. Overplotted are expected density curves for planet compositions equivalent to 100% water, silicates, and iron. Arrows represent the lower limits on the planetary mass from solely transit observations.

3930 exoplanets out of the 5178 confirmed exoplanets to date\(^1\). The abundance and diversity of exoplanets yielded by *Kepler* enabled the first studies of exoplanets as a population. Importantly, *Kepler* enabled studies of system architecture (*Fabrycky* et al. (2014)) and demographics (*Fulton* et al. (2017)), revealing tantalizing imprints of planet formation and evolution processes across the exoplanet population. On the other hand close inspection of some individual exoplanets with peculiar properties began to provide indications of new and previously unexpected physical and chemical regimes of planetary atmospheres and interiors (e.g., ultra-hot Jupiters (*Arcangeli* et al. (2018))).

Parallel advancements were made both in new ways to detect and characterize exoplanets. Spectrographs on *Hubble Space Telescope* and infrared photometers on *Spitzer Space Telescope* were harnessed to zoom into more detailed properties of exoplanets. This started with the first detection of the atmosphere of HD 209458 (*Charbonneau* et al. (2002)) followed by detection of its escaping atmosphere (*Vidal-Madjar* et al. (2003)), thermal emission (*Charbonneau* et al. (2005)), and scattering due to atmospheric haze (*Lecavelier Des Etangs* et al. (2008), *Sing* et al. (2009), *Désert* et al. (2011)). A novel instrument mode for the *Wide Field Camera 3* on HST (HST/WFC3 spatial scan, *McCullough & Bushouse* (2008), *Deming* et al. (2013)) enabled observations of bright exoplanet host stars, which resulted in the detection of water vapour in the atmosphere of a multitude of transiting exoplanets.

\(^1\)As of September 2022, NASA Exoplanet Archive
1.2 Towards studying Earth-like planets

In the context of time series spectroscopy for atmospheric characterization, ground-based spectrographs, both in low (Bean et al. (2010)) and high-resolution (Snellen et al. (2010)), have been adapted in the last decade through the development of new observational data analysis techniques to reach the high precision needed for characterizing planetary atmospheres.

In summary, the confluence of new instruments, discovery, and characterization of interesting new exoplanetary systems, and new observational data analysis techniques has been the primary driver behind the tremendous acceleration of the field in the first three decades of its existence. This thesis is an attempt to contribute to these three important axes of observational exoplanetary science.

1.2 Towards studying Earth-like planets

One of the most coveted goals in exoplanet science has been to find and characterize planets similar to Earth. In terms of similarity in size, only 150 transiting exoplanets have been discovered to date between the size of 0.9 to 1.1 $R_\oplus$, most of them by *Kepler*. These constitute a diverse collection of exo-Earths ranging from scorching lava planets like Kepler-78b (Sanchis-Ojeda et al. (2013)) with equilibrium temperature in excess of 2000 K, to a few temperate earth analogues like Kepler-186f (Quintana et al. (2014)) in the habitable zone of their host star. The major challenge in detecting exo-Earths is the small transit depth of $\sim 80$ ppm expected for an earth size planet around an exact sun like star. *Kepler* was successful in detecting such small transit depth signals for faint stars (around $V_{mag} \sim 15$) but in a small part of sky, whereas TESS surveying a larger part of sky is only sensitive to such signals for bright stars (around $V_{mag} \sim 10$). Upcoming space telescope *PLATO* (Rauer et al. (2014)) will survey a much larger fraction of sky and is expected to yield many more Earth-sized planets, about a dozen of which could be in habitable zones around solar-type stars (Heller et al. (2022)).

Follow-up characterization of earth-sized planets is no less challenging than detecting them. Two main obstacles to determining their precise size and mass are instrumental systematics and stellar activity of the host star. The latter contaminates both the radial velocity and transit signals. Nevertheless, large ground-based telescopes can play a pivotal role in improving the precision on both planetary size and mass through precise transit depth and transit timing variation (TTV) measurements respectively.

The ultimate goal is to characterize the atmospheres of small rocky exoplanets. To put in perspective the magnitude of this challenge, we segue into the details of atmospheric characterization of exoplanets in general in the next section.

1.3 Atmospheric characterization of transiting exoplanet atmospheres

Characterizing an exoplanet’s atmosphere opens the doors to investigation of its detailed chemical composition in the form of molecular absorbers and cloud condensates, and eventually forms a piece of the puzzle to trace its path of formation and evolution. The geometry
of a transiting exoplanet provides a favourable vantage point for an observer to examine the planet throughout its orbit. A transit or eclipse when observed spectroscopically allows a more detailed inspection of an exoplanet’s atmosphere.

While the primary transit observation is sensitive to the size of the exoplanet, the secondary eclipse observation where the exoplanet is occulted by the host star is sensitive to both its size and the ratio of its dayside brightness temperature to the stellar brightness temperature. Wavelength dependent primary transit depths hence measured by an observer constitute a planet’s transmission spectrum which contains features due to prominent chemical absorbers in its atmosphere (Seager & Sasselov (2000)). Wavelength dependent secondary eclipse depths constitute a planet’s emission spectrum which contains information on the planet’s albedo and the atmospheric thermal and compositional structure (Burrows (2005)).

Phase curves observed spectroscopically provide additional information needed to measure the global three-dimensional atmospheric composition and temperature structure independently (e.g., Stevenson et al. (2014), Arcangeli et al. (2019)). Monitoring the brightness of a planetary system through the course of a planet’s tidally locked orbit, also referred to as a phase curve, can be used to infer a map of its longitudinal temperature and atmospheric dynamics (Knutson et al. (2007), Snellen et al. (2009), Borucki et al. (2009)).

One of the main themes of this thesis is to measure the transmission spectrum of exoplanets from ground, so we focus the discussion in subsequent sections to the observations of primary transits of exoplanets unless specified otherwise.

### 1.3.1 Spectroscopy of transiting exoplanets at low and at high spectral resolution

There are broadly two observational routes to obtaining the spectra of transiting exoplanets. One way is to measure the transit depth at multiple wavelengths, which is done by measuring the full transit or eclipse of an exoplanet and extracting the transmission or emission spectra, separating the planet from the star in time (e.g., Charbonneau et al. (2002)). Another way is to observe the star + planet system at high spectral resolution, and use the Doppler shift of the planetary atmospheric lines to disentangle them spectrally from the star (e.g., Snellen et al. (2010)). We develop on these two main pathways towards observing transiting exoplanet atmospheres, starting with the low-resolution spectroscopy.

### 1.3.2 Low-resolution spectroscopy

The first detection of HD 209458b by Charbonneau et al. (2002) was obtained by using the G750L grism of HST/STIS to measure the low resolution spectra of the star during and out of the planet’s transit. Taking the ratio of the time averaged in-transit and out-of-transit spectra led to the detection of excess absorption due to atmospheric Na around 589 nm. A variation of this classical approach to measuring the low-resolution transmission spectrum is to actually measure the absolute transit depth in wavelength integrated light curves. The optically thin part of a transiting planet’s atmosphere imparts a wavelength dependent signa-
Figure 1.2: Models of broadband exoplanet transits calculated using batman demonstrating the orders of magnitude changes in the expected transit signal strengths for archetypes of an exo-Earth (black), an exo-Neptune (green), and an exo-Jupiter (red) orbiting a solar type star (left panel) and an M-dwarf (right panel) with a common orbital period of 3 days in all cases. The Y axes for both left and right panels show the decrease in stellar flux during the transit of a planet and the X axes for both panels show the time from mid-transit hours.

As illustrated in Figure 1.2, in terms of just broadband transit measurement, exoplanetary transit depths typically range from $\sim 0.01\%$ (100 ppm) for an Earth size planet, $\sim 0.1\%$ (1000 ppm) for a Neptune size planet, and $\sim 1\%$ (10000 ppm) for a Jupiter size planet transiting a solar type star. Planets around smaller stars relative to the Sun, like M dwarf, exhibit higher transit depth signal. However, the wavelength dependent variation in transit depths i.e. feature amplitudes in the transmission spectrum, depends on both the size of the planet and the characteristics of its atmosphere. To estimate the order of magnitude of expected signal from a planet’s atmosphere when observed in transmission, one can assume a hydrostatic equilibrium to obtain the scale height of the planet’s atmosphere, expressed as

$$H = \frac{k_B T}{\mu_m g}$$

(1.1)

where $H$ is the scale height, $k_B$ is the Boltzmann constant, $\mu_m$ is the mean molecular weight of the atmosphere in amu (2.33 for a hydrogen helium dominated atmosphere, ), and $g$ is the surface gravity of the planet.

Transit depth, which itself is proportional to $\delta = \frac{R_P}{R_\star}$, changes with respect to wavelength and its order of magnitude variation due to the optically thin layers of the atmosphere spanning $N$ scale heights can be then approximated as

$$\Delta \delta \approx 2N_H \delta \left( \frac{H}{R_P} \right)$$

(1.2)
Figure 1.3: Change in planetary transit depth ($\Delta \delta$ in ppm on the Y axis) with respect to wavelength (in $\mu$m on X axis) calculated from the platon forward transmission spectrum models for archetype hot Jupiter (left panel) and hot Neptune and hot Earth (right panel) exoplanets for varying mean-molecular weights of the atmospheres ($\mu_m$) and the planetary equilibrium temperatures ($T$). As demonstrated in Equation 1.2, the models here show that the amplitudes of the features in the transmission spectrum vary inversely with $\mu_m$ and directly with $T$. The prominent spectral features in the visible bandpass (0.5 to 1 $\mu$m) are due to H$_2$-He collision induced absorption, alkali atoms (Na, K), and in the infrared to near-infrared bands are due to molecular absorption by H$_2$O, CH$_4$, CO$_2$, and CO.

where $\Delta \delta$ is the change in transit depth with respect to wavelength, $N_H$ is the number of scale heights, $\delta$ is the transit depth, and $R_p$ is the radius of the planet up to which the planet is optically thick in all wavelengths.

As it can be seen from Equation 1.2, the order of magnitude change in transit depth $\Delta \delta$ with respect to wavelength, which is an estimate of the typical amplitude of a feature in the transmission spectrum, is proportional to the scale height of the atmosphere. Assuming equilibrium chemistry and a cloud free atmosphere, a typical hot-Jupiter with 1200 K equilibrium temperature and with likely hydrogen and helium dominated atmosphere around a sun-like star would exhibit a $\sim$300 ppm signal for $N_H = 5$ in the transmission spectrum. A hot Neptune of similar temperature would have transmission spectral features of $\sim$200 ppm. Going down to hot sub-Neptunes at similar temperature but higher mean-molecular weight atmosphere (e.g. $\mu = 13$ amu) the signal reduces to $\sim$10 ppm. A precise earth analog with earth like heavy mean molecular atmosphere ($\mu = 28$) and temperature ($T_{eq} = 255$ K) around a sun like star would show $\sim$1 ppm atmospheric signal in transmission. Toy forward models for transmission spectrum of archetype rocky to gas-giant atmospheres are illustrated in Figure 1.3.

Keeping in mind the typical atmospheric signal strengths expected for different subpopulations of exoplanets helps put in perspective the current state-of-the-art of the exoplanet atmospheric studies, specifically in transmission. Near infrared bands containing the water and methane absorption features have been sought after primarily using the G102 and G141 grisms of HST/WFC3. Some examples of recent noteworthy results in this context are the detection of water vapour in both transmission and emission for the hot-Jupiter WASP-43b (Kreidberg et al. (2014)), in transmission for the hot-Neptune HAT-P-26b (Wakeford et al. (2017)), and in transmission for K2-18b (Tsiaras et al. (2019)), an exoplanet straddling the
super-Earth-mini-Neptune boundary. Preliminary transmission spectrum of the hot Saturn WASP-96b from its early release science observations using JWST NIRISS-SOSS has also confirmed the previously observed NIR water features for the planet from HST/WFC3\(^1\). Previous indication of presence of CO\(_2\) via broadband photometry from Spitzer for the hot-Jupiter WASP-39b was unambiguously confirmed through detection of the spectrally resolved molecular feature from 4 to 4.5 \(\mu\)m using JWST NIRSpec (The JWST Transiting Exoplanet Community Early Release Science Team et al. (2022)).

Detection of atmospheric features from high mean-molecular atmospheres, in particular on smaller rocky planets, remains challenging and beyond the limit of HST/WFC3. However, hydrogen dominated atmospheres for some well known earth size planets like TRAPPIST-1 b and c (Gillon et al. (2017)) have been ruled out using HST/WFC3 (de Wit et al. (2016)). Rocky planets including and similar to those in the TRAPPIST-1 system around smaller stars will be prime targets for atmospheric characterization using JWST in its first year of operation (e.g., Rathcke et al. (2021), Kreidberg et al. (2021)).

Low-resolution transmission spectrum in the visible wavelengths (400 to 900 nm) contains absorption features prominently due to Na (around 589 nm), Li (around 670 nm), K (around 768 nm), TiO/VO (600 to 800 nm), and an overall slope due to clouds or Rayleigh scattering. In the last decade, ground-based low resolution spectroscopy efforts using multi-object spectrographs (pioneered by Bean et al. (2010)) and long slit spectrographs (e.g., Nortmann et al. (2016)) on large ground-based telescopes have played a key role in probing the transmission spectra of exoplanets in the visible wavelength range. With a typical resolution of R \(\sim 10–100\), instruments on 8-10 m class telescopes including Gemini/GMOS (e.g., Stevenson et al. (2014)), VLT/FORS2 (e.g., Sedaghati et al. (2017)), Magellan/IMACS (e.g., Espinoza et al. (2019)), and GTC/OSIRIS (e.g., Chen et al. (2017)) have been used to measure transmission spectra in 5 to 20 nm wide wavelength bins. Low-resolution transmission spectra in the visible wavelength range is useful in constraining the atmospheric continuum pressure level. In contrast, high resolution spectroscopy observations, as described in more detail in the next section, typically lose this information about the continuum. This is where a low-resolution transmission spectrum comes in handy and can be critical in breaking the well-known degeneracy between the continuum pressure level and atmospheric chemical abundances (Heng & Kitzmann (2017)).

Numerous recent results demonstrate the capabilities of ground-based low-resolution spectroscopy for characterizing transiting exoplanet atmospheres. These span a range of exoplanets including the measurement of absolute Na abundance in the hot Saturn WASP-96b (Nikolov et al. (2018)), simultaneous detection of Na, Li, and K in the hot Neptune WASP-127b (Chen et al. (2018)), and evidence against low mean molecular weight atmosphere for the small rocky planet GJ 1132b (Diamond-Lowe et al. (2018)). Such visible band transmission spectra will be crucial in complementing near-infrared JWST observations which will only be limited to 600 nm at the blue end.

\(^1\)https://www.esa.int/ESA_Multimedia/Images/2022/07/Exoplanet_WASP-96_b_NIRISS_transmission_spectrum
Ground-based low-resolution spectrophotometry observations inherently yields high precision transit light curves without large gaps. Hence, besides measuring the transmission spectrum, they can be used to improve the precision on the planetary transit parameters. This includes transit depth precisions of the $\sim$100-200 ppm comparable to that from HST/STIS (e.g., Todorov et al. (2019)), and mid-transit times at the precision of $\sim$10 seconds which can be instrumental in constraining TTVs (e.g., Huitson et al. (2017), Bouma et al. (2019)).

We next take a look at ground-based high-resolution spectroscopy measurements before going back into further details about the challenges and future prospects of ground-based low-resolution spectroscopy in Section 1.4 onwards.

### 1.3.3 High-resolution spectroscopy

Soon after the detection of the first transiting hot Jupiter HD 209458b, attempts were made using ground-based moderate to high resolution spectrographs ($\lambda/\Delta\lambda \sim$ 10000 to 100000) to measure its transmission spectrum in the visible bandpass (e.g., Bundy & Marcy (2000)). However, the first ground based detection of atmospheric features in an exoplanet’s optical transmission spectrum was reported in 2008 by Redfield et al. (2008) who observed the Na I doublet in the transmission spectrum of the hot Jupiter HD 189733b from time series high resolution spectra measured using the High Resolution Spectrograph on the 9.2 metres Hobby Eberly Telescope. Following this, Snellen et al. (2008) detected Na I doublet in the atmosphere of HD 209458b using the High Dispersion Spectrograph on the 8.2 metres Subaru Telescope.

The technique employed by these early high-resolution spectroscopy measurements was the same as the one used by Charbonneau et al. (2002) albeit at lower resolution using HST/STIS: take the ratio of the time averaged in-transit and out-of-transit spectra and look for excess absorption due to the lines of atmospheric absorbers in the wavelength region of interest. Recently, transmission spectrum measured through this technique has yielded multiple detections of Na I in hot Jupiters (e.g., Casasayas-Barris et al. (2017)) and hot-Neptunes (e.g., Seidel et al. (2022)). Spectrally resolving the Na I doublet, which is possible due to the high spectral resolution of these observations, is useful to constrain the thermospheric temperature from the core of the lines (e.g., Pino et al. (2018)) as well as global wind patterns from the broadening of the line wings (e.g., Seidel et al. (2021)). Extended upper atmospheres have also been characterized using this technique by measuring excess absorption in the Balmer H$\alpha$ line (e.g., KELT-9b, Yan & Henning (2018)) and the metastable helium triplet (Oklopčić & Hirata (2018), e.g., HAT-P-11, Allart et al. (2018)).

An alternative technique based on high-resolution spectroscopy is to use the Doppler shift of the multiple weak planetary atmospheric lines to disentangle them from the stellar spectral lines and telluric lines. Early studies used this concept to put upper limits on the reflected light from the hot Jupiter $\tau$ boo b (Charbonneau et al. (1998, 1999)). The first definitive detection, however, that paved the way for this technique was reported only about a decade ago by Snellen et al. (2010) who used CRIRES on VLT to detect CO lines during a transit of HD 209458b. Since then several detections of a multitude of atomic and molecular species for
hot Jupiters have been secured, both in transmission (e.g., Birkby et al. (2013), Hoeijmakers et al. (2018)) and emission (e.g., Pino et al. (2020)). Recently this technique has also been used to measure asymmetric atmospheric chemistry between day and nightside of an ultra-hot Jupiter (WASP-76b, Ehrenreich et al. (2020)).

Continuous efforts in both theoretical interpretation frameworks and instrumentation are being made to improve the Doppler spectroscopy technique. Recent advancement in atmospheric retrieval framework in this context has led to the measurement of absolute abundance of water and CO in the atmosphere of a hot Jupiter, which was then used to constrain the C/O ratio of the planet at extraordinary precision (WASP-77 A b using Gemini/IGRINS, Line et al. (2021)). The next big challenge for the coming decade is to push the sensitivity of the technique down to the level of signal expected from rocky exoplanet atmospheres. Future high resolution spectrographs like ELT/HIRES and ELT/METIS with larger photon collecting area could be used to detect atmospheric signatures of small rocky exoplanets (e.g., Serindag & Snellen (2019)).

1.4 Challenges to ground-based low resolution spectroscopy of transiting exoplanets

The ground-based low-resolution method of measuring exoplanetary transmission spectra is akin to that used by the earliest space based low-resolution spectrophotometry (e.g., using HST/WFC3, Deming et al. (2013)). However, ground-based low-resolution spectroscopy is also unique with its own set of advantages and challenges. This thesis is extensively aimed towards addressing these challenges and leveraging the most from such observations. Hence, in order to put the main themes of this thesis in context, we describe these challenges in more detail in the next sections before going into the existing methods to tackle these them and future prospects of such observations.

1.4.1 Earth’s atmosphere

All ground-based optical observations, even those conducted at observatories on high altitude summits like Mauna Kea, must inevitably deal with the effect of Earth’s atmosphere. The earth’s atmosphere offers the most optimal observing windows for ground-based optical spectroscopy in the visible wavelength range (400 to 900 nm) and some narrow bands in the near-infrared. For broadband transit photometry, since the earliest transit detection surveys (e.g., Henry (1999)), telluric contamination has been well recognized as one of the main sources of impedance to the required photometric precision. Time series low-resolution spectroscopy of exoplanet transits, also referred to as transit spectrophotometry, suffers from the same time dependent effect as broadband photometry, but also incurs additional wavelength dependent effects. Together these form a significant fraction of the noise budget. It is helpful to zoom into the different ways telluric effects manifest as in order to design strategies to mitigate them.
Figure 1.4: Absorption spectrum of the sky at Cerro Paranal for nominal conditions calculated using ESO SkyCalc (Noll et al. (2012), Jones et al. (2013)). The prominent telluric transmission features are labelled in the wavelength range commonly targeted by ground-based low-resolution spectroscopy observations.

1.4.1.1 Atmospheric extinction

The dominant effect on incoming stellar flux is atmospheric extinction, which is effectively dimming of the star due to absorption or scattering of incident flux by the line of sight atmospheric column. The observed broadband ‘white’ light curve through the course of night suffers from extinction which implies that the observed stellar flux varies exponentially with airmass. This first order atmospheric extinction is quantified by the wavelength dependent extinction coefficient of the bandpass of the observation. The wavelength dependence essentially stems from Rayleigh scattering which at a given airmass, causes a bluer bandpass to suffer more extinction as compared to a redder bandpass (see Allen (1963) for approximate expression for wavelength dependent atmospheric extinction coefficients). Hence, when conducting transit spectrophotometry, which effectively means doing transit photometry in many 5-20 nm wide bandpasses simultaneously, it is important to keep in mind the wavelength dependence of atmospheric extinction. Practically for transit light curves, this manifests as a low-frequency wavelength dependent trend at a timescale of few hours which is nearly the same as a typical duration of transit for a close-in planet and must be accounted for when measuring the spectroscopic transit depths.

The second order atmospheric extinction originates from the colour or the spectral energy distribution (SED) of the star which, albeit small, can be an issue when conducting differential spectrophotometry (Young et al. (1991)) as we discuss in more detail in Section 1.5.1.1.
1.4 Challenges to ground-based low resolution spectroscopy of transiting exoplanets

1.4.1 Telluric transmission and emission

In addition to the broad effect of atmospheric extinction, gases and aerosols in the earth’s atmosphere also impart absorption and emission features on the incoming stellar flux (see e.g., Smette et al. (2015)). In Figure 1.4 we show the prominent telluric absorption features in the range of 300 to 1000 nm which is the primary wavelength range of observations discussed in this thesis. These include the strong Huggins ozone bands starting blueward of 400 nm, weaker Chappius ozone bands from 500 to 700 nm, molecular oxygen bands (γ, B and A), and water bands redward of 800 nm. The water absorption bands, which are even stronger in the near infrared, arise due to precipitable water vapour (PWV) in the atmosphere which can also vary through a night. This is especially problematic and needs to be accounted for following up M-dwarf planets whose host star SEDs peak in the near-infrared (e.g., Berta et al. (2012), Baker et al. (2017)).

Along with telluric transmission, sky radiance due to terrestrial light sources, moonlight, and telluric OH\textsuperscript{−} emission lines contribute to the background. The standard approach to mitigate this is to take an annulus around the observed spectral trace to estimate the wavelength dependent background and subtract it during optimal extraction (Horne (1986)).

1.4.1.3 Atmospheric differential refraction

Atmospheric differential refraction poses an issue particularly for spectrophotometric observations (Simon (1966), Filippenko (1982)). Through the night, stellar flux traverses air columns with refractive indices that vary with respect to both wavelength and airmass. As a result, the observed stellar spectral trace suffers from a ‘stretch’ causing the same CCD cross-dispersion pixels to sample different regions of the stellar spectrum for different exposures. There are a choice of ways to mitigate this which, to name a few, include accounting for the theoretically calculated shifts e.g., using the model proposed by Filippenko (1982), using an atmospheric dispersion corrector at the spectrograph itself (e.g., LADC on FORS2, Avila et al. (1997)), empirical correction by measuring the ‘stretch’ through cross-correlation of spectral features (e.g., Huitson et al. (2017)), or in particular for spectroscopic transit light curves letting a systematics model account for it (e.g., Chapter 3 in this thesis).

1.4.1.4 Atmospheric scintillation

Atmospheric scintillation, colloquially known as twinkling, is the intensity fluctuation caused by the high-altitude turbulence in the earth’s atmosphere which causes the stellar wavefront to cross layers with varying refractive indices (see Young (1967) and the series of papers Dravins et al. (1997a, 1997b, 1998) for detailed discussions). Scintillation noise adds to the photon noise in quadrature (Osborn et al. (2015)). It can be seen from the updated analytic approximation suggested by Osborn et al. (2015) that the amplitude of scintillation noise typically decreases with increasing exposure time and telescope diameter. Several methods have been suggested and experimented with in literature to minimize scintillation noise in
Introduction

ground-based high-precision photometry e.g., conjugate-plane photometry (Osborn et al. (2011)), beam-shaping diffusers (Stefansson et al. (2017)), and differential photometry.

1.4.2 Telescope and instrumental noise

The telescope and the spectrograph can contribute a significant proportion of the noise budget. We describe the prominent sources of noise at the telescope and the spectrograph here in brief. At the detector there are contributions from readout noise and flat fielding. As the telescope tracks the field of view during the night, the telescope rotator moves to compensate for the rotation of the field of view causing the gravity vector of the spectrograph to change. This causes a flexure which can deteriorate the spatial stability of the spectral trace falling on the detector. Active flexure compensation systems (e.g., for Gemini/GMOS Hook et al. (2004)) have been designed to correct for the effect of flexure. Nevertheless, the flexure effects have been seen to persist given the high precision of the observations for transiting exoplanets (see e.g., Stevenson et al. (2014)). The instrument flexure due to the Cassegrain rotator causes two types of systematics seen in transit light curves which could also be wavelength dependent: a low-frequency trend at timescales of few hours, and an abrupt kink usually around the meridian where the rate of change of Cassegrain rotator position angle is the fastest. The ADC used for compensating atmospheric dispersion as described in Section 1.4.1.3 can also be a source of additional systematics as seen in the case of VLT/FORS2 (Moehler et al. (2010)). Finally, for some spectrographs, the time delay of the instrument shutter blades can lead to a systematic ‘odd-even’ offset in alternate exposures at timescales of a few minutes, as reported by Stevenson et al. (2014) for Gemini/GMOS.

1.4.3 Stellar variability

Transit observations are essentially indirect observations of exoplanets as they rely on stellar flux and its variation to infer information about the planet. Hence, intrinsic variability of the host star can be an impediment to the accuracy and precision of both broadband transit depths and transmission spectra. In addition to photometric variability, chromospheric activity index log R’ $R'_{HK}$ which measures the emission in Ca II H and K lines is commonly used to quantify the activity of main sequence stars (Noyes et al. (1984)).

In the case of broadband transit photometry, photospheric spots either occulted or un occulted by the planet can lead to erroneous measurement of transit parameters (e.g., Miller-Ricci et al. (2008))). For transit spectrophotometry, bright and dark stellar spots can manifest as wavelength dependent offsets in the transmission spectrum (e.g., McCullough et al. (2014), Espinoza et al. (2019)). The effect of stellar activity on a transmission spectrum is exacerbated for FGKM host stars. The contribution to the transmission spectrum due to unocculted spots or faculæ can be expressed as

$$
\delta_{\text{obs}}(\lambda) - \delta(\lambda) = \delta(\lambda) \left( \frac{1}{1 - f_{\text{spot/fac}} \left( \frac{F_{\lambda,\text{spot/fac}}}{F_{\lambda,\text{phot}}} \right)} - 1 \right)
$$

(1.3)
1.4 Challenges to ground-based low resolution spectroscopy of transiting exoplanets

Figure 1.5: Contribution to the observed transmission spectrum of WASP-19b in the visible bandpass due to unocculted stellar high contrast (blue), and low contrast (red) spots, and faculae (yellow). The solid lines and the shaded regions for each case show the corresponding contributions due to the mean and ±1σ covering fractions respectively for each case independently. The covering fractions for each case are consistent with the measured spot and facula properties measured from their respective crossings by WASP-19b in the transit light curves by Espinoza et al. (2019).

where $\delta_{\text{obs}}(\lambda)$ is the observed transmission spectrum, $\delta(\lambda)$ is the true transmission spectrum without stellar contribution, $f_{\text{spot/fac}}$ is the spots or faculae covering fraction, $F_{\lambda,\text{spot/fac}}$ is the spectrum of the spots or faculae, and $F_{\lambda,\text{phot}}$ is the spectrum of the quiescent stellar photosphere.

For comparison, the Sun has a variability amplitude (peak to trough) of 0.1 to 0.3 % due to sunspots (e.g., Kopp et al. (2005)). FGK dwarfs in general have spot induced stellar variability amplitudes ranging from 0.1 to 0.6 % as measured by McQuillan et al. (2014) for the large sample of main sequence stars observed by Kepler. M dwarfs in comparison can have larger peak to trough variability amplitudes in the range of 1 to 2% (Newton et al. (2017)).

Just in the visible bandpass from 400 to 1000 nm, the resultant contamination in the form of wavelength dependent offsets (or a slope) on the transmission spectrum can have amplitudes ranging from +0.2 to -6 % for FGK dwarfs and +20 to -3 % for M dwarfs (see the detailed calculations by Rackham et al. (2017b), Rackham et al. (2019)). An example of wavelength dependent stellar contamination due to stellar spots and faculae is shown in Figure 1.5 for the ultra-hot Jupiter WASP-19b transiting an active solar-type star. This level of stellar contamination slope in a transmission spectrum is comparable or more than the typical contribution from the planetary atmosphere alone as shown in Figure 1.3, and hence adds another layer of degeneracy for interpretation of exoplanet spectra. Another problem, in parallel to stellar activity at a single epoch, is the change in stellar photosphere over multiple epochs of observations which hinders combination of multi-epoch spectra to increase the signal-to-noise.
For host stars with early stellar spectral type, stellar pulsations like $\delta$ Scuti pulsations can be a prominent source of stellar contamination. They can be a nuisance in broadband photometry (e.g., WASP-33b, Herrero et al. (2011)), in the time series of Balmer lines (e.g., KELT-9b, Wyttenbach et al. (2020)), and, given their wavelength dependent amplitude, likely in the transmission spectra as well.

We explore methods to correct for stellar contamination, in particular multi-epoch stellar variability in Chapter 4.

1.5 Pathways to mitigate the challenges to ground-based low-resolution transit spectroscopy

In view of the challenges posed by low-resolution ground-based transit spectrophotometry, as discussed in Section 1.4, we now delve into the common strategies employed by observers to tackle them. A central goal of this thesis is to improve these strategies in order to increase the scope of such observations and overall improve the accuracy and precision of the measured exoplanet properties, including their spectra. The context of the discussion in this section are observations obtained using low-resolution spectrographs mounted on 8 m class large ground-based telescopes, but several aspects can be extended to similar observations with instruments at smaller (2 to 4 m class) telescopes.

1.5.1 Multi-object spectroscopy: a solution to obtaining low-resolution spectra of transiting exoplanets

Differential photometry has been a well established method to correct for a variety of sources of errors hindering ground-based optical observations of stars. It typically involves observing one or more comparison stars in the field of view of the instrument to correct for telluric and instrumental effects on the target star flux. The same technique applied to time series spectroscopy is referred to as differential spectrophotometry.

In the context of exoplanet transit observations, it is necessary to collect flux from the target and the comparison star simultaneously to maximize the duty cycle of the observations. One simple way to do so is using a long-slit spectrograph and aligning a relatively wide ($\sim$10 arcseconds) and long slit along the target and a comparison star. However, this approach provides less flexibility and choice in terms of the comparison stars that can be used. Another way to do this is, which is technically more versatile, is to use a multi-object spectrograph (e.g., Gemini/GMOS, VLT/FORS2) that allows using a mask with custom slits for the chosen target and comparison stars in a field of view of usually up to 10 arcminutes. The slits chosen are usually wide ($\sim$10 arcseconds) to limit the losses due to changes in seeing and telescope tracking and ensure good background sampling. Since the slits used are wide, the wavelength resolution of such observations is seeing limited. An instrument de-rotator ensures the alignment of slits on the chosen stars throughout the night. Ideally, the resultant spectral traces of target and comparison stars fall onto the detector such that each cross-dispersion direction pixel column samples the same wavelength region of spectra.
1.5 Pathways to mitigate the challenges to ground-based low-resolution transit spectroscopy

![Figure 1.6: An example of a typical MOS observational setup from the GMOS observation of the transit of hot-Jupiter XO-2 N b. The panel (a) shows the field of view with the target (top source, XO-2 N) and the comparison (bottom source, XO-2 S) stars with the positions of the masks marked by red squares. The panel (b) is one raw frame exposure showing the 2D spectral trace on the detector, with the top and bottom traces corresponding to XO-2 N and XO-2 S respectively. The PA of the instrument is set to match the PA between the target and comparison star so that the spectral traces follow simultaneous cross dispersion columns.](image)

for each star. This way the differential photometric corrections of telluric and instrumental systematics in the target star spectra can be done for the simultaneous wavelength range covered for both target and comparison stars. An example of a raw 2D frame from such observational setup is shown in Figure 1.6.

The first demonstration of multi-object spectroscopy (MOS) for measuring low-resolution transmission spectrum for an exoplanet was done by Bean et al. (2010) who used VLT/-FORS2 to obtain the transmission spectrum of the super-Earth GJ 1214b orbiting an M dwarf. Bean et al. (2010) showed that MOS observations are capable of correcting for the myriad of errors that affect ground-based spectrophotometry and can yield transit depth precisions (few 100 ppm transit depth precision in 20 nm wide bins) necessary for constraining atmospheric signatures in the exoplanet spectra. This kick-started a decade of ground-based MOS observation campaigns to obtain ground-based low-resolution transmission and emission spectra for a wide range of exoplanets, ranging from inflated ultra-hot Jupiters to small rocky worlds. We refer the reader to the introduction section of Chapter 3 for a detailed list of prominent MOS observations and results of exoplanets obtained from different instruments.

While the MOS technique took a first major step in reloading the photon collecting power of large ground-based telescopes for measuring exoplanet spectra, there are some limitations and caveats that inhibit the MOS approach from reaching its full potential. We discuss them in the next subsection, followed by a new data-driven methodology aimed at overcoming them.

1.5.1.1 Shortcomings of conventional multi-object spectroscopy

The reliance of the MOS technique on comparison stars spawns the majority of its shortcomings. Consider the use of one or more comparison stars to correct the telluric effects on the target star flux. Studies investigating the selection of optimal comparison stars for precise
Introduction

ground-based differential photometry (e.g., Young et al. (1991), Broeg et al. (2005)) have indicated that target and comparison stars are in general not affected in the same manner by atmospheric extinction. Differential photometry, in fact, is most optimal when the spectral type and brightness of target and comparison stars are most similar. This is true for MOS observations as well, where it has often been seen that having target and comparison stars both as visual binaries is an ideal case for differential spectrophotometry. We show this in Chapter 2 where we present the Gemini/GMOS transmission spectrum of the hot Jupiter XO-2N b orbiting the star XO-2N with a nearby visual binary XO-2S of identical brightness and spectral type.

One commonly used strategy is to average over many comparison stars to create an ‘optimal’ comparison star that can be used to perform differential spectrophotometry. However, differences in the SEDs or colours between target and comparison stars (or an optimal comparison star) can lead to residual second order differential atmospheric extinction effect after doing the conventional differential photometry. This was also noticed in the first MOS observations from Bean et al. (2010) who reported a residual long term trend in the target/comparison light curve arising from colour difference between the target and the comparison star.

Another issue is the differential effect of atmospheric scintillation. As per the study by Dravins et al. (1997a) (see Section 9 of this paper), high-altitude turbulence causing scintillation typically has a spatial coherence scale of the order of few arcseconds. This is much smaller than the typical separations of few arcminutes between the target and nearby comparison stars (as also noted by Osborn et al. (2011)), which, depending on the telescope size and exposure times, can lead to suboptimal correction of scintillation noise and even introduce additional errors when doing differential photometry.

The differences in the instrumental systematics described in Section 1.4.2 between the target and comparison star fluxes can cause further undermine the target/comparison correction. As elaborated in more detail in Chapter 4, these can lead to dilution of transit signal and introduction of trends originally not seen in the target transit light curve. Another limitation is that if the brightness difference between target and comparison stars is too large, for the same exposure times as the target star the result will be the comparison star either being saturated when it is too bright or having low signal-to-noise ratio when it is too faint. Figure 1.7 shows an example of the systematic trends in the target and comparison light curves individually, and the residual trends arising from differences between target and comparison star light curves in the target/comparison light curve for a transit of the warm-Neptune HAT-P-26b observed using Gemini/GMOS. Such trends are often modelled simultaneously with a transit model using either parametric (e.g., polynomial combinations of various regressors) or non-parametric models (e.g. Gaussian Process regression, see Section 1.5.2.1). Note that target/comparison correction assumes a linear relationship between the target and comparison star light curves. However, it is likely that this relationship is non-linear given the multiple sources of errors inherent in ground-based observations. Hence, the effectiveness of modelling the residual trend in target/comparison light curve ultimately hinges on how well the differential spectrophotometry performs.
1.5 Pathways to mitigate the challenges to ground-based low-resolution transit spectroscopy

In terms of astrophysical systematics, variability of comparison stars can also be the source of uncertainty. This is a well known hurdle for ground-based transit photometry searches as noted by many early efforts e.g., Henry (1999). For transit spectrophotometry, this also adds a wavelength dependent uncertainty on top of the existing wavelength dependent effect from the target star (as described in Section 1.4.3).

Finally, the fundamental observational design of MOS requiring the availability of comparison stars shuns the follow-up of exoplanets orbiting bright host stars with no suitable comparison stars nearby. This leads to exclusion of many interesting candidates orbiting bright host stars e.g., those discovered by TESS.

We summarize the potential sources of systematics and transit signal contamination posed by ground-based transit spectrophotometry observations in Figure 1.7 and 1.8.
Figure 1.8: Illustration of the systematics affecting the spectroscopic light curve for the same transit observation shown in Figure 1.7. In particular, it can be seen that the effect of atmospheric extinction with changing airmass varies significantly from blue to red wavelengths.
1.5 Pathways to mitigate the challenges to ground-based low-resolution transit spectroscopy

For the ground-based low-resolution spectroscopy of exoplanets to take the next leap with the current large ground-based telescopes and the future ELTs, it is imperative to develop solutions that can overcome the current limitations of these observations. One such solution is to improve the analysis methods used to correct for the errors inherent in ground-based spectroscopic observations. This is the main goal of this thesis, but we describe the key ideas of such methods in brief in the next subsection.

1.5.2 Data driven methods to model the systematics in ground-based transit light curves

Since the classic methods centred around differential spectrophotometry have several pitfalls, we take a step back from analysing the target/comparison light curves and focus on analysing the target light curves themselves. The goals of the new method developed in this context are twofold: one, to measure the ‘white’ light curve (integrated over the whole wavelength range) transit depth from the target star white light curve, and two, to measure the transmission spectrum from the target star spectroscopic light curves. As we discuss in Section 1.5.1.1 and in more detail in Chapter 3, the new method should be able to account for the non-linear relationship between the target star light curves (both white and spectroscopic) and any time-series used as regressors to model its systematics. Multivariate distributions in the form of Gaussian processes (GPs) provide an intuitive and powerful data-driven Bayesian framework to infer general non-linear functional relationships. GPs have been extensively used in disentangling planetary signals from instrumental and astrophysical systematics in the context of both transit light curves (e.g. Gibson et al. (2012)) and radial velocity measurements (e.g. Rajpaul et al. (2015)). In the next subsection, we describe in brief the concept of GPs and their application to MOS transit light curves. This is a useful technical background for the new method and its various flavours for analysing MOS observations which we develop in more detail in Chapter 3.

1.5.2.1 A brief introduction to Gaussian Processes and their application to fitting MOS light curves

Parametric polynomial functions of various regressors like time, airmass, width of the spectral trace PSF etc. have been traditionally used to model complex systematics in transit light curves. The drawback of this approach is that the more flexibility you want to introduce in the model, e.g., by increasing the polynomial order, the larger the function space becomes, and it becomes difficult to regularize it i.e. avoid overfitting. The framework of GPs provides both better flexibility and efficient regularization that make them an attractive option for use in fitting transit light curve systematics.

A Gaussian process is essentially a probability distribution of functions that all share some properties based on the parameters that define the distribution (see Rasmussen & Williams (2006) for a thorough discussion). Analogous to a univariate Gaussian distribution that describes the distribution of a single random variable with a mean and standard deviation, a Gaussian process describes a distribution of a random vector of say length N with a mean
vector of length N and an $N \times N$ covariance matrix. In the context of transit light curves, since we are interested in modelling the systematics using GPs, the mean vector or the mean function is the transit model and the information about the systematics is encoded in the covariance matrix. Hence, the model fit to the light curve can be described as:

$$f = \mathcal{GP}(T(t, \phi), \Sigma(X, \theta))$$  \hspace{1cm} (1.4)$$

where $T(t, \phi)$ is the transit model calculated for time $t$ and transit parameters $\phi$, and $\Sigma$ is the covariance matrix computed for the set of inputs or regressors $X$ and set of kernel hyperparameters $\theta$. The covariance matrix is the heart of the GP which governs the nature of correlated behaviour a function sampled by this GP would have. A common practice to construct the covariance matrix is by defining what’s known as a kernel function $k(x_i, x_j)$. The kernel function, for every pair of points $x_i$ and $x_j$ in the domain $X$, computes the $ij$’th member of the covariance matrix $\Sigma$. An example of a kernel function is the squared exponential kernel function:

$$k(x_i, x_j | (A, \eta)) = A \exp\left(-\frac{(x_i - x_j)^2}{\eta^2}\right)$$  \hspace{1cm} (1.5)$$

where $A$ and $\eta$ are the amplitude and length scale hyperparameters. The physical motivation behind using a squared exponential (or its variations e.g., the Matérn 3/2 kernel) is that the correlations between function values corresponding to two regressor points $x_i$ and $x_j$ should diminish with the distance between the regressor points. The expression for kernel function shown in Equation 1.5 is for just one GP regressor, but any number of GP regressors can be included by either adding kernel functions for each or by adding the corresponding distance term $(x_i - x_j)/\eta$ in quadrature. The latter type of combination results in having a single amplitude hyperparameter $A$ and a length scale hyperparameter $\eta$ for each GP regressor. While the hyperparameter $A$ quantifies the amplitude of the correlations, $\eta$ quantifies the rate at which the correlations for a pair of points of the corresponding GP regressor diminish with the distance between them. This is demonstrated in the Figure 1.9.

The transit light curve fitting hence involves simultaneously optimizing both the GP hyperparameters and the transit parameters. Gaussian process likelihood is a multivariate extension of the the univariate Gaussian likelihood, and it is defined as:

$$\log \mathcal{L}(r|X, \phi, \theta) = -\frac{1}{2} r^T \Sigma^{-1} r - \frac{1}{2} \log |\Sigma| - \frac{N}{2} \log (2\pi)$$  \hspace{1cm} (1.6)$$

where $r$ is the vector of residuals $y - T(t, \phi)$ between the light curve $y$ and the transit model $T(t, \phi)$, $\Sigma$ is the covariance matrix with $|\Sigma|^{-1}$ signifying its inverse and $|\Sigma|$ its determinant, and $N$ is the number of data points in the light curve $y$. The GP likelihood can be combined with a prior on the transit and GP hyperparameters to compute the Bayesian posterior probability. Standard optimization routines or MCMC samplers like emcee can then
Figure 1.9: Covariance kernel functions and the GP samples derived from them. In the top panel we show the actual values of the exponential squared kernel function described in Equation 1.5 for various combinations of the amplitude ($A$) and the length scale hyperparameter ($\eta$) computed for a range of values of $\tau = (x_i - x_j)$. Note that for smaller values of $\eta$ the value of kernel function falls faster as compared to larger values (comparing red and black lines respectively). The Matérn 3/2 kernel function, which we use throughout this thesis, shows a smoother fall as compared to the exponential squared function (dashed ochre line). In the bottom panel we show random samples derived from each of the kernels shown in the top panel, specifically showing the variations due to different length scales.
be used to find the optimal set of parameters that best describe the observed light curve, as we demonstrate in several analyses presented in this thesis.

This method of using GPs to fit target/comparison transit light curves obtained through ground-based MOS observations was first demonstrated by Gibson et al. (2013). In Chapter 3, we develop new extensions to this method and demonstrate that it performs remarkably well in fitting the target transit light curves themselves, without the need to perform differential photometry in the first place. One interesting aspect of this approach is that the comparison star light curves can now be used as one of the GP regressors. Doing so prevents direct introduction of systematics from the comparison star light curves which are not present originally in the target transit light curves. Moreover, extending this method to also the spectroscopic transit light curves, it lets the GP infer the non-linear functional relationship between the target transit light curves and any regressor used to detrend them, and is ultimately a more general and robust approach than simply doing differential photometry or common-mode corrections in case of spectroscopic transit light curves. We refer the reader to Chapter 3 for a more thorough development of this new method and its application to both white and spectroscopic transit light curves. The GP model we have described until now is one dimensional in the sense that it is capable of modelling covariance in the data in one dimension which i.e., time. It can be applied to fit each spectroscopic transit light curves but only independently for each wavelength bin. Recently a 2D GP model for jointly modelling covariance in both time and wavelength was introduced by Gordon et al. (2020) who combined the time dimension covariance matrix $\Sigma$ and wavelength dimension covariance matrix $\mathbf{R}$ through a Kronecker product of both. We show in Chapter 5, which is the first application of a 2D GP model to ground-based transit light curves, that the 2D GP model is adept and even better than 1D GP models at fitting the correlated systematics and inferring the white light curve transit depth accurately without the need of using comparison star light curves.

1.6 Strengths of applying Gaussian Processes regression to measuring exoplanet spectra

The immediate practical advantage of using GPs is that it provides a flexible and data-driven Bayesian tool for disentangling the transit signal from the systematics in transit light curves. It effectively removes the need to arbitrarily define parametric functions of regressors to fit the systematics in transit light curves. By sensing the similarity between the correlated behaviour of the data and the GP regressors, a GP model is capable of sampling a wide range of possible functional relationships between the data and the regressors while also allowing efficient regularization to avoid overfitting. A GP model also has a simple prescription in the form of the length scale hyperparameter $\eta$ to model transit light curve systematics at various timescales, e.g., ranging from the odd-even effect at minutes time scales to the airmass dependent atmospheric extinction lasting for several hours.

In summary, GPs have consistently proven to be a powerful tool for tackling the numerous challenges faced by ground-based MOS observations of exoplanets described in Section 1.4. We use them throughout this thesis to measure the accurate and precise transit signals in
the presence of overwhelming systematics in target transit light curves. From the outcomes of our analyses in all the chapters in this thesis that in total amount to $\sim$300 hours of MOS observations of exoplanet transits, there are several important implications which are described in detail in each chapter, but we also summarize them in the next subsections in brief.

1.6 Strengths of applying Gaussian Processes regression to measuring exoplanet spectra

1.6.1 Widening the scope of low-resolution transit spectroscopy to exoplanets orbiting bright stars

We demonstrate in Chapter 3 that it is possible to attain required precision in ground-based MOS transmission spectra by foregoing use of comparison star. This has important implications for the scope of ground-based observations of low-resolution exoplanet spectra, specifically in terms of the planets that can be feasibly observed.

The availability of nearby comparison stars with similar brightness and spectral type has been a limiting criterion for judging the feasibility of observing a target. This traditional strategy has led to very few planets orbiting bright stars with $V_{\text{mag}} \leq 9$ being followed up for low-resolution transmission spectroscopy from ground. Among these, many likely do not have a suitable comparison star needed for the conventional differential spectrophotometry. Most exoplanet low-resolution spectroscopy surveys from large ground-based telescopes have hence largely targeted planets orbiting stars with $V_{\text{mag}} \sim 10$. For observing very bright stars one solution is to use a smaller telescope aperture (e.g., 55 Cnc e, de Mooij et al. (2014)), or use narrowband spectroscopy (e.g., Sing et al. (2011), Colón et al. (2012)). However even these alternative solutions rely on the availability of suitable comparison stars.

Our method introduced in Chapter 3 enables efficient correction of transit light curve systematics for planets around bright stars with no suitable comparison stars. This opens up the possibility of atmospheric follow-up of exoplanets orbiting bright stars e.g., those discovered by TESS, using large ground-based telescopes. This is also transformative because very bright targets have been known to be difficult to follow-up with even space-based telescopes. For example in the case of HST/WC3, before the spatial scan mode became the norm for bright targets, it was known that the prominence of certain light curve systematics are correlated with the fluence levels (Wilkins et al. (2014)). The current upper limits of brightness for JWST are currently unknown but will be tested by the Bright Star Program of the Early Release Science observations of transiting exoplanets (Bean et al. (2018)). Besides the high fluence expected to cause systematics to worsen, an additional challenge for JWST observations of very bright stars will be also to maintain high duty cycle while preventing saturation. As an example, Cycle 1 JWST programs to observe the NIR to MIR secondary eclipse of the super-earth 55 Cnc e orbiting a $K_{\text{mag}} = 4.02$ star will be restricted to just one mode that would not saturate (Brandeker et al. (2021)) and another mode that will saturate partially (Hu et al. (2021)). Hence, ground-based MOS observations in visible and near-infrared armed with the new analysis methods developed in this thesis could be a powerful way to fill the gap of conducting atmospheric and general follow-up studies of exoplanets transiting bright stars.
1.6.2 The bright star opportunity for ground-based follow-up of transiting rocky exoplanets

In addition to exoplanet spectra, a natural outcome of ground-based MOS observations from large ground-based telescopes is very high precision broadband transit light curves without large gaps. Theoretically, such light curves can be used to improve the precision on even small broadband planetary transit depths and TTVs in case of multi-planetary systems. However, traditional low-resolution MOS — because of its reliance on comparison stars — is unable to achieve this goal for bright stars. Techniques to spread the flux from bright stars over more pixels by defocusing (e.g., Croll et al. (2011)) or using beam-shaping diffusers (Stefansson et al. (2017)) have been used to achieve higher observing efficiency before saturation for bright stars. However, even these techniques are ultimately limited by the availability of suitable comparison stars (e.g., Vissapragada et al. (2020)).

The time and wavelength dependent 2D GP method demonstrated in Chapter 5 is adept at overcoming this obstacle by leveraging the common information about systematics across spectroscopic transit light curves of just the target star. This method unlocks the potential of low-resolution spectroscopy observations from large ground-based telescopes like Gemini and VLT to precisely follow-up transit signals as small as that of an earth sized planet ($\sim 80$ ppm) around a sun like star with $V_{\text{mag}} \leq 9$. An important implication of this result is that rocky planets transiting bright stars being discovered by TESS and in future PLATO could be followed up from ground-based facilities for more precise planet size and mass (via TTVs) characterization. This would ultimately contribute to the growing efforts of planetary mass and radius measurements for small planets to constrain planet formation and evolution models (Fulton et al. (2017)).

1.7 This thesis

The chapters of this thesis are distributed as follows. We start with Chapter 2, in which we present the ground-based low-resolution transmission spectrum of the hot-Jupiter XO-2b obtained using Gemini/GMOS. This chapter demonstrates an ideal case for ground-based MOS observations where the comparison star is a visual binary of the target star. In Chapter 3 we introduce a new method of analysing ground-based MOS observations that can be used to obtain transmission spectrum in cases when the comparison star observations are either not available or not suitable. We demonstrate the application of this method to six Gemini/GMOS observations of the warm Neptune HAT-P-26b. In Chapter 4 we apply this method to eight Gemini/GMOS observations of the ultra-hot Jupiter WASP-19b transiting its active solar type star. In addition to mitigating potential effects from variable comparison star, we introduce a new method to correct for the effect of stellar variability (as discussed in Section 1.4.3) on multi-epoch transmission spectra of exoplanets. In Chapter 5, we employ a 2D time and wavelength dependent GP model to a transit of the ultra-hot Jupiter WASP-18b orbiting a $V_{\text{mag}} = 9.3$ star. Through this analysis we show that transit depth precisions required for the transit follow-up of rocky exoplanets orbiting bright stars without comparison stars can be achieved using ground-based low-resolution spectroscopy observations. Finally, in Chapter
In this thesis, we present ghalib, a Python library of tools implementing the methods discussed in this thesis and applied to a long term survey ground-based survey in visible and near-infrared spread of transiting gas giants spanning a wide range of bulk properties.
Enhanced metallicity of an exoplanet atmosphere from precise alkali lines measurements:

*Ground-based low-resolution detection of Na and K in the atmosphere of a cloud-free gas-giant*

Vatsal Panwar, Jean-Michel Désert, Michael R. Line, Jacob L. Bean, Timothy M. Brown et al.

To be submitted to the Astrophysical Journal

Abstract

We present the cloud-free atmosphere of the hot Jupiter XO-2Nb inferred from 7 transits observed using Gemini/GMOS. The transmission spectrum of XO-2Nb in the 420 to 920 nm range exhibits clearly identifiable strong pressure-broadened alkali lines of Sodium (Na) and Potassium (K). We retrieve the metallicity of the planet using the Na and K lines ([M/H] = 0.8 ± 0.38), and provide, for the first time for a cloud-free exoplanet, a measurement of the ratio [Na/K] = 1.26 ± 0.50. We find no evidence for the presence of clouds, hazes or ionization, and the transmission spectrum can be explained using an isothermal temperature structure. Despite the high metallicity of its host star, we find that XO-2Nb follows the trend observed in the Solar System that less massive planets are metal enriched. Our finding that the atmosphere of the exoplanet XO-2Nb is metal-enriched supports the idea that the XO-2 binary stellar systems formed in high metallicity molecular clouds, with especially the gas-giant XO-2Nb maintaining a significant fraction of refractory materials in its atmosphere over time.
2.1 Introduction

The knowledge of their masses and radii allows an estimate of the bulk composition of transiting exoplanets. Measuring abundances of the main chemical species in their atmospheres provides another level of constraint on their physical nature, more specifically from metals, since heavy elements play a key role in the formation of close-in giant planets (Adibekyan (2019)). Pressure broadened spectral absorption features of neutral alkali metals Na and K have long been predicted to dominate the visible band spectrum of atmospheres of cool sub-stellar objects (Burrows et al. (2000)) and in the transmission spectrum of cloudless hot-Jupiters Seager & Sasselov (2000), Hubbard et al. (2001), Brown (2001). For cold sub-stellar objects it is theoretically expected that condensation and rain out processes should clear most of the metals from the optically thin layers of the atmospheres while retaining the less refractory alkali ones (Burrows & Sharp 1999; Lodders 1999; Burrows et al. 2002; Sudarsky et al. 2003). Thus, analogously, the optical transmission spectra of hot-Jupiters at similar temperatures should also be dominated by strong absorption from pressure-broadened sodium Na i and potassium K i lines. The core and wings of Na and K together efficiently absorb the incoming stellar radiation leading to low geometric albedos at visible wavelengths (see e.g., HD209458b Brandeker et al. (2022)). Ionized alkali atoms at pressures less than 300 bars are also expected to be the source of electrons necessary for explaining the inflation of hot-Jupiters through ohmic dissipation mechanisms (Batygin & Stevenson 2010).

The first detection of spectral features due to an exoplanet atmosphere was obtained in the form of the Na absorption line in the transmission spectrum of the first transiting hot-Jupiter HD209458b (Charbonneau et al. 2002). However, the measured Na abundance by Charbonneau et al. (2002) was ∼3× smaller than predicted from equilibrium chemistry. This led to numerous speculations, including clouds or hazes, ionization, low global Na abundance, rain out, as well as non-local thermal equilibrium effects (Charbonneau et al. 2002; Barman et al. 2002; Fortney et al. 2003; Barman 2007). The Na broadband absorption feature, especially in the Na line wings, was confirmed later by low-resolution transit spectroscopy using HST/STIS (Sing et al. (2008a), Sing et al. (2008b)) and chromatic Rossiter-McLaughlin effect (Santos et al. (2020)). However, absorption in the core of Na lines remains contentious in light of their non-detection in high-resolution observations from VLT/ESPResso (Casasayas-Barris et al. (2020); Casasayas-Barris et al. (2021)).

Ground-based Multi-Object Spectroscopy (MOS) provides a way to obtain the transmission spectra at low-resolution across the visible and NIR bandpasses (Bean et al. 2010, 2011) where the alkali lines like Na and K have the strongest opacities. A host of spectrographs on medium (4 to 6 metre class; e.g. WHT (Kirk et al. 2020), Magellan (Espinoza et al. 2019)) to large telescopes (8-10 metre class; e.g. Gemini (Huitson et al. 2017) and GTC (e.g. Chen et al. (2017); Sing et al. (2011)) have used the MOS technique to compensate for telluric and instrumental systematics in the transit light curves in order to measure the planetary spectra. However, despite numerous campaigns by independent groups, clear detection of alkali absorption features remains rare. Majority of MOS spectra obtained for even gas-giants with large scale heights are featureless (e.g. WASP-29b Gibson et al. (2013)), suffer from stellar activity effects (e.g. WASP-19b Espinoza et al. (2019); Panwar et al. (2022a), WASP-103b
2.1 Introduction

Kirk et al. (2021)), or muting due to clouds and hazes (e.g. WASP-52b Chen et al. (2017)). The strongest detection of cloud-free atmosphere with well resolved Na wings from ground-based MOS technique has been for the hot Saturn WASP-96b using VLT/FORS2 (Nikolov et al. 2018; Nikolov et al. 2022; McGruder et al. 2022).

An additional issue in this context for detecting K from ground is the overlap of the deep telluric O$_2$ A band at 765 nm with the peak of K absorption feature at 768.2 nm. Hence, correcting telluric contamination in the region around K lines has been subject to the availability of most suitable comparison stars for differential spectrophotometry. Remarkably, to our knowledge the only conclusive detection of the wings of K in the hazy atmosphere of the hot-Jupiter WASP-96b using GTC-OSIRIS was made possible using a comparison star of low ∆r’$_{mag}$ = 1 and relatively close by (∼40.5") as compared to the target star. Recent work by Chen et al. (2022) also detected Na and K in the atmosphere of hot-Jupiter HAT-P-1b. Another marginal detection of wings of K was obtained by Sedaghati et al. (2016) for the hot-Jupiter WASP-17b, but the observations in this context were also limited by the corrections around the telluric O$_2$ A band. It is also worth noting that the strongest ground-based Na detections for gas-giants have been obtained using relatively suitable comparison stars with either less colour difference between the target and comparison stars (e.g. WASP-96b Nikolov et al. 2018) or using a visual binary companion as a comparison star (e.g. WASP-94 A b (Ahrer et al. 2022)).

2.1.1 The XO-2 System

The XO-2 system (XO-2N and XO-2S) is a system of wide solar-type binary stars with identically similar brightness and separated by only 31.2". This makes the two binaries excellent comparison stars for each other, and consequently the transiting hot-Jupiter XO-2Nb ($T_{eq}$ = 1300 K), a prime candidate for atmospheric studies from ground-based MOS observations. Yet another interesting feature of the XO-2 system is that both the stars host planets, the hot-Jupiter transiting XO-2N (Burke et al. 2007) and a multiple system of non-transiting Saturn and Jupiter mass planets orbiting XO-2S (Desidera et al. 2014). While the two stars are almost twins and Solar-like, with $T_{eff}$ = 5500 K, the planetary systems around them are different in nature. This configuration makes the XO-2 system a unique laboratory for understanding the origin and diversity of planetary systems, in particular through the study of the relative chemical composition of these stellar hosts (Ramírez et al. 2010; Teske et al. 2013, 2015; Ramírez et al. 2015). There are several reported values for the high metallicity of XO-2N, which are all consistent within 1σ and significantly higher compared to other planet-hosting stars (Burke et al. 2007; Ammler-von Eiff et al. 2009; Torres et al. 2012; Teske et al. 2015).

Connections between planet formation and observed metallicities of solar twin host stars have been hinted for long (Gonzalez 1997; Meléndez et al. 2009). In particular, solar twin binaries with significant differences in refractory material in their photospheres could be an outcome of planet formation. Core condensation before gas-accretion onto the star can deprive the stellar photosphere of the refractories, whereas engulfment of protoplanetary
cores can enrich the star with the same. XO-2 system is one of the few known binary systems e.g. 16 Cyg A-B (Ramírez et al. (2011)), WASP-94 A-B (Teske et al. (2016)), WASP-160 A-B (Jofré et al. (2021)), which host planets and the stellar hosts show differential refractory content. For XO-2N, we adopt the most conservative value of $[\text{Fe/H}] = 0.39 \pm 0.14$ from Teske et al. (2013). The metallicity of XO-2S is slightly lower ($0.28 \pm 0.14$). While the differential between the metallicity and refractory content of some binary systems, including XO-2N and XO-2S, is well established, the metallicity of planets in these systems and their connection with the host star metallicity remains unexplored.

In this work we measure the metallicity of the hot-Jupiter XO-2Nb through its transmission spectrum, and put it in the context of metallicity of its host star and the binary system. Previous transit observations detected signatures of Na and K in XO-2Nb, but at lower precision and different spectral resolution (Sing et al. 2011, 2012), which both prevented detection of pressure broadened wings of alkali lines. Recent observation of XO-2Nb by Pearson et al. (2019) resolved the wings of Na absorption but couldn’t conclusively rule out the presence of clouds and provided only a lower limit to Na abundance due to lower precision of their spectrum. In this paper, we present the most precise transmission spectrum of XO-2Nb which exhibits signs of cloud-free atmosphere and also shows strong Na and K spectral features, in particular the expected pressure-broadened wings of Na.

The paper is distributed as follows. In Section 2.2 we describe the GMOS observations of XO-2Nb and the data reduction. In Section 2.3 we discuss the analysis of transit light curves to obtain the transmission spectrum of the planet. In Section 2.4 we present the retrieval analysis and interpretation of the transmission spectrum. In Section 2.5 we discuss our results and put the measured metallicity and chemical abundances of the atmosphere of XO-2Nb in the context of expectations from planet formation theories.

### 2.2 Observations and Data Reduction

#### 2.2.1 Observations

We observed eight complete transits of XO-2Nb using the Gemini North telescope located at Mauna Kea, Hawaii. The observations were taken as part of a survey program of hot Jupiter atmospheres from Gemini/GMOS (Proposal ID: 2012B-0398; PI: J.-M Désert). The observations used the same technique and similar setup as described in (Huitson et al. 2017; Todorov et al. 2019; Panwar et al. 2022b, a). For each observation, we used the MOS mode of GMOS to observe time series spectrophotometry of XO-2N and XO-2S simultaneously. Each observation lasted approximately 5.5 hours. Our MOS mask had wide slits of 10 arcsec width and 30 arcsec long for each star. The PA between the two stars was set to 432 deg. E of N, to ensure that both stars have the same wavelength coverage.

Five transits were observed in the red optical with the R150 grism, covering a wavelength range of 525-900 nm with resolving power $R = 631$. A further three transits were observed in the blue optical with the B600 grism, covering a wavelength range of 400-650 nm with resolving power $R = 1688$. For all observations, we used the grisms in first order. For the R150
observations, the requested central wavelength was 620 nm, and we used the OG515_G0330 filter to block light below 515 nm. The blocking filter was used to avoid contamination from light from higher orders. For B600 observations, the requested central wavelength was 530 nm and no blocking filter was needed.

For all observations, we windowed regions of interest (ROI) on the detector in the cross-dispersion direction to reduce readout time. The target and reference star are only separated by $\sim 30$ arcsec, and so we used only one ROI for both slits. We binned the output $1 \times 2$, binning in the cross-dispersion direction, to reduce readout time. The detector was read out with 6 amplifiers and with gains of approximately $2 e^-/ADU$. Exposure times were chosen to keep count levels between 10,000 and 30,000 peak ADU and well within the linear regime of the CCDs. Table 2.1 shows the observation log for each transit. The numbers given under ‘Transit No’ in the table are the numbers by which we will refer to each transit in this paper. Transit 4 is a partial transit with no after-transit baseline, so we do not use this observation for our final analysis of transmission spectrum. The majority of other transit observations are complete, with no significant stoppages in the time series. Transit 5 in particular suffered from large seeing variations and high winds towards the final third of the observation, which increases the uncertainties for the parameters extracted from this light curve.

2.2.2 Data reduction

We use our custom pipeline for reducing the GMOS data described in Huitson et al. (2017). The steps performed by the pipeline are to: correct for gain and bias level, combine images across multiple amplifiers, use the time series of each pixel to flag cosmic rays, remove bad columns and background, apply a 1D optimal extraction of spectra, perform a wavelength calibration, and correct for time- and wavelength-dependent spectral shifts, the dominant source of systematic noise in the GMOS spectra.

In the present paper, the cosmic-ray removal flagged a few percent of pixels per observation. The pipeline also flagged 0.7-2.3 % of columns as bad columns depending on the observation. The background flux level is found to be 1-2 % of the stellar flux level for all observations. The wavelength calibration is based on CuAr spectra secured at high resolution using narrow slits of 1 arcsec MOS mask. The final uncertainties in the wavelength solution are approximately 1 nm for all observations, corresponding to 10 % of the bin widths used in the final transmission spectrum.

Following Huitson et al. (2017) we correct for wavelength-and -time dependent shifts. We correct for differences in the dispersion-direction offset between the target and comparison star spectra. This occurs because the PA of the instrument is not exactly the same as the PA between the target and comparison star. We use cross-correlation to measure the offset, which was between 0.68 and 1.00 pixels for all observations. We correct for this drift and interpolate the comparison star’s spectrum onto the target star’s wavelength grid. Finally, we use multiple spectral features for cross-correlation as a function of time to correct for any wavelength-dependent shifts. When constructing the light curves for each spectral bin, we use the time series cross-correlated to the closest spectral feature in wavelength to that bin,
Table 2.1: Observing dates (in DD-MM-YYYY format) and modes for the eight transits of XO-2Nb secured with Gemini-North/GMOS.

<table>
<thead>
<tr>
<th>No.</th>
<th>Obs ID</th>
<th>UT Date</th>
<th>Grism</th>
<th>Exposure Time (s)</th>
<th>No. Exposures</th>
<th>Duty Cycle (%)</th>
<th>Seeing (arcsec)</th>
<th>Airmass</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GN-2013A-Q-38</td>
<td>01/13/2013</td>
<td>R150</td>
<td>20</td>
<td>531</td>
<td>54</td>
<td>0.36-1.02</td>
<td>1.16-2.00</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>GN-2013B-Q-53</td>
<td>11/12/2013</td>
<td>R150</td>
<td>23</td>
<td>453</td>
<td>54</td>
<td>0.43-0.95</td>
<td>1.16-1.42</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>GN-2014A-Q-32</td>
<td>25/02/2014</td>
<td>R150</td>
<td>15</td>
<td>630</td>
<td>42</td>
<td>0.29-0.58</td>
<td>1.16-1.84</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>GN-2014B-Q-12</td>
<td>07/12/2014</td>
<td>R150</td>
<td>20</td>
<td>513</td>
<td>53</td>
<td>0.36-3.42</td>
<td>1.16-1.75</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>GN-2014B-Q-52</td>
<td>10/01/2015</td>
<td>B600</td>
<td>150</td>
<td>139</td>
<td>41</td>
<td>0.47-0.87</td>
<td>1.16-2.43</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>GN-2014B-Q-52</td>
<td>18/01/2015</td>
<td>B600</td>
<td>150</td>
<td>139</td>
<td>41</td>
<td>0.29-0.91</td>
<td>1.16-1.68</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>GN-2014B-Q-52</td>
<td>31/01/2015</td>
<td>B600</td>
<td>80</td>
<td>202</td>
<td>78</td>
<td>0.29-0.47</td>
<td>1.16-1.90</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>GN-2014B-Q-52</td>
<td>31/01/2015</td>
<td>B600</td>
<td>80</td>
<td>202</td>
<td>78</td>
<td>0.29-0.47</td>
<td>1.16-1.90</td>
<td></td>
</tr>
</tbody>
</table>
2.3 Analysis

We use the calibrated and differential spectrophotometry corrected time series for each wavelength bin of observations to measure the transit parameters for XO-2Nb. We fit the white transit light curves with the transit model batman (Kreidberg et al. 2014) to measure the orbital inclination, $i$, the system scale, $a/R_*$, central transit time, $T_0$, orbital period, $P$, radius contrast, $R_p/R_*$, and a linear limb-darkening coefficient, $u$. We parameterize any residual systematics in the target/comparison light curves with a linear model of planet orbital phase $\times$ baseline stellar flux for all observations.

Figure 2.1: Normalized and time-averaged 1D stellar spectra for two out of the total eight observed transits of XO-2Nb observed using the GMOS-B600 (left panel) and GMOS-R150 (right panel) grisms. For both panels in red is the spectrum for XO-2N (target star) and in blue is the spectrum for XO-2S (comparison star). Note that both the stars, being visual binaries, have nearly identical spectrum which provides an ideal differential spectrophotometric correction required to correct the systematics in transit light curves. This is especially crucial for corrections around the O$_2$ A band (765 nm) which enables measurement of the K absorption feature expected for the planetary atmosphere around 768.2 nm. Vertical dashed lines in both panels mark the boundaries of the wavelength bins we use to construct the spectroscopic light curves to measure the transmission spectrum. The gaps in the stellar spectrum are due to detector gaps and cross-dispersion columns of shifted charge that we omit in our analysis.

thus reducing the effects of the shifts. Moreover, the bin sizes we use for spectroscopic light curves (10 to 25 nm) are also significantly large as compared to the wavelength dependent shifts hence reducing their effect on the final transmission spectrum.

For each observation, we use these calibrated and corrected spectral time series to produce white and wavelength dependent transit light curves. We sum flux spectrally over each wavelength for each exposure to produce light curves for the host and for the target stars. We then divided each light curve of XO-2N by the corresponding one of XO-2S in order to correct for these time series of instrumental effects and of telluric variations during these observations. We choose to perform differential spectrophotometry instead of adopting alternative techniques of fitting the target light curves directly, as introduced by (Panwar et al. 2022b,a) because of the availability of an ideal comparison star in form of XO-2S which is close to the target XO-2N both in terms of spatial distance on the sky and brightness.
We report the transit depths as a function of wavelength to derive transmission spectra for each observation, and found that they are consistent with each other within $1\sigma$. We thus co-add these spectra by median-combining them to obtain a combined transmission spectrum for the R150 and B600 grism respectively. The transmission spectra in each grism show excellent agreement in the wavelength domain where they overlap (530-630 nm). We combined the B600 and R150 transmission spectra into one single spectrum (Figure 2.5), and compare it to atmospheric models.

### 2.4 Atmospheric modelling and interpretation of the transmission spectrum

To interpret the GMOS transmission spectrum we use the thermo-chemically consistent version Kreidberg (2015) of the CHIMERA retrieval code Line et al. (2013b,a). The model computes the thermo-equilibrium molecular and condensate abundances, assuming rain out
Figure 2.2: Detrended Target/Comparison white transit light curves for B600 (top panel) and R150 (bottom panel) observations of transits of XO-2Nb. All three transit light curves for observations using the gratings B600 (Transit 6, 7, and 8) and four transit light curves for R150 (Transit 1, 2, 3, and 5) have been overplotted together in the top and bottom panels respectively. The residuals for all observations are also overplotted with the average RMS across the B600 and R150 observations annotated.
Figure 2.3: Detrended Target/Comparison spectroscopic light curves for the three transits (Transit 6, 7, and 8) observed using B600 grating. For each bin the spectroscopic transit light curves for all three transits have been overplotted together. Also annotated are the wavelength bin boundaries for each spectroscopic light curves, and the average RMS of the residuals across the three transits for each bin.
Figure 2.4: Similar to Figure 2.3 for the four transit light curves observed using R150 grating (Transit 1, 2, 3, and 5).
Na and K absorption in the atmosphere of XO-2Nb

(Burrows & Sharp (1999)), given various scalings to the solar elemental abundances Lodders et al. (2009) and a parameterized temperature-pressure profile (Parmentier & Guillot (2014)). The molecular abundances for H$_2$O, CH$_4$, CO, CO$_2$, NH$_3$, H$_2$S, TiO, VO, Na, K, HCN, C$_2$H$_2$, FeH, and H$_2$-H$_2$/He CIA/Rayleigh Scattering are then combined with R 200 correlated-K (CK) opacities computed from the Freedman et al. (2014) line-by-line cross-section database (see also Table 2 in Lupu et al. (2014)) to compute the transmission spectrum, which is binned down to the data resolution, corresponding typically to 2-3 model CK points per data wavelength bin.

We combine the various parameterized atmospheres with the PyMultinest routine ? to solve the parameter estimation and Bayesian model comparison problems (e.g., Line & Parmentier (2016)). The fiducial atmospheric model is assumed to be cloud free and parameterized with a Guillot (2010) analytic gray planet-averaged radiative-equilibrium TP profile, metallicity ([M/H]), sodium-to-potassium ratio [Na/K], and 10-bar radius scaling. For simplicity, we fix the TP profile shape (IR and visible opacities set to 10$^{-2}$ cm$^2$/g, and 5 x 10$^{-3}$ cm$^2$/g respectively) but allow for a scaling to the incident flux to account for an unknown albedo and redistribution. Thus, the model has a total of 4 parameters.

We explore various atmospheric parameterizations physically motivated by the data to test the robustness of the retrieval results for [M/H] and [Na/K]. We tested the impact of the C/O ratio on the fitted models. We tested the impact of a power law haze with two additional parameters, a power-law slope and an amplitude (Lecavelier Des Etangs et al. (2008)). We also tested whether the presence of a thermospheric temperature gradient could be encrypted in the data. To do so, we look for either increasing or decreasing temperature with altitude, and parameterized with a gradient and base pressure free parameters. Finally, photo-ionization can be an important process that depletes alkali metals in atmospheres (Vidal-Madjar et al. (2011); Lavvas et al. (2014)). Thus, we also tested whether photo-ionization of the alkalis as parameterized via a base pressure, down to 1 mbar, and a mixing ratio gradient (e.g., Lavvas et al. (2014)). Remarkably, none of these effects listed significantly influenced the parameter estimates for [M/H] and [Na/K] nor were these more complex parameterizations justified by the Bayes factor comparison between the models (Δln(B) < 0.5 for more complex models). Hence, we choose to show the parameter estimation results for, and thus base our conclusions on, the fiducial model. Our parameter estimation results are shown in Figure 2.6 with only the posteriors for the two main parameters of interest that are constrained by our spectrum – [M/H] and [Na/K] – out of total four free parameters in the fiducial model.

2.5 Results and Discussion

2.5.1 Results

The values of $a/R_\star$, $i$ and $P$ given in Table 2.2 agree with those measured by Damasso et al. (2015) within the 1σ measurement uncertainties. In addition, our analysis of 7 transits simultaneously improves on the precision obtained in Damasso et al. (2015) by roughly a factor of 2. Our measured values are also consistent within 2σ with earlier values measured by Fernandez et al. (2009), Crouzet et al. (2012) and Kundurthy et al. (2013). The better
Figure 2.5: Gemini/GMOS transmission spectrum (wavelength dependent $R_p/R_\star$ with their $\pm 1\sigma$ uncertainties, black diamonds) of XO-2Nb combined from the three B600 (421 to 646 nm) and four R150 (531 to 922 nm) transits along with the best fit CHIMERA model and $\pm 1\sigma$ model uncertainties shown in the shaded red region. The blue points show the binned value of the best fit CHIMERA model for the wavelength bins for which we measure the transmission spectrum.

agreement with the later values is consistent with the expected improvement in precision of later studies.

The combined GMOS-North optical transmission spectrum of XO-2Nb shown in Figure 2.5 does exhibit clear spectral features expected in this bandpass from alkali lines of Na and K. The pressure-broadened features of sodium wings are seen over 1000 Å wide range as expected from theoretical models for a cloud-free atmosphere. Our measurements of Na and K confirm the previous detections of these lines in XO-2Nb by Sing et al. (2011, 2012); Pearson et al. (2019) that were obtained at different spectral resolutions and precisions. The best fit CHIMERA model presented in Figure 2.5 provides a reduced chi-squared ($\chi^2_\nu$) close to 1 suggesting that it matches the data well without overfitting. We find that the GMOS spectrum is best explained by contribution from only the alkali opacities, water, and continuum H/He Rayleigh scattering. According to the Bayes factor comparisons between models with specific species turned ‘on and off’, sodium, potassium, and water are detected at $10.5\sigma$, $3.7\sigma$, and $1.87\sigma$ respectively, but we do not detect Lithium (Li) or any other molecules.

Since the Na and K absorption lines are detected at a high significance level, we can measure the metallicity of the atmosphere at high precision. We measure the planet’s atmospheric metallicity to be $[\text{M/H}]=0.8 \pm 0.38$, and $[\text{Na/K}]=1.26 \pm 0.50$; we emphasize that these values are relative to Solar metallicity by definition. The elemental abundances in the sun’s
photosphere and in CI-chondrites provide a Solar [Na/K] = 1.230 ± 0.009 (Lodders et al. (2009)). This implies that the abundance ratio Na/K is between 100 and 850 in the atmosphere of XO-2Nb, which is significantly higher (5 to 50 times) higher than the value for the Solar System (~17).

Based on these values, we derive the thermo-equilibrium volume mixing ratios for Na = 24.6±10.4 ppmv (parts per million volume) and for K = 0.08±0.025 ppmv at 1σ level. Forcing the [M/H] and [Na/K] ratios to solar value results in a worse fit, and in that comparison the non-solar values from the best fit are preferred 2.6σ. We also tested various extreme C/O values (low and high) and find that such changes have no effect on the M/H and Na/K. We also tested TP profiles with the presence of a thermosphere and effects of photoionization, and find that both effects are unconstrained by the data, likely because they happen at lower pressure than we probe (less than 1 mbar). Furthermore, we find no evidence for the presence of Lithium absorption feature at 670.8 nm. We stress that these constraints are robust against the numerous atmospheric parameterizations discussed in Section 2.4.

2.5.2 Impact of stellar variability on the transmission spectrum

We do not observe any stellar spot crossing in the transit light curves and hence only consider the effects of unocculted spots or faculae on the transmission spectra. Photometric monitoring of XO-2N in the visible bands by Zellem et al. (2015) revealed that the host star shows a peak-to-peak flux variability at amplitude of ∼0.4 % in the R band. Based on the empirical relations between spot coverage and stellar variability amplitude by Rackham et al. (2019), this would correspond to a low spot coverage fraction (f_{\text{spot}} ≤ 0.5 % at 1σ. Calculations by Rackham et al. (2019) for 1% transit depth, which is the same as the transit depth of XO-2Nb, show that the visual offset in the transmission spectrum in the 400 to 900 nm range due to stellar photospheric heterogeneity for a K0 host star (spectral type of XO-2) can range from +30 ppm (only due to spots) to -200 ppm (due to both spots and faculae). For the Na and K lines specifically, line offsets due to stellar variability are within ±30 ppm for K0 type stars (Rackham et al. (2019)). In comparison, the amplitude of Na and K absorption feature in the transmission spectrum we observe is 400 ppm and 200 ppm respectively, which is an order of magnitude larger than the expected offset in these lines due to stellar heterogeneity. Nevertheless, we minimize any potential contamination in the transmission spectrum due to overall broadband offsets to the transmission spectrum at each epoch by median subtracting the spectrum at each epoch before combining to construct the final spectrum shown in Figure 2.5.

2.5.3 Implications of the high metallicity of XO-2Nb

We find that the metallicity of XO-2Nb’s atmosphere as traced by the abundances of alkali metals is high, and that the abundances of refractories Na and K, and their ratio, are high when compared to Solar values. It is worth noting that our metallicity constraint for the planet – [M/H] = 0.8±0.38 – is remarkably consistent within 1σ with the independent measurement of the stellar metallicity by Teske et al. (2015) who find [Fe/H]=0.4±0.14.
2.5 Results and Discussion

Figure 2.6: Summary of atmospheric retrieval of the GMOS transmission spectrum of XO-2Nb using CHIMERA. Top left panel shows the combined GMOS transmission spectrum along with the best fit CHIMERA model (dark blue line, blue points showing the values binned in the wavelength bins of the data) and its ±1σ uncertainties in the shaded red region. Also, overplotted are models that show individual contributions from CIA (H/He), Na, K, and H$_2$O. Top right panel shows the posterior distributions for the metallicity [M/H] and [Na/K] for derived from the CHIMERA retrieval on the GMOS spectrum using a fiducial model as described in Section 2.4. The constraints on both are to within a factor of 3. Bottom left panel shows the limb transmittance indicating at what pressure levels the GMOS spectrum probes. Where the transmittance rapidly changes from “1” (red) to “0” (blue) is where the observations are most sensitive (around where the limb optical depth is 1, yellow). The observations probe (generously) between 0.1 and $10^{-5}$ bars. The bottom right panel shows the best fit atmosphere temperature pressure (TP) profile (indicated by the top axis) and gas mixing ratio profiles (bottom axis). The species that influence the spectrum are bold (Na, K, and H$_2$O), but other absorbers with mixing ratios > 1 ppb (parts per billion) are shown as dashed curves. Equilibrium condensate profiles (dot-dashed) are also shown for Na$_2$S, KCl, and MgSiO$_3$. The nominal TP profile sits between possible cloud producing condensates. Depletion of alkali’s due to condensation is unlikely to occur in this atmosphere as it is too hot.
We emphasize that in our work the approach of constraining [M/H] through both Na and K, is different from what has been done so far for the majority of exoplanets where the planetary metallicity has been derived using H$_2$O feature in the HST/WFC3 bandpass (e.g. Barstow et al. (2016); Pinhas et al. (2019)). This is another fundamental difference, which makes the picture more complicated when comparing the metallicities between many planets (those estimated with refractories, and those estimated with volatiles see e.g. Welbanks et al. (2019)). Only a handful of other exoplanets: WASP-127b (Chen et al. (2018); Spake et al. (2021)), WASP-39b (Wakeford et al. (2018)), HD 209458b (Deming et al. (2013); Sing et al. (2016)), and HD 189733b (McCullough et al. (2014)) are known where all three species, Na, K, and H$_2$O, have been detected simultaneously and have been used to measure the planetary metallicity. Metallicity measurements via only volatile species like H$_2$O inherently suffer from a degeneracy with the C/O ratio, and additional Na and K measurements can be useful to break this degeneracy (Welbanks et al. (2019)). XO-2Nb could join this list through the measurement of volatile molecular bands of water, CO, CO$_2$, and methane using HST/WFC3 or JWST.

Our findings of super-Solar [M/H] and [Na/K] provide a direct evidence that the hot-Jupiter XO-2Nb must have formed in a region of the disk rich in refractory metals, and reinforce the scenario that the binary stellar systems XO-2 would have formed in a high metallicity molecular cloud (Teske et al. (2013)). Indeed, the high estimates of [Fe/H] for the XO-2 system are exceptional when compared to the metallicity distribution of nearby stars since only 2% of the nearby G dwarfs have [Fe/H] > 0.2, as determined from early photometric measurements (Rocha-Pinto & Maciel (1996)). Interestingly, this is also very close to the frequency of occurrence of hot-Jupiters (Dawson & Johnson (2018)). The paucity of planets around metal poor stars is consistent with the findings of Santos et al. (2004) and Fischer & Valenti (2005) who show that the fraction of stars with gas-giant planets increases with stellar metallicity, and conclude that planetary systems have higher likelihood to be born around stars from higher metallicity molecular clouds.

This is important because a higher metallicity may lead to the formation of a more massive disc, which might be more likely to cause the formation of giant planets and subsequently their migration (Lin et al. (1996)). For XO-2Nb, the high planetary metallicity provides an important constraint on the planet’s formation and suggest large amounts of heavy elements must be retained within the planetary H/He envelope. In the context of the metal enriched planet XO-2Nb, we demonstrate that the stellar metallicity is consistent with the observed inverse relationship between the mass of the planet and the metal enrichment ($Z_{pl}/Z_{star}$), and as computed by Miller & Fortney (2011) for the non-inflated planet. We also show the position of XO-2Nb in the mass-metallicity diagram in Figure 2.7.

### 2.5.4 Connection with the differential chemical abundances of XO-2N and XO-2S

We now discuss our results in the context of differential chemical abundances measured for the stars XO-2N and XO-2S. Stars born in multiple systems like XO-2 are expected to have
similar composition in the absence of accretion effects. Ramírez et al. (2015) conducted differential measurements of the chemical composition of the photospheres of XO-2N and S, and find that the differences between the two stars correlate strongly with an element’s dust condensation temperature (shown in Figure 5 Ramírez et al. (2015)). In XO-2N, refractories are significantly overabundant while volatiles are only marginally enhanced. In particular, Ramírez et al. (2015) find that the Na and K are overabundant in XO-2N compared to XO-2S by $0.035 \pm 0.009$ and $0.008 \pm 0.027$ dex, respectively. If some fractionation occurred during the accretion process in the XO-2S system, or if previously fractionated material was accreted onto the star, then one would expect a correlation of the measured photospheric abundances of the elements with their condensation temperatures. One possible explanation could be that rocky material equivalent of $\sim 20$ M$_{\oplus}$ could have accreted onto XO-2N as a result of the migration of the hot Jupiter XO-2Nb (Ramírez et al. 2015). Alternatively, XO-2S is metal-poor relative to XO-2N, which could be because more disk material was used to make the small planets observed in XO-2S. However, our observation of high metallicity of the hot-Jupiter XO-2Nb favours the previous scenario where the XO-2N must have accreted the disc material after the migration of XO-2Nb. This might have increased the envelope metallicity of the star significantly. The accreted material is mixed throughout the convective region of the envelope, which typically comprises only about 3% of the mass of the star near the main sequence.

Guillot et al. (2006) and Miller & Fortney (2011) find that the mass of heavy elements in planets is a function of the stellar metallicity and predict 100-150 Earth masses of heavy elements for planets orbiting a star with three times the solar metallicity. Guillot et al. (2006) use evolution models and constraints on the stellar ages to derive the mass of heavy elements present in exoplanets. Our observational result is consistent with their calculations given that XO-2N’s metallicity is between 1.8 and 3.5 times Solar, and that metal enrichment of XO-2Nb is between 2 to 10 times Solar. However, Thorngren et al. (2016) find a weaker correlation between stellar metallicity and planetary metal-enrichment, and derive a relationship between the planetary metal-enrichment relative to the parent star and the planetary mass. They conclude that significant amounts of heavy elements should be in the H/He envelopes of giant planets, rather than in their cores, such as we see in XO-2Nb. Our findings on XO-2Nb follows this latter trend predicted by Thorngren et al. (2016), which is consistent with the core-accretion planet formation scenario.

Nevertheless, it is also expected that a 150 M$_{\oplus}$ planet such as XO-2Nb should have a smaller radius than the one measured, which is at odds with our findings of the planet’s radius. Indeed, XO-2Nb has an inflated radius despite its high atmospheric metallicity, which implies that the mechanism that leads to the inflation must be acting in opposition to the effect of adding heavy elements to the planet (Miller & Fortney 2011; Thorngren et al. 2016).

### 2.5.5 The large Na/K and the depletion of K

Our retrieved Na abundance is consistent with the metallicity [M/H], and consistent with the Solar value scaled to that metallicity. However, the retrieved K abundance is $\sim 20$ times
smaller than expected for this metallicity when scaled to Solar value. This implies that K is depleted relative to Na for this metallicity, which explains the enhanced Na/K. Below, we explore various scenarios that would explain such the large Na/K and the depletion of K observed in XO-2Nb; these include primordial hypothesis, and depletion through ionization and condensation. The final composition of a planet is mostly determined by the chemical and dynamical evolution of the disk during the formation and growth of planetesimals. It is expected that the relative abundances of Na and K trace back to planet formation location and processes within the protoplanetary disk (e.g., Mordasini et al. (2016)). Assuming that the Na/K measured in XO-2Nb is primordial, then this implies that the region of the disk where the planet formed should have been depleted in K compared to Na. This means that either the primordial composition of the disk is depleted in K, or that K needs to be lost from the gas phase but also its condensates or molecules need to be redistributed at a different radial location of the disk from where the planet is formed. Otherwise, gravitational heating (and the possible migration of the planet closer to the star) would have released the condensed alkalis back into the atmosphere (Lavvas et al. 2014). However, Na and K have similar chemical behaviour, so their condensates should form at similar planet formation region, and should not have very different responses to the differentiation mechanisms. Consequently, the Na/K ratio should not vary too much across the protoplanetary disk. Moreover, any differentiation process will have to occur in a short timescale, before the disk is accreted to the star, in order to allow for the migration of the planet at its current short orbital distance.

Hence, it is difficult to explain the observations of large Na/K based on chemical and physical processes in the protoplanetary disk at the time of planet formation. Beside the entire disc being depleted in K, a remaining possibility would be that XO-2Nb formed at a place in the disk depleted in K, likely by photo-ionization, in the optically thin part of the disk, but where Na is not affected. Furthermore, the large Na/K may be linked to the high stellar metallicity. Indeed, observations of stellar photospheres with various metallicity show that the Na/K ratio ranges between 0.1 to 100 for stars of low and high metallicity Shimansky et al. (2003); Takeda & Takada-Hidai (2012). We highlight that the Na/K we derive for XO-2Nb is significantly above these values.

We investigate whether later atmospheric processes could explain the large Na/K and K depletion. Molecular chemistry cannot explain the low K abundance, for example through KCl, whose abundance is expected to increase at pressures higher than 10 mbar Lavvas et al. (2014), but again that’s unlikely since Na and K have similar chemical behaviour. There is a possibility of condensation of K into potassium chloride clouds (Gao & Benneke (2018)). However, based on the condensation curve of KCl shown in Figure 2.6, the XO-2Nb would have to be much cooler (T_{eq} = 600-900 K) for condensation to be a sink at the pressures of 1-100 mbar probed by the GMOS data.

Our models show that thermal ionization and photo-ionization are not mechanism that are required to fit the data (see Sect. 2.4). The test of the presence of photo-ionization provides a very low Bayesian evidence of 0.26σ. This is because these processes are not effective at the level of pressure probed by the current data (1-100 mbar) on this planet.
2.5 Results and Discussion

Figure 2.7: Mass-metallicity plot of the solar system planets (black square) including known exoplanet metallicity estimates for previous planets (in grey and black circular points), and XO-2Nb (red point) with respect to solar metallicity. The dashed line shows at first a power law fit to the solar system planets, with a plateau at solar metallicity.

Interestingly, photo-ionization would favour a depletion of K over Na, due to the lower ionization level of K, but this process happens at lower pressure than probed by our data Lavvas et al. (2014). Thermal ionization is included in our modelling, but it is effective at pressure higher than 100 mbar (Lavvas et al. (2014)), and thus we should not expect a depletion of K through this mechanism in our data. This was also investigated by Pearson et al. (2019) who found that for thermal ionization to be more dominant mechanism at the pressures of 1-100 mbar probed by the GMOS spectrum, the planet’s equilibrium temperature ($T_{eq}$) needs to be around 1700 K, whereas the planet’s $T_{eq}$ is $\sim$1300 K.

Atmospheric escape could favour a depletion of Na compared to K given their relative masses, but that would require additional fine-tuning of the escape mechanism and rate.

In summary, none of the atmospheric processes listed above can explain the K depletion.

Our models show that clouds or haze are not required to fit the data (see Section 2.4), and that the atmospheric layer probed by our data (1-100 mbar) is cloud free, but condensation could happen below. At the level of pressures and temperatures probed by our data, we expect the TiO/VO to condense (below 1700-1800 K) so the lack of clouds at this altitude point towards rain out of TiO/VO condensates. However, there are likely condensates deeper in the atmosphere (e.g., enstatite, iron, corundum, perovskite) that form clouds. Our observation also suggests that we do not detect TiO$_2$ seed particles at high altitudes that are expected to be present (Helling et al. (2008)), indicating a limit on the vigor of vertical mixing that would loft condensate particles to the upper atmosphere.
2.6 Conclusions

In this work we have presented the ground-based low-resolution visible band (420 to 920 nm) transmission spectrum of the hot-Jupiter XO-2Nb using transit spectrophotometry observations obtained through Gemini/GMOS. We summarize below our key conclusions from the observational analysis, observed transmission spectrum and its theoretical interpretation:

1) Owing to the existence of its host star's visual binary companion, XO-2Nb is an ideal case for conducting low-resolution ground-based transit spectrophotometry across the bandpass. Identical brightness and spectral type of the comparison star XO-2S implies that we obtain excellent correction of telluric and instrumental systematics across the bandpass — including in the deep telluric O$_2$ absorption feature around the planetary K line — by simply performing differential spectrophotometry and fitting for the stellar flux baseline using a linear model of planetary phase.

2) We obtain a precision of $\sim$60 ppm per 10-20 nm bin in the combined transmission spectrum of XO-2Nb from 7 GMOS transits. We observe strong absorption features in the transmission spectrum due to Na I and K I at amplitudes of $\sim$400 and $\sim$200 ppm respectively.

3) Through atmospheric retrieval performed using CHIMERA, we find that the GMOS transmission spectrum of XO-2Nb is most consistent with a cloud-free atmosphere with no significant evidence for clouds or hazes. We detect pressure-broadened absorption features in the transmission spectrum of XO-2Nb due to Na and K at 10.5 and 3.7 $\sigma$ respectively. H$_2$O absorption at the red end of the transmission spectrum is detected tentatively at 1.87 $\sigma$.

4) We obtain a super-solar planetary atmosphere metallicity of [M/H] = 0.8±0.38 dex, and planetary [Na/K] = 1.26±0.5. Our measurements are robust against more complex parametrizations for atmospheric models, and based on past stellar photometry of the host star we rule out the effect of stellar activity on the amplitude of Na and K features observed in the transmission spectrum.

5) Super solar [M/H] and [Na/K] for XO-2Nb, consistent with the previously measured super metallicity of the host star, indicate that the planet could have formed in a region of disk rich in refractory material before migrating to its current location.

6) The super-solar metallicity of XO-2Nb indicating high refractory content in its birth environment could be consistent with late accretion of disc material by the host star XO-2N, which could be an explanation behind the previously measured high refractory content of XO-2N as compared to its binary twin XO-2S.

7) The large Na/K ratio of XO-2Nb remains difficult to explain even after accounting for known potential sources that could cause depletion of K in its atmosphere.

Overall, our results provide an observational evidence that is consistent with extrasolar giant planets, as a class, being enhanced in heavy elements. In light of our results, future atmospheric characterization of transiting gas giants in multiple stellar systems such as XO-2 could enable a comparative characterization between stellar and planetary chemical abun-
dances. Further insight can be gained about XO-2Nb from future transit spectroscopy observations of XO-2Nb with JWST could also reveal the abundances of volatile species like H$_2$O, CO, and CO$_2$, and further constraint the refractory to volatile ratio of the atmosphere of the planet. Such studies of exoplanets in similar systems as XO-2 could both be ideal for ground-based spectrophotometry, and when conducted for a statistically large sample, could be instrumental in unveiling the impact of the properties of the host star, its birth environment, and its youth on the outcomes of planet formation.

2.7 Acknowledgements

Based on observations obtained at the Gemini Observatory (acquired through the Gemini Observatory Archive and Gemini Science Archive), which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil). Based in part on Gemini observations obtained from the National Optical Astronomy Observatory (NOAO) Prop. ID: 2012B-0398; PI: J.-M Désert. We thank Christiane Helling, Saugata Barat, and Bob Jacobs for useful discussions, Eliza Kempton for providing feedback on the manuscript. J.M.D acknowledges support from the Amsterdam Academic Alliance (AAA) Program, and the European Research Council (ERC) European Union’s Horizon 2020 research and innovation programme (grant agreement no. 679633; Exo-Atmos). J.L.B. acknowledges support from the David and Lucile Packard Foundation. This work is part of the research programme VIDI New Frontiers in Exoplanetary Climatology with project number 614.001.601, which is (partly) financed by the Dutch Research Council (NWO). This material is based upon work supported by the NWO TOP Grant Module 2 (Project Number 614.001.601). This material is based upon work supported by the National Science Foundation (NSF) under Grant No. AST-1413663. This research has made use of NASA’s Astrophysics Data System. The authors also acknowledge the significant cultural role and reverence the summit of Mauna Kea has within the indigenous Hawaiian community. This research has made use of Astropy,$^1$ a community-developed core Python package for Astronomy Astropy Collaboration et al. (2013, 2018), NumPy Harris et al. (2020), matplotlib Hunter (2007), SciPy Virtanen et al. (2020) and IRAF Tody (1986) distributed by the NOAO, which is operated by AURA under a cooperative agreement with the NSF.

---

$^1$http://www.astropy.org
A new method to measure the spectra of transiting exoplanet atmospheres using multi-object spectroscopy

Vatsal Panwar, Jean-Michel Désert, Kamen O. Todorov, Jacob L. Bean, Kevin B. Stevenson, C. M. Huitson, Jonathan J. Fortney, Marcel Bergmann


Abstract

Traditionally, ground-based spectrophotometric observations probing transiting exoplanet atmospheres have employed a linear map between comparison and target star light curves (e.g. via differential spectrophotometry) to correct for systematics contaminating the transit signal. As an alternative to this conventional method, we introduce a new Gaussian Processes (GP) regression-based method to analyse ground-based spectrophotometric data. Our new method allows for a generalised non-linear mapping between the target transit light curves and the time series used to detrend them. This represents an improvement compared to previous studies because the target and comparison star fluxes are affected by different telluric and instrumental systematics, which are complex and non-linear. We apply our method to six Gemini/GMOS transits of the warm (T_{eq} = 990 K) Neptune HAT-P-26b. We obtain on average ~20 % better transit depth precision and residual scatter on the white light curve compared to the conventional method when using the comparison star light curve as a GP regressor and ~20 % worse when explicitly not using the comparison star. Ultimately, with only a cost of 30% precision on the transmission spectra, our method overcomes the necessity of using comparison stars in the instrument field of view, which has been one of the limiting factors for ground-based observations of the atmospheres of exoplanets transiting bright stars. We obtain a flat transmission spectrum for HAT-P-26b in the range of 490-900
nm that can be explained by the presence of a grey opacity cloud deck, and indications of transit timing variations, both of which are consistent with previous measurements.

3.1 Introduction

Low-resolution transmission spectroscopic observations of transiting gas giant exoplanets have been extensively used to probe their atmospheric compositions. The multi-object spectroscopy (MOS) technique (Bean et al. 2010, 2011) has produced spectrophotometric measurements of exoplanet atmospheres at low-resolution ($R \sim 10-100$) with various ground-based observatories from optical to near infrared (Gemini/GMOS: see e.g. Crossfield et al. 2013), Gibson et al. (2013), Stevenson et al. (2014), Huitson et al. (2017), Todorov et al. (2019), Wilson et al. (2021); VLT/FORS2: see e.g. Bean et al. (2010), Sedaghati et al. (2017), Nikolov et al. (2016), 2018, Carter et al. (2019), Wilson et al. (2020); Magellan/MMIRS and IMACS: see e.g. Bean et al. (2011, 2013), Rackham et al. (2017a), Espinoza et al. (2019), Weaver et al. (2020), McGruder et al. (2020); LBT/MODS: see e.g. Mallon & Strassmeier (2016), Yan et al. (2020)) and long-slit spectrographs at low-resolution (GTC/OSIRIS: see e.g. Sing et al. (2012), Murgas et al. (2014, 2019), Nortmann et al. (2016), Chen et al. (2017, 2018, 2020, 2021)) which have resulted in the detection of spectral features due to Rayleigh scattering, atomic and molecular absorption, and/or grey opacity clouds (in the form of flat or featureless spectra). The detection of pressure broadened profile of Na I doublet weakly in the atmosphere of WASP-4b (e.g. Huitson et al. 2017), significantly in the atmosphere of the hot Saturn WASP-96b (Nikolov et al. 2018) consistent with a cloud-free atmosphere, and the detection of Na, Li, and K absorption along with the signatures of scattering due to haze in the atmosphere of hot Neptune WASP-127b (Chen et al. 2018) are some examples that demonstrate that ground-based MOS observations are capable of estimating absolute abundances in some of the gas giants with clear atmospheres.

Notably, transit observations using large ground-based telescopes like Gemini and VLT have yielded transit depth precision comparable to space-based observations from HST in their white light curves (e.g. Bean et al. (2010), Todorov et al. (2019)). Observations for the same planet repeated over multiple epochs and from different instruments have helped in ascertaining the robustness of results (e.g. WASP-4b May et al. (2018), Bixel et al. (2019); WASP-19b Sedaghati et al. (2017), Espinoza et al. (2019)) and in interpreting and mitigating the transit light source effect due to stellar photospheric heterogeneity which has the strongest observable effect on a transmission spectrum in the optical wavelength range (Rackham et al. 2018)). Furthermore, ground-based MOS observations have pushed the limits of atmospheric characterisation down to terrestrial planets (Diamond-Lowe et al. 2018, 2020a, 2020b) for which the optical transmission spectra have been able to rule out the presence of clear and low mean molecular weight atmospheres.

Spectrophotometric observations obtained using ground-based multi-object spectrographs are affected by telluric and instrumental systematics at levels comparable or even more than the amplitude of variations due to the planetary atmosphere in the transmission spectrum that we aim to measure. The conventional technique to compensate for systematics in
ground-based low resolution spectra has been to simultaneously observe one or more reference or comparison stars in the instrument’s field of view (Bean et al. 2010) and use that to correct for the systematics similarly affecting the target star light curve through differential photometry. The Rossiter-McLaughlin effect based observations measure changes in line shape for detecting transits (van Sluijs et al. 2019) and deriving low-resolution transmission spectra (Oshagh et al. 2020, Di Gloria et al. 2015). Such observations follow a parallel approach to measure transmission spectrum that does not need a comparison star but typically yield low resolution transmission spectra at sub-optimal precision.

All the aforementioned ground-based MOS studies, however, have always used comparison stars to deal with systematics in the transit light curves and extract the signals of planetary atmosphere. The leftover systematics after Target/Comparison star light curve normalisation, arising due to brightness or differences in spectral types between the target and comparison stars, are then conventionally modelled by parametric models constructed using polynomials based on a set of decorrelation parameters (e.g. Gibson (2014), Stevenson et al. (2014), Todorov et al. (2019)), or a non-parametric approach using Gaussian Processes (GP) regression (e.g. Gibson et al. (2012, 2013)).

Some previous works, instead of dividing the target star light curve by the comparison star light curve, fit the target star light curves directly by using comparison star light curves or a PCA components of multiple comparison stars as linear regressors (see e.g. Jordán et al. 2013, Espinoza et al. 2019). The Divide-White method introduced by Stevenson et al. (2014) extracts the transmission spectrum from target star light curves by using non-analytic models of wavelength dependent systematics derived from comparison star light curves. Note that all the aforementioned approaches: doing differential spectrophotometry, using comparison star light curves as linear regressors or non-analytic models of systematics, assume a linear relationship between the systematics in target and comparison star flux variations. Differential photometric corrections in particular perform best when the comparison stars are similar to the target stars in brightness and spectral type (Broeg et al. 2005, Croll et al. 2015).

In most cases, it is likely that light from the target and comparison stars may not have travelled through the same column of atmosphere, especially in scenarios where the separation between the target and comparison stars in the sky is comparable or more than the typical spatial scale of variations in atmospheric turbulence. Systematics at the instrumental level and stellar variability in the comparison star can further cause complex non-linear variations between the target and comparison star fluxes. This implies that the linear functional forms of mapping between the two assumed by conventional methods are sub-optimal and may even be a source of additional systematics.

The conventional strategy of MOS observations has relied on the availability of suitable close-by comparison stars which presents some issues. In situations when comparison stars are fainter than the target star or of a different spectral type, the Target/Comparison normalisation is photon-limited by the brightness of the comparison stars (in the whole bandpass or a range of wavelengths where the spectral shape and relative brightness of the comparison and target star differs the most, see e.g. Diamond-Lowe et al. (2020a, 2020b)). On the
other hand, if the comparison stars are brighter than the target stars (as it happens to be the case for comparison star for the GMOS observations of HAT-P-26b presented in this paper), the duty cycle of the observations gets limited. Moreover, if the target star is in a sparse field, which is often the case for bright host stars, then there is less choice of an optimal comparison star given the limited instrument’s field of view which has been a limiting factor in ground-based high precision spectrophotometric follow-up of exoplanets orbiting bright stars.

In view of these several limitations, there is a need for a more generalised and robust approach to marginalise systematics in ground-based spectrophotometric light curves which accounts for non-linear relationship between target and comparison star fluxes, and does not explicitly rely on the availability of comparison stars.

We present a novel alternative method in this paper which takes a more generalised approach when using a set of auxiliary time series (e.g., comparison star light curves, target star PSF width, airmass, etc.) to model systematics in the target star transit light curves. Our new method in essence lets a Gaussian Process model explore the underlying unknown and likely non-linear functional form between the regressors used to model the systematics in the target star transit light curves. This can be achieved for both the integrated white light curve and spectroscopic light curves. Through our method, we also demonstrate that remarkably precise wavelength-dependent transit depth measurements of exoplanet spectra can be reached when not using the comparison star light curves at all. We describe the method and its application in detail to our observations of the warm Neptune HAT-P-26b observed by Gemini/GMOS in Section 3.4.

The paper is distributed as follows: in Section 3.2 we describe in detail our observational setup for the 6 transits of HAT-P-26b observed by GMOS to which we apply the new method we introduce in this paper. In Section 3.3, we describe the data reduction steps to extract stellar spectra from raw data, and in Section 3.4 we discuss the analysis to model the GMOS transit light curves. Specifically, in Section 3.4.1 we introduce our new method to model the telluric and instrumental systematics directly in the target star light curves. In Section 3.5.1 we compare our new analysis method with the conventional approach, and discuss its caveats and implications for future ground-based observations of exoplanet atmospheres. In Section 3.5.2 we interpret the optical to infrared transmission spectrum for HAT-P-26b from combined GMOS, HST, and Spitzer measurements using atmospheric models. We discuss the indications of transit timing variations for the planet in Section 3.5.3, and in Section 3.6 we present our conclusions.

3.2 Observations

3.2.1 The warm Neptune HAT-P-26b

HAT-P-26b is a low density warm ($T_{eq} = 990$ K) Neptune discovered by Hartman et al. (2011) orbiting its chromospherically quiet K1 host star in a close orbit of period $\sim 4.23$ days. Given its large scale height the planet has been the subject of multiple atmospheric
characterisation studies, including those constraining its atmospheric metallicity. Constraining the atmospheric metallicity of exo-Neptunes is crucial for tracing the dominant planet formation scenarios governing the formation of these planets, and distinguishing between scenarios of core accretion (Pollack et al. 1996) and in situ formation (Bodenheimer et al. 2000). Both of these scenarios can lead to significantly different metal enrichment of the atmosphere of a Neptune mass planet like HAT-P-26b. Initial studies of HAT-P-26b using Magellan/LDSS-3C and Spitzer by Stevenson et al. (2016) indicated tentative evidence of water vapour features in the red optical. Wakeford et al. (2017) reported a strong detection of the 1.4 μm water vapour feature muted by a grey opacity cloud as evident from the near-infrared and visible observations from HST/WFC3 and STIS respectively. From these observations Wakeford et al. (2017) retrieve a near solar metallicity atmosphere with a high altitude cloud deck suppressing the transit spectral features. MacDonald & Madhusudhan (2019) further combined the observations from Stevenson et al. (2016) and Wakeford et al. (2017) to perform a comprehensive retrieval analysis reporting the presence of several species of metal hydrides with absorption features ranging from optical to near infrared and a tentative hint of Rayleigh scattering.

In this paper, we present 6 Gemini/GMOS transit observations to measure the transmission spectrum of HAT-P-26b in the visible from 490 to 900 nm, extending the wavelength coverage of the transmission spectrum published by Wakeford et al. (2017) further towards blue optical. The primary motivations of our study are to investigate the exoplanet spectrum in the optical, expanding the wavelength coverage blueward, and independently test the presence of clouds and Rayleigh scattering. Additionally, from the precise mid-transit times obtained from our high SNR GMOS transit light curves we also investigate the transit timing variations (TTVs) for the planet previously indicated by Stevenson et al. (2016) and von Essen et al. (2019).

3.2.2 GMOS Transmission spectroscopy

We observed a total of six transits of HAT-P-26b using the Gemini North telescope located at Mauna Kea, Hawaii, and the Gemini South telescope located at Cerro Pachon, Chile. Three transits were observed using Gemini North and three transits were observed using Gemini South. The observations used the same technique and setup as described in Huitson et al. (2017) (hereafter referred to as H17), which is similar to that of previous observations using GMOS (e.g. Bean et al. (2010, 2011, 2013), Gibson et al. (2013), Stevenson et al. (2014)). All transits were observed as part of two survey programs of hot Jupiter atmospheres from GMOS North and South (P.I. J.-M. Désert) described in H17 (see Table 3.1 for program numbers).

For each observation, we used the MOS mode of GMOS to observe the time series spectrophotometry of HAT-P-26b and a comparison star TYC 320-426-1, simultaneously. HAT-P-26 and the comparison star are separated by ~ 3.8 arcmin. HAT-P-26 has a V magnitude of 11.76 and TYC 320-426-1 has a V magnitude of 11.08, and similar spectral type from visual inspection of prominent stellar spectral features. Each observation lasted approximately 5 to
5.5 hours. To avoid slit losses, our MOS mask had wide slits of 10 arcsec width for each star. The slits were 30 arcsec long to ensure adequate background sampling for each star.

In order to provide similar wavelength coverage between HAT-P-26 and the comparison star, the PA of the MOS mask needs to be as close as possible to the PA between the two stars. The PA for HAT-P-26 and the comparison star is 23 deg. E of N. However, at this PA, no suitable guide stars fell into the patrol field of Gemini’s guider, the On Instrument Wave Front Sensor (OIWFS). We therefore used the Peripheral Wavefront Sensor (PWFS) instead for three of our observations from Gemini South (see Table 3.1 for details). The PWFS has a larger patrol field, but a lower guiding precision and so is used as a backup option if there are no suitable guide stars available for OIWFS. This setup enabled us to orient the instrument so that the instrument PA matched the PA between the two stars.

However, from the initial analysis, we found that the photometric precision was lower when using the PWFS than for our previous survey observations obtained using the OIWFS due to higher dispersion direction drift in case of PWFS as compared to OIWFS. The dispersion direction drift over a night is ∼15 pixels for PWFS as compared to ∼1 pixel for OIWFS. We therefore modified the setup for three of our observations at Gemini North (see Table 3.1) to be able to continue using OIWFS. In this new setup, we selected the PA of the MOS mask to be 7 deg. E of N. While this meant that the wavelength coverage was different for both stars, it meant that we could orient the GMOS field of view such that a suitable guide star fell within the range of the OIWFS. We therefore achieved improved guiding in exchange for the loss of approximately 1/3rd of the wavelength coverage.

Three transits were observed in the red optical with the R150 grating, covering a wavelength range of 530-900 nm with ideal resolving power $R = 631$. Two transits were observed in the blue optical with the B600 grating, covering a wavelength range of 490-680 nm with ideal resolving power $R = 1688$. The ideal resolving powers assume a slit width of 0.5 arcsec. In our case, due to using a wide slit, our resolution was seeing limited. Given the range of seeing measured in Table 3.1, our resolution is up to 4× lower than the ideal value depending on observation. For each observation, we used the gratings in first order. For the R150 observation, the requested central wavelength was 620 nm and we used the OG515_G0330 filter to block light below 515 nm. The blocking filter was used to avoid contamination from light from higher orders. For B600 observation, the requested central wavelength was 520 nm and no blocking filter was needed.

For all observations, we windowed regions of interest (ROI) on the detector to reduce readout time. We used one ROI for each slit, with each ROI covering the whole detector in the dispersion direction and approximately 40 arcsec in the cross-dispersion direction. We binned the output $1 \times 2$, binning in the cross-dispersion direction, to further reduce readout time. For the observations at Gemini South, the detector was read out with 3 amplifiers. For the observations at Gemini North, the detector was read out with 6 amplifiers. All amplifiers had gains of approximately 2 $e^-$/ADU. Exposure times were chosen to keep count levels between 10,000 and 30,000 peak ADU and well within the linear regime of the CCDs. Table 3.1 shows the observation log for each transit, as well as which observations were obtained.
at Gemini South and which at Gemini North. The numbers given under ‘No’ in the table are the numbers by which we will refer to each transit observation in this paper.

### 3.3 Data Reduction

#### 3.3.1 GMOS data

We used our custom pipeline designed for reducing the GMOS data, the steps for which are described in more detail in H17. We extract the 1D spectra and apply corrections for additional time- and wavelength-dependent shifts in the spectral trace of target and comparison stars on the detector due to atmospheric dispersion and airmass. In this section, we describe the main points of the pipeline and the additional corrections we apply to the data before extracting and analysing the transit light curves.

For the R150 grating, we only use 2/3 of the detector in the dispersion direction. For all observations we use a moving boxcar median of 20 frames in time for each pixel to compare its value in the frames immediately before and after. We flag pixels deviating more than 5 times the boxcar median value as cosmic rays and replace it with the median boxcar value. The cosmic-ray removal flagged a few percent of pixels per observation. Our pipeline flagged 1.8-3.8 % of columns as bad depending on observation. The majority (80 %) of flagged columns are consistent between the transits for each detector. For observations 1 and 3, these include columns of shifted charge occurring mostly in the transition regions between amplifiers, as discussed in H17. These columns are not present on the GMOS-North detector (observations 2, 4, and 5) and are also not present in observation 6, which was taken after a detector upgrade at GMOS-South.

We tested our extraction with and without flat-fielding and find that flat-fielding does not significantly affect the scatter of the resulting transit light curves. For this reason, and since flat-fielding did not improve the scatter blueward of 700 nm, we chose not to perform flat-fielding for all transit observations. We notice no slit tilt in the spectra of HAT-P-26 and the comparison star unlike as seen in H17 and Todorov et al. (2019). The sky lines in the frames for all transits are parallel to the pixel columns. Thus, we choose to not perform any tilt correction.

We subtracted the background while performing optimal extraction (Horne 1986), and found that taking the median background value in each cross-dispersion column provided the best fit to the background fluxes compared to using fits to the flux profile in each cross-dispersion column. The background fluxes were 1-10 % of the stellar flux for the R150 observations and 2-20 % of the stellar flux for the B600 observations, depending on the wavelength range and exposure number.

After spectral extraction, we performed wavelength calibration using CuAr lamp spectra taken on the same day as each science observation. To obtain the CuAr spectra at high resolution, we used a separate MOS mask to that used for science, which had the same slit position and slit length as the science mask of only 1 arcsec width. We used the same grating and filter setup as the corresponding science observation.
Table 3.1: Observing Conditions for GMOS Runs. The numbers in the first column are the numbers by which we will refer to each transit observation throughout the rest of the paper. Observation IDs starting with “GS” were observed at Gemini South using the ideal PA and the PWFS, while those starting with “GN” were observed at Gemini North using the non-ideal PA and OIWFS (see Section 3.2 for more details).

<table>
<thead>
<tr>
<th>No.</th>
<th>Program ID</th>
<th>UT Date</th>
<th>Grating</th>
<th>Guider and PA</th>
<th>Exposure Time (s)</th>
<th>No. of Exposures</th>
<th>Duty Cycle (%)</th>
<th>Seeing (arcsec)</th>
<th>Airmass Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GS-2013A-Q-27</td>
<td>2013 Mar 20</td>
<td>R150</td>
<td>PWFS, ideal PA</td>
<td>50</td>
<td>226</td>
<td>63</td>
<td>0.6 - 1.8</td>
<td>1.21 - 2.06</td>
</tr>
<tr>
<td>2</td>
<td>GN-2013A-Q-38</td>
<td>2013 Apr 10</td>
<td>R150</td>
<td>OIWFS, non-ideal PA</td>
<td>15</td>
<td>574</td>
<td>45</td>
<td>0.3 - 0.9</td>
<td>1.04 - 1.97</td>
</tr>
<tr>
<td>3</td>
<td>GS-2014A-Q-59</td>
<td>2014 May 09</td>
<td>R150</td>
<td>PWFS, ideal PA</td>
<td>20</td>
<td>318</td>
<td>40</td>
<td>0.3 - 1.0</td>
<td>1.21 - 2.00</td>
</tr>
<tr>
<td>4</td>
<td>GN-2016A-LP-6</td>
<td>2016 Mar 12</td>
<td>B600</td>
<td>OIWFS, non-ideal PA</td>
<td>25-40</td>
<td>299</td>
<td>47</td>
<td>0.6 - 1.0</td>
<td>1.21 - 1.96</td>
</tr>
<tr>
<td>5</td>
<td>GN-2016A-LP-6</td>
<td>2016 Apr 15</td>
<td>B600</td>
<td>OIWFS, non-ideal PA</td>
<td>90-150</td>
<td>161</td>
<td>85</td>
<td>0.4 - 1.6</td>
<td>1.04 - 1.74</td>
</tr>
</tbody>
</table>
We used the IDENTIFY task in the Gemini IRAF package to identify spectral features in the CuAr spectra. A wavelength solution was then constructed by a linear fit to the pairs of wavelength vs. pixel number in each ROI and then refined by comparison with known stellar and telluric feature locations. The final uncertainties in the wavelength solution are approximately 1 nm for all observations, which is ∼5% of the bin widths used to construct the final transmission spectrum.

Before generating the transit light curves, we performed further reduction of the extracted 1D spectra. This is because, as in H17, we found that there is a dispersion-direction shift of the spectra on the detector during each observation that is a function of time and of wavelength, such that the spectra ‘stretch’ over time. The result is that wavelength bins identified in fixed pixel space will not sample the same wavelength in each exposure. The effect therefore needs to be accounted for in order to build transit light curves that sample a constant wavelength region over time. Failure to account for this effect can introduce spurious slopes in the transmission spectra, as discussed in H17. In H17, we found that a model for differential atmospheric refraction explained the shifts well for our previous observations. This is consistent with the fact that GMOS has no atmospheric dispersion compensator, and so we expect an effect from differential atmospheric refraction. However, the differential atmospheric refraction model does not adequately fit the shifts observed in the HAT-P-26 data studied here. We therefore use the alternative method developed in H17, in which we use multiple spectral features for cross-correlation as a function of time to account for the wavelength-dependent shifting empirically.

However, instead of simply using the shifted spectra corresponding to the measured shift value from cross-correlation with respect to the nearest feature for constructing light curves for each spectral bin (as done in H17), we proceed further to use the information from cross-correlation with the spectral features to apply corrective shifts to each pixel in the 1D spectrum. From the measured shifts of the spectral features for an exposure, we estimate the shifts in the spectra for pixels in between and away from the features by a linear interpolation between the shift values for three features used for the cross-correlation. The interpolated shift values thus obtained for each pixel for an exposure are then applied to the whole 1D spectrum. We then repeat this step for every exposure so that in the end we have the same wavelength solution for the spectra across all exposures. We repeat this step for the comparison star as well, using the same set of spectral features as those used for the target star spectrum for cross-correlation. Finally, we interpolate the comparison star’s spectrum from all exposures onto the wavelength solution of the target star, omitting detector gaps and bad columns, which ensures that both the target and comparison star spectra for every exposure have the same wavelength solution within the uncertainty of our estimates on the shifts derived from cross-correlation.

As final step in our reduction process, we also correct for the dispersion direction offset between the target and comparison star spectra. This occurs because the PA of the instrument is not exactly the same as the PA between the target and comparison star for observations taken using OIWFS guider. We used cross-correlation to measure the offset, which was between -18.2 and -16.0 pixels for the southern observations. For the northern observations,
the offset was between -830 and -600 pixels due to the non-deal PA. We then interpolated the comparison star's spectrum onto the target star's wavelength grid, while omitting bad columns for both spectra (which are the same columns on the detector but are at different wavelengths for each star).

We show the final wavelength-calibrated 1D spectra for an arbitrarily chosen exposure for the target and comparison star in Figure 3.1 for all the observations. Note that some residual shifts (at the order of a few pixels) still remain between the target and comparison star spectra, especially towards the redder end for R150 observations (beyond 710 nm where fringing also becomes strong) as seen in Figure 3.1. This is because the shifts between target and comparison star spectra vary both in time and wavelength, and constant offsets followed by interpolation to a common wavelength grid does not entirely correct for it. We refrain from further empirical corrections at this stage and to minimize the effects of any residual shifts we choose to use broad 20 nm wide bins for our spectroscopic light curves. This wavelength width is significantly larger than spatial scale of the shifts between the target and comparison star spectra. We nevertheless do not include the spectra beyond 730 nm for Observation 1 and 3 for computing the transmission spectrum due to excessive fringing in that region. We also emphasize that the residual shifts between the comparison star and HAT-P-26 spectra are not an issue for the new method we introduce in the paper of using only the target star to extract the transmission spectrum (see Section 3.4.1).

### 3.4 Transit Light Curve Analysis

We now describe our light curve analysis methods that we apply to the 6 transit observations of HAT-P-26b. We first briefly discuss the noise models that we use to correct for the systematics in the light curves in Section 3.4.1. In this section, we also introduce and motivate a new method to directly model the systematics in the target star light curves.

We have summarised the the conventional method used to date and the new method introduced in this paper and their various types of applications to white and spectroscopic light curves in Table 3.2.

The novel aspect of the new method in the context of both the white light curves and the spectroscopic light curves is that instead of assuming a linear functional form, we explore a distribution of functions (described by a GP) to explore the likely non-linear functional form of the mapping between the target transit light curves and one or more decorrelation time series (e.g. comparison star light curves). The new method in the context of all its applications is an alternative to the applications of the conventional linear method to fit for systematics in MOS transit light curves as described in Table 3.2. We describe the shortcomings of the conventional method and motivate the need for the new method in Sections 3.4.1.1 and 3.4.1.2 respectively.

The conventional method to fit the white light curves specifically has two different types of applications: 1) Conv1:WLC : two step method of first performing differential spectrophotometry (normalising the target star light curve by the comparison star light curve) and then fitting the resultant light curve with a GP, and 2) Conv2:WLC : one step method of using
3.4 Transit Light Curve Analysis

Figure 3.1: Optimally extracted spectra for HAT-P-26 and the comparison star from an arbitrarily chosen exposure, corrected for dispersion direction shifts and normalised by their exposure time for the 6 GMOS observations of HAT-P-26. Each panel shows one exposure for each observation, and the observation numbers correspond to the programs described in Table 3.1. For all observations, especially the GMOS-N observations taken using non-ideal PA, the comparison star spectrum has been shifted to the same wavelength grid as the target star spectrum using prominent common stellar features in the spectra which have been marked by the black dashed vertical line. The green vertical lines show the wavelength range considered for obtaining the transmission spectrum for each observation. The gaps in the spectra correspond to physical gaps in the CCD and bad columns.
a linear model with one or more comparison star light curves or their PCA components as one of the regressors to fit the target transit light curves. Conv2:WLC is especially suited for when there are more than one comparison stars available, which is not the case in this paper. In Section 3.4.2 we apply the conventional method Conv1:WLC, and the new methods New:WLC and New:WLC;No_Comp to fit white-light curves for each observation.

In the context of fitting spectroscopic light curves, a frequently used method to correct for wavelength independent systematics in particular, in addition to using the comparison star light curves, is to perform a ‘common-mode correction’. This is the approach of the Conv1:λLC method which subtracts a white light curve derived common-mode trend from each wavelength binned light curve. However, this approach also assumes a linear relationship between the common-mode trend and the spectroscopic light curves. Our new method New:λLC explores the likely non linear relationship in this context (e.g. arising from wavelength dependent effects with changing airmass) by using the common-mode trend as a GP regressor. We fit the spectroscopic light curves using Conv1:λLC and New:λLC in Section 3.4.3.2 and 3.4.3.3 respectively.

3.4.1 Modelling systematics in transit light curves

In the following sections, we model the instrumental and telluric time-dependent systematics in the HAT-P-26 transit light curves by following both the conventional method and the new method we introduce in this paper. We describe them both in the next two subsections to compare them and motivate the need for the new method.

3.4.1.1 Conventional method: using comparison star or common-mode trend as linear regressor

The conventional method involves first dividing the target star light curve by comparison star light curve and then fitting the transit signal and systematics in the Target/Comparison light curve simultaneously using a transit model and a GP respectively. In the case of spectroscopic light curves, there is an additional step of removing the common-mode trend before fitting with a GP. The GP model takes as regressors, or inputs, a set of decorrelation time series which include e.g., time (time stamps of individual exposures), width (FWHM) and spatial shifts of the traces of target and comparison stars on the detector (e.g. Nikolov et al. (2018), Diamond-Lowe et al. (2020b)). However, the step of doing differential spectrophotometry itself in this approach raises concerns on the relevance of decorrelation parameters derived from the individual target and comparison star spectral traces in the context of modelling the differential Target/Comparison light curve. In general, the step of doing differential spectrophotometry, assumes that the target and comparison star fluxes are affected by the same or linearly related time and wavelength-dependent systematics. Subtracting common-mode trend also assumes a linear relationship between the white light curve and the spectroscopic light curves. Given the complex nature of both instrumental and telluric systematics this is likely not the case. Considering the transit depth precisions (∼ 100-500 ppm per ∼ 20 nm bins) we are aiming for, dividing the target star light curve by the comparison star light
Table 3.2: Summary of the conventional and new methods used to model the systematics in white light curves (WLC) and spectroscopic light curves (λLC) in this paper in Section 3.4.1. The ‘Application’ column specifies the different ways of applying the methods with a more detailed description in the column ‘Description’. ‘Abbreviation’ specifies how we refer to each of these applications in this paper. 

<table>
<thead>
<tr>
<th>Method</th>
<th>Application</th>
<th>Description</th>
<th>Abbreviation</th>
<th>Example References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential spectrophotometry using</td>
<td>Target/Comparison WLC: fit with GP</td>
<td>Conv1:WLC</td>
<td></td>
<td>Gibson et al. (2013)</td>
</tr>
<tr>
<td>comparison star LCs</td>
<td>Target/Comparison λLC:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>common-mode subtracted, fit with GP</td>
<td>Conv1:λLC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison star(s) LC as linear regressor</td>
<td>Target WLC: fit with linear model</td>
<td>Conv2:WLC</td>
<td></td>
<td>Espinoza et al. (2019)</td>
</tr>
<tr>
<td></td>
<td>including comparison star(s) white</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LC or their PCA as regressors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target λLC: fit with linear model</td>
<td>Conv2:λLC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>including comparison star(s) λLC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>or their PCA as regressors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New method</td>
<td>Comparison star LC as GP regressor</td>
<td>Target WLC: fit with comparison star as a GP regressor</td>
<td>New:WLC</td>
<td>This work</td>
</tr>
<tr>
<td>No comparison stars</td>
<td>Target WLC: fit with GP regressors</td>
<td>New:WLC,No_Comp</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>excluding comparison star LC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>common-mode trend as GP regressor</td>
<td>Target λLC: fit with common-mode</td>
<td>New:λLC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>trend as a GP regressor</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
curve or subtracting common-mode trend can potentially propagate unwanted systematics and deteriorate the light curve SNRs which can be difficult to correct for when fitting the Target/Comparison light curve. An example of this are the B600 observations of HAT-P-26b we present in this paper where the target and comparison star light curves have systematics significantly different from each other. In this context, simply normalising the target star by the comparison star contaminates the transit signal originally present in the Target star light curve (see white light curves for observation 4 and 5 in Figure 3.2).

In cases when the instrument’s field of view is large, in the order of ~ 10 arcminutes, recent works (Jordán et al. 2013, Espinoza et al. 2019) have used Principal Component Analysis (PCA) to optimally use information from multiple comparison stars in the field of view. This approach (Conv2:WLC and Conv2:λLC) relies on the availability of multiple comparison stars, and involves using the PCA components of more than one comparison star in log-space as regressors in a linear regression model to fit the systematics in the target star light curve. Specifically Espinoza et al. (2019), and other studies analysing IMACS/Magellan observations e.g., Weaver et al. (2020), McGruder et al. (2020), Kirk et al. (2021) use the model averaging scheme for linear regression models outlined by Gibson (2014) to incorporate the number of relevant PCA components as an additional uncertainty in their model. Since we only observed one comparison star, we do not test the PCA based approach in this work.

### 3.4.1.2 New method: using comparison star or common-mode trend as a GP regressor

The intrinsic limitations of differential spectrophotometry using one or more comparison stars to correct for systematics in the light curves (also described in more detail in Section 3.1) narrows down the set of exoplanets around bright host stars that can be followed up for atmospheric characterisation from ground-based multi-object spectrographs. Hence, there is a need for a new method that doesn’t explicitly rely on the comparison stars and can model the transit light curve systematics and extract the transmission spectra solely from the target star.

With the new method we introduce in this paper, we present a way to directly fit the target star light curves using a set of time series recorded at the same time as target light curve as GP regressors. This includes the comparison star light curve (when fitting the white light curves, described in more detail in Section 3.4.2), and common-mode trend when fitting the spectroscopic light curves (see Section 3.4.3). This is essentially the novel aspect of our method: for both white and spectroscopic light curves, we use the set of time series, which have traditionally been used linearly to correct them either as simple normalising factors or as linear regressors, directly as regressors in a GP model. This allows for letting the GP itself explore an exhaustive set of non-linear mappings between the target transit light curve and regressors like comparison stars or the common-mode trend. This approach is more capable of incorporating the complex differences in which the target and comparison stars are affected by systematics during any observations. Our method also provides more accurate uncertainties by propagating them through the Bayesian framework of GPs.
The underlying GP framework we use for our new method is the same as that introduced by Gibson et al. (2012) to model the transit signal and systematics simultaneously in the wavelength integrated white light curves and the wavelength binned light curves for each transit. In our new method, instead of fitting the Target/Comparison light curves with a GP model, we model the target star light curves directly as a numerical transit model combined with an additive GP model to account for systematics affecting the light curve. This means that we skip the step of dividing the target star light curve by one or multiple comparison star light curves, and instead use the comparison star light curve as one of the GP regressor. In the case of fitting spectroscopic target light curves, we use the common-mode trend derived from the white light curve as a GP regressor (see Section 3.4.3). We describe the GP formalism used for both the conventional and new methods in more detail in the next section.

### Gaussian Process regression model

A Gaussian Process model to account for the systematics in a transit light curve means that we model the observed transit light curve (which for the conventional method is Target/Comparison and for the new method just the target star light curve) as a multivariate Gaussian Process distribution with the mean function as the numerical transit model, and a covariance matrix $\Sigma$:

$$f = \mathcal{GP}(T(t, \phi), \Sigma(X, \theta))$$  \hspace{1cm} (3.1)

where $f$ is the flux time series representing a transit light curve, $t$ is time, $\phi$ is the set of planet transit parameters, $T(t, \phi)$ is the astrophysical transit light curve model, and $\Sigma(X, \theta)$ is the covariance matrix described by a kernel function for a set of regressors or input parameter vectors $X$ and hyperparameters $\theta$:

$$\Sigma_{ij} = k(x_i, x_j | \theta)$$  \hspace{1cm} (3.2)

Note that here we assume that the systematics we are attempting to model using the GP are additive, and we could just modify Equation 3.1 to instead have the GPs model multiplicative systematics, which subsequently gives identical results within the precision of our data as we tested and was also reported by Gibson et al. (2013). The kernel function takes a set of input parameter vectors $X$ ($x_1, x_2, x_3, ... x_P$) (each vector $x_p$ of the same length as the number of points ($N$) in the light curve) which could be time, Cassegrain Rotator Position Angle (CRPA), the airmass, FWHM of the PSF of the spectra trace of the target star, and the measured position of the spectral trace corresponding to each exposure (averaged across the dispersion direction). This is analogous to using these time series quantities as decorrelation parameters to construct parametric models. In particular, in our new method we additionally also test the use of the comparison star light curve as one of the regressors to the GP.
We choose to use the Matérn 3/2 kernel function as it is known to provide a good prescription for time correlated noise at the time scales typically observed in GMOS transit light curves (Gibson et al. 2012):

\[ k(x_i, x_j|\theta) = A \left( 1 + \sqrt{3} R_{ij} \right) \exp \left( -\sqrt{3} R_{ij} \right) + \delta_{ij} \sigma_w^2 \] (3.3)

where \( A \) is the hyperparameter specifying the amplitude of covariance, \( \sigma_w \) is the white noise term (which we fit for) and \( \delta \) is the Kronecker delta. We emphasize that keeping the white noise term \( \sigma_w \) free when fitting the light curves is an important aspect of our proposed method in this paper. The best fit value of \( \sigma_w \) represents the combined noise variances in the target star light curve and in the individual decorrelation parameters used as GP regressors, assuming no heteroscedasticity in our observed light curves (see Equation 6 in Mchutchon & Rasmussen (2011)). When we use comparison star light curve as one of the GP regressors, the best fit value of \( \sigma_w \) represents the combined noise variance from the comparison star light curve and the target star light curves. We highlight that this is a way to propagate the relevant uncertainties from the comparison stars within the Bayesian framework of GPs (which we use to fit for \( \sigma_w \) as described below) in contrast to just adding them in quadrature as done in the case of differential photometry. This is analogous to fitting for a jitter term in the methods that use comparison stars as an input to linear regression models (e.g. Espinoza et al. 2019).

The term \( R_{ij} \) in the Equation 3.3 is a quadrature summation of pairwise difference between regressor points (\( \eta \) being the inverse length scale hyperparameter corresponding to each input vector). The \( R_{ij} \) term for \( P \) number of input vectors can be described as:

\[ R_{ij} = \sum_{p=1}^{P} \left( \frac{x_{p,i} - x_{p,j}}{\eta_p} \right)^2 \] (3.4)

This is one of the few ways in which information from multiple input parameters or regressors can be combined to describe the covariance matrix of the GP, and involves a single amplitude hyperparameter \( (A) \) and length scale hyperparameters (\( \eta \)) for each of the input parameters respectively. We also considered and tested another type of combination where we take the kernel in Equation 3.3 for each regressor, and construct the final kernel as the sum of kernels for each regressor (similar to the approach followed by Aigrain et al. (2016), k2sc). This combination leads to using more number of hyperparameters as each GP regressor now also has a respective amplitude hyperparameter in addition to a length scale hyperparameter. For all the observations we analyse in this paper, we find that the first type of kernel combination (described in Equation 3.4) performs consistently better in terms of root mean square (RMS) of the residuals and consistency of best fit transit parameters with the literature values as compared to the other combinations.
The joint GP posterior probability distribution we marginalise over to estimate the transit parameters and hyperparameters corresponding to the best fit to the observed light curves is:

\[ p(f|t, \phi, \theta) = \pi(\phi, \theta) \times L[\mathcal{GP}(T(t, \phi), \Sigma(X, \theta))] \]  (3.5)

where \( \pi(\phi, \theta) \) encodes the prior probability on the transit model parameters \( \theta \) and hyperparameters \( \phi \), and \( L[\mathcal{GP}(T(t, \phi), \Sigma(X, \theta))] \) is the GP likelihood, written in form of log-likelihood as:

\[ \log L(r|X, \phi, \theta) = -\frac{1}{2} r^T \Sigma^{-1} r - \frac{1}{2} \log |\Sigma| - \frac{N}{2} \log(2\pi) \]  (3.6)

where \( r \) is the vector of residuals of the observed light curve from the mean function \( f - T(t, \phi) \) and \( N \) is the number of data points in the light curve.

We used the transit modelling package \texttt{batman} (Kreidberg 2015 which is an implementation of the formalism of Mandel & Agol (2002)) to calculate the numerical transit model \( T(t, \phi) \), and the package \texttt{george} (Ambikasaran et al. 2015) for constructing and computing the GP kernels and likelihoods.

### 3.4.2 Analysis of White Transit Light Curves

#### 3.4.2.1 Constructing white light curves

For each observation, we constructed the target and comparison star white light curves by summing the measured flux over 530 to 700 nm for observations 1 and 3, 530 to 900 nm for observations 2 and 6 (as these R150 observations do not show fringing redward of 700 nm), and 490 to 680 nm for observations 4 and 5. We then normalise the total flux in each exposure by the corresponding exposure time for both the target and comparison star. The white transit light curves thus obtained are shown in Figure 3.2. The white light curves for each observation contain information on the dominant time-dependent systematics affecting all the wavelength channels, and analysing them prior to fitting the wavelength dependent light curves is an important step to constraining transit parameters and understanding the sources of systematics that can affect the final transmission spectrum.

#### 3.4.2.2 Fitting the white transit light curves

We obtain the best fits for each transit observation independently using both the conventional method \texttt{Conv1:WLC} and the two applications of the new method \texttt{New:WLC} and \texttt{New:WLC;No_Comp} as described in 3.4.1. For both methods and for each transit white light curve, we fix the orbital period \( P \) and eccentricity \( e \) to literature values, and fit for the orbital inclination \( i \), orbital separation \( a/R_\star \), mid transit time \( T_0 \), and planet to star radius ratio \( R_P/R_\star \). For \( i \) and \( a/R_\star \), we put a Gaussian prior with the mean and
Table 3.3: Summary of priors and fixed values for the parameters (transit model and GP hyperparameters) used to fit the transit light curves of HAT-P-26b. We fixed the planet orbital period (P) and eccentricity (e) for all fits. \( U \) represents a uniform prior applied within the specified range, and \( \mathcal{N} \) represents a Gaussian prior with the specified mean and standard deviation. \( T_c \) is the predicted mid transit time for each epoch using the ephemeris from Hartman et al. 2011. For the linear limb darkening coefficient mean of the Gaussian prior is taken as the theoretically calculated value from PyLDTk (Parviainen & Aigrain 2015) for the B600 and R150 wavelength ranges.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior/Fixed value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>P [d]</td>
<td>4.2345</td>
<td>Hartman et al. (2011)</td>
</tr>
<tr>
<td>e</td>
<td>0.124</td>
<td>Hartman et al. (2011)</td>
</tr>
<tr>
<td>( i ) (^{[\circ]} )</td>
<td>( \mathcal{N} ) (88.09, 1.5)</td>
<td>Wakeford et al. (2017)</td>
</tr>
<tr>
<td>( R_P/R_\star )</td>
<td>( U ) (0, 1)</td>
<td>–</td>
</tr>
<tr>
<td>( a/R_\star )</td>
<td>( \mathcal{N} ) (11.89, 1.2)</td>
<td>Wakeford et al. (2017)</td>
</tr>
<tr>
<td>( T_0 ) [d]</td>
<td>( U ) (( T_c )-0.001, ( T_c )+0.001)</td>
<td>Hartman et al. (2011)</td>
</tr>
<tr>
<td>( u_1 ) [B600]</td>
<td>( U ) (0.603, 0.03)</td>
<td>PyLDTk</td>
</tr>
<tr>
<td>( u_1 ) [R150]</td>
<td>( U ) (0.73, 0.03)</td>
<td>PyLDTk</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior/Fixed value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln (A) )</td>
<td>( U ) (-100, 100)</td>
<td>–</td>
</tr>
<tr>
<td>( \ln (\eta_p) )</td>
<td>( U ) (-100, 100)</td>
<td>–</td>
</tr>
<tr>
<td>( \sigma_w )</td>
<td>( U ) (0.00001, 0.005)</td>
<td>–</td>
</tr>
</tbody>
</table>

We summarise all the priors we use in this paper in Table 3.3. We also fit for the white noise hyperparameter \( \sigma_w \) as described in Section 3.4.1 which lets the GP model fit for the white noise variance in the target star light curves along with contributions from the variance in the GP regressors (e.g. the comparison star light curve). This is one of the key advantages and an important feature of our method as instead of propagating the variance from comparison star light curve by simply adding in quadrature (as is the case when the target star light curve is normalised by the comparison star light curve), our method provides a way to propagate uncertainties from the comparison star light curve to our fit of the target star light curve within the Bayesian framework of GPs described in.

standard deviation as the mean and 3 times the \( 1 \sigma \) uncertainty values measured by Stevenson et al. (2016) respectively. We use truncated wide uniform priors for \( R_P/R_\star \) and for the mid transit time (\( T_0 \)) around the values predicted by a linear ephemeris. We adopt a linear stellar limb darkening law and calculate the limb darkening coefficients and uncertainties on them (stemming from uncertainties in stellar parameters) for the wavelength range integrated to obtain the white light curve, and the wavelength bins we adopt for spectroscopic light curves (see Section 3.4.3) using PyLDTk (Parviainen & Aigrain 2015), which uses the spectral library in (Husser et al. 2013), based on the PHOENIX stellar models. We put a Gaussian prior on the linear limb darkening coefficient with the mean value and the standard deviation as the mean and 3 times the \( 1 \sigma \) uncertainty calculated from PyLDTk respectively.
Section 3.4.1. We further emphasize that fitting for $\sigma_w$ is crucial for allowing the GP model to capture the white noise in the target star light curves.

For the conventional method application Conv1:WLC, we perform fits for both R150 and B600 observations using all combinations of following GP regressors common to both target and comparison star light curves: time, CRPA, and airmass. For the new method application New:WLC we use all combinations of the following GP regressors: comparison star light curve, time, CRPA, airmass, PSF full width at half maxima (FWHM) of the spectral trace for every exposure (averaged across the dispersion direction) for the target star. For New:WLC;No_Comp we use all GP regressors as for New:WLC except comparison star light curve to demonstrate the performance of fits without using the comparison star at all. We determine the GP regressor combination that best describes the systematics for all the methods in Sections 3.4.2.3 and 3.B.

For all applications of the conventional and new methods, we first find the Maximum a-Posteriori (MAP) solution by optimising the GP posterior using the Powell optimiser in the SciPy python package. We put wide uniform priors on the GP hyperparameters and sample them logarithmically. The logarithmic sampling of hyperparameters effectively puts a shrinkage prior on the hyperparameters which pushes them to smaller values if the corresponding GP regressor truly does not represent the correlated systematics in the data (Gibson et al. 2012, Gibson 2014). Using the MAP solution as the starting point we marginalise the GP posterior over all the hyperparameters and transit model parameters through an MCMC using the package emcee, a pure-Python implementation of the affine-invariant Markov chain Monte Carlo (MCMC) ensemble sampler (Goodman & Weare 2010, Foreman-Mackey et al. 2013). We use 50 walkers for 10000 steps and check for the convergence of chains by estimating the integrated auto-correlation times for each walker following the method described in (Goodman & Weare 2010). We ensure that the total length of our chains are at least 50 times the auto-correlation times to make sure our samples are effectively independent and have converged.

We discard the first 1000 samples as burn-in. We judge the final goodness of fit based on the consistency of best fit transit parameters with the literature values, and the model selection criteria described in Section 3.4.2.3 for each combination of GP regressors and the various forms of kernel combinations (described in Section 3.4.1 ). We also tested the robustness of our fits using a nested sampler dynesty (Speagle 2020) and obtain posteriors consistent with those from emcee. We measure the best fit transit parameters as the median of the corresponding posteriors and $\pm 34$ percentile from the median as their $1\sigma$ uncertainties.

3.4.2.3 Selecting the best GP regressor combinations for white light curve fits

We select the best combination of GP regressors for both new and conventional methods of fitting the white light curves separately by comparing the Bayesian Information Criterion (BIC, Schwarz 1978) and Bayesian evidence estimated using dynesty. For each GP regressor combination we calculate the BIC corresponding to the GP likelihood computed for the best fit parameters. BIC computed using the GP likelihood takes into account the covariance
Figure 3.2: Raw wavelength integrated white target and comparison star light curves of GMOS observations of HAT-P-26b first normalised by the exposure times of the individual exposures and then by their out of transit median flux. Note the low frequency trend present in both the target and comparison star light curves due to the changing airmass (shown in grey) through the night. Observations 2, 4, and 5 were taken using non-ideal PA which is reflected here in the deviating trends between the target and comparison star light curves of the corresponding observations which subsequently contaminates the transit signal in the Target/Comparison light curve and are examples of sub-optimal results from Target/Comparison star light curve normalisation. Observation 6, which was taken with the newly installed Hamamatsu detector on Gemini South, also shows a similar deviating trend between the the target and comparison star light curves.
Figure 3.3: White transit light curves for HAT-P-26b obtained using the GMOS-R150 grism integrated in the range of 530 to 700 nm for observations 1 and 3, and in the range of 530 to 900 nm for the observations 2 and 6. Purple points show the comparison star light curve, black points show the target star (HAT-P-26) light curve overplotted with the best fit NewWLC model in red and the corresponding residuals plotted in the bottom panel of each observation, and green points show the detrended target star light curve overplotted with the batman transit model corresponding to the best fit transit parameters in blue. For observation 4, in pink is overplotted the PSF width time series for the spectral trace of target star. Note that the target and comparison star light curve are affected by the known odd-even pattern in GMOS datasets due to unequal travel times of the GMOS shutter blades which are known to differ slightly with the direction of motion (Stevenson et al. 2014, Jorgensen 2009) as seen significantly in observations 1 and 3.
Figure 3.4: White transit light curves for HAT-P-26b obtained using the GMOS-B600 grism integrated in the range of 490 to 680 nm for observations 4 and 5. Purple points show the comparison star light curve, black points show the target star (HAT-P-26) light curve overplotted with the best fit New WLC model in red and the corresponding residuals plotted in the bottom panel of each observation, green points show the detrended target star light curve overplotted with the batman transit model corresponding to the best fit transit parameters in blue. For observation 4, in pink is overplotted the PSF width time series for the spectral trace of target star.
structure in the data through the covariance matrix (see Equation 3.6). We discuss the model selection threshold in more detail in Appendix 3.B.

We have highlighted the best GP regressor combinations in the Tables 3.9 to 3.14 for the following applications of new and conventional methods that we compare further in Section 3.5.1:

1) New:WLC - Target LC fit with Time and comparison LC, and additional regressors if that is favoured by higher $\log_e Z$ in some cases,

2) New:WLC;No_Comp - Target LC fit without comparison LC as a regressor (Time and/or an additional regressor),

3) Conv1:WLC - Target/Comparison LC fit using the best regressor combination.

For each of the three cases above, we perform the model selection by separately comparing the $\log_e Z$ for the set of GP regressor combinations applicable to each case. Also note that since the new and conventional methods are not fitting the same light curves exactly, we do not use $\log_e Z$ or BIC to perform comparison between the methods themselves but only to choose the best GP regressor combinations for each of them.

3.4.2.4 Odd-even effect in GMOS light curves

The consecutive exposures in the GMOS light curves have been known to suffer from an odd-even effect due to the unequal travel times of the GMOS shutters (Jorgensen 2009) with respect to the direction of motion. This has also been previously observed by Stevenson et al. (2014). We observe the level of this effect for our HAT-P-26b observations to be as high as 700 ppm just for the target star light curves, and as high as 200 ppm for the Target/Comparison light curves, varying with the observation and the corresponding exposure time, and observed most significantly in the R150 observations 1, 2, and 3 (Figures 3.2 and 3.3). The comparison star light curve also suffers from the same odd-even effect at similar time scales as the target star as confirmed from the Lomb Scargle periodograms of both the light curves. Normalising the target light curve by the comparison light curve does not correct for this effect entirely as can be seen for observation 2 (which has the shortest exposure time among all observations) in Figure 3.2 where the odd even effect is still visible in the Target/Comparison light curve. This shows that the odd-even effect prevalent in GMOS observations doesn’t affect the target and comparison light curves in the same manner from one exposure to the next and hence cannot be corrected for completely through a linear method like differential spectrophotometry. This is especially true for observations with shorter exposure times. Instead, the odd-even effect is superimposed on existing high frequency noise in the Target/Comparison light curves due to other variations between systematics affecting the target and comparison light curves individually. This further motivates the need for methods alternative to performing differential spectrophotometry to correct for the effect in the target star light curves directly, which is what our new method does. In particular when considering the residual RMS for observation 2, which has the shortest exposure time of all observations (and hence the largest amplitude of odd-even effect difference between the target and com-
The new method New:WLC performs much better at modelling the odd-even effect in the target star light curves compared to the conventional method Conv1:WLC. In New:WLC the odd-even effect is accounted for by using the comparison star as one of the GP regressors.

Stevenson et al. (2014) use different flux offsets on odd and even frames respectively to correct for this effect in another target HAT-P-7 in their survey, but the method they ultimately use for WASP-12 in Stevenson et al. (2014) is Divide-White which corrects for this effect automatically. Essentially Stevenson et al. (2014) use a linear mapping between the Target/Comparison light curves and an analytical functional form (different offsets for alternative exposures) or a non-analytical form (the Divide-White method) to correct for this effect. In this paper we correct for the effect in the target star light curve directly by letting the GP model do the non-linear mapping between the target star light curves and the odd-even effect information in the comparison star light curve. For spectroscopic light curves the white light curve derived common-mode trend when used as a GP regressor accounts for this effect for New:λLC as described in Section 3.4.3.3.

It should be noted that it is not just because of the presence of differential odd-even effect in the data that makes our new method more effective than the conventional method. We performed a simple transit injection and retrieval test by applying both methods to a pair of synthetic target and comparison star light curves both sharing the same correlated systematics but different levels of white noise. We find that when the comparison star light curve has higher level of white noise than the target star light curve, our new method performs much better than the conventional method in terms of both accuracy and precision of retrieving injected transit parameters.

We highlight that besides the odd-even effect, there are additional possible sources of instrumental and atmospheric systematics that can affect the comparison and target star fluxes differently, which would be potentially be present in data from other multi-object spectrographs as well. These effects can range from low-frequency trends e.g., due to changing CRPA through the night, or high and low frequency telluric absorption variations. The latter effect could be even more significant in near-infrared bands due to second-order colour-dependent extinction effect (e.g. Blake & Shaw 2011, Young et al. 1991).

After performing fits to the white transit light curves and gleaning information about the dominant time-dependent systematics affecting each of our observations, we fit the spectroscopic light curves to obtain the transmission spectrum, as described in more detail in the following section.

3.4.3 Analysis of Spectroscopic Light Curves

3.4.3.1 Construction of spectroscopic light curves

We constructed the spectroscopic transit light curves (λLC) for both target and comparison stars by summing the flux in ~ 20 nm wide bins within the same wavelength range as the respective white light curves. We normalise each exposure in the individual target and com-
Table 3.4: Best fit transit parameters obtained from the fits to white transit light curves of 6 GMOS observations of HAT-P-26b presented in this work. Three sub-rows for each observation number (specified in the first column) show the best fit transit parameters and residual RMS from the applications of our new method (two sub-rows marked New:WLC and New:WLC,No_Comp) and from the conventional method (third sub-row Conv1:WLC). New:WLC and New:WLC,No_Comp are applied to Target LC and Conv1:WLC is applied to Target/Comparison LC. The third column shows the combination of regressors for the GP noise model in Section 3.4.1, where ‘Time’ is the time of each exposure in the light curves, ‘Comp’ is the comparison star light curve, and ‘PSF’ is the full width half maxima of the spectral trace PSF. The bottom section of the table shows the weighted average of transit parameters measured for R150 and B600 observations from the applications of both new and conventional methods, and the transit parameters (weighted average of New:WLC from B600 and R150) eventually used to derive the common-mode trend used to fit the spectroscopic light curves (ΔLC) in Section 3.4.3.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>GP regressors</th>
<th>( R_p/R_\star )</th>
<th>( T_{0,[BJD_{TDB}]-2450000} )</th>
<th>( a/R_\star )</th>
<th>( i ,[^\circ] )</th>
<th>( u_1 )</th>
<th>( \sigma_w )</th>
<th>RMS [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>New:WLC</td>
<td>Time, Comp, PSF</td>
<td>0.0725 ± 0.0023</td>
<td>6371.74717 ± 2.4e-4</td>
<td>11.35 ± 0.48</td>
<td>87.18 ± 0.49</td>
<td>6011 ± 0.25</td>
<td>268 ± 249</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Time</td>
<td>0.0685 ± 0.0023</td>
<td>6371.73731 ± 2.4e-4</td>
<td>11.23 ± 0.44</td>
<td>87.03 ± 0.43</td>
<td>6005 ± 0.29</td>
<td>1500 ± 1254</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time</td>
<td>0.0730 ± 0.0034</td>
<td>6371.74728 ± 2.9e-4</td>
<td>11.21 ± 0.49</td>
<td>86.99 ± 0.47</td>
<td>6007 ± 0.23</td>
<td>299 ± 281</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>New:WLC</td>
<td>Time, Comp</td>
<td>0.0703 ± 0.0034</td>
<td>6392.92016 ± 2.4e-4</td>
<td>12.18 ± 0.45</td>
<td>88.09 ± 0.58</td>
<td>6022 ± 0.26</td>
<td>380 ± 366</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Time, PSF</td>
<td>0.0713 ± 0.0042</td>
<td>6392.91988 ± 2.7e-4</td>
<td>11.98 ± 0.47</td>
<td>87.82 ± 0.54</td>
<td>6022 ± 0.26</td>
<td>347 ± 332</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time</td>
<td>0.0703 ± 0.0045</td>
<td>6392.91946 ± 3.8e-4</td>
<td>11.28 ± 0.61</td>
<td>87.18 ± 0.66</td>
<td>6003 ± 0.26</td>
<td>569 ± 557</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>New:WLC</td>
<td>Time, Comp</td>
<td>0.0694 ± 0.0032</td>
<td>6786.27282 ± 3.2e-4</td>
<td>10.88 ± 0.56</td>
<td>86.59 ± 0.52</td>
<td>5955 ± 0.24</td>
<td>412 ± 398</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Airmass</td>
<td>0.0625 ± 0.0011</td>
<td>6786.27284 ± 3.8e-4</td>
<td>11.67 ± 0.75</td>
<td>86.76 ± 0.86</td>
<td>6003 ± 0.26</td>
<td>1567 ± 1536</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time</td>
<td>0.0696 ± 0.0029</td>
<td>6786.27270 ± 3e-4</td>
<td>11.07 ± 0.72</td>
<td>86.81 ± 0.68</td>
<td>6111 ± 0.29</td>
<td>431 ± 420</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>New:WLC</td>
<td>Time, Comp</td>
<td>0.0683 ± 0.0048</td>
<td>6837.54218 ± 3.8e-4</td>
<td>12.12 ± 0.58</td>
<td>87.85 ± 0.70</td>
<td>5959 ± 0.025</td>
<td>580 ± 560</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Time</td>
<td>0.0706 ± 0.0044</td>
<td>6837.54214 ± 3.4e-4</td>
<td>12.05 ± 0.59</td>
<td>87.77 ± 0.67</td>
<td>598 ± 0.26</td>
<td>589 ± 562</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time</td>
<td>0.0671 ± 0.0047</td>
<td>6837.54211 ± 4.3e-4</td>
<td>11.93 ± 0.62</td>
<td>87.54 ± 0.69</td>
<td>6022 ± 0.26</td>
<td>609 ± 590</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>New:WLC</td>
<td>Time, Comp</td>
<td>0.0654 ± 0.0045</td>
<td>7460.01335 ± 3.6e-4</td>
<td>11.82 ± 0.55</td>
<td>87.71 ± 0.58</td>
<td>73 ± 0.026</td>
<td>125 ± 101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Time, PSF</td>
<td>0.0673 ± 0.0057</td>
<td>7460.01315 ± 6.2e-4</td>
<td>11.50 ± 0.72</td>
<td>87.44 ± 0.74</td>
<td>733 ± 0.027</td>
<td>193 ± 166</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time, Airmass</td>
<td>0.0735 ± 0.0053</td>
<td>7460.01402 ± 4.6e-4</td>
<td>11.29 ± 0.54</td>
<td>87.26 ± 0.54</td>
<td>732 ± 0.025</td>
<td>194 ± 165</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>New:WLC</td>
<td>Time, Comp</td>
<td>0.0685 ± 0.0053</td>
<td>7493.89006 ± 3.1e-4</td>
<td>11.97 ± 0.55</td>
<td>87.81 ± 0.64</td>
<td>733 ± 0.026</td>
<td>164 ± 133</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New:WLC,No_Comp</td>
<td>Time, Airmass, PSF</td>
<td>0.0726 ± 0.0021</td>
<td>7493.89000 ± 2.4e-4</td>
<td>12.56 ± 0.44</td>
<td>88.46 ± 0.69</td>
<td>721 ± 0.024</td>
<td>283 ± 264</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conv1:WLC</td>
<td>Time, Airmass</td>
<td>0.0764 ± 0.0089</td>
<td>7493.89887 ± 4.8e-4</td>
<td>11.64 ± 0.68</td>
<td>87.57 ± 0.73</td>
<td>730 ± 0.029</td>
<td>218 ± 181</td>
<td></td>
</tr>
</tbody>
</table>

For \( \Delta LC \) fits

\[ 0.0701 ± 0.0013 \]

\[ 11.74 ± 0.25 \]
Transmission spectrum of HAT-P-26b using a new method for MOS

Comparison \( \lambda \)LC by the corresponding exposure times. Similar to our white light curve analyses, we fit the \( \lambda \)LC for each observation using the conventional method \( \text{Conv1; } \lambda \)LC and the new method \( \text{New; } \lambda \)LC as described in Section 3.4.1.

3.4.3.2 Fitting the spectroscopic light curves using conventional method

We first describe the application \( \text{Conv1; } \lambda \)LC of the conventional method of fitting \( \lambda \)LCs. We divide the target \( \lambda \)LCs by the corresponding comparison star \( \lambda \)LCs. GMOS observations, like many other ground-based MOS observations, have been conventionally corrected for wavelength-independent systematics through common-mode correction (e.g. Stevenson et al. (2014), H17, Todorov et al. (2019), Wilson et al. (2021)), which leverages the information about the time dependent systematics contained in the white light curve to correct the individual wavelength bins. However, while performing common-mode correction using the white light curve provides an effective way to remove dominant time-dependent systematics, it also implies that we effectively lose information on the absolute value of transit depths and obtain relative transit depths, which is nevertheless useful to search for dominant features in the transmission spectrum.

In \( \text{Conv1; } \lambda \)LC, we follow the conventional common-mode correction approach and derive the common-mode trend as the residuals obtained by subtracting the transit model computed using the weighted averaged transit parameters for the white light curves (last row of Table 3.4) from the white light curves for all observations. Except the limb darkening coefficient we use the same transit parameters for both B600 and R150 observations to construct the transit model. We perform the common-mode correction by subtracting the white light curve derived common-mode trend from the corresponding Target/Comparison \( \lambda \)LC. Note that using the weighted averaged transit parameters to derive the common-mode trend across all observations is valid here as HAT-P-26 is known to be inactive and any potential contamination from stellar activity for the individual epochs that could affect the transit depth is below the precision of our measurements (discussed in more detail in Section 3.5.2).

We then fit the common-mode corrected \( \lambda \)LCs with the model described in 3.4.1, using only time as a GP regressor. This is mainly to account for wavelength-dependent trends not removed by the common-mode correction and arising likely due to wavelength-dependent differential atmospheric extinction between the target and comparison stars with changing airmass through the night (discussed in more detail in Section 3.5.1). Since our main goal is to measure the wavelength-dependent transit depths, we fix the orbital inclination (\( \iota \)), orbital separation (\( a/R_\star \)), and mid transit time (\( T_0 \)) to the best fit values for the corresponding white light curve in Section 3.4.2 (see Table 3.4), and orbital period and eccentricity to literature values. We use a linear limb darkening law for each wavelength bin and fix the limb darkening coefficients to pre-calculated values by PyLDTk (approximating a top hat transmission function for each wavelength bin).

We find that doing common-mode correction prior to fitting the Target/Comparison \( \lambda \)LCs improved the precision of measured transit depths in R150 observations by \( \sim 15\% \) on average per wavelength bin compared to when not performing common-mode correction. The
Target/Comparison R150 λLCs along with their best fit models, detrended light curves, and the residuals are shown in the top three panels of Figures 3.5, 3.6, 3.7, and 3.8.

For the B600 observations, the target and comparison star light curves suffer from significantly different trends through the night as already discussed in Section 3.4.2. Hence, doing Target/Comparison normalisation contaminates the transit signal. This can be noticed by visually inspecting the B600 white light curves in Figure 3.2 and also in the B600 λLCs. Nevertheless, same as done for R150 λLCs, we apply Conv1;λLC to B600 λLCs by doing common-mode correction followed by fitting common-mode corrected Target/Comparison light curves. The resultant fits, detrended light curves, and the residuals are shown in Figures 3.9 and 3.10.

### 3.4.3.3 Fitting the spectroscopic light curves using the new method

We now describe the application New:λLC of our new method to fitting the λLCs. One of the motivations behind our new method is that as we observe the planetary transit through a range of airmass during a night, the differential atmospheric extinction between the target and the comparison star across the optical wavelength range due to the difference in brightness and/or spectral type between the stars implies that simple normalisation of the target λLCs by the comparison λLCs introduces wavelength dependent systematics in the light curves. This is also evident as residual trends in the λLCs after conventional common-mode correction as done Section 3.4.3.2. When inspecting the individual target and comparison star spectra, we find that for our GMOS observations the bluest end of the stellar spectrum suffers from ∼ 5 to 10 % more extinction at high airmasses compared to the reddest end. For the B600 observations specifically, the difference between the atmospheric extinction between the target (fainter than the comparison star) and the comparison star is 5 % more at the bluest end than the reddest end of the spectrum.

The conventional method to mitigate this residual wavelength-dependent noise remaining after common-mode correction as mentioned in Section 3.4.3.2 is to fit the common-mode corrected Target/Comparison λLCs using a linear or quadratic function of airmass or time as the baseline, or a GP model with time as a regressor. In this conventional method, however, there is no straightforward way to ascertain additional systematics propagated to the light curves during the division by comparison λLCs and then common-mode correction. The linear approach of conventional method is also sub-optimal due to non-linear wavelength dependent difference between target and comparison λLCs, lack of wavelength dependent information present in the common-mode trend, and other potential non-linear differences between the target λLCs and common-mode trend.

With our new method to fit the λLCs, we propose to neither perform normalisation by the comparison star λLCs nor perform common-mode correction to the λLCs. We instead use the information of the time dependent systematics contained in the white light curves as one of the regressors in the GP noise model described in Section 3.4.1 for fitting the corresponding λLCs. This is possible through two different combinations of GP regressors:
1) Time and GP noise model of the white light curve (from Section 3.4.2),
2) Time and normalised residuals between the white light curve and its best fit transit model (these residuals are same as the conventional common-mode trend).

The first combination effectively still uses information from the comparison star (which was used to fit the white target light curve and obtain the GP noise model in Section 3.4.2). The second combination however for both R150 and B600 observations doesn’t rely on the comparison star directly and simply leverages the information contained in the common-mode trend to inform the GP systematics model for each $\lambda$LC. This combination is in part analogous to the combination employed by Divide-White method (Stevenson et al. 2014) which uses the white target light curve residuals as a common-mode correction factor in combination with non-analytic models of wavelength dependent systematics derived from the comparison star $\lambda$LCs. In contrast to the conventional Divide-White method, we do not use any information from the comparison star $\lambda$LCs and simply subtract the transit model from the white target light curve and use the residuals or the common-mode trend hence obtained as a regressor in the GP model for the individual $\lambda$LCs. We eventually use the second combination (time and common-mode trend) to fit the target $\lambda$LCs.

It should be noted that the white light curve transit parameters we obtain from not using comparison stars at all (New:WLC;No_Comp sub-row in Table 3.4) are consistent with those obtained from using comparison star as one of the GP regressor (New:WLC sub-row in Table 3.4, see detailed comparison in Section 3.5.1.1). Hence, the derived common-mode trend is consistent between whether we use the comparison stars or not to fit the white light curves, and hence the common-mode trend is not a function of the comparison star light curve in our new method.

Similar to the conventional method described in Section 3.4.3.2, we use the same weighted averaged transit parameters (except the limb darkening coefficient) for both the B600 and R150 observations and the respective transit models used to obtain the common-mode trend from the white light curves. The common-mode trend is then used as a GP regressor to fit the target $\lambda$LCs. Similar to the conventional method, when fitting $\lambda$LCs we keep all the transit model parameters except the transit depth fixed to the best weighted average values derived from the white light curve, and also fix the linear limb darkening coefficients to the pre-computed values from PyLDTk for each spectral bin.

Using the common-mode trend as a GP regressor to fit target $\lambda$LCs as an alternative to subtracting it from $\lambda$LCs is a novel approach and we test its robustness using a transit injection and recovery test described in detail in Appendix 3.A. We find from this test that using the common-mode trend as a GP regressor yields transmission spectra consistent with and on average 25 % better precision than that obtained from the conventional common-mode correction.

Through our transit injection test in Appendix 3.A (see right panel of Figure 3.15) we also demonstrate the choice of using time as a GP regressor in addition to the common-mode trend to fit the target $\lambda$LCs. The common-mode trend by itself models the high frequency systematics in the target $\lambda$LCs which also includes the odd-even effect described in Section
3.4.2.4. Time as an additional GP regressor models the wavelength dependent low frequency trend across the λLCs. It is possible to use additional GP regressors to fit λLCs, but since we independently fit each λLC in this paper, it is not possible to perform model selection for all the λLCs together as done for the white light curves in Section 3.4.2.3. Hence, we stick to the simplest choice of using only time as the additional GP regressor to model the wavelength-dependent trend. It would be possible for a future study into joint modelling of systematics for all λLCs in both time and wavelength dimension to comprehensively explore the use of additional regressors.

The Target λLCs for both B600 and R150 observations along with their best fit models from the new method, detrended light curves, and the residuals are shown in the bottom three panels of Figures 3.5, 3.6, 3.7, 3.8, 3.9, and 3.10. The resulting transmission spectra are tabulated in Tables 3.5 to 3.7, and shown in Figure 3.11 and 3.12.

We compare the transmission spectrum of HAT-P-26b constructed from the best fit wavelength-dependent transit depths for each observation obtained from the conventional and the new method introduced in this paper, and interpret and discuss them in the context of previous transmission spectroscopy measurements of HAT-P-26b.

3.5 Results and Discussion

3.5.1 Comparison of the two Methods and Implications

3.5.1.1 Comparing the white light curve fits

We first compare the performance of the conventional and new methods applied to fitting the white light curves. We compare the three cases Conv1:WLC, New:WLC, and New:WLC;No_Comp used to fit the white transit light curves for each observation highlighted in Table 3.4.

From Table 3.4 we find that the new method (New:WLC and New:WLC;No_Comp) provides similar results compared to the conventional method Conv1:WLC at a precision better than 2σ level.

New:WLC yields on average lower residual RMS compared to Conv1:WLC for all observations. For observations 1, 2, 4, and 5 New:WLC also yields marginally smaller (by ~20% on average) uncertainties on $R_P/R_*$ as compared to the Conv1:WLC. With New:WLC;No_Comp when not using the comparison star at all, we achieve marginally larger (by ~10-20% on average) $R_P/R_*$ uncertainties for all observations except observations 1, 3, and 5. For observation 5, New:WLC;No_Comp gives ~80% smaller uncertainty on $R_P/R_*$. For observations 1 and 3, New:WLC;No_Comp leads to an order of magnitude larger uncertainty on $R_P/R_*$. This is because for these two R150 observations the odd-even effect is particularly high and using comparison star light curve, either as a GP regressor or linearly (as in Conv1:WLC), is crucial to account for the odd-even effect in the target light curve.

For the B600 observations (4 and 5) specifically, the comparison star light curves have time dependent trends significantly different from the target star light curve due to the non-ideal
Table 3.5: New λLC, R150 : Wavelength dependent transit depths (in ppm) for the individual GMOS-R150 observations (marked by the columns) and combined from all observations obtained using the new method described in Section 3.4.3.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>5301 - 5501</td>
<td>4530 ± 473</td>
<td>4813 ± 850</td>
<td>4897 ± 526</td>
<td>4404 ± 995</td>
<td>4681 ± 301</td>
</tr>
<tr>
<td>5501 - 5701</td>
<td>4530 ± 406</td>
<td>5038 ± 809</td>
<td>4684 ± 385</td>
<td>5034 ± 922</td>
<td>4682 ± 252</td>
</tr>
<tr>
<td>5701 - 5999</td>
<td>4958 ± 339</td>
<td>5073 ± 578</td>
<td>5199 ± 397</td>
<td>4440 ± 593</td>
<td>4971 ± 216</td>
</tr>
<tr>
<td>5999 - 6199</td>
<td>5099 ± 277</td>
<td>4773 ± 464</td>
<td>4797 ± 99</td>
<td>4608 ± 576</td>
<td>4830 ± 90</td>
</tr>
<tr>
<td>6199 - 6600</td>
<td>4867 ± 183</td>
<td>4655 ± 332</td>
<td>5177 ± 176</td>
<td>4663 ± 414</td>
<td>4982 ± 106</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>5308 ± 364</td>
<td>5047 ± 191</td>
<td>4962 ± 503</td>
<td>5356 ± 316</td>
<td>5148 ± 138</td>
</tr>
<tr>
<td>6800 - 7000</td>
<td>5037 ± 609</td>
<td>-</td>
<td>4642 ± 369</td>
<td>-</td>
<td>4752 ± 310</td>
</tr>
<tr>
<td>6799 - 7399</td>
<td>-</td>
<td>4896 ± 91</td>
<td>-</td>
<td>4937 ± 183</td>
<td>4903 ± 87</td>
</tr>
<tr>
<td>7799 - 7999</td>
<td>-</td>
<td>4968 ± 211</td>
<td>-</td>
<td>5018 ± 411</td>
<td>4978 ± 184</td>
</tr>
<tr>
<td>7999 - 8201</td>
<td>-</td>
<td>4831 ± 273</td>
<td>-</td>
<td>5205 ± 364</td>
<td>4954 ± 211</td>
</tr>
<tr>
<td>8201 - 8801</td>
<td>-</td>
<td>4836 ± 240</td>
<td>-</td>
<td>4740 ± 391</td>
<td>4809 ± 204</td>
</tr>
</tbody>
</table>
Table 3.6: Conv1:λLC, R150 : Wavelength dependent transit depths (in ppm) for the individual GMOS-R150 observations (marked by the columns) and combined from all observations obtained using the conventional method described in Section 3.4.3.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>5301 - 5501</td>
<td>3947 ± 1112</td>
<td>4637 ± 673</td>
<td>4886 ± 766</td>
<td>4853 ± 331</td>
<td>4762 ± 272</td>
</tr>
<tr>
<td>5501 - 5701</td>
<td>5145 ± 413</td>
<td>4951 ± 474</td>
<td>4850 ± 81</td>
<td>4420 ± 287</td>
<td>4837 ± 77</td>
</tr>
<tr>
<td>5701 - 5999</td>
<td>5396 ± 311</td>
<td>4921 ± 290</td>
<td>4875 ± 109</td>
<td>4729 ± 228</td>
<td>4892 ± 100</td>
</tr>
<tr>
<td>5999 - 6199</td>
<td>5457 ± 220</td>
<td>4546 ± 202</td>
<td>4776 ± 110</td>
<td>4907 ± 334</td>
<td>4867 ± 82</td>
</tr>
<tr>
<td>6199 - 6600</td>
<td>5222 ± 183</td>
<td>4601 ± 159</td>
<td>4716 ± 65</td>
<td>4678 ± 203</td>
<td>4760 ± 54</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>5178 ± 171</td>
<td>4936 ± 243</td>
<td>4491 ± 177</td>
<td>4722 ± 199</td>
<td>4841 ± 98</td>
</tr>
<tr>
<td>6800 - 7000</td>
<td>4943 ± 1096</td>
<td>-</td>
<td>4400 ± 353</td>
<td>-</td>
<td>4441 ± 321</td>
</tr>
<tr>
<td>6799 - 7399</td>
<td>-</td>
<td>4569 ± 111</td>
<td>-</td>
<td>4989 ± 93</td>
<td>4844 ± 74</td>
</tr>
<tr>
<td>7799 - 7999</td>
<td>-</td>
<td>5074 ± 193</td>
<td>-</td>
<td>4709 ± 275</td>
<td>4976 ± 140</td>
</tr>
<tr>
<td>7999 - 8201</td>
<td>-</td>
<td>5115 ± 223</td>
<td>-</td>
<td>4667 ± 293</td>
<td>4975 ± 161</td>
</tr>
<tr>
<td>8201 - 8801</td>
<td>-</td>
<td>4654 ± 214</td>
<td>-</td>
<td>4271 ± 289</td>
<td>4525 ± 164</td>
</tr>
</tbody>
</table>
Table 3.7: New \( \lambda \)LC, B600 : Wavelength dependent transit depths (in ppm) for the individual GMOS-B600 observations (marked by the columns) and combined from all observations obtained using the new method described in Section 3.4.3.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>4</th>
<th>5</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>4900 - 5100</td>
<td>4925 ± 75</td>
<td>4266 ± 550</td>
<td>4913 ± 75</td>
</tr>
<tr>
<td>5100 - 5300</td>
<td>5197 ± 418</td>
<td>4755 ± 448</td>
<td>4992 ± 306</td>
</tr>
<tr>
<td>5300 - 5500</td>
<td>4879 ± 324</td>
<td>4713 ± 289</td>
<td>4787 ± 216</td>
</tr>
<tr>
<td>5500 - 5700</td>
<td>4925 ± 255</td>
<td>4819 ± 393</td>
<td>4894 ± 214</td>
</tr>
<tr>
<td>5700 - 6000</td>
<td>4735 ± 197</td>
<td>4784 ± 212</td>
<td>4757 ± 145</td>
</tr>
<tr>
<td>6000 - 6200</td>
<td>5026 ± 41</td>
<td>4401 ± 223</td>
<td>5006 ± 40</td>
</tr>
<tr>
<td>6200 - 6400</td>
<td>4830 ± 187</td>
<td>4759 ± 381</td>
<td>4816 ± 168</td>
</tr>
<tr>
<td>6400 - 6600</td>
<td>4583 ± 226</td>
<td>5091 ± 307</td>
<td>4761 ± 182</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>4760 ± 410</td>
<td>5196 ± 398</td>
<td>4984 ± 286</td>
</tr>
</tbody>
</table>
Figure 3.5: Spectroscopic light curves for observation 1 (R150) fit using the conventional method (Conv1: λLC, three panels on this page) and the new method introduced in this paper (New: λLC, three panels in the continued Figure on the next page). The leftmost panel for each method shows the best fit to the light curves for each wavelength bin, the middle panel shows the detrended light curves, and the rightmost panel shows the corresponding residuals, their histograms, and the RMS of their scatter. The target λLCs show a wavelength dependent low frequency trend due to changing airmass through the night.
Figure 3.5: Continued.
Observation 2

**Figure 3.6:** Same as Figure 3.5 for observation 2 (R150).
Transmission spectrum of HAT-P-26b using a new method for MOS

Figure 3.6: Continued.
Figure 3.7: Same as Figure 3.5 for observation 3 (R150).
Figure 3.7: Continued.
Figure 3.8: Same as Figure 3.5 for observation 6 (R150).
Figure 3.8: Continued.
3.5 Results and Discussion

Figure 3.9: Spectroscopic light curves for observation 4 (B600) fit using the conventional method (conv1: λLC, three panels on this page) and the new method introduced in this paper (New: λLC, three panels in the continued Figure on next page). The leftmost panel for each method shows the best fit to the light curves for each wavelength bin, the middle panel shows the detrended light curves, and the rightmost panel shows the corresponding residuals, their histograms, and the RMS of their scatter. The target λLCs show a wavelength dependent low frequency trend due to changing airmass through the night.
Figure 3.9: Continued.
Figure 3.10: Same as Figure 3.9 for observation 5 (B600).
Figure 3.10: Continued.
Table 3.8: Conv1:λLC, B600 : Wavelength dependent transit depths (in ppm) for the individual GMOS-B600 observations (marked by the columns) and combined from all observations obtained using the new method described in Section 3.4.3.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>4</th>
<th>5</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>4900 - 5100</td>
<td>2577 ± 1280</td>
<td>4572 ± 1373</td>
<td>3504 ± 937</td>
</tr>
<tr>
<td>5100 - 5300</td>
<td>4151 ± 791</td>
<td>5174 ± 774</td>
<td>4674 ± 553</td>
</tr>
<tr>
<td>5300 - 5500</td>
<td>4199 ± 553</td>
<td>4510 ± 429</td>
<td>4393 ± 339</td>
</tr>
<tr>
<td>5500 - 5700</td>
<td>3776 ± 1130</td>
<td>4974 ± 294</td>
<td>4898 ± 284</td>
</tr>
<tr>
<td>5700 - 6000</td>
<td>4825 ± 229</td>
<td>4868 ± 246</td>
<td>4845 ± 168</td>
</tr>
<tr>
<td>6000 - 6200</td>
<td>4976 ± 166</td>
<td>5041 ± 127</td>
<td>5017 ± 101</td>
</tr>
<tr>
<td>6200 - 6400</td>
<td>4945 ± 283</td>
<td>4968 ± 212</td>
<td>4960 ± 170</td>
</tr>
<tr>
<td>6400 - 6600</td>
<td>4819 ± 238</td>
<td>4906 ± 237</td>
<td>4862 ± 168</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>4860 ± 383</td>
<td>4903 ± 260</td>
<td>4889 ± 215</td>
</tr>
</tbody>
</table>

PA of the observational setup, which significantly contaminates the transit signal in the resulting Target/Comparison light curves (as seen in Figure 3.2). From a visual inspection of the B600 light curves in Figure 3.2, Target/Comparison corrects for the odd-even effect in the transit light curve but adds additional low frequency trend not present in the original target transit light curve. This effect especially strong before and during the transit. Nevertheless, the conventional method Conv1:WL for fitting the B600 Target/Comparison light curves using a GP with time and airmass as regressors retrieves transit parameters consistent with R150 observations. Notably, New:WL achieves lower uncertainties on $R_p/R_*$ as compared to Conv1:WL. Not using the comparison star with New:WL;No_Comp yields lower (observation 5) or marginally larger but comparable (observation 4) uncertainties on $R_p/R_*$.

We conclude from comparison across the aforementioned three cases of fitting the white light curves that both applications of our proposed new method perform consistently and even better in some instances compared to the conventional method. We also conclude that in most instances it is possible to detrend the target light curves and achieve decent precision on transit parameters even without using the comparison stars at all.

3.5.1.2 Comparing the spectroscopic light curve fits

We now compare the transmission spectra obtained from the conventional method Conv1:λLC and the new method New:λLC to fit λLCs. We emphasize here that when not using comparison star at all to fit the white target light curves we obtain consistent transit parameters and hence the common-mode trend as when we use the comparison star (also see Section 3.4.3.2). Hence all conclusions below about the new method are valid whether the common-mode trend is derived by using the comparison star indirectly as in New:WL or not using it at all in New:WL;No_Comp.
For both B600 and R150 observations, the transmission spectra shown in Figures 3.11 and 3.12 and corresponding wavelength-dependent transit depth values in Tables 3.5, 3.6 reveal that the individual and combined transmission spectra across observations from the Conv1:λLC and New:λLC are on average consistent within their 2σ uncertainties. New:λLC on average yields ~ 40% smaller RMS of the residuals per wavelength bin (see Figures 3.5 to 3.10).

The per wavelength bin uncertainties on the transmission spectra are on average ~ 30% larger from New:λLC as compared to that from Conv1:λLC for the R150 observations. For the B600 observation 4 New:λLC yields ~ 50% smaller uncertainties, especially for the bluest wavelength bins. New:λLC also performs similarly well in terms of precision for the three bluest bins for the B600 observation 5, but yields nearly ~ 30% larger uncertainties for the redder bins. This difference in uncertainties on the transmission spectra points towards fundamental differences between the two methods in their approach of dealing with the systematics which we elaborate on. One clear difference is the number of free hyperparameters used for the GP models in both methods. For Conv1:λLC, the GP model uses only one regressor (time) and hence two hyperparameters (amplitude and length scale, see Equation 3.3 and 3.4 in Section 3.4.1). New:λLC in comparison uses two regressors (time and common-mode trend) and hence three hyperparameters (common amplitude, and one length scale hyperparameters for each of the regressors). Using a more flexible model with more hyperparameters is one of the reasons behind larger uncertainties in the transmission spectra from New:λLC.

Note that Conv1:λLC before fitting the GP model also involves two additional steps: dividing by comparison λLCs and subtracting the common-mode trend. Both of these steps are linear corrections which do not explicitly propagate uncertainties arising from non-linear differences between the target λLCs and the comparison λLCs or common-mode trend. It can be seen from the target λLCs in Figures 3.5 to 3.10 that target star light curves suffer from a low frequency trend in time that varies with wavelength due to wavelength-dependent extinction that is changing with the airmass. These low frequency trends still remain after division by comparison λLCs and subtracting the common-mode trend, as seen in the Target/Comparison λLCs in Figures 3.5 to 3.10. There is also a high frequency trend (e.g. odd-even effect described in Section 3.4.2.4) which affects every wavelength bin in a similar manner.

Our new method New:λLC for fitting target λLCs doesn’t use comparison λLCs and accounts for trends at both frequencies in addition to accounting for their wavelength dependence in one step. The Bayesian framework of GPs propagates the uncertainties in the information from the common-mode trend as relevant to each target λLCs. Specifically, the common-mode trend helps in accounting for the high frequency trends while time accounts for the low frequency trend varying with respect to wavelength (as demonstrated in Appendix 3.A). Moreover, when not using comparison star λLCs, using common-mode correction as a GP regressor can potentially provide better precision as compared to conventional common-mode correction as we demonstrate in Appendix 3.A. By not using the comparison star λLCs, we prevent possible introduction of additional systematics due to different instrumental systematics or differential atmospheric extinction between the target and comparison stars with
changing airmass. This is supported by the superior performance of New:λLC for λLCs for the bluest bins of observation 4 as compared to Conv1:λLC in terms of precision and accuracy of transit depths.

3.5.1.3 Implications of measuring transmission spectra without using comparison stars

In the previous subsection, we show that decent precision on fits to both white transit light curves and λLCs can be achieved even when not using the comparison star at all. We highlight the usefulness of this aspect of our method for cases when the comparison star is not a suitable reference for systematics in the target star light curve either due to large differences in brightness or spectral type, or issues with the observational setup, as is the case of our GMOS-B600 observations. In fact, our new method essentially removes the transit signal from the white target light curves and uses the information in the residuals (common-mode trend) to fit the target λLCs. In this context, a further step could be that we may not need to fit the white target light curves and we can rely on using the previously measured planet transit parameters from other observatories e.g. TESS, HST/STIS in the bandpass significantly overlapping with GMOS and theoretical priors on the limb darkening for the star to compute the transit signal used to obtain the common-mode trend which can be used to fit λLCs. Caveats of this approach of bypassing the fit of the white light curve are observations of planets with variable broadband transit depths due to e.g., stellar host variability over multiple epochs. In such cases, it would be essential to fit the white light curve transit depths for individual epoch first to be able to obtain common-mode trend with normalisation to the transit depth for that epoch leading to accurate absolute transmission spectra. In particular, for active host stars, instead of using the same transit depth to derive the common-mode trend across all epochs (as we do for HAT-P-26 in this paper), we advise using the individual best fit white light curve transit depths for each epoch.

Our new method of extracting transmission spectra from solely target star light curves has further implications for ground-based follow-up atmospheric observations of exoplanets orbiting bright host stars especially those discovered by TESS. In particular, the majority of TESS stellar host stars are bright in optical, with median $V_{mag} \sim 11$ as indicated by simulations from Barclay et al. (2018), and may not have a choice of comparison stars with similar brightness and spectral type in the limited field of view of up to 10 arcminutes for most ground based multi object spectrographs. We recommend using New:WLC,No_Comp followed by New:λLC for obtaining ground-based transmission spectra of such exoplanets orbiting bright stars. Furthermore, another strength of our new method is that it can potentially mitigate significant second order colour dependent extinction effects arising due to differences in target and comparison star spectral types (Blake & Shaw 2011).

3.5.2 Interpretation of the Optical to NIR Transmission Spectrum

We generate the combined transmission spectrum by weighted averaging the wavelength-dependent transit depths across common wavelength bins covered by individual R150 and B600 observations, taking the squared reciprocal of the transit depth uncertainties as weights
Transmission spectrum of HAT-P-26b using a new method for MOS

**Figure 3.11:** Transmission spectra for the GMOS-R150 observations obtained using Conv\(1:\lambda\perp\)LC and New\(:\lambda\perp\)LC (slightly shifted in wavelength for clarity) described in Section 3.4.3. The average GMOS optical transit depth (corresponding to the weighted average white light curve \(R_p / R_\star = 0.0701^2 = 4914\) ppm), which is consistent with the median HST STIS/G750L transit depth from Wakeford et al. (2017), is marked by the dashed line. For each observation, in black are shown the spectra obtained through conventional method Conv\(1:\lambda\perp\)LC of fitting Target/Comparison \(\lambda\perp\)LCs using a GP model with time as a regressor. In red are the transmission spectra obtained by using the new method New\(:\lambda\perp\)LC to extract transmission spectra only from Target \(\lambda\perp\)LCs: using a GP model with time and common-mode trend as regressors are shown in red. Overplotted is the observed stellar spectrum for the target star (HAT-P-26b) in green.
Figure 3.12: Transmission spectra for the GMOS-B600 observations obtained using Conv1:λLC and New:λLC (slightly shifted in wavelength for clarity) as described in Section 3.4.3. The average GMOS optical transit depth (corresponding to the weighted average white light curve \( R_\mathrm{P}/R_\star \approx 0.0701 \times 4914 \) ppm), which is consistent with the median HST STIS/G750L transit depth from Wakeford et al. (2017), is marked by the dashed line. For each observation, in black are shown the spectra obtained through conventional method Conv1:λLC of fitting Target/Comparison λLCs using a GP model with time as a regressor. In red are the transmission spectra obtained by using the new method New:λLC to extract transmission spectra only from Target λLCs : using a GP model with time and common-mode trend as regressors are shown in red. Both the B600 observations were obtained using non-ideal PA which manifests as widely different time dependent trends in the target and comparison λLCs. This leads to contamination of the transit signal in Conv1:λLC (black points), especially for the bluest wavelength bins as seen here for both B600 observations. Overplotted is the observed stellar spectrum for the target star (HAT-P-26b) in green.

for the respective observations. The combined transmission spectrum values from both methods for R150 observations are shown in Tables 3.5 and 3.6, and for B600 observations in Tables 3.7 and 3.8. Since for the B600 observations, New:λLC performs much better than Conv1:λLC, we only consider the combined transmission spectra obtained from New:λLC for further comparison with atmospheric models.

We use the open source atmospheric modelling code platon (Zhang et al. 2019, 2020) based on ExoTransmit (Kempton et al. 2017) to conduct a simple retrieval analysis for the atmosphere of HAT-P-26b to interpret our combined GMOS observations in conjunction with the near infrared transmission spectra measurements from HST and Spitzer reported by Wakeford et al. (2017). For the self-consistent retrieval framework of platon we consider equilibrium chemistry models for three cases: 1) both metallicity and C/O fixed to solar values, and 2) both metallicity and C/O free to fit, 3) metallicity free and C/O fixed to solar value. For all three cases we also let free the pressure level of a grey opacity cloud deck.

Since early measurements of chromospheric activity indicator \( S_{HK} \) index (Hartman et al. 2011) and subsequent photometric follow-up observations by von Essen et al. (2019) show no signs of activity or significant spot modulated variability of stellar photospheric brightness, we do not include contributions from transit light source effect (Rackham et al. 2018) in our retrieval analysis. From the stellar photometry reported by von Essen et al. (2019) no signatures of spot modulations of stellar flux are observed and the upper limit on V band photometric variability for HAT-P-26 (a K1 dwarf) is 2.3 parts per thousand or 0.23 % (which is the maximum scatter in the light curves). Referring to empirical relationship between the peak to peak optical variability amplitude vs spot covering fraction for K dwarfs from Rack-
ham et al. (2019), we note that 0.2 % variability would correspond to less than 1 % spot covering fraction. Considering the upper limit of 1 % spot covering fraction, we use Equation 3 from Rackham et al. (2019) to estimate the upper limit on the amplitude of wavelength dependent stellar contamination factor on the transmission spectrum and find it to be 0.9901. Considering the average transit depth of HAT-P-26b to be around 5000 ppm, this would correspond to a maximum offset of 50 ppm to the transmission spectrum, which is about a factor 5 to 10 less as compared to the average precision of the transmission spectrum in the individual epochs. We hence conclude that given the precision of our observations, we would not be able to detect the offsets due to stellar contamination corresponding to the available upper limits from stellar photometry.

Note that our GMOS-B600 observation 4 was taken at the same time (on 12/03/2016 UT) as one of the HST/WFC3 observations of Wakeford et al. (2017) and the consistency of the median wavelength-dependent transit depth between both observations taken simultaneously from two different further underscores the suitability of combining them. Hence, we do not introduce any vertical offset between the measurements from GMOS, HST, and Spitzer in further analysis.

We find that the transmission spectrum of HAT-P-26b from the combined GMOS, HST, and Spitzer measurements are best explained by a model corresponding to a solar metallicity and solar C/O atmosphere with a grey opacity cloud deck at $\log_{10} P \text{(bar)} = -2.5^{+0.53}_{-0.28}$, which is consistent with the pressure level of the cloud deck constrained by Wakeford et al. (2017) ($\log_{10} P \text{(bar)} \sim -2$) using STIS/G750L observations. The $\chi^2_{red}$ for the best fit model with a grey opacity cloud is 1.68 compared to 17.4 for a cloud-free model. The resulting best fit model along with the cloud-free model for comparison is shown in Figure 3.13. Given the lack of coverage at the bluest optical end of the transmission spectrum due to the drop in throughput of GMOS observations blueward of 490 nm our observations cannot constrain the signatures of tentative Rayleigh scattering predicted by MacDonald & Madhusudhan (2019). We also do not confirm or rule out the $\sim 400$ ppm TiH feature at 0.54 $\mu$m predicted by MacDonald & Madhusudhan (2019) due to our precisions around this region (see Tables 3.5 to 3.7) being comparable to the amplitude of the feature as well as our seeing limited resolution restricting us to 20 nm wide wavelength bins.

### 3.5.3 Transit Timing Variations

Ground based transit observations from multi object spectrographs like GMOS can provide high precision (of the order of 10s of seconds) on the mid-transit time as a result of the high signal to noise nature of observations and continuous sampling of the transit including the ingress and egress without gaps. An example is the mid transit times from the Gemini/GMOS observations of WASP-4b (H17) which when combined with other timing measurements including those from TESS by Bouma et al. (2019, 2020) have been used to study the transit timing variations of the planet at high precision.

For HAT-P-26b, we obtain an average precision of $\sim 25$ seconds on the mid-transit times across the 6 GMOS transit observations as shown in Table 3.4 (mid-transit times from New:WLC).
3.5 Results and Discussion

Figure 3.13: Combined optical transmission spectrum from the 4 GMOS-R150 (red points) and 2 GMOS-B600 (blue points) observations obtained from New:WLc presented in Section 3.4.3.3 along with the previous measurements in the optical and near infrared from HST/STIS-G750L and HST/WFC3-G102 and WFC3-G141, and in infrared from Spitzer as presented by Wakeford et al. (2017). Overplotted is the best fit transmission spectroscopy model obtained using platon which has a cloud deck at 3.5 millibar ($10^{-2.5}$ bar), in solid green, and a cloud free model in dotted green for comparison.

We combine our mid-transit times from New:WLC with those compiled by von Essen et al. (2019). The mid-transit time measured from GMOS for observation 4 is consistent with that measured from the simultaneous HST/WFC3 transit observation from Wakeford et al. 2017 within the 1 σ uncertainty. Taking the zeroth epoch same as that considered by von Essen et al. (2019) we compute the observed minus calculated (O – C) for the GMOS mid-transit times assuming a linear ephemeris for the calculated or predicted mid-transit times. To these O – C values combined with the measurements from von Essen et al. (2019) we then fit a sinusoidal model with three free parameters: amplitude of the TTVs $A_{TTV}$, period (P, in number of epochs), and a phase value ($\phi_{TTV}$) using emcee. The resulting best fit and fits from random samples from the posteriors computed by emcee are shown in Figure 3.14.

Our best fit sinusoidal fit has an amplitude of $A_{TTV} = 1.21_{-0.039}^{+0.040}$ minutes, with period $P = 366.016_{-14.19}^{+14.76}$ epochs and $\phi_{TTV} = -2.74_{-0.37}^{+0.38}$. The reduced chi-squared value (with a degree of freedom 22) for the sinusoidal fit to the O – C including the GMOS and von Essen et al. (2019) measurements is $\sim 5$ as compared to $\sim 288$ for O – C = 0 which is the case when the measured O – C values would be consistent with a linear ephemeris. This is consistent with the indication of TTVs for HAT-P-26b previously reported by von Essen et al. (2019) and also indicated by Stevenson et al. (2016), and motivates future follow up using both transit and secondary eclipse measurements to determine the physical explanation behind the TTVs.
Figure 3.14: Observed minus calculated mid transit times (O – C, from a linear ephemeris) from the mid transit times presented by von Essen et al. (2019) (black points, including a compilation of all the previously published mid transit times and those measured by them) and those presented in this paper (red and blue points, numbers corresponding to observation number in Table 3.1). Overplotted in dashed black line is the best fit sinusoidal model to only the O – C values from von Essen et al. (2019), in solid black is the best fit sinusoidal model fit to O – C values from the von Essen et al. (2019) and the GMOS observations, and in orange are the randomly sampled fits from the MCMC posteriors.
3.6 Conclusions

We have introduced a new method to model systematics in ground based spectrophotometric observations that allows for a generalised non-linear mapping between the target star transit light curves and the time series used as regressors to detrend them. We test and demonstrate the performance of the new method in comparison to the conventional method by applying both methods to ground-based optical transmission spectra of the warm Neptune HAT-P-26b from 6 transits observed by Gemini/GMOS as part of our ground-based survey of exoplanet atmospheres in the optical.

We summarise the key aspects and conclusions for the new method we introduce in this paper:

1) With the new method, we fit the systematics and transit signal in the target star white light curves directly by using a GP regression model conditioned with various combinations of regressors which include the simultaneously observed comparison star white light curve. This is a generalisation of conventional linear methods which have used comparison star white light curves as a linear regressor. The new method when using comparison star white light curves as a GP regressor lets the GP determine the underlying non-linear mapping between the comparison and target star light curves. This approach utilises the information about systematics from the comparison star light curves without introducing additional uncertainties as often is the case when doing differential photometry. It also propagates uncertainties appropriately within the Bayesian framework of GPs when using the comparison star light curve as a GP regressor.

2) The application of the new method New:WLC;No_Comp to fit the target white light curves without using the comparison light curves emulates a scenario when suitable comparison stars may not be available. We show that even in the absence of suitable comparison stars, accurate transit parameters with comparable precisions can be obtained from the white target transit light curve fit using our new method.

3) The new method when applied to λLCS lets the GP determine the non-linear mapping between the white target light curve derived common-mode trend and the individual target λLCS. We show by application to observed and transit injected λLCS that this approach without needing to perform normalization by comparison λLCS is robust and achieves accurate transmission spectra. From the transit injection test, we conclude that using common-mode trend as a GP regressor achieves ~ 20% better precision on the transmission spectra compared to that from conventional common-mode correction.

4) Except for the bluest bins in B600 observations, the new method yields marginally higher uncertainties on the transmission spectra. We interpret this increase in uncertainties as an outcome of fitting for both low and high frequency systematics in λLCS in one step and propagating the uncertainties in the process. In contrast, the conventional linear method with multiple steps of dividing by comparison λLCS and subtracting the common-mode trend
does not explicitly propagate uncertainties at each step.

5) In the context of bluest bins in B600 observations, where in addition to effects due to non-ideal PA we also expect largest differential atmospheric extinction between the target and comparison star spectra due to changing airmass, we show that our new method is able to extract the transmission spectra for scenarios when the conventional Target/Comparison normalisation strongly contaminates the transit signal.

6) We demonstrate that just the target white light curve itself can be used to model the time and wavelength-dependent systematics in the spectroscopic target light curves, albeit at the cost of ~30% larger uncertainties on the transmission spectra. This approach can ultimately be used for future optical and near infrared ground-based atmospheric characterisation of exoplanets orbiting bright host stars with little or no available choice of comparison stars with similar brightness and spectral type in the instrument field of view.

7) The current prescription of the new method as applied to λLCs in this paper fits each λLC independently and hence does not explicitly model potential covariance in the wavelength dimension. A future possible extension to our method, especially when applied to medium resolution spectrophotometric observations, is to jointly model the λLCs accounting for potential covariance due to systematics in wavelength dimension.

Based on our analyses, we obtain the following conclusions about the atmosphere of HAT-P-26b:

1) Through equilibrium chemistry retrieval analysis of combined GMOS optical observations with near infrared HST and Spitzer observations, we conclude that the terminator of HAT-P-26b is consistent with solar metallicity and C/O atmosphere with a grey opacity cloud layer at \( \log_{10} P \text{ (bar)} = -2.5^{+0.53}_{-0.28} \) obscuring the alkali absorption features in optical and suppressing the water absorption features in the near infrared, consistent with the findings of Wakeford et al. (2017). The low resolution nature of our observations and comparatively low precision on the transit depths preclude confirmation of presence of metal hydride features predicted by MacDonald & Madhusudhan (2019).

2) Based on the mid transit times constrained by the GMOS transits we find further indications of TTVs for HAT-P-26b in agreement with previous studies. This warrants future follow up primary and secondary eclipse observations of the planet to investigate the physical origin of TTVs.

Finally our results add to the growing library of optical transmission spectra of exoplanets obtained using ground-based low-resolution spectrographs. The precision and accuracy of our measurements combined with the repeatability of the observations over multiple epochs
emphasize the importance of optical ground-based observations in complementing the up-
coming observations of transiting exoplanets in the infrared using JWST.

3.7 Acknowledgements

Based on observations obtained at the Gemini Observatory (acquired through the Gemini
Observatory Archive and Gemini Science Archive), which is operated by the Association
of Universities for Research in Astronomy, Inc. (AURA), under a cooperative agreement
with the NSF on behalf of the Gemini partnership: the National Science Foundation (United
States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia,
Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e
Inovação (Brazil). Based in part on Gemini observations obtained from the National Optical
Astronomy Observatory (NOAO) Prop. ID: 2012B-0398; PI: J.-M Désert. We thank the anony-
ymous reviewer for their thorough comments and feedback which helped in improving the
paper. V.P. acknowledges stimulating discussions with Lorenzo Pino, Claire Baxter, Jacob
Arcangeli, Nilooofar Khorshid, Bob Jacobs, Saugata Barat, Ernst de Mooij, Neale Gibson, and
Dan-Foreman Mackey which helped in shaping the paper. V.P. also acknowledges inspiring
discussions on Gaussian Processes with Mehmet Deniz Aksulu. J.M.D acknowledges sup-
port from the Amsterdam Academic Alliance (AAA) Program, and the European Research
Council (ERC) European Union’s Horizon 2020 research and innovation programme (grant
agreement no. 679633; Exo-Atmos). This work is part of the research programme VIDI New
Frontiers in Exoplanetary Climatology with project number 614.001.601, which is (partly)
financed by the Dutch Research Council (NWO). This material is based upon work sup-
ported by the NWO TOP Grant Module 2 (Project Number 614.001.601). This material is
based upon work supported by the National Science Foundation (NSF) under Grant No. AST-
1413663. This research has made use of NASA’s Astrophysics Data System. The authors also
acknowledge the significant cultural role and reverence the summit of Mauna Kea has within
the indigenous Hawaiian community. This research has made use of Astropy,\(^1\) a community-
developed core Python package for Astronomy Astropy Collaboration et al. (2013, 2018),
NumPy Harris et al. (2020), matplotlib Hunter (2007), SciPy Virtanen et al. (2020) and
IRAF Tody (1986) distributed by the NOAO, which is operated by AURA under a cooperative
agreement with the NSF.

\(^1\)http://www.astropy.org
3.A Testing robustness of using common-mode trend as a GP regressor for $\lambda$LC

Correcting for time only dependent systematics in $\lambda$LCs has been conventionally done by dividing or subtracting each $\lambda$LC by a common-mode trend derived from the white light curve. One of the novel aspects of the method we introduce in this paper is to use this common-mode trend as a GP regressor instead of subtracting it from each $\lambda$LC. In this section, we perform a transit injection and recovery test to assess the robustness of using the common-mode trend as a GP regressor to fit $\lambda$LCs and deriving the transmission spectrum.

We take the observation 4 B600 comparison star white light curve and additively inject to it a transit signal with known transit parameters and linear limb darkening coefficient fixed to the PyLDTk for HAT-P-26b. We inject the same transit signal to each of the 20 nm spectroscopic light curves keeping the limb-darkening coefficient of the injected model same across wavelength bins.

We first fit the synthetic white transit light curve (referred to as WLC for brevity) with the injected signal using a batman model for the transit plus a GP with time as a regressor for the systematics. We subtract the best fit batman transit model thus obtained from the WLC to obtain the residuals (i.e. common-mode trend) to be used for the next steps. For fitting the $\lambda$LCs, we test 4 different cases: 1) Using time and WLC residuals as GP regressors to fit $\lambda$LC, 2) Subtracting the WLC residuals from each $\lambda$LC (conventional common-mode correction) and fitting the common-mode corrected $\lambda$LC using time as GP regressor, 3) Using only WLC residuals as a GP regressor to fit $\lambda$LC, and 4) Using only time as a GP regressor to fit $\lambda$LC. The first two cases are the ones that we eventually use in the paper (in Section 3.4.3). We discuss the latter two cases to demonstrate the individual contributions from time and WLC residuals as GP regressor, respectively. We show the resulting transmission spectra for each case in Figure 3.15. Note that the case 2) here involving conventional common-mode correction does not involve dividing $\lambda$LCs by any corresponding comparison star $\lambda$LCs, and hence is not exactly the same as the conventional method used in the paper (in Section 3.4.3.2 where we do divide the targets $\lambda$LCs by comparison star $\lambda$LCs.

We outline below the conclusions from our transit injection test below:

1) The new method introduced by us in the paper of using the WLC residuals as a GP regressor along with time to fit the $\lambda$LCs robustly retrieves the injected transit signal in the
Figure 3.15: Transmission spectra from different GP regressor combination cases used to test the robustness of using common-mode trend as a GP regressor to fit the \( \lambda \)LCs with the injected transit signal (horizontal green line in both panels) as described in Section 3.A. The dashed black line in both panels shows the best fit transit depth for the WLC, obtained from fitting it using a GP model with only time as a regressor. Left panel shows the spectra resulting from fitting \( \lambda \)LCs using 1) New method: Time and WLC residuals as a GP regressor (pink) and 2) Subtracting WLC residuals and fitting using Time as a GP regressor (black). The two cases are consistent with each other and with the best fit WLC transit depth within 1\( \sigma \). The new method results in 25 \% smaller uncertainties on average. Right panel shows the spectra from fitting \( \lambda \)LCs with a GP with regressor as 3) only WLC residuals (blue) and 4) only Time (orange). WLC residuals only case yields high precision and less accurate spectra, and vice versa for Time only case. This shows that WLC residuals are better at fitting high frequency wavelength independent systematics while Time helps in fitting the lower frequency wavelength dependent trend. Both are complementary to achieve better accuracy and precision.

The mean values of the transit depths across the bins from this method (in pink in the Figure 3.15) are consistent within \( \pm \) 1 \( \sigma \) with those retrieved from the conventional common-mode correction (subtract the WLC residuals from each bin and fit using a time-dependent GP model, grey in the Figure 3.15).

2) The mean values from both the methods are centred around the best fit WLC transit depth within \( \pm \) 1 \( \sigma \) (black dashed line) and deviate by almost 2 to 3 \( \sigma \) from the injected transit depth. This is a potential pitfall of both the new and conventional methods of using the common-mode trend, and shows that the accuracy of both the methods depends on the accuracy of WLC fit.

3) The uncertainties from the new method (GP regressors: time and WLC Residuals) is on average 25 \% lower as compared to that from the conventional common-mode correction followed by fitting using time as a GP regressor. This shows that the new method of using the common-mode trend as a GP regressor performs better than conventional common-mode correction in terms of retrieved precision on transmission spectra.

4) We also show the results from two additional cases: using only WLC residuals as GP regressor, and using only time as a GP regressor performs. We find that using only WLC
residuals as a GP regressor performs poorly in terms of the accuracy of the retrieved transit depths (blue points in Figure 3.15). The WLCs suffer from a wavelength dependent low-frequency trend due to the changing airmass through the night. The shape of this trend varies across the wavelength bins due to wavelength dependent atmospheric extinction. This effect can also be seen in the target star spectroscopic light curves shown in the paper in Figures 3.5 to 3.10. The WLC residuals by themselves when used as a GP regressor model the high frequency systematics but are unable to take this low frequency wavelength dependent effect into account. Using time as an additional GP regressor helps to take this account as shown in Figure 3.15 (pink points). On the other hand using only time as the GP regressor (orange points), while performing well in terms of overall accuracy of transit depths, performs poorly in terms of precision. We interpret this as the inability of the time only GP regressor model to account for the high frequency systematics in \( \lambda \)LCs, which is the reason the uncertainties on the corresponding transit depths are larger.

3.B Model selection criteria for GP regressor combinations for fitting white transit light curves

We summarize the BIC and log Bayesian evidence values for each GP regressor combination for all observations corresponding to the two methods used for fitting white light curves (Section 3.4.2) in Tables 3.9 to 3.14. The combinations are shown in decreasing order of \( \log_e Z \), which is broadly consistent with increasing order of BIC values.

We use the prescription of Kass & Raftery 1995 to define the threshold of \( \Delta \text{BIC} \) and difference of log Bayesian evidences to estimate the evidence in favour of a GP regressor combination against other combinations. According to this prescription, for two models \( M_1 \) and \( M_0 \),

\[
\Delta \text{BIC} = \text{BIC}_1 - \text{BIC}_0 \geq 10 \text{ implies a strong evidence in favour of the model } M_0. \]

In terms of Bayesian evidences, \( \log_e Z_0 - \log_e Z_1 = \log_e \left( \frac{Z_0}{Z_1} \right) \geq 5 \) implies a strong evidence in favour of the model \( M_0 \) with log Bayesian evidence \( \log_e Z_0 \). The BIC values and \( \log_e Z \) for each GP regressor combination for both methods are shown in Tables 3.9 to 3.14. The GP regressor combinations in these tables are shown in decreasing order of \( \log_e Z \) which is broadly consistent with the increasing order of BIC.

Note that for each of the three cases mentioned in Section 3.4.2.3, during model comparison we neglect the regressor combinations for which one or more GP length scale hyperparameter (\( \eta_p \)) are unconstrained despite having higher Bayesian evidence. We check the posteriors for the corresponding combination sampled by dynesty and emcee for each \( \eta_p \) to confirm if they are constrained. We also confirm that for models with \( \Delta \log_e Z \) less than our threshold of 5, we obtain consistent transit parameters among all models. In such a case of multiple equally good models, we choose the model with the least number of GP regressors in the combination.
Table 3.9: BIC calculated using the GP likelihood and log Bayesian evidences (log, Z) from dynesty for all possible combinations of GP regressors used to fit Target star light curve alone (top panel of the table labelled 'Target LC') using New:WLC and New:WLC;No_Comp, and to fit the Target divided by the Comparison star light curves (bottom panel 'Target/Comparison LC') using Conv1:WLC. The combinations are shown in decreasing order of log, Z which is broadly consistent with increasing order of BIC. The best GP regressor combination we choose for the three cases we fit the HAT-P-26 transit white light curves for in Section 3.4.2 - New:WLC, New:WLC;No_Comp, and Conv1:WLC, are highlighted in bold below with the corresponding case in brackets. The best fit transit and GP parameters for each of these cases are detailed in Table 3.4.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time, Comp, PSF (New:WLC)</td>
<td>-2752.45</td>
<td>1360.13</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-2745.64</td>
<td>1358.46</td>
</tr>
<tr>
<td></td>
<td>Comp, PSF, CRPA</td>
<td>-2751.09</td>
<td>1357.8</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, PSF, CRPA</td>
<td>-2746.95</td>
<td>1357.26</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, PSF</td>
<td>-2748.32</td>
<td>1352.35</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, CRPA</td>
<td>-2741.66</td>
<td>1352.11</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, PSF</td>
<td>-2742.03</td>
<td>1351.56</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass</td>
<td>-2716.01</td>
<td>1349.68</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, CRPA</td>
<td>-2710.99</td>
<td>1347.52</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass</td>
<td>-2724.78</td>
<td>1347.45</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, CRPA</td>
<td>-2712.1</td>
<td>1346.79</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, CRPA</td>
<td>-2720.65</td>
<td>1346.04</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, CRPA</td>
<td>-2715.72</td>
<td>1345.09</td>
</tr>
<tr>
<td></td>
<td>Comp, CRPA</td>
<td>-2717.5</td>
<td>1344.86</td>
</tr>
<tr>
<td></td>
<td>Comp, PSF</td>
<td>-2687.77</td>
<td>1329.57</td>
</tr>
<tr>
<td></td>
<td>Comp</td>
<td>-2630.46</td>
<td>1305.13</td>
</tr>
<tr>
<td></td>
<td>CRPA</td>
<td>-1920.39</td>
<td>961.67</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, CRPA</td>
<td>-1909.55</td>
<td>961.37</td>
</tr>
<tr>
<td></td>
<td>Airmass, CRPA</td>
<td>-1915.0</td>
<td>961.37</td>
</tr>
<tr>
<td></td>
<td>Time, CRPA</td>
<td>-1915.05</td>
<td>961.33</td>
</tr>
<tr>
<td></td>
<td>Time, PSF, CRPA</td>
<td>-1909.54</td>
<td>961.24</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, PSF, CRPA</td>
<td>-1907.83</td>
<td>960.82</td>
</tr>
<tr>
<td></td>
<td>PSF, CRPA</td>
<td>-1914.9</td>
<td>960.38</td>
</tr>
<tr>
<td></td>
<td>Airmass, PSF, CRPA</td>
<td>-1913.19</td>
<td>960.27</td>
</tr>
<tr>
<td></td>
<td>Time (New:WLC;No_Comp)</td>
<td>-1886.92</td>
<td>947.31</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass</td>
<td>-1883.75</td>
<td>946.45</td>
</tr>
<tr>
<td></td>
<td>Time, PSF</td>
<td>-1887.37</td>
<td>946.39</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, PSF</td>
<td>-1881.06</td>
<td>945.04</td>
</tr>
<tr>
<td></td>
<td>Airmass</td>
<td>-1741.57</td>
<td>867.93</td>
</tr>
<tr>
<td></td>
<td>Airmass, PSF</td>
<td>-1735.42</td>
<td>866.36</td>
</tr>
<tr>
<td></td>
<td>PSF</td>
<td>-1163.66</td>
<td>584.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Airmass</td>
<td>-2751.16</td>
<td>1366.1</td>
</tr>
<tr>
<td></td>
<td>Time, CRPA</td>
<td>-2744.1</td>
<td>1365.45</td>
</tr>
<tr>
<td></td>
<td>Time (Conv1:WLC)</td>
<td>-2749.47</td>
<td>1363.63</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass</td>
<td>-2750.97</td>
<td>1363.13</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, CRPA</td>
<td>-2745.71</td>
<td>1361.75</td>
</tr>
<tr>
<td></td>
<td>Airmass, CRPA</td>
<td>-2750.96</td>
<td>1360.75</td>
</tr>
<tr>
<td></td>
<td>CRPA</td>
<td>-2728.94</td>
<td>1353.64</td>
</tr>
</tbody>
</table>
### Table 3.10: Same as Table 3.9 but for observation 2.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, Comp, Airmass, PSF</td>
<td>-7317.6</td>
<td>3644.81</td>
<td></td>
</tr>
<tr>
<td><strong>Time, PSF (New:WLC;No Comp)</strong></td>
<td><strong>-7305.5</strong></td>
<td><strong>3644.04</strong></td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-7315.53</td>
<td>3643.67</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF, CRPA</td>
<td>-7307.65</td>
<td>3643.45</td>
<td></td>
</tr>
<tr>
<td>Time, PSF, CRPA</td>
<td>-7298.95</td>
<td>3643.43</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF</td>
<td>-7313.89</td>
<td>3643.35</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF, CRPA</td>
<td>-7311.87</td>
<td>3642.73</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF</td>
<td>-7315.44</td>
<td>3642.21</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF</td>
<td>-7316.14</td>
<td>3641.99</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF, CRPA</td>
<td>-7310.97</td>
<td>3641.41</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF, CRPA</td>
<td>-7309.27</td>
<td>3641.0</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF</td>
<td>-7262.97</td>
<td>3610.49</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, CRPA</td>
<td>-7213.6</td>
<td>3601.23</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF, CRPA</td>
<td>-7220.49</td>
<td>3599.57</td>
<td></td>
</tr>
<tr>
<td><strong>Time, Comp (New:WLC)</strong></td>
<td><strong>-7220.01</strong></td>
<td><strong>3599.03</strong></td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, CRPA</td>
<td>-7223.57</td>
<td>3598.4</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass</td>
<td>-7220.62</td>
<td>3598.38</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, CRPA</td>
<td>-7217.23</td>
<td>3598.13</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass</td>
<td>-7211.17</td>
<td>3595.11</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-7202.29</td>
<td>3593.22</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-7192.99</td>
<td>3592.55</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass</td>
<td>-7197.75</td>
<td>3592.31</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-7186.79</td>
<td>3592.02</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-7195.98</td>
<td>3591.63</td>
<td></td>
</tr>
<tr>
<td>Comp, CRPA</td>
<td>-7159.79</td>
<td>3572.72</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-7157.87</td>
<td>3567.11</td>
<td></td>
</tr>
<tr>
<td>PSF, CRPA</td>
<td>-7102.58</td>
<td>3546.44</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-7032.34</td>
<td>3515.26</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>-6491.33</td>
<td>3242.77</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF</td>
<td>-6491.87</td>
<td>3242.14</td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-3717.88</td>
<td>1884.53</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, CRPA</td>
<td>-6823.07</td>
<td>3409.41</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-6816.64</td>
<td>3408.92</td>
<td></td>
</tr>
<tr>
<td><strong>Time (Conv1:WLC)</strong></td>
<td><strong>-6829.35</strong></td>
<td><strong>3407.31</strong></td>
<td></td>
</tr>
<tr>
<td>Time, Airmass</td>
<td>-6823.18</td>
<td>3401.17</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-6811.13</td>
<td>3392.76</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-6774.79</td>
<td>3381.19</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-6116.87</td>
<td>3059.6</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.11: Same as Table 3.9 but for observation 3.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time, Comp, PSF</td>
<td>-3885.21</td>
<td>1938.94</td>
</tr>
<tr>
<td></td>
<td>Comp, CRPA</td>
<td>-3891.73</td>
<td>1937.57</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, CRPA</td>
<td>-3881.97</td>
<td>1937.51</td>
</tr>
<tr>
<td><strong>Time, Comp (New:WLC)</strong></td>
<td><strong>-3890.89</strong></td>
<td><strong>1936.67</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time, Comp, CRPA</td>
<td>-3885.91</td>
<td>1936.38</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-3880.05</td>
<td>1936.18</td>
</tr>
<tr>
<td></td>
<td>Comp, PSF, CRPA</td>
<td>-3885.92</td>
<td>1934.91</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, CRPA</td>
<td>-3885.92</td>
<td>1934.88</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass</td>
<td>-3885.14</td>
<td>1933.38</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, PSF, CRPA</td>
<td>-3874.19</td>
<td>1932.66</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, Airmass, PSF</td>
<td>-3879.2</td>
<td>1932.05</td>
</tr>
<tr>
<td></td>
<td>Time, Comp, PSF, CRPA</td>
<td>-3880.22</td>
<td>1931.76</td>
</tr>
<tr>
<td></td>
<td>Comp</td>
<td>-3867.11</td>
<td>1924.29</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass</td>
<td>-3860.11</td>
<td>1924.24</td>
</tr>
<tr>
<td></td>
<td>Comp, Airmass, PSF</td>
<td>-3855.28</td>
<td>1922.76</td>
</tr>
<tr>
<td></td>
<td>Comp, PSF</td>
<td>-3861.72</td>
<td>1921.71</td>
</tr>
<tr>
<td></td>
<td><strong>Airmass (New:WLC;No_Comp)</strong></td>
<td><strong>-3098.4</strong></td>
<td><strong>1548.87</strong></td>
</tr>
<tr>
<td></td>
<td>Airmass, PSF</td>
<td>-3092.41</td>
<td>1548.68</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass</td>
<td>-3093.0</td>
<td>1546.51</td>
</tr>
<tr>
<td></td>
<td>Airmass, PSF, CRPA</td>
<td>-3086.93</td>
<td>1546.41</td>
</tr>
<tr>
<td></td>
<td>Airmass, CRPA</td>
<td>-3092.73</td>
<td>1545.9</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, CRPA</td>
<td>-3086.94</td>
<td>1545.77</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, PSF</td>
<td>-3087.04</td>
<td>1545.47</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, PSF, CRPA</td>
<td>-3080.92</td>
<td>1545.39</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>-3074.6</td>
<td>1537.76</td>
</tr>
<tr>
<td></td>
<td>Time, PSF</td>
<td>-3068.33</td>
<td>1536.61</td>
</tr>
<tr>
<td></td>
<td>Time, CRPA</td>
<td>-3068.48</td>
<td>1536.52</td>
</tr>
<tr>
<td></td>
<td>Time, PSF, CRPA</td>
<td>-3026.42</td>
<td>1535.77</td>
</tr>
<tr>
<td></td>
<td>CRPA</td>
<td>-3038.38</td>
<td>1520.9</td>
</tr>
<tr>
<td></td>
<td>PSF, CRPA</td>
<td>-3032.53</td>
<td>1520.04</td>
</tr>
<tr>
<td></td>
<td>PSF</td>
<td>-1541.36</td>
<td>797.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CRPA</td>
<td>-3914.13</td>
<td>1951.6</td>
</tr>
<tr>
<td><strong>Time (Conv1:WLC)</strong></td>
<td><strong>-3910.8</strong></td>
<td><strong>1950.74</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Airmass, CRPA</td>
<td>-3908.2</td>
<td>1950.73</td>
</tr>
<tr>
<td></td>
<td>Time, CRPA</td>
<td>-3908.19</td>
<td>1949.33</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass</td>
<td>-3904.99</td>
<td>1948.04</td>
</tr>
<tr>
<td></td>
<td>Time, Airmass, CRPA</td>
<td>-3902.58</td>
<td>1946.02</td>
</tr>
<tr>
<td></td>
<td>Airmass</td>
<td>-3876.99</td>
<td>1929.5</td>
</tr>
</tbody>
</table>
### Table 3.12: Same as Table 3.9 but for observation 6.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log.(Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, Comp (New:WLC)</td>
<td>-3470.86</td>
<td>1729.74</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, CRPA</td>
<td>-3463.71</td>
<td>1729.59</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF</td>
<td>-3464.93</td>
<td>1729.01</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF</td>
<td>-3465.22</td>
<td>1728.93</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, CRPA</td>
<td>-3465.2</td>
<td>1728.82</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass</td>
<td>-3469.46</td>
<td>1728.36</td>
<td></td>
</tr>
<tr>
<td>Time, PSF</td>
<td>-3458.14</td>
<td>1728.31</td>
<td></td>
</tr>
<tr>
<td>Time (New:WLC;No Comp)</td>
<td>-3462.02</td>
<td>1728.25</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF, CRPA</td>
<td>-3450.06</td>
<td>1727.93</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-3456.38</td>
<td>1727.9</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass</td>
<td>-3458.23</td>
<td>1727.72</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF, CRPA</td>
<td>-3459.56</td>
<td>1727.66</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF</td>
<td>-3456.33</td>
<td>1726.52</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-3464.58</td>
<td>1725.28</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-3458.3</td>
<td>1724.88</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass</td>
<td>-3465.22</td>
<td>1724.86</td>
<td></td>
</tr>
<tr>
<td>Time, PSF, CRPA</td>
<td>-3451.36</td>
<td>1724.82</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF, CRPA</td>
<td>-3456.38</td>
<td>1723.2</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, CRPA</td>
<td>-3469.22</td>
<td>1722.99</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-3451.81</td>
<td>1721.35</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF, CRPA</td>
<td>-3459.28</td>
<td>1721.35</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF</td>
<td>-3464.72</td>
<td>1720.7</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF, CRPA</td>
<td>-3449.66</td>
<td>1720.28</td>
<td></td>
</tr>
<tr>
<td>Comp, CRPA</td>
<td>-3455.24</td>
<td>1720.15</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-3440.12</td>
<td>1717.28</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF</td>
<td>-3451.3</td>
<td>1717.02</td>
<td></td>
</tr>
<tr>
<td>PSF, CRPA</td>
<td>-3434.29</td>
<td>1716.03</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-3436.46</td>
<td>1713.5</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF</td>
<td>-2049.26</td>
<td>1067.31</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>-2049.62</td>
<td>1025.21</td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-2003.18</td>
<td>1009.93</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log.(Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, Airmass</td>
<td>-3457.24</td>
<td>1728.58</td>
<td></td>
</tr>
<tr>
<td>Time (Conv1:WLC)</td>
<td>-3462.83</td>
<td>1728.26</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-3451.18</td>
<td>1728.22</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-3457.09</td>
<td>1724.8</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-3453.4</td>
<td>1724.07</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-3456.81</td>
<td>1723.7</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-3357.8</td>
<td>1694.99</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.13: Same as Table 3.9 but for observation 4.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, Comp, Airmass, PSF</td>
<td>-2192.71</td>
<td>1082.94</td>
<td></td>
</tr>
<tr>
<td><strong>Time, Comp, PSF (New:WLC)</strong></td>
<td><strong>-2197.27</strong></td>
<td><strong>1082.68</strong></td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF, CRPA</td>
<td>-2192.41</td>
<td>1079.94</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF, CRPA</td>
<td>-2181.83</td>
<td>1076.69</td>
<td></td>
</tr>
<tr>
<td>Time, Comp</td>
<td>-2164.5</td>
<td>1071.75</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, CRPA</td>
<td>-2159.48</td>
<td>1071.53</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass</td>
<td>-2159.49</td>
<td>1071.45</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, CRPA</td>
<td>-2154.54</td>
<td>1069.72</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-2186.96</td>
<td>1069.65</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, CRPA</td>
<td>-2155.06</td>
<td>1056.87</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF</td>
<td>-2104.68</td>
<td>1048.43</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF, CRPA</td>
<td>-2117.88</td>
<td>1048.11</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF</td>
<td>-2117.39</td>
<td>1045.55</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF</td>
<td>-2133.95</td>
<td>1045.51</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF, CRPA</td>
<td>-2106.45</td>
<td>1043.38</td>
<td></td>
</tr>
<tr>
<td><strong>Time, PSF (New:WLC;No Comp)</strong></td>
<td><strong>-2098.13</strong></td>
<td><strong>1041.09</strong></td>
<td></td>
</tr>
<tr>
<td>Time, PSF, CRPA</td>
<td>-2092.94</td>
<td>1040.37</td>
<td></td>
</tr>
<tr>
<td>Comp, CRPA</td>
<td>-2093.71</td>
<td>1039.24</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF, CRPA</td>
<td>-2098.9</td>
<td>1037.5</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass</td>
<td>-2007.12</td>
<td>988.35</td>
<td></td>
</tr>
<tr>
<td>PSF, CRPA</td>
<td>-1879.64</td>
<td>934.87</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-1865.92</td>
<td>932.78</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-1864.82</td>
<td>931.93</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass</td>
<td>-1860.93</td>
<td>931.8</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-1862.81</td>
<td>931.07</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-1869.76</td>
<td>927.73</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-1788.32</td>
<td>892.26</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-1776.6</td>
<td>888.34</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF</td>
<td>-957.97</td>
<td>479.72</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>-903.7</td>
<td>455.17</td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-727.93</td>
<td>440.63</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log (Z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-2081.55</td>
<td>1033.03</td>
<td></td>
</tr>
<tr>
<td><strong>Time, Airmass (Conv1:WLC)</strong></td>
<td><strong>-2086.72</strong></td>
<td><strong>1032.28</strong></td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-2084.06</td>
<td>1030.49</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-2064.95</td>
<td>1029.94</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-2059.88</td>
<td>1028.8</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-1757.82</td>
<td>886.95</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-1548.24</td>
<td>769.83</td>
<td></td>
</tr>
</tbody>
</table>
### Table 3.14: Same as Table 3.9 but for observation 5.

<table>
<thead>
<tr>
<th>Target LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time, Comp (New:WLC)</strong></td>
<td>-1705.44</td>
<td>839.69</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, CRPA</td>
<td>-1700.34</td>
<td>839.08</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass</td>
<td>-1702.14</td>
<td>838.48</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF</td>
<td>-1700.46</td>
<td>838.41</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF</td>
<td>-1705.33</td>
<td>837.62</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, PSF, CRPA</td>
<td>-1700.09</td>
<td>837.25</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, Airmass, CRPA</td>
<td>-1698.0</td>
<td>835.85</td>
<td></td>
</tr>
<tr>
<td>Time, Comp, PSF, CRPA</td>
<td>-1695.61</td>
<td>835.45</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF, CRPA</td>
<td>-1702.25</td>
<td>830.41</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, CRPA</td>
<td>-1698.15</td>
<td>827.49</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, PSF, CRPA</td>
<td>-1645.89</td>
<td>806.17</td>
<td></td>
</tr>
<tr>
<td><strong>Time, Airmass, PSF (New:WLC;No Comp)</strong></td>
<td>-1650.67</td>
<td>805.84</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF, CRPA</td>
<td>-1646.15</td>
<td>805.54</td>
<td></td>
</tr>
<tr>
<td>Time, PSF</td>
<td>-1619.75</td>
<td>797.54</td>
<td></td>
</tr>
<tr>
<td>Comp, CRPA</td>
<td>-1622.41</td>
<td>797.52</td>
<td></td>
</tr>
<tr>
<td>Time, PSF, CRPA</td>
<td>-1614.88</td>
<td>797.38</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF, CRPA</td>
<td>-1617.76</td>
<td>796.39</td>
<td></td>
</tr>
<tr>
<td>Airmass, PSF</td>
<td>-1621.21</td>
<td>796.22</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass, PSF</td>
<td>-1616.27</td>
<td>794.23</td>
<td></td>
</tr>
<tr>
<td>Comp, Airmass</td>
<td>-1566.87</td>
<td>779.17</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-1489.8</td>
<td>739.88</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass</td>
<td>-1492.2</td>
<td>739.64</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-1491.95</td>
<td>738.85</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-1487.15</td>
<td>736.49</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-1477.59</td>
<td>735.4</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-1472.75</td>
<td>735.0</td>
<td></td>
</tr>
<tr>
<td>PSF, CRPA</td>
<td>-1332.67</td>
<td>657.14</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-1253.85</td>
<td>632.24</td>
<td></td>
</tr>
<tr>
<td>Comp, PSF</td>
<td>-603.51</td>
<td>376.41</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>-607.97</td>
<td>309.19</td>
<td></td>
</tr>
<tr>
<td>PSF</td>
<td>-610.43</td>
<td>309.04</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target/Comparison LC</th>
<th>GP regressors</th>
<th>GP BIC</th>
<th>log, Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time, Airmass (Conv1:WLC)</strong></td>
<td>-1641.27</td>
<td>807.24</td>
<td></td>
</tr>
<tr>
<td>Airmass, CRPA</td>
<td>-1635.99</td>
<td>805.33</td>
<td></td>
</tr>
<tr>
<td>Time, Airmass, CRPA</td>
<td>-1636.33</td>
<td>804.86</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-1602.97</td>
<td>796.05</td>
<td></td>
</tr>
<tr>
<td>Time, CRPA</td>
<td>-1598.17</td>
<td>795.92</td>
<td></td>
</tr>
<tr>
<td>Airmass</td>
<td>-1462.59</td>
<td>745.79</td>
<td></td>
</tr>
<tr>
<td>CRPA</td>
<td>-1199.12</td>
<td>705.79</td>
<td></td>
</tr>
</tbody>
</table>
A new method to correct for host star variability in multi-epoch observations of exoplanet transmission spectra

Vatsal Panwar, Jean-Michel Désert, Kamen O. Todorov, Jacob L. Bean, Kevin B. Stevenson, C. M. Huitson, Jonathan J. Fortney, Marcel Bergmann


**Abstract**

Transmission spectra of exoplanets orbiting active stars suffer from wavelength-dependent effects due to stellar photospheric heterogeneity. WASP-19b, an ultra-hot Jupiter ($T_{\text{eq}} \sim 2100$ K), is one such strongly irradiated gas-giant orbiting an active solar-type star. We present optical (520-900 nm) transmission spectra of WASP-19b obtained across eight epochs using the Gemini Multi-Object Spectrograph (GMOS) on the Gemini-South telescope. We apply our recently developed Gaussian Processes regression based method to model the transit light curve systematics and extract the transmission spectrum at each epoch. We find that WASP-19b’s transmission spectrum is affected by stellar variability at individual epochs. We report an observed anticorrelation between the relative slopes and offsets of the spectra across all epochs. This anticorrelation is consistent with the predictions from the forward transmission models, which account for the effect of unocculted stellar spots and faculae measured previously for WASP-19. We introduce a new method to correct for this stellar variability effect at each epoch by using the observed correlation between the transmission spectral slopes and offsets. We compare our stellar variability corrected GMOS transmission spectrum with previous contradicting MOS measurements for WASP-19b and attempt to reconcile them. We also measure the amplitude and timescale of broadband stellar variability of WASP-19 from TESS photometry, which we find to be consistent with the effect
observed in GMOS spectroscopy and ground-based broadband photometric long-term monitoring. Our results ultimately caution against combining multi-epoch optical transmission spectra of exoplanets orbiting active stars before correcting each epoch for stellar variability.

4.1 Introduction

Ground-based observations using low resolution multi-object spectroscopy (hereafter referred to as MOS) on large telescopes (Bean et al. 2010, 2011) have yielded precise optical and near-infrared transmission spectra which have helped to constrain the atmospheric properties of exoplanets ranging from transiting hot gas giants (e.g. Nikolov et al. 2018) to smaller and cooler rocky exoplanets (e.g. Diamond-Lowe et al. 2018). Conventionally, the ground-based MOS technique has been restricted to exoplanets transiting host stars with comparison stars of similar brightness and spectral type in the instrument’s field of view, which can be used for differential spectrophotometry. However, recent development in the techniques of modelling telluric and instrumental systematics in this context (Panwar et al. 2022b, hereafter referred to as P22) have also extended the application of the MOS technique to exoplanets orbiting host stars, including bright stars targeted by TESS planet detection campaigns, with no suitable nearby comparison stars.

A long-standing issue in transit spectroscopy of exoplanets has been the contamination of the planetary spectrum due to stellar variability stemming from stellar photospheric heterogeneity. The amplitude of such contamination can be comparable to the desired precision of the transmission spectra (e.g. Rackham et al. (2017a)). The level of contamination is particularly significant for active solar type stars observed in the optical wavelength range typically probed by ground-based MOS observations using instruments like VLT/FORS2 or Gemini/GMOS. This wavelength dependent effect (Pont et al. 2008, McCullough et al. 2014) stems from unocculted or occulted magnetically active regions like spots and faculae in the stellar photosphere and has recently come to be commonly referred to as the transit light source effect (Rackham et al. 2018). Several observations and in-depth modelling (Rackham et al., 2018, 2019) have revealed this wavelength-dependent effect as imprinted on the transmission spectra of exoplanets orbiting active stars (e.g. Espinoza et al. 2019, Kirk et al. 2021, Sedaghati et al. 2021, Nikolov et al. 2021).

The transit light source effect has been observed and modelled in the MOS observations of many transiting exoplanets in recent years. The framework of (Rackham et al., 2018, 2019) was first implemented in a Bayesian atmospheric retrieval code AURA by Pinhas et al. (2018) and was recently also used by Nikolov et al. (2021) to model the transit light source effect in the transmission spectrum of WASP-110b. Other Bayesian retrieval codes like POSEIDON introduced by MacDonald & Madhusudhan (2017) (e.g. applied to WASP-103b observations in Kirk et al. 2021) and platon (e.g. applied to the VLT/FORS2 observations of WASP-19b in Sedaghati et al. (2021)) also fit for the stellar photospheric heterogeneity parameters together with the planetary atmosphere.

WASP-19b (Hebb et al. 2010) is one such example of a transiting gas giant exoplanet orbiting an active G dwarf with significant stellar variability. Active FGK dwarfs have been known to
produce prominent features in the transmission spectra (Rackham et al. 2019) due to stellar activity. Hence, it is pertinent to account for the effect of stellar activity when studying the atmosphere of WASP-19b. WASP-19b also falls in the class of ultra-hot Jupiters \((T_{\text{eq}} \gtrsim 2000\, \text{K})\); e.g., Arcangeli et al. 2018, Lothringer et al. 2018, Kitzmann et al. 2018), which have recently been the subject of atmospheric characterization in the optical through low resolution MOS (e.g. Stevenson et al. 2014, Wilson et al. 2021) and high resolution spectroscopy (e.g. Hoeijmakers et al. 2019, Pino et al. 2020, Ehrenreich et al. 2020).

Studies presenting discrepant optical transmission spectrum of WASP-19b have been published recently using VLT/FORS2 (Sedaghati et al. 2017), VLT/ESPRESSO (Sedaghati et al. 2021), and Magellan/IMACS by Espinoza et al. (2019). This motivated us to follow up the system using Gemini/GMOS.

The paper is distributed as follows. We first review the state of the art in atmospheric studies of WASP-19b in Section 4.2. In Section 4.3 we describe our observations of WASP-19b from Gemini/GMOS. In Section 4.4 we describe our data reduction of these observations and in Section 4.5 we discuss the analysis of transit light curves to obtain the transmission spectrum. In Section 4.6 we discuss the interpretation of the transmission spectrum, especially in the context of the host star’s activity. We also compare the results from GMOS observations with forward transmission spectrum models accounting for the effect of stellar variability. We introduce a new empirical approach to correct for the effect of stellar variability in the transmission spectrum at individual epochs before constructing the final combined transmission spectrum. We discuss the implications for the atmosphere of WASP-19b from the combined GMOS and HST/WFC3 transmission spectrum in Section 4.6.3. We further put in context the effect of stellar variability observed in the broadband transit depths measured from TESS photometry of 58 transits of WASP-19b observed over two sectors. We describe our analysis of TESS and ground based photometric follow-up from Las Cumbres Observatory Global to monitor the stellar variability of WASP-19 in the Appendix 4.A. Specifically, we use the long-term photometry of the system from TESS covering several transits to understand the effect of stellar variability on the broadband optical transit depth, and compare it with the relative variations seen between the GMOS transmission spectra at multiple epochs. In Section 4.7 we present our conclusions.

### 4.2 The case of WASP-19b

WASP-19b, one of the shortest period Jupiter mass giant exoplanets known (orbiting a G8V star in just 18.9 hours), is situated in the ”sub-Jupiter” desert in the mass vs orbital period distribution of the population of hot Jupiters which shows a pileup around orbital period of 3 – 4 days (Szabó & Kiss 2011, Hellier et al. 2011). It is also an ideal candidate for atmospheric characterization on multiple accounts. With the high level of stellar irradiation and resultant equilibrium temperature of 2100 K, and low surface gravity \((\log_{10}g [\text{m/s}^2] = 2.15)\), WASP-19b is expected to have TiO and VO at gas phase equilibrium in the upper atmosphere that, if present in a cloud free atmosphere, will absorb the incident optical stellar flux and could cause thermal inversion (Hubeny et al. 2003, Fortney et al. 2008). The host
star WASP-19 is also known to be active, with the optical stellar flux varying peak to trough 2 to 3 % at a period of $\sim 10.5$ days (Hebb et al. 2010, Huitson et al. 2013, Espinoza et al. 2019). The chromospheric Ca II H & K line emission ratio of WASP-19 quantified by $\log(R'_{HK}) = -4.5 \pm 0.03$ (Anderson et al. 2013, Knutson et al. 2010) quantifies the high level of chromospheric activity of the star. Table 4.1 shows the properties of the host star WASP-19 from the literature.

With a dayside temperature of $2240 \pm 40$ K (inferred from TESS and previous secondary eclipse depth measurements, Wong et al. 2020), WASP-19b is on the cusp of transition of hot to ultra-hot Jupiters (Parmentier et al. 2018, Baxter et al. 2020), at which point atmospheric opacities, molecular dissociation, H$-$ opacity, latent heat and thermal inversion begin to become relevant (Arcangeli et al. 2018, Lothringer et al. 2018, Kitzmann et al. 2018). Retrieval analysis of emission spectra including secondary eclipse depth measurements from Spitzer and TESS secondary eclipse observations (Wong et al. 2016, 2020) indicate an atmosphere with no dayside thermal inversion and moderately efficient day-night circulation. However, in contrast to these findings, Rajpurohit et al. (2020) interpret the excess eclipse depth in the Spitzer 4.5 $\mu$m band as due to CO in emission and thus as an evidence of thermal inversion in the atmosphere of WASP-19b.

Using transmission spectroscopy of WASP-19b, Huitson et al. (2013) have detected absorption features due to water in the 1.1-1.7 $\mu$m range HST/WFC3 G141 observations, which is consistent with a solar abundance atmosphere with no or only low level of clouds. There is evidence that high levels of UV flux from active stars could be responsible for the dissociation of molecular absorbers like TiO (Knutson et al. 2010). Huitson et al. (2013) hypothesize this to be one of the possible reasons behind non-detection of TiO in their HST/STIS optical transmission spectrum. The presence or absence of TiO in the atmosphere can affect the overall energy budget of WASP-19b, drives thermal inversion in the atmosphere, and ultimately affects the inferences about the atmospheric metallicity and C/O which hold potential clues to the formation and evolution history of gas giants (Madhusudhan 2012, Mordasini et al. 2016, Eistrup et al. 2018).

The picture in the optical wavelength range of the transmission spectrum of WASP-19 is mired with a discrepancy due to two different studies reporting contrasting results. Sedaghati et al. (2017) from their observations obtained using VLT/FORS2 first reported the detection of TiO features in the optical transmission spectrum with a strong scattering slope due to hazes towards the blue end and a water feature towards the red end at high significance. However, Espinoza et al. (2019) detect a featureless optical transmission spectrum from their observations using Magellan/IMACS, with no significant TiO features and no slope due to hazes. This is consistent with the picture apparent from low resolution optical transmission spectrum from HST/STIS reported by Huitson et al. (2013). Sedaghati et al. (2021) use high resolution spectroscopic observations from VLT/ESPRESSO to search for signatures of atomic and molecular species in the optical via cross-correlation analysis and report a tentative indication of TiO at $\sim 3 \sigma$ confidence. Through chromatic, Rossiter-McLaughlin effect analysis Sedaghati et al. (2021) also report a strong scattering slope towards the blue wave-
4.2 The case of WASP-19b

lengths, consistent with the findings of Sedaghati et al. (2017) at low-resolution and in contrast with the flat spectrum presented by Espinoza et al. (2019).

Activity and variability of the host star WASP-19 contaminates the transmission spectrum of the planet via the transit light source effect (Rackham et al. 2017b). Espinoza et al. (2019) observed occultations of stellar spots and plages and used them to put constraints on the spot size and spot temperature contrast with respect to the stellar photosphere. Interestingly, the transmission spectrum from one of the six epochs analysed by Espinoza et al. (2019) shows a significantly steeper slope compared to those from other epochs due to stellar activity. Espinoza et al. (2019) perform retrievals accounting for stellar activity on the transmission spectra from all epochs independently. They find that the epoch showing a steep slope can be best explained by strong stellar contributions from stellar activity. However, all the other five epochs show no statistically significant contribution from stellar activity contamination and are most consistent with a flat line. Espinoza et al. (2019) eventually reject the spectrum with steep slope when they construct the combined transmission spectrum from the mean subtracted transmission spectra of the other five epochs. They also do not apply any additional slope corrections to the individual spectra before combining them.

Sedaghati et al. (2021) in their reanalysis of the VLT/FORS2 observations of Sedaghati et al. (2017) analyse the effect of stellar surface heterogeneity on WASP-19b’s transmission spectrum through a POSEIDON (MacDonald & Madhusudhan 2017) retrieval analysis of the transmission spectra from the three epochs. Each VLT/FORS2 epoch was observed in a different wavelength range, going from blue to red optical. Sedaghati et al. (2021) from their retrieval analysis find that the VLT/FORS2 spectrum is best explained by an atmosphere with $100\times$ sub-solar TiO. They also find that after accounting for stellar activity, the significance of TiO detection in the VLT/FORS2 spectrum goes from $7.7\sigma$ to $4.7\sigma$. The stellar spot contrast and covering fractions retrieved by Sedaghati et al. (2021) from their VLT/FORS2 spectrum are consistent with those measured by Espinoza et al. (2019) from their Magellan/IMACS spectrum. Additionally, Sedaghati et al. (2021) also perform a retrieval on the Magellan/IMACS combined transmission spectrum from Espinoza et al. (2019) and find a marginal preference for the model with TiO ($\Delta \ln Z = 0.5$) compared to a flat line or models with only contributions from stellar activity.

In summary, both FORS2 and IMACS spectra have confirmed the significant effect of stellar activity in the transmission spectrum of WASP-19b. Both spectra have different morphologies and an agreement between them still remains at a marginal threshold as indicated by the retrieval of the IMACS spectrum by Sedaghati et al. (2021). This tension in the observations of WASP-19b’s atmosphere, including the presence or absence of TiO, motivated us to further investigate its optical transmission spectrum, which we present in this paper. In this paper, we present a study of WASP-19b’s transmission spectrum from 8 epochs observed using Gemini/GMOS in the wavelength range of 520 to 900 nm. We present a new approach to analyse and correct the effect of stellar variability at each epoch by looking at its two broad manifestations: the slope and the offset of the transmission spectrum. The new data analysis method introduced in P22 mitigates potential systematics due to non-linear differences
Multi-epoch stellar variability effect in the transmission spectrum of WASP-19b

Table 4.1: Stellar parameters of WASP-19.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M⋆ (M⊙)</td>
<td>0.904 ± 0.04</td>
<td>Tregloan-Reed et al. (2013)</td>
</tr>
<tr>
<td>R⋆ (R⊙)</td>
<td>1.001 ± 0.035</td>
<td>Gaia Collaboration et al. (2018)</td>
</tr>
<tr>
<td>Protation (days)</td>
<td>10.50 ± 0.2</td>
<td>Bonomo et al. (2017)</td>
</tr>
<tr>
<td>V⋆ (mag)</td>
<td>12.31 ± 0.04</td>
<td>Zacharias et al. (2013)</td>
</tr>
<tr>
<td>Teff,⋆ (K)</td>
<td>5460 ± 90</td>
<td>Doyle et al. (2013)</td>
</tr>
<tr>
<td>SpT⋆</td>
<td>G8V</td>
<td>Hebb et al. (2010)</td>
</tr>
<tr>
<td>L⋆ (log10L⊙)</td>
<td>−0.09 ± 0.005</td>
<td>Gaia Collaboration et al. (2018)</td>
</tr>
<tr>
<td>log(g⋆)</td>
<td>4.45 ± 0.05</td>
<td>Torres et al. (2012)</td>
</tr>
<tr>
<td>[Fe/H]⋆</td>
<td>0.15 ± 0.07</td>
<td>Torres et al. (2012)</td>
</tr>
<tr>
<td>Distance (pc)</td>
<td>270.41 ± 1.46</td>
<td>Gaia Collaboration et al. (2018)</td>
</tr>
<tr>
<td>log(R′HK)</td>
<td>−4.5 ± 0.03</td>
<td>Anderson et al. (2013)</td>
</tr>
</tbody>
</table>

between the target and comparison star light curves and enables accurate measurement of the slopes and offsets of the transmission spectrum at each epoch.

4.3 Multi-epoch transit observations of WASP-19b

4.3.1 Gemini/GMOS transit observations of WASP-19b

We observed eight transits of WASP-19b (Table 4.2) in the red optical using GMOS on the Gemini South telescope located at Cerro Pachon, Chile. Since the host star WASP-19 is known to be active (e.g. spot crossing events seen in the observations by Espinoza et al. 2019), we spread the observations over a period of 2 years. All 8 transits were observed as part of a survey program of hot Jupiter atmospheres from Gemini/GMOS (Proposal ID: 2012B-0398; PI: J.-M Désert) and described in more detail in Huitson et al. (2017) (referred to as H17 hereafter). The observations were performed using the same set-up as described in H17, which is similar to that of previous exoplanet atmospheric observations using GMOS (e.g. Bean et al. 2010, 2011; Gibson et al. 2013). For each observation, we used the multi-object spectroscopy mode of GMOS-South to obtain time series spectrophotometry of WASP-19 and 2 nearby comparison stars (described in more detail below) simultaneously. All the 8 transits were observed in the red optical using R150 + G5306 grating combination, covering a wavelength range of 525-900 nm with an ideal resolving power R ∼ 600. The ideal resolving powers assume a slit width of 0.5 arcsec. We used masks with 10 arcsec wide slits on each star, and obtained a seeing-limited spectral resolution. Given the range of seeing measured during our observations (Table 4.2) our resolution is approximately 2 − 3 × lower than the ideal value depending on observation.

For all the observations, we used the grating in first order. The requested central wavelength was 620 nm, and we used the OG515_G0330 filter to block light blueward to 515 nm. The blocking filter was used to avoid contamination from light from higher orders. For all observations, we windowed the regions of interest (ROI) on the detector in the cross-dispersion direction to reduce the readout time and improve the duty cycle. We used one ROI for each
4.4 Data Reduction

4.4.1 Data reduction of the Gemini/GMOS observations

We used our custom pipeline designed for reducing the GMOS data, the steps for which are described in more detail in \( H17 \). We extract the 1D spectra and apply corrections for addi-

\[ \text{https://www.gemini.edu/sciops/instruments/gmos/imaging/detector-array/gmosn-array-hamamatsu?q=node/10004} \]


### 4.4.2 Flat-Fielding

We acquired 100 flat frames for transits 1 through 7 and 200 flat frames for transit 8 with median count levels of $\sim 10,000$ ADU. For transits 1 through 7, both the flat field and science frames show fringing at the 10% amplitude. We construct a master flat by median combining the series of flats for each observation. We found that the scatter in the transit light curves redward of 700 nm was $10 \times$ photon noise without flat-fielding, which is marginally higher when performing flat-fielding. On inspection of the frames, we found that noise is added by flat-fielding because the phase, period and amplitude of the fringe pattern are significantly different between the flat fields and the science frames. The fringe pattern in the science frames also changes by several times the photon noise during each transit observation. For transit 8, the fringe amplitude is an order of magnitude lower than in the other transits. However, flat-fielding still increased the scatter redward of 700 nm by 10-20%. We attribute this to low-levels of fringing still being present in the transit 8 observations taken with the new detector. Moreover, flat-fielding should not be a major issue since we measure the transit depth for the same set of pixels relative to time. However, changes in the gravity vector of the instrument due to changing pointing through the night can cause the spectral trace to drift to different sets of detector pixels during the observation. We tested our extraction with and without flat-fielding and find that flat-fielding does not significantly affect the scatter of the resulting transit light curves. We found that the flat fielding changed the light curve scatter on average by 40 ppm across all 8 transit light curves, which is about 10 times lower than the typical photon noise for a GMOS transit light curve of WASP-19b. For this reason, and since flat-fielding did not improve the scatter blueward of 700 nm, we chose not to perform...
Data Reduction

4.4.3 Spectral extraction

We follow the steps outlined in more detail in H17 to detect and mask cosmic ray hits and bad pixel columns (mainly due to shifted charge) from all science frames. We then subtracted the background while performing the optimal extraction (Horne 1986), and found that subtracting the median value in each cross-dispersion column provided the best fit to the background fluxes compared with performing fits to the flux profile in each cross-dispersion column. We also found that the precision in the light curves was $\sim 2 \times$ better when using the median value for each column rather than a fit.

To test the degree to which background subtraction affects our resultant transmission spectra, we also extracted spectra in which the background subtraction was multiplied by 10. We found that all $R_p/R_\star$ values in the final transmission spectra deviated by much less than 1-$\sigma$ between the two cases, indicating that the final results are robust to uncertainties in background subtraction. The background flux level was 1.4-10 % of the stellar flux for WASP-19, 5-25 % of the stellar flux for comparison star 1 and 0.5-3 % of the stellar flux for the comparison star 2.

We additionally also extract the average PSF width of the spectral trace, which we use later for transit light curve modelling in Section 4.5.2. For each exposure, we first bin the 2D spectral trace in the raw science frames at an arbitrary interval of 10 pixels in the dispersion direction. To this binned spectral trace for each column, we then fit Gaussians along the cross-dispersion. We then take the average best fit FWHM of all the Gaussians to obtain the PSF width for each exposure.

4.4.4 Wavelength calibration

After spectral extraction, we performed wavelength calibration using CuAr lamp spectra taken on the same day as each science observation, following similar steps as described in H17 and P22. The final uncertainties in the estimated wavelength solution are approximately 1 nm for all observations, which is $\sim$ 5% of the size of wavelength bins (20 nm) we use in the final transmission spectrum for all transits in Section 4.5.3. This level of uncertainty in the wavelength solution is smaller than our resolution element ($\sim$ 4 nm in R150 e2v detectors) and is insufficient to cause systematic effects in the final wavelength-dependent light curves.

4.4.4.1 Dispersion-direction shifts of the stellar spectra

The wavelength solution for GMOS data is known to shift and stretch with time because of the absence of atmospheric dispersion compensator (ADC). These shifts and stretches vary
Figure 4.1: One of the extracted 1D spectra for target and comparison stars for GMOS-R150 observation 1 and observation 8 of WASP-19b. The two comparison star spectra shown here are both for observation 1 to illustrate the relative brightness difference between the target and comparison stars. All spectra were extracted at a similar airmass and normalized by the exposure time. Prominent stellar and telluric features which we use for measuring the shifts in spectral trace with time are labelled. It can be seen that the fringing gets significant at wavelengths longer than 720 nm for observation 1 (and similarly for observations 2 to 7) but is reduced in observation 8 taken with the new detector. The gaps in wavelength coverage are due to the physical gaps between individual CCDs in the detector, and deviation from the PA of the telescope for comparison star 2.
both in time and wavelength and manifest as a slope in the measured transmission spectrum of the planet if not corrected for, as demonstrated by H17. In a recent study, Pearson et al. (2018) introduced a method to measure Gemini/GMOS spectral shifts by computing the cross power spectrum of the stellar spectra in the Fourier space, also known as phase-only correlation algorithm. This is equivalent to performing cross-correlation of the stellar spectra in the wavelength space, as done by H17. We follow the H17 approach which is also described in detail in P22, and select three features (Na, Hα, and O2) in the stellar spectrum of WASP-19 (also labelled in Figure 4.1). In brief, we measure the spectral shifts for each feature in time by cross-correlation of each of the 1D spectra with a reference spectrum obtained around the mid-transit. The measured spectral shifts are then used to apply interpolated corrective shifts to every pixel for each exposure. We repeat this step for both the comparison stars as well, using the same set of spectral features as the target star spectrum. To correct for shifts between the target and comparison stars themselves, we then interpolate the comparison stars’ spectra for each exposure onto the interpolated common wavelength solution of the target star, omitting detector gaps and bad columns. This results in a common wavelength solution for both the target and comparison star spectra. We apply these corrections to each observation.

However, for all observations we find that the transmission spectrum we obtain from the GP based methods we use in P22 are consistent within 1σ whether we perform the spectral shift and stretch corrections or not. This indicates that the GP model from P22 which we eventually use to fit the spectroscopic light curves (described in more detail in Section 4.5.3) mitigates the effects of stellar spectral shifts and stretch on the final transmission spectrum. Hence, we opt to use the optimally extracted spectra without any shift and stretch corrections. Moreover, we eventually use only the target spectroscopic light curves to extract the transmission spectrum, which prevents the effects of shifts between the target and comparison star spectra. We additionally also use a wavelength bin size of 20 nm, which is significantly larger than the average amplitude of spectral shifts.

### 4.4.5 Extracting the light curves

After extracting the time series of the 1D spectra for the target and comparison stars for each transit observation, we proceed to construct the corresponding light curves. We construct the white light curves for both the target and comparison stars by summing the flux for each exposure spectrally over the wavelength range of 520 to 720 nm for transit 1 to transit 7, and from 520 to 900 nm for transit 8. Since the exposure time in general was not fixed throughout the night, we also normalized the total flux in each exposure by the corresponding exposure time. We then normalize the comparison star light curve by its median, and the target star light curve by the median of the out of transit exposures. For constructing the spectroscopic light curves, we repeat the same process for each of the 20 nm wide wavelength bins.
4.5 Analysis

We now describe our light curve analysis as applied to the 8 GMOS transit observations of WASP-19b with the goal to obtain the planet’s transmission spectrum. We first discuss the analysis of the white transit light curves in Section 4.5.2 for which we use two independent methods: the conventional method that fits for the Target/Comparison light curves, and the new method recently introduced by P22 of fitting the target star light curves directly using the comparison star light curve as one of the GP regressors. In Section 4.5.3 we describe the analysis of the wavelength binned light curves also using the conventional method and the new method from P22 to obtain the transmission spectrum for each observation.

4.5.1 Modelling systematics in GMOS transit light curves

We model the instrumental and telluric time-dependent systematics in the WASP-19 transit light curves constructed in Section 4.4.5 following the conventional as well as the new method introduced and described in more detail in P22. The conventional method involves a linear approach of first normalizing the target light curve by the comparison star light curve to correct for systematics commonly affecting both the target and comparison star light curves. The resulting Target/Comparison light curve typically still suffers from residuals systematics arising from non-linear differences at the telluric level e.g. brightness or spectral type between the target and comparison stars (leading to different telluric systematics in their respective light curves). This differences can also arise at the instrument level e.g. due to a non-ideal PA, unequal travel times of the instrument shutter common in GMOS observations. The Target/Comparison light curve is then fit with a transit model added to a parametric or a non-parametric (e.g. Gaussian Processes (GP) Gibson et al. 2012) systematics model.

The new method introduced by P22 (New:WLC followed by New:λLC, as described in Table 2 in P22) fits the target transit light curve directly, using the comparison star light curve as one of the regressors in a GP systematics model. The advantage of the new method is that it avoids introducing unwanted systematics to the target light curve as a result of normalization by a non-ideal comparison star light curve that behaves differently during the night. The new method in fact allows using the comparison star light curve as one of the regressors to the GP model (for the white target light curves) and let the GP glean the likely non-linear mapping between the systematics common to both the target and comparison light curves. In this process, it also propagates the uncertainties within a Bayesian framework instead of simple addition by quadrature (as is the case when doing Target/Comparison normalization). This approach is further relevant to our observations of WASP-19 as the only comparison star we have at our disposal is significantly fainter (∼ 1.22 magnitude fainter) as compared to WASP-19. Moreover, we are already dealing with a host star whose stellar variability has a significant effect on the transmission spectra (as discussed in Section 4.6.1. Additional stellar variability of the comparison star can lead to further contamination of the final transmission spectrum due to wavelength dependent effect present in the comparison star spectroscopic light curves themselves. For the comparison stars observed using Gemini/GMOS we do ob-
serve stellar variability in their TESS light curves albeit at lower amplitudes as compared
to WASP-19 (described in more detail in Appendix 4.A.1.2). Hence, it is important to not
directly use the comparison star spectroscopic light curves when measuring the final trans-
mittance spectrum. Our new method only uses the comparison star white light curve to fit
the target star white light curve and then uses the target star common-mode trend to fit the
spectroscopic target light curves, as we describe in more detail in Section 4.5.3.

Both the methods have the common aspect of fitting the transit light curve as a systematics
model added to a numerical transit model. The main difference between the two methods is
that the new method uses the GP framework of Gibson et al. (2012) to model the systematics
directly in the target star light curves, accounting for the non-linear differences between the
target and comparison star light curves. In this method, the comparison star light curves are
essentially used as a control sample to check that the noise is efficiently modelled.

The GP model we use for modelling the systematics for both methods (i.e., for modelling
both the Target/Comparison and Target star light curve respectively) is the same as that
described in more detail in P22. In brief, we use a Matérn 3/2 kernel function to construct
the GP covariance matrix, with a single amplitude hyperparameter and a length scale hy-
perparameter for each of the inputs to the GP.

### 4.5.2 Analysis of White Transit Light Curves

We describe some steps and details common to both methods (i.e., both Target/Comparison
and Target light curves ) mentioned in Section 4.5.1 as applied to the transit white light
curves for all the 8 transits, and mention specifically the points at which the two methods
differ.

We use the transit modelling package batman (Kreidberg 2015) to calculate the numerical
transit model $T(t, \phi)$ (where $t$ are the time stamps of each exposure and $\phi$ is the set of orbital
transit parameters), and the package george (Ambikasaran et al. 2015) for constructing and
computing the GP kernels and likelihoods. In Table 4.3 we summarize the parameters we
fix and the priors we employ in our fitting procedure. We fix the orbital period ($P$) and ec-
centricity (e), and fit for the orbital inclination ($i$), orbital separation ($a/R_*$), central transit
time ($T_0$), planet to star radius ratio ($R_p/R_*$). We employ a linear limb darkening law and
fit for the linear limb darkening coefficient $u_1$. We choose to use linear limb darkening law
because given the precision and time resolution of our light curves, multiple free parame-
ters describing the limb darkening e.g. in case of quadratic or non-linear limb darkening law,
are difficult to constrain. Hence, the linear law in this context is the simplest choice to fit
for. Recent work by Patel & Espinoza (2022) demonstrates that specifically quadratic limb
darkening coefficients for sun-like or cooler stars often suffer from discrepancies between
the theoretical and empirical methods used to estimate them.

For each transit model parameter except the linear limb darkening coefficient we put trun-
cated uniform priors within $\sim 10 \sigma$ bounds around their literature values. We calculate the
linear limb darkening coefficient $u_1$ for the wavelength range integrated to obtain the white
light curve, and the wavelength bins we adopt for spectroscopic light curves (see Section
Figure 4.2: White GMOS-R150 target light curves for WASP-19b and their best fits obtained using the new method from observations 1 to 4. For all observations the black points are the target light curves overplotted with the best fit transit + GP systematics model in red, purple points show the comparison star light curve, and green points are the detrended target light curve overplotted with the best transit model in blue. The detrended light curve and the respective best fit transit model have been offset for clarity. Note that for all observations we observe significant odd-even effect in both the target and comparison star light curves, which are efficiently modelled by the GP model in the new method using the comparison star as one of the GP regressors. The gap in the light curve for observation 4 is due to outliers in the light curve around the inflection point for the Cassegrain rotator which happens when the target reaches zenith.
Figure 4.3: Same as Figure 4.2 but for observations 5 to 8. Note that observation 8 is noisier as compared to other observations because it was taken using a different detector (as described in 4.3.1) and setup as compared to all the other observations. The gaps in the light curves for observations 5 and 8 are due to outliers in the light curve around the inflection point for the Cassegrain rotator which happens when the target reaches zenith.
using PyLDTk (Parviainen & Aigrain 2015), which uses the spectral library in (Husser et al. 2013), based on the stellar atmosphere modelling code PHOENIX. We put a Gaussian prior on the linear limb darkening coefficient with the mean value and the standard deviation as the mean and 3 times the 1 $\sigma$ uncertainty calculated from PyLDTk respectively. We also fit for the white noise parameter $\sigma_w$ which lets the GP model fit for the white noise variance in the target star light curves and also includes contribution from the variance inherent in a noisy GP input itself (e.g. the comparison star light curve). This is in fact an important feature of the new method and provides a natural way to propagate uncertainties from the comparison star light curve to our fit of the target star light curve within a Bayesian framework. We emphasize that fitting for $\sigma_w$ is crucial for letting the GP model capture the white noise in the target star light curves.

We perform the white light curve fits using all possible combinations of GP input parameters used to construct a combined Matérn 3/2 kernel function (described in more detail in P22). For the conventional method, we use time, CRPA (Cassegrain Rotator Position Angle), and airmass as GP regressors. For the new method, we use the same set of GP regressors as the conventional method but additionally also the point spread function (PSF) width of the target spectral trace for every exposure, and the comparison star light curve.

We put wide uniform priors on the logarithm of the GP hyperparameters which include the covariance kernel function amplitude $A$ and the length scales for each GP regressor $\eta_p$. We effectively sample the amplitude and length scale hyperparameters logarithmically as shown in Table 4.3. The logarithmic sampling of hyperparameters effectively puts a shrinkage prior on them, which pushes them to smaller values if the corresponding input vector truly does not represent the covariance in the time series (Gibson et al. 2012).

We first find the Maximum a-Posteriori (MAP) solution by optimizing the GP posterior (see P22 and Gibson et al. 2012 for more detail) using the Powell optimizer in the SciPy python package.

Using the MAP solution as the starting point, we marginalize the GP posterior over all hyperparameters and transit model parameters through an MCMC using the package emcee (Goodman & Weare 2010, Foreman-Mackey et al. 2013). We use 50 walkers for 10000 steps and check for the convergence of chains by using the integrated autocorrelation times for each emcee walker following the method described in (Goodman & Weare 2010). We ensure that the total length of our chains is greater than 50 times the integrated autocorrelation time which indicates that our samples are effectively independent and have converged. We also tested the robustness of our posteriors from a nested sampler using the package dynesty (Speagle 2020) and obtain posteriors consistent with those from emcee well within 1 $\sigma$. We estimate the best fit parameters by taking the 50th percentile and their $+1\sigma$ and $-1\sigma$ uncertainties by taking the 84th and 16th percentile respectively of the MCMC posteriors. We show and compare the best fit transit parameters (corresponding to the combination of GP inputs that perform best for both methods) and their $\pm 1\sigma$ uncertainties in Table 4.4. In Figures 4.2 and 4.3 we show the best fits to the target star light curve obtained from the new method for all 8 observations.
Table 4.3: Summary of priors and fixed values for the parameters (batman transit model and GP hyperparameters) used to fit the transit light curves of WASP-19b. For all the fits we fixed the planet orbital period (P) and eccentricity (e). \( \mathcal{U} \) shows a uniform prior applied within the specified range, and \( \mathcal{N} \) represents a Gaussian prior with the mean and standard deviation respectively. \( T_c \) is the predicted mid-transit time for each epoch using the ephemeris from Hartman et al. 2011. For the limb darkening we use a Gaussian prior around the mean linear limb darkening coefficient theoretically calculated by PyLDTk (Parviainen & Aigrain 2015) corresponding to the stellar parameters in Table 4.1 and for the R150 wavelength range of 520 to 900 nm.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior/Fixed value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) [d]</td>
<td>0.7888390</td>
<td>Lendl et al. (2013)</td>
</tr>
<tr>
<td>( e )</td>
<td>0.0046</td>
<td>Hellier et al. (2011)</td>
</tr>
<tr>
<td>( i ) [°]</td>
<td>( \mathcal{N} ) (79.5, 1.5)</td>
<td>Lendl et al. (2013)</td>
</tr>
<tr>
<td>( R_p / R_\star )</td>
<td>( \mathcal{U} ) (0, 1)</td>
<td>–</td>
</tr>
<tr>
<td>( a / R_\star )</td>
<td>( \mathcal{N} ) (3.573, 1.5)</td>
<td>Lendl et al. (2013)</td>
</tr>
<tr>
<td>( T_0 [d] )</td>
<td>( \mathcal{U} ) (( T_c )-0.001, ( T_c )+0.001)</td>
<td>Hellier et al. (2011)</td>
</tr>
<tr>
<td>( u_1 ) [R150]</td>
<td>( \mathcal{N} ) (0.63, 0.03)</td>
<td>PyLDTk</td>
</tr>
</tbody>
</table>

GP model hyperparameters

| \( \ln (A) \)      | \( \mathcal{U} \) (-100, 100) | – |
| \( \ln (\eta_p) \) | \( \mathcal{U} \) (-100, 100) | – |
| \( \sigma_w \)     | \( \mathcal{U} \) (0.00001, 0.005) | – |

We select the best GP regressor combination for both methods independently using two criteria: 1) Bayesian evidence \( \log_e Z \) estimate from dynesty and 2) the Bayesian Information Criterion (BIC, Schwarz 1978) computed using the GP likelihood corresponding to the best fit transit model parameters and hyperparameters. We use the \( \Delta \)BIC and \( \Delta \log_e Z \) threshold prescribed by Kass & Raftery (1995) and also used in P22 to choose the best GP regressor combination for the two methods individually. We find that the GP regressor combination selection based on both the criteria (BIC and \( \log_e Z \)) always agree within their model selection thresholds as prescribed in Kass & Raftery (1995). We show the best GP regressor combination and the best fit transit parameters and their 1\( \sigma \) uncertainties for both the conventional and new methods in Table 4.4. The best fit GP hyperparameters and their uncertainties are shown in Table 1 in the supplementary material. In Table 4.5 we list the transit parameters measured by previous studies and this work using TESS photometry as described in Section 4.A.1.1.

For the new method, using the comparison star as one of the GP regressors gives the best fit for most of the observations. Specifically, for the new method applied to observations 3 and 6, we find that using only the comparison star light curve as the GP regressor performs best. For all other observations, using time or airmass as a regressor in addition to the comparison star light curve helps to model the lower frequency variations in the target star light curve which are not present in the comparison star light curve.
For the conventional method, using just time as a GP regressor gives the best fit for most observations. We find that the new method for most of the observations, and in particular observation 8, gives comparable or better fits compared to the conventional method when considering the transit depth precisions and the residual RMS. The new method yields on average 10 to 20 % smaller RMS on the residual scatter for the best fit as compared to the conventional method.

4.5.2.1 Correcting for the odd-even effect in the light curves

The consecutive exposures in the GMOS light curves suffer from an odd-even effect due to unequal travel times of the GMOS blade-shutters with respect to the direction of motion, and have been previously observed and corrected for in P22 and Stevenson et al. (2014). We estimate the level of this effect for our WASP-19b observations to be around ~300 ppm for both the target and comparison star light curves. This effect is most significantly observed in observations 2, 3, 4, 5, and 6 (Figures 4.2 and 4.3). Note that the amplitude of this odd-even effect is not exactly the same for both the target and comparison star light curves (since it depends on the direction of motion of blade-shutters). This difference in the amplitudes was in fact observed for one of the transits of HAT-P-26b in P22 (labelled as observation 2 in that paper). It was observed that due to the difference in the timing of this odd-even effect for the target and comparison light curves respectively, simply dividing the target by the comparison light curves as done during the conventional method doesn't correct for this effect and instead exacerbates it. This is one of the examples of a non-linear relationship between how the same source of systematics affect the target and comparison star light curves. The new method resolves this by letting the GP determine this non-linear mapping. Note that the timescale of the odd-even effect is the same for both target and comparison star light curves as we confirm from their individual Lomb Scargle periodograms. Using the comparison star light curve as a GP regressor as in the new method is able to efficiently model this effect, as can be observed in the best fit models and the residuals in Figures 4.2 and 4.3.

Once we retrieve the best fit transit parameters for each observation from the respective white transit light curve, we use this information to fit the spectroscopic light curves and obtain the transmission spectrum as described in more detail in Section 4.5.3.

4.5.3 Analysis of Spectroscopic Light Curves

We now describe the analysis of the spectroscopic light curves (hereafter referred to as λLC) constructed by integrating the 1D stellar spectrum in 20 nm wide bins (as mentioned in Section 4.4.5. We chose the bin width of 20 nm as it is a few times the seeing limited resolution of ~ 4 nm for our observations. This is similar to the previous R150 Gemini/GMOS observations from our survey program published by H17, Todorov et al. (2019), and P22. For inspecting especially the bins centred around the 589 nm Na doublet we also construct spectroscopic light curves in 10 nm wide bins to sample the core and wings of the Na doublet.
Table 4.4: Best fit transit parameters obtained from the fits to white transit light curves of eight GMOS-R150 observations analysed in this work. Two rows for each observation as specified in the second column compare the best fit transit parameters and residual RMS resulting from the new method from P22 of fitting the Target white light curve and the conventional method of fitting the Target divided by the Comparison light described in more detail in Section 4.5.1. The third column specifies the best GP regressor combination for the GP model for both methods for each observation, as described in more detail in Section 4.5.2. ‘Time’ refers to the time stamps of the exposures in the observation, and ‘Comp’ refers to the comparison star light curve. $\sigma_u$ values are the median white noise value quantifying the diagonal of the GP covariance matrix with a measured uncertainty. The RMS values are the standard deviation of the residuals between the light curve and the transit model and the predicted mean of the GP systematics model corresponding to the median of the posteriors. Both of these quantities can be viewed as two ways of estimating the white noise level of the light curves.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>GP regressors</th>
<th>$R_p/R_*$</th>
<th>$T_0[BJD_{TDB}]$</th>
<th>$a/R_*$</th>
<th>$i[\degree]$</th>
<th>$u_1$</th>
<th>$\sigma_u$ [ppm]</th>
<th>RMS [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>New</td>
<td>Time, Comp</td>
<td>0.1449$^{+0.0037}_{-0.0035}$</td>
<td>2456316.730224$^{+0.00194}_{-0.00184}$</td>
<td>3.6$^{+0.05}_{-0.05}$</td>
<td>79.88$^{+0.42}_{-0.41}$</td>
<td>0.63$^{+0.03}_{-0.02}$</td>
<td>383$^{+35}_{-28}$</td>
<td>337</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1389$^{+0.0012}_{-0.0012}$</td>
<td>2456316.729882$^{+0.00237}_{-0.00237}$</td>
<td>3.52$^{+0.07}_{-0.06}$</td>
<td>79.11$^{+0.51}_{-0.49}$</td>
<td>0.64$^{+0.02}_{-0.03}$</td>
<td>395$^{+39}_{-30}$</td>
<td>349</td>
</tr>
<tr>
<td>2</td>
<td>New</td>
<td>Time, Comp</td>
<td>0.1451$^{+0.0009}_{-0.0009}$</td>
<td>2456327.773539$^{+7.0e-05}_{-7.0e-05}$</td>
<td>3.57$^{+0.03}_{-0.03}$</td>
<td>79.23$^{+0.22}_{-0.21}$</td>
<td>0.59$^{+0.02}_{-0.02}$</td>
<td>373$^{+25}_{-25}$</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1451$^{+0.0012}_{-0.0012}$</td>
<td>2456327.773395$^{+0.000124}_{-0.0000101}$</td>
<td>3.58$^{+0.04}_{-0.03}$</td>
<td>79.29$^{+0.26}_{-0.25}$</td>
<td>0.61$^{+0.02}_{-0.02}$</td>
<td>431$^{+27}_{-23}$</td>
<td>416</td>
</tr>
<tr>
<td>3</td>
<td>New</td>
<td>Comp</td>
<td>0.1408$^{+0.0009}_{-0.0009}$</td>
<td>2456335.662255$^{+8.1e-05}_{-8.1e-05}$</td>
<td>3.52$^{+0.04}_{-0.03}$</td>
<td>79.1$^{+0.28}_{-0.24}$</td>
<td>0.58$^{+0.02}_{-0.02}$</td>
<td>345$^{+24}_{-21}$</td>
<td>327</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time</td>
<td>0.1398$^{+0.0022}_{-0.0023}$</td>
<td>2456335.662113$^{+0.000124}_{-0.000127}$</td>
<td>3.53$^{+0.04}_{-0.04}$</td>
<td>79.19$^{+0.30}_{-0.27}$</td>
<td>0.61$^{+0.02}_{-0.02}$</td>
<td>338$^{+24}_{-22}$</td>
<td>316</td>
</tr>
<tr>
<td>4</td>
<td>New</td>
<td>Time, Comp, Airmass</td>
<td>0.1434$^{+0.0021}_{-0.0021}$</td>
<td>2456667.763054$^{+0.000116}_{-0.000113}$</td>
<td>3.59$^{+0.04}_{-0.04}$</td>
<td>79.5$^{+0.3}_{-0.3}$</td>
<td>0.61$^{+0.03}_{-0.02}$</td>
<td>353$^{+25}_{-23}$</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1391$^{+0.0034}_{-0.0029}$</td>
<td>2456667.763072$^{+0.000175}_{-0.000176}$</td>
<td>3.61$^{+0.06}_{-0.05}$</td>
<td>79.74$^{+0.43}_{-0.41}$</td>
<td>0.63$^{+0.02}_{-0.02}$</td>
<td>379$^{+25}_{-23}$</td>
<td>346</td>
</tr>
<tr>
<td>5</td>
<td>New</td>
<td>Time, Comp</td>
<td>0.1488$^{+0.0035}_{-0.0033}$</td>
<td>2456697.739132$^{+0.000155}_{-0.000157}$</td>
<td>3.6$^{+0.04}_{-0.04}$</td>
<td>79.64$^{+0.3}_{-0.29}$</td>
<td>0.6$^{+0.03}_{-0.03}$</td>
<td>326$^{+24}_{-22}$</td>
<td>294</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1416$^{+0.0039}_{-0.0025}$</td>
<td>2456697.739098$^{+0.000201}_{-0.000216}$</td>
<td>3.54$^{+0.05}_{-0.05}$</td>
<td>79.17$^{+0.36}_{-0.41}$</td>
<td>0.6$^{+0.02}_{-0.02}$</td>
<td>380$^{+28}_{-26}$</td>
<td>346</td>
</tr>
<tr>
<td>6</td>
<td>New</td>
<td>Comp</td>
<td>0.1469$^{+0.0012}_{-0.0011}$</td>
<td>2456727.715017$^{+9.8e-05}_{-9.6e-05}$</td>
<td>3.52$^{+0.04}_{-0.04}$</td>
<td>79.05$^{+0.32}_{-0.32}$</td>
<td>0.62$^{+0.02}_{-0.02}$</td>
<td>402$^{+30}_{-25}$</td>
<td>379</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Time</td>
<td>0.1442$^{+0.0013}_{-0.0014}$</td>
<td>2456727.715003$^{+0.000123}_{-0.000123}$</td>
<td>3.52$^{+0.04}_{-0.04}$</td>
<td>79.05$^{+0.33}_{-0.33}$</td>
<td>0.62$^{+0.02}_{-0.02}$</td>
<td>395$^{+30}_{-25}$</td>
<td>378</td>
</tr>
<tr>
<td>7</td>
<td>New</td>
<td>Time, Comp</td>
<td>0.1487$^{+0.0035}_{-0.0035}$</td>
<td>2456757.690677$^{+0.000173}_{-0.000167}$</td>
<td>3.61$^{+0.05}_{-0.05}$</td>
<td>79.68$^{+0.3}_{-0.36}$</td>
<td>0.62$^{+0.03}_{-0.03}$</td>
<td>375$^{+39}_{-31}$</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1427$^{+0.0023}_{-0.0025}$</td>
<td>2456757.690571$^{+0.000165}_{-0.000124}$</td>
<td>3.59$^{+0.05}_{-0.05}$</td>
<td>79.57$^{+0.39}_{-0.35}$</td>
<td>0.63$^{+0.01}_{-0.01}$</td>
<td>403$^{+42}_{-36}$</td>
<td>356</td>
</tr>
<tr>
<td>8</td>
<td>New</td>
<td>Time, Comp</td>
<td>0.1482$^{+0.0033}_{-0.0038}$</td>
<td>2457022.740594$^{+0.000252}_{-0.000237}$</td>
<td>3.54$^{+0.03}_{-0.03}$</td>
<td>79.45$^{+0.24}_{-0.22}$</td>
<td>0.59$^{+0.01}_{-0.01}$</td>
<td>753$^{+65}_{-58}$</td>
<td>702</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Airmass</td>
<td>0.1371$^{+0.0059}_{-0.0054}$</td>
<td>2457022.740269$^{+0.000505}_{-0.000542}$</td>
<td>3.58$^{+0.05}_{-0.05}$</td>
<td>79.35$^{+0.25}_{-0.24}$</td>
<td>0.59$^{+0.01}_{-0.01}$</td>
<td>1051$^{+91}_{-81}$</td>
<td>958</td>
</tr>
</tbody>
</table>
We fit the spectroscopic light curves to extract the transmission spectrum using both the conventional method and the new method as introduced in P22 and described here in brief in the next two subsections. For both methods to fit the spectroscopic light curves, we follow the same procedure as the white light curves in Section 4.5.2 to sample the posterior and obtain the best fit parameters and their uncertainties using emcee and dynesty.

### 4.5.3.1 Conventional method using common-mode correction

We first describe in brief the conventional method of fitting λLCs. We divide each target λLC by the corresponding comparison star λLC. GMOS λLCs are known to suffer from wavelength-independent systematics which are conventionally corrected for using common mode corrections (Stevenson et al. 2014, H17, Todorov et al. 2019). We essentially use the GP noisemodel from the best fits to the Target/Comparison white light curves for each observation obtained in Section 4.5.2 to do a conventional common-mode correction and remove time-dependent systematics common across all wavelength bins.

For the conventional method, we derive the common-mode trend by subtracting the best fit white light curve transit model from the observed Target/Comparison white light curves. For each observation, this transit model is constructed using the corresponding best fit transit parameters obtained using the conventional method for the respective white Target/Comparison light curve as mentioned in Table 4.4. We then normalize the Target/Comparison λLC by their respective median out-of-transit flux and subtract the common-mode trend from each of them. We find that doing common-mode correction prior to fitting the Target/Comparison λLCs improved the precision of measured transit depths by ∼ 15% on average per wavelength bin as compared to when we do not perform common mode correction. However, performing common-mode correction also implies that we effectively lose information on the absolute value of transit depths and the transmission spectra relative to the white light curve transit depth which was used to derive the common-mode trend.

We fit the common-mode corrected Target/Comparison λLCs independently with the model described in 4.5.1 as also used for white light curves in Section 4.5.2, using only time as a GP regressor. Using time as a GP regressor at this stage helps in accounting for residual
wavelength-dependent systematics in the $\lambda$LCs after common-mode correction, likely due to wavelength-dependent differential atmospheric extinction between the target and comparison stars with changing airmass through the night.

Since our main goal with the spectroscopic light curves is to measure the transit depth in each wavelength bin, we fix the orbital inclination ($i$), orbital separation ($a/R_*$), and mid-transit time ($T_0$) to the best fit values for the corresponding white light curve in Section 4.5.2 (see Table 4.4), and orbital period and eccentricity to literature values. We use a linear limb darkening law and employ a Gaussian prior for the limb darkening coefficients around the PyLDTk pre-calculated values for each wavelength bin (approximating a top hat transmission function for each wavelength bin).

### 4.5.3.2 New method using the common-mode trend as a GP regressor

We now describe the application of the new method introduced in P22 to fit the Target $\lambda$LCs directly. One of the motivations behind application of this approach is the large difference in brightness ($\sim 1.22$ mag in $V_{\text{mag}}$) between the target and comparison star, which makes the correction for time and wavelength dependent systematics through differential spectrophotometry suboptimal and a source of additional uncertainties. Instead of dividing the target star $\lambda$LCs by the comparison star $\lambda$LCs and performing the conventional common-mode correction, we use the common-mode trend derived from the white target light curve as a GP regressor to fit the systematics and transit depth in the target $\lambda$LC. We do this by first deriving the common-mode trend in the same way as done for the conventional method in Section 4.5.3.1, but now using the white target light curve. For this, we use the transit model corresponding to the best fit transit parameters obtained using the new method for the respective white target light curve for each observation.

We then use the common-mode trend and time both as GP regressors to fit the individual $\lambda$LCs independently. The common-mode trend helps in modelling the largely wavelength independent high frequency systematics including the known odd-even effect described in Section 4.5.2.1. Using time as an additional GP regressor helps in modelling the smoother low frequency trend in the Target $\lambda$LCs which is due to the changing airmass through the night and is wavelength dependent. Similar to the conventional method described in Section 4.5.3.1, for the new method as well we keep all the transit model parameters except the transit depth and the limb darkening coefficient for each $\lambda$LC fixed to the best fit values derived from the new method for the corresponding white target light curve (tabulated in Table 4.4).

It should be noted that both the new and conventional methods of fitting the spectroscopic light curve use the common-mode trend which means the resultant transmission spectra from both the methods are relative to the white light curve transit depth used to derive the common-mode. We also considered two further possible GP regressor combinations excluding the common-mode trend: 1) $\lambda$LC and 2) time and $\lambda$LC. We show the resultant transmission spectra overlapped with those from using common-mode and time as GP regressor in the Figure 1 in supplementary material. Figure 2 in supplementary material shows the
Multi-epoch stellar variability effect in the transmission spectrum of WASP-19b

difference in per bin BIC (derived from the GP likelihood) between the respective methods used to fit target $\lambda LC$. Based on the average per bin $\Delta$BIC for all observations as seen in Figure 2 of supplementary material, we conclude that using common-mode and time are the best favoured GP regressor combination for fitting the target $\lambda LC$.

In principle, using comparison spectroscopic light curves as GP regressors would be preferable over using the common-mode as a GP regressor, but this would concretely depend on the comparison star itself. In the precise case of WASP-19b observations in this paper, the comparison star is 1.22 magnitude fainter as compared to WASP-19 which leads to worse fits (higher BIC as seen in Figure 2 of supplementary material). Hence, this forces us to go for the next best option, which is using the common-mode and time as GP regressors. We note that the transmission spectra for all the observations obtained from the new method using common-mode trend are consistent with those obtained using the best GP regressor combination excluding the common-mode trend: comparison $\lambda LC$ and time.

We show the $\lambda LC$s and their best fits obtained from both the conventional and the new method in Figures 3 to 10 in the supplementary material. The respective transmission spectra for each observation from both the methods are plotted in Figure 4.4 and tabulated in Table 4.9 and 4.8.

4.5.3.3 Comparison between the transmission spectrum from the conventional and the new method

We compare the transmission spectrum at each epoch derived from two methods: the conventional method in which we use the comparison star $\lambda LC$s followed by common-mode trend subtraction, and the new method in which we do not use the comparison star $\lambda LC$s and use the common-mode trend as a GP regressor. The transmission spectrum for each epoch obtained from both the methods are plotted for comparison in Figure 4.4. The average per bin precision on the transmission spectrum from the two methods are comparable. However, particularly in the case of observation 1 and observation 8 (the noisiest observations in our dataset), the new method yields 30 % and 50 % respectively smaller average transit depth uncertainties compared to the conventional method. The RMS of the residuals in $\lambda LC$ from the new method across all observations are smaller by a factor of 3 on average as compared to those from the conventional method (annotated in $\lambda LC$ Figures 1 to 8 in the supplementary material). The better transit depth precision yielded by the new method is an outcome of both not using the fainter and hence noisier comparison star $\lambda LC$s and a generalized non-linear mapping of the white light curve common-mode trend with the individual $\lambda LC$s. This advantage of the new method was also demonstrated in P22. In essence, the transmission spectra from the new method are less susceptible to additional uncertainties and bias introduced in the conventional method by simply dividing the target spectroscopic light curves by spectroscopic light curves of a significantly fainter comparison star. Since we don’t use the comparison star spectroscopic light curves in the new method, the transmission spectra hence obtained are not affected by wavelength dependent changes in the stellar spectra due to potential variability of the comparison star itself which could
Figure 4.4: Comparison of the transmission spectra of WASP-19b for each of the eight GMOS-R150 observations extracted using the conventional method (black points) and the new method (red points) from P22 described in Section 4.5.3.1 and 4.5.3.2 respectively. Observation numbers of each epoch are the same as in Table 4.2. Observed GMOS-R150 stellar spectrum of WASP-19 for an arbitrary exposure is shown for each epoch in green. As described in Section 4.5.3.3 we eventually use the transmission spectra from the new method for subsequent interpretation in the paper.
complicate our study and correction of the host star variability on the transmission spectrum of WASP-19b in Section 4.6.2. Hence, in the subsequent sections in the paper, we consider only the transmission spectra from the new method for further interpretation.

In Section 4.6 we also discuss the effect of stellar variability on the transmission spectrum during each epoch and introduce a new way to correct them before constructing the combined transmission spectrum and comparing it with previous studies and atmospheric models.

4.6 Discussion

4.6.1 Effect of stellar activity on the transmission spectrum of WASP-19b

WASP-19 is known to vary at a level of \( \sim 2 \% \) peak to trough as seen from the TESS photometry and our ground based monitoring from LCO telescopes (Figure 4.13), which translates to \( \sim 2 \% \) variation in white light curve transit depth from GMOS observations. We do not identify any spot crossing events in our GMOS observations like those observed by Espinoza et al. (2019) and Mancini et al. (2013) despite the precision of the GMOS transit light curves. However, spots and faculae are also expected to significantly affect the transmission spectrum via the transit light source effect (Rackham et al., 2018, 2019) especially in the visible wavelength range covered by our GMOS observations. Hence, it is necessary to correct for this effect of stellar activity in the transmission spectrum at each epoch before combining them and producing the transmission spectrum.

We estimate the impact on the transmission spectrum from unocculted stellar heterogeneity in a semi-empirical way. We use the estimates on temperature contrast and covering fraction of spots and faculae reported by Espinoza et al. (2019) based on the spot and faculae crossing events observed in their Magellan/IMACS light curves and previously by Mancini et al. (2013). Espinoza et al. (2019) use the PHOENIX model stellar photospheres (Husser et al. 2013) and the observed spot and faculae contrasts to derive the estimates on spot and faculae covering fraction ranges that correspond to the 2 % amplitude of stellar flux variability of WASP-19 seen in the visible bandpass. In Table 4.6 we summarize the spot and faculae properties from Espinoza et al. (2019) and Mancini et al. (2013) which we use in this paper to compute the effect of stellar activity on the transmission spectrum of WASP-19b.

Now that we have an estimate of the properties of the stellar inhomogeneities on the stellar surface corresponding to the visible stellar flux variability, we estimate their impact on the theoretical transmission spectrum of WASP-19b. To do so, we use the open source atmospheric modelling code platon (Zhang et al. 2019, 2020) based on ExoTransmit (Kempton et al. 2017) to calculate forward models for the transmission spectra corresponding to solar metallicity and C/O, accounting for the effect of unocculted stellar photospheric heterogeneity. The wavelength dependent effect of stellar variability implemented by platon (Equation 4 in Zhang et al. 2019) is the same as the one described in McCullough et al. (2014) and Rackham et al. (2018). We use platon to calculate the forward models for three independent cases using the parameters listed in Table 4.6: high contrast spots, low contrast spots,
Table 4.6: Stellar photospheric heterogeneity parameters for WASP-19 corresponding to the peak to trough V band variability amplitude of 2 %.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\text{phot}}$</td>
<td>Immaculate stellar photosphere temperature</td>
<td>5460 K</td>
<td>Doyle et al. (2013)</td>
</tr>
<tr>
<td>$T_{\text{spot, high}}$</td>
<td>High contrast spot temperature</td>
<td>4780 K</td>
<td>Mancini et al. (2013)</td>
</tr>
<tr>
<td>$T_{\text{spot, low}}$</td>
<td>Low contrast spot temperature</td>
<td>5270 K</td>
<td>Espinoza et al. (2019)</td>
</tr>
<tr>
<td>$T_{\text{fac}}$</td>
<td>Faculae temperature</td>
<td>5600 K</td>
<td>Espinoza et al. (2019)</td>
</tr>
<tr>
<td>$f_{\text{spot, high}}$</td>
<td>High contrast spot covering fraction</td>
<td>$2^{+2.4}_{-0.7}$ %</td>
<td>Espinoza et al. (2019)</td>
</tr>
<tr>
<td>$f_{\text{spot, low}}$</td>
<td>Low contrast spot covering fraction</td>
<td>$10^{+3.0}_{-3.0}$ %</td>
<td>Espinoza et al. (2019)</td>
</tr>
<tr>
<td>$f_{\text{fac}}$</td>
<td>Faculae covering fraction</td>
<td>$19^{+31}_{-10}$ %</td>
<td>Espinoza et al. (2019)</td>
</tr>
</tbody>
</table>

and faculae. Realistically, the stellar photosphere would be a combination of the three cases with spots and faculae contributing opposing effects. However, we consider the effect of each case separately to inspect the overall range of effect on the transmission spectrum due to stellar activity.

In Figure 4.5 we show the platoni forward models normalized to the HST/WFC3 spectrum from Huitson et al. 2013 and the individual GMOS transmission spectra from each of the eight observations overplotted. We notice from Figure 4.5 that unocculted stellar spots and faculae corresponding to the contrasts and covering fraction ranges for WASP-19 as estimated by Espinoza et al. 2019 can lead to an offset of up to $\sim$ 3000 ppm in the GMOS-R150 wavelength range of 520-900 nm. We measured this offset range from the same platoni models overplotted in Figure 4.5 which account for the effect due to unocculted high and low contrast spots and faculae with respect to the contrasts and covering fractions mentioned in Table 4.6. We emphasise that this is consistent with the observed spread $\sim$ 4000 ppm seen in the mean levels of the transmission spectra from our GMOS observations (see Figure 4.5).

We also inspect the TESS photometry for the variation of WASP-19b’s transit depths across the 58 transits with respect to the stellar flux. We first compare the variation in absolute TESS SAP (simple aperture photometry) flux measured for WASP-19b and the two comparison stars and find that all three stars show a similar offset in their absolute SAP flux levels from Sector 9 to Sector 36. This implies that the change in SAP flux of WASP-19 between two sectors is not astrophysical. Hence, we conduct the transit depth vs out of transit flux comparison for the two sectors independently by normalizing each sector’s photometry independently. Specifically we normalize each of the two orbits for both sectors by their respective median SAP flux, as shown in Figure 4.A.1. The resultant normalized out of transit flux vs transit depth comparison for both the sectors is shown in Figure 4.7. Within the individual sectors themselves WASP-19’s flux varies by $\sim$ 2 % which we interpret as due to rotational modulation by spots and faculae as also evident from the Lomb Scargle periodograms of both the sectors in Figure 4.A.1. Both the sectors show a scatter of $\sim$ 4000 ppm which is consistent with the spread in mean transmission spectra level seen in the eight GMOS-R150 observations (Figure 4.6), and the 6 Magellan/IMACS transmission spectra from Espinoza.
This is expected because both the GMOS-R150, Magellan/IMACS, and TESS observations have a significant overlap in wavelength range.

We find that Sector 9 photometry shows an anticorrelation between the out of transit stellar flux and the transit depth as expected from unocculted spots. The Pearson correlation coefficient of $-0.31$ at two tailed $p$ value of 0.1 indicates that the anticorrelation is not significant. We speculate that spot or faculae occultations by the planet during the transits observed by TESS could be responsible for this deviation from the anticorrelation expected from the stellar brightness variations due to only unocculted spots or faculae. Both spot and faculae occultations have been observed by Espinoza et al. (2019) with photometric amplitudes $\sim 3000$ ppm in the transit light curves. The average RMS we obtain from the TESS light curves is of the same order of $\sim 3000$ ppm as shown in Figures 13 and 14 in the supplementary material. Hence, from our fits of each TESS transit light curve in this work it is not possible to detect and fit for the signatures of spot or faculae occultations along with the transit signal. Hence, we speculate that the TESS transit depths are overall affected by spot and faculae occultations.

Unocculted spots and faculae impart not only an offset but also a slope to the optical transmission spectrum which can vary significantly across multiple epochs due to stellar variability. We demonstrate this in the forward models plotted in Figure 4.5. On average, high and low contrast spots impart a positive offset and a negative slope, while faculae, on the other hand impart a negative offset and a positive slope on the spectrum (Rackham et al. 2017b, Espinoza et al. 2019). To measure this effect at first order on the GMOS transmission spectrum at each epoch, we fit a linear slope to the transmission spectrum from each observation and compare the best fit linear slope value to the corresponding mean transmission spectrum level for each epoch. The linear slopes and mean transmission spectrum level for each epoch are plotted in Figure 4.6. A visual illustration of how we construct Figure 4.6 is given in Figure 4.14. We find that the GMOS observations with larger mean transmission spectrum level have a negative slope and vice-versa. We measure an anticorrelation between the transmission spectra mean level and their slopes with Pearson correlation coefficient $-0.61$ at 2-tailed $p = 0.1$. This anticorrelation is expected from the theoretical forward models accounting for the effect of stellar variability as we demonstrate next.

For comparison with predictions from theoretical models, we compute the platon forwards models for the transmission spectrum of WASP-19b while also accounting for the effect of unocculted spots and faculae. These are the models plotted in Figure 4.5. To obtain the expected slopes and offsets from the platon forward models we follow the same approach as that applied to the observed transmission spectra. In the GMOS wavelength range of 520 to 720 nm (common to all 8 epochs) we fit a linear slope to the platon forward models corresponding to mean and $\pm 1\sigma$ and $\pm 2\sigma$ properties of the three cases from Table 4.6: high-contrast spots, low-contrast spots and faculae. The predicted mean and $\pm 1\sigma$ and $\pm 2\sigma$ slope and transmission spectrum mean level for all three cases are shown as shaded regions in Figure 4.6.

We find that the transmission spectra mean level vs slope trend in GMOS observations is broadly consistent with the predictions from the forward models that account for the stellar
variability due to spots and faculae, as shown by the shaded region in Figure 4.6. We note that the model predicted trends deviate from the best fit linear trend to the data, especially at the faculae end, as shown in Figure 4.6. This is because the models we consider describe end-member effect from a purely spot or faculae dominated stellar photosphere. Realistically, WASP-19’s stellar photosphere is more likely to host a mix of both spots and faculae. This explains the deviation between the slope vs offset trend predicted by the models as shown by the shaded region in Figure 4.6 and the trend measured in the data. Nevertheless, the trend in slope vs offset space across all epochs can have implications on the morphology of the final transmission spectra combined from multiple epochs. Corrections for both the slope and offset at each epoch need to be applied before combining the transmission spectra. We use the observed trend in transmission spectral slopes and offsets to combine multi-epoch spectra; we describe this new empirical approach to correct for the effect of stellar variability in the following section.

4.6.2 A new empirical approach to correct for stellar variability across multiple epochs

Conventionally, transmission spectra obtained at different epochs have been combined by first applying an offset with respect to a reference level, e.g., as done for WASP-19b by Espinoza et al. (2019). This offset is constant with respect to wavelength. For comparison, we also first apply this constant offset correction to our GMOS observations. We choose a reference transit depth $T_{D,ref} = 21686$ ppm which is the value on the Y axis projected by the linear fit to the transmission spectral mean and slope corresponding to zero transmission spectrum slope on the X axis. This projection is marked as black dot dashed lines in Figure 4.6. Note that the $T_{D,ref}$ measured from the linear fit to the measured slopes and offsets of the data do not coincide with the analogous ‘zero-point’ indicated by the predicted slopes and offsets from the platon models marked by the shaded region in Figure 4.6. This is in part related to the normalization of the platon models to the HST/WFC3 data. Another reason for this deviation is, as we mention in Section 4.6.1, the platon models we use represent only spot or only faculae dominated photosphere. Either a mixture of both spot and faculae on the photosphere or a change in normalization of the models, or both can change the nature of the linear trend between the transmission spectral slopes and offsets.

Following an empirical approach we choose to use the $T_{D,ref}$ measured from the data. We first apply a constant offset to the GMOS transmission spectrum from each epoch with respect to $T_{D,ref}$ before weighted median combining them to obtain the combined transmission spectrum. The combined transmission spectrum after constant offset correction is shown as black points in Figure 4.8.

However, it is clear from the anticorrelation observed between the transmission spectra means and slopes in Figure 4.6 that just a constant offset correction is not enough as it does not remove the different slopes imparted by stellar variability at each epoch. Hence, instead of a constant offset correction, we introduce here a new empirical approach that also corrects for the slope. For a transmission spectrum we compute the difference between
Figure 4.5: Optical transmission spectra of WASP-19b from the individual 8 transit observations from Gemini/GMOS in this work in context of HST/WFC3 spectra from Huitson et al. (2013). Overplotted for comparison are platon forward models matching the water absorption feature in HST/WFC3 for a cloud free atmosphere with solar metallicity and C/O and including the mean (dashed lines) and 1σ range (shaded) of the correction factor due to unocculted high contrast spots (in blue), low contrast spots (red) and faculae (yellow) as described in more detail in Section 4.6.1. All the platon forward models here have been normalized to the HST/WFC3 observations.
4.6 Discussion

Figure 4.6: Slope of the transmission spectra (X axis) obtained from 8 GMOS transit observations vs the mean level of each transmission spectra (Y axis). We observe an anticorrelation between the slope and the transmission spectral mean level with the corresponding Pearson correlation coefficient of $-0.61$ at 2-tailed $p$ value of 0.1. The linear fit to the slope and transmission spectra mean is shown as black dashed line along with grey lines showing fits from random samples from emcee posterior of the linear fit. The black dot-dash lines show the projection of the linear fit we use to obtain $T_{D_{\text{ref}}}$ in Section 4.6.2. Overplotted in shaded regions are the $\pm 1\sigma$ (dark shaded) and $\pm 2\sigma$ (faint shaded) slope vs transmission spectrum mean derived from the platon models plotted in Figure 4.5 which account for the effect of unocculted high contrast spots (blue), low contrast spots (red) and faculae (yellow).
Figure 4.7: Best fit TESS transit depths vs median out of transit flux for the 58 transits observed by TESS in Sector 9 (left panel) and Sector 36 (right panel). The transit depths were obtained by fitting the individual transit light curves as described in Section 4.4.1. The median out of transit light curve flux was measured after normalizing the two orbits of each sector by their respective median flux to mitigate any effect of the systematic offset between the orbits of individual sector. The dashed black line shows the linear fit to the points and the orange lines show randomly drawn samples from the MCMC posteriors of the linear fit. In the inset for each sector are the Pearson correlation coefficients and two tailed p values for the correlation between the transit depths and out of transit flux.
its linear fit and the TD$_{ref}$ for each wavelength bin. The wavelength dependent offset hence obtained for each epoch can now be used to correct both the slope and offset in a transmission spectrum. We apply the wavelength-dependent offset to the transmission spectra at each epoch and then weighted median combine the slope corrected spectra to obtain a combined transmission spectrum, shown in Figure 4.8 as green points.

We emphasize that there are some major caveats to this empirical approach of slope and offset correction. Our approach is agnostic to the spectral slope present in the spectra due to the planetary atmosphere itself. If a spectral slope intrinsically due to the planetary atmosphere exists (e.g. due to haze scattering), it would vanish after our slope correction. Hence, our approach cannot resolve the discrepancy between the Magellan/IMACS and VLT/FORS2 data with respect to presence or absence of hazes. Moreover, given the wavelength coverage of the GMOS-R150 data, the GMOS transmission spectrum in this work is less sensitive to a slope due to hazes which impart a much stronger signature blueward of 400 nm. However, what our approach of slope and offset correction preserves is any prominent spectral features in the individual transmission spectra. In other words, our slope correction would remove any planetary atmospheric slope, but will retain spectral features, e.g. due to Na and K or TiO/VO molecular bandheads if present. Especially in the wavelength range of 520 to 720 nm probed by our GMOS-R150 observations, we expect a stronger contribution from spectral features from Na/K or TiO/VO as compared to scattering due to hazes. We next compare the stellar variability corrected (for both slope and offset) combined GMOS transmission spectrum with atmospheric forward models. Another caveat of our approach is that the linear approximation of the impact of unocculted spots in terms of an offset and slope works well given the precision and wavelength span of our data. McCullough et al. (2014) show that the impact of spots when modelling the spot and quiescent photosphere spectra as blackbodies is in general non-linear with respect to wavelength. Hence, we recommend exploring other parametric approximations e.g. a quadratic polynomial, which might be better suited for modelling the effect of unocculted stellar spots in the transmission spectra obtained from datasets spanning different wavelength ranges and precisions. In summary, we recommend correcting the impact of heterogeneous stellar photosphere in the transmission spectrum at each individual epoch before combining them.

4.6.3 Optical and near-infrared transmission spectrum of WASP-19 and comparison with forward atmospheric models

We now discuss the combined GMOS transmission spectrum which has been obtained after correcting for stellar variability in conjunction with the HST/WFC3 spectrum and their comparison to the forward models for WASP-19b’s atmosphere computed using platon. We restrict our comparison to the 520 to 720 nm range for the GMOS transmission spectrum as the only data points we have beyond 720 nm are from observation 8 with much larger uncertainties for any meaningful model comparison. We expect that given the limited wavelength range and resolution of the transmission spectrum per epoch, an atmospheric retrieval wouldn’t be able to meaningfully resolve the degeneracies between the contribution due to stellar variability which causes the offset between the GMOS and HST/WFC3 data, and that
Figure 4.8: Comparison of transmission spectrum of WASP-19b obtained from weighted median combining the transmission spectrum from 8 epochs in three different ways described in more detail in Section 4.6.2 and shown here in top three panels: 1) without applying any slope or offset corrections to the individual epochs, 2) applying a constant offset correction to the individual epochs, and 3) applying a wavelength dependent offset (slope) correction to the individual epochs. Coloured points in all top three panels show the transmission spectrum measured at each epoch and the weighted median combined spectrum in black points. The combined transmission spectrum from 1), 2), and 3) are overplotted for comparison in the bottom panel in black, blue, and green points respectively. Note that not applying any stellar variability corrections leads to spurious features in the transmission spectrum as seen in the black points, which are corrected to some degree by constant offset correction as seen in the blue points, and to a better degree with wavelength dependent offset correction as seen in the green points.
from the planetary atmosphere. Hence, we choose to perform only forward model comparisons which we expect are sufficient for our goal of testing the presence or absence of TiO and Na features.

We normalize the HST/WFC3 transmission spectrum from Huitson et al. (2013) to the $\text{TD}_{\text{ref}}$ calculated in Section 4.6.2. We construct forward models using platon for five different cases, each with equilibrium chemistry and solar C/O: 1) Solar metallicity, 2) Solar metallicity and TiO abundance suppressed $100 \times$, 3) Solar metallicity and TiO abundance suppressed $1000 \times$ 4) Solar metallicity and no TiO, and 5) Solar metallicity, no TiO, and no Na. We apply a similar treatment of slope and offset removal to all the models in the GMOS bandpass as done for the GMOS transmission spectrum in Section 4.6.2. For each platon model, we perform a linear fit to the model in the 520 to 720 nm range and calculate wavelength-dependent offsets with respect to the median of the model. We then apply these wavelength dependent offsets to the models additively in exactly the same manner as done for the GMOS spectra. We subsequently use these slope corrected transmission models for comparison with the observed transmission spectrum.

Since the HST/WFC3 spectrum was obtained at a different epoch relative to the GMOS data, the stellar activity level is likely different between these epochs. Therefore, the relative offset between the GMOS and HST/WFC3 spectrum is arbitrary and needs to be accounted for when comparing the GMOS and HST/WFC3 spectrum together with the forward models. We leverage the shape of the spectral features in HST/WFC3 transmission spectra for comparison with the models. Each of the models we consider are consistent with the shape of the HST/WFC3 water absorption spectral feature from Huitson et al. (2013). Hence, we first anchor all the platon models to match the HST/WFC3 points. Next, to compare the GMOS spectrum with the models, we apply a range of constant offsets (with respect to wavelength) in steps of 10 ppm to compute the minimum reduced chi-squared ($\chi^2_{\nu}$) between the GMOS spectrum and each model. Considering 11 degrees of freedoms (10 data points and 1 vertical direction offset), we find the minimum $\chi^2_{\nu}$ values for the five forward models as shown in Table 4.7 which we further use for model comparison.

Based on the $\chi^2_{\nu}$, we rule out solar metallicity atmosphere with solar TiO abundance as compared to no TiO case at 5$\sigma$. A solar metallicity atmosphere with $1000 \times$ or completely depleted TiO best explains the shape of the GMOS transmission spectrum. This is 10 times lower TiO abundance reported from the FORS2 observations by Sedaghati et al. (2021). The TiO depletion could be because of cold-trapping processes condensing TiO at the terminator as discussed by Parmentier et al. (2013) which could also explain the non-detection of Fe by Sedaghati et al. (2021). As compared to models with no Na, we favour the models with solar abundance Na by 3$\sigma$ when considering the transmission spectrum for smaller 10 nm wide bins near the Na feature. A zoom of the models around the Na 589 nm doublet showing this tentative detection of Na is shown in Figure 4.10. Interestingly, Sedaghati et al. (2017) also obtain a 3.4$\sigma$ Na detection in the VLT/FORS2 spectra, however the amplitude of the tentative Na absorption we detect in the GMOS spectrum is smaller than the VLT/FORS2 spectrum by $\sim 500$ ppm.
Figure 4.9: The GMOS combined transmission spectrum of WASP-19b corrected for stellar variability and the HST/WFC3 transmission spectrum from Huitson et al. (2013) normalized to the median of the GMOS spectrum as compared to various platon forward transmission models binned in the GMOS and HST/WFC3 wavelength bins. We also show the higher resolution platon forward models for the 'No TiO' and 'Solar' scenarios. Similar to the stellar variability correction applied to the GMOS transmission spectrum, wavelength dependent offsets with respect to a linear fit only in the GMOS bandpass (520 to 720 nm) have been applied to the platon transmission models as well for a fair comparison. The HST/WFC3 transmission spectrum amplitude is consistent with solar H$_2$O abundance which is the same for all models plotted here. The shape of the GMOS spectrum is inconsistent with the shape predicted by the solar abundance models (dashed blue line) at more than 3$\sigma$ as compared to the models with suppressed TiO abundance, as discussed in more detail in Section 4.6.3.
Table 4.7: Model comparison criteria for arbitrarily offset GMOS data and the platon forward models shown for various cases shown here in the column ‘Model’ and plotted in Figure 4.9. The top and bottom part of the table are for transmission spectrum with 20 nm and 10 nm wide bins respectively. The columns ‘$\chi^2_{\nu}$’ and ‘BIC’ show the minimum reduced chi-squared and the corresponding Bayesian Information Criterion respectively obtained by offsetting the GMOS transmission spectrum by varying amounts with respect to the platon forward model normalized to HST/WFC3. ‘N σ...’ shows the number of sigmas by which each model is preferred over the Solar metallicity case. All the models have the metallicity and C/O ratio fixed to the solar value.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2_{\nu}$</th>
<th>N σ from Solar BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar metallicity</td>
<td>4.298</td>
<td>– 36.408</td>
</tr>
<tr>
<td>Solar metallicity and TiO abundance suppressed 100×</td>
<td>1.794 4</td>
<td>21.189</td>
</tr>
<tr>
<td>Solar metallicity and TiO abundance suppressed 1000×</td>
<td>0.954 4.5</td>
<td>18.264</td>
</tr>
<tr>
<td>Solar metallicity and no TiO</td>
<td>0.876 5</td>
<td>17.869</td>
</tr>
<tr>
<td>Solar metallicity, no TiO, and no Na</td>
<td>1.654 3</td>
<td>23.995</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2_{\nu}$</th>
<th>N σ from Solar BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10 nm) Solar metallicity and no TiO</td>
<td>2.771 –</td>
<td>49.873</td>
</tr>
<tr>
<td>(10 nm) Solar metallicity, no TiO, and no Na</td>
<td>3.380 3</td>
<td>63.800</td>
</tr>
</tbody>
</table>

Figure 4.10: A zoom-in of the GMOS combined transmission spectrum around the Na doublet at 0.589 μm. The green points and black points show the transmission spectrum computed for 20 nm and 10 nm wide bins respectively. We find that the 10 nm bin GMOS transmission spectrum favours a model with Na at $\sim3\sigma$ as compared to models with no Na indicating a tentative detection of Na absorption. We show the platon forward models used for comparison here in red for the model with Na and in yellow for the model with Na (thicker line showing the binned spectrum and thinner line showing the higher resolution model).
4.6.4 Comparison of the GMOS transmission spectrum of WASP-19b with previous studies

In this section, we compare our GMOS transmission spectrum of WASP-19b with that from Magellan/IMACS and VLT/FORS2. We first briefly summarize the results on the optical transmission spectrum of WASP-19b from two previous studies. The VLT/FORS2 transmission spectrum of WASP-19b from Sedaghati et al. (2017) was obtained by combining 3 transits observed at 3 different epochs. They detect significant features due to absorption by Na, H$_2$O, and substantially sub-solar TiO abundance, and a steep slope due to haze scattering towards the blue optical end of the spectrum. With further reanalysis of the same dataset using retrieval models accounting for the effect of stellar activity Sedaghati et al. (2021) arrive at a similar detection of 100× sub-solar TiO albeit at a lower significance (7.7σ → 4.7σ) of detection as compared to Sedaghati et al. (2017). From their ESPRESSO high resolution spectroscopy observations Sedaghati et al. (2021) find a tentative 3σ detection of TiO consistent with 1000× sub-solar TiO abundance. However, as also noted by Sedaghati et al. (2021), this isn’t a confirmation of the VLT/FORS2 result due to low statistical significance.

In a separate study of WASP-19b, Espinoza et al. (2019) observe 6 transits in the optical using Magellan/IMACS and do not find any signatures of absorption due to TiO, Na, or a scattering slope towards blue optical. Espinoza et al. (2019) observe spots and faculae crossings in two of their observations, along with a slope in the transmission spectrum from one of their transits attributed to the transit light source effect due to unocculted stellar spots. Espinoza et al. (2019) exclude the spectrum with the steep slope and construct the combined transmission spectrum from the other five epochs. Using a semi-analytical retrieval, Espinoza et al. (2019) conclude that the combined Magellan/IMACS transmission along with HST/WFC3 and Spitzer measurements are best explained by an atmosphere with solar composition water and sub-solar TiO and Na.

We suggest that tension between the Magellan/IMACS spectrum (Espinoza et al. (2019)) and the VLT/FORS2 spectrum (Sedaghati et al., (2017, 2021)), can be explained by the stellar variability of the host star. Espinoza et al. (2019) and Sedaghati et al., (2017, 2021) study transit observations taken at different epochs with likely varying levels of stellar variability. Note that given the wider wavelength range coverage of both Magellan/IMACS and VLT/FORS2 observations as compared to our GMOS observations, both studies account for the stellar variability contribution at individual epochs using a Bayesian retrieval that fits for both planetary atmospheric parameters and stellar photospheric heterogeneity parameters. Espinoza et al. (2019) find a statistically stronger contribution from stellar activity compared to a flat line in only one out of their six epochs which they eventually omit when constructing the final combined transmission spectrum using a constant offset correction. Sedaghati et al. (2021) from their POSEIDON retrieval of the three VLT/FORS2 epochs independently as well as jointly confirm the significant contribution from the stellar variability.

Our GMOS observations exceed in both the telescope collecting area and the total number of epochs of the planet probed as compared to Espinoza et al. (2019) and Sedaghati et al. (2017, 2021). We obtain a median precision of 100 ppm per 20 nm bins, compared to 250
ppm per 10 nm bins for both Magellan/IMACS and VLT/FORS2 transmission spectra. Despite the smaller wavelength coverage, our GMOS observations spread over eight epochs probe a larger range of stellar variability, and at higher precision on the transmission spectrum both at individual epochs and the combined transmission spectrum. We leverage these aspects in the application of our empirical approach of applying relative corrections to individual epochs as described in Section 4.6.2. We compare our stellar variability corrected spectrum with the Magellan/IMACS and VLT/FORS2 spectrum in Figure 4.11. We observe that the GMOS spectrum without stellar correction (black points in Figure 4.11) shows a feature in the 550 to 650 nm range at similar amplitudes to that seen in Sedaghati et al. (2017). However, the feature diminishes significantly when we correct for stellar variability as seen in the green points in Figure 4.11.

From our stellar variability corrected spectrum, we conclude that we do not observe significant absorption features due to TiO, which is consistent with the findings of Espinoza et al. (2019) from their Magellan/IMACS spectrum. However, our tentative detection of Na absorption is inconsistent with Espinoza et al. (2019) and more consistent with Sedaghati et al. (2017). Given our weak evidence for Na absorption, this doesn’t confidently resolve the tension between Espinoza et al. (2019) and Sedaghati et al. (2017) with respect to Na. Additionally, our GMOS spectrum favours 1000× or lower sub-solar TiO scenarios and rules out the solar TiO scenario (see Section 4.6.3) at high level of significance. This is consistent with the findings of both Espinoza et al. (2019) and Sedaghati et al., (2017, 2021) who also rule out the solar TiO scenario from their Magellan/IMACS and VLT/FORS2 spectrum respectively. Interestingly, similar to our GMOS data, the ESPRESSO high resolution observation of Sedaghati et al. (2021) is also consistent with 1000× sub-solar TiO. We emphasize that despite the differences in spectral morphology, the depletion of TiO at the terminator of WASP-19b is a scenario that explains the transmission spectrum from the GMOS, Magellan/IMACS, VLT/FORS2, and ESPRESSO observations.

4.7 Conclusions

We present the optical transmission spectrum for the ultra-hot Jupiter WASP-19b from eight transits of the planet observed using GMOS on Gemini South spread over a span of two years. The main conclusions from our study are:

1) To extract the transmission spectrum of WASP-19b, we mitigate the effects of a fainter comparison star using the method to analyse MOS data that is presented in Panwar et al. (2022b). Using this method, we obtain an average factor of three improvement in the RMS of the spectroscopic light curves compared to the conventional method. We measure the transmission spectra without introducing additional uncertainties from a faint comparison star. We use a Bayesian framework for propagating uncertainties when modelling the systematics in the target star light curves.

2) We find that the transmission spectra of WASP-19b obtained at different epochs vary suggestively in terms of their slopes and relative offsets which impedes the process of combining them to construct the final transmission spectrum. We interpret this as a result of the impact
Figure 4.11: Comparison of the mean subtracted combined GMOS transmission spectrum with and without any stellar variability corrections (black, blue, and green points) of WASP-19b with the mean subtracted combined transmission spectrum from Magellan/IMACS reported by Espinoza et al. (2019) (red points) and VLT/FORS2 reported by Sedaghati et al. (2017) (grey points). The stellar variability corrected GMOS spectrum (green and blue points) does not show any significant TiO absorption features in the 550 to 720 nm which is consistent with the Magellan/IMACS spectrum but inconsistent with the the VLT/FORS2 spectrum. The Gemini/GMOS spectra shows tentative evidence of Na absorption at lower amplitude as compared to Sedaghati et al. (2017). However, the Gemini/GMOS, Magellan/IMACS, and VLT/FORS2 spectra all rule out the presence of solar TiO and are consistent with sub-solar TiO scenario at the terminator of WASP-19b.
of stellar variability on the transmission spectra at individual epochs. WASP-19b orbits an active solar type star which shows a stellar flux variability of 2% in the optical as confirmed by both TESS and ground-based broadband photometry. Hence, relative corrections for stellar variability at each epoch need to be applied before constructing the combined transmission spectrum.

3) We observe that the effect of stellar variability manifests broadly in two ways: a slope and an offset to the transmission spectra, both of which need to be measured and accounted for when co-adding multiple epochs. We compute these effects for WASP-19b’s transmission spectrum using the spot (positive slope) and faculae (negative slope) temperature contrasts and covering fractions corresponding to the host star’s amplitude of variability as measured by previous studies.

4) For the eight GMOS observations of WASP-19b presented in this paper, the offsets between the transmission spectra from each epoch span a range of $\sim$4000 ppm, and a slope -0.4 to 0.6 ppm per Angstrom. The trend in the transmission spectra slope vs offset broadly matches in amplitude and sign that predicted by forward atmospheric models accounting for the effect of stellar spots and faculae.

5) We introduce a new empirical approach for correcting the stellar variability in transmission spectra across multiple epochs by using the measured spectral slopes and offsets to apply relative corrections between epochs and to construct the combined transmission spectrum.

6) Our stellar variability corrected GMOS spectrum rules out the solar TiO scenario at $5\sigma$, and is most consistent with the 1000× sub-solar TiO or lower TiO abundance scenario. Significant depletion of TiO could point towards condensation or cold trapping of TiO at the terminator as predicted by Parmentier et al. (2013).

7) After accounting for different bin sizes, we obtain on average $\sim$40 % better precision in the GMOS spectrum compared to the previous MOS optical transmission spectrum observed by Espinoza et al. (2019) and Sedaghati et al. (2017). In terms of the spectral morphology, our non-detection of TiO features is consistent with the findings of Espinoza et al. (2019) but inconsistent with Sedaghati et al. (2017). We tentatively detect Na absorption albeit at lower amplitude than that detected by Sedaghati et al. (2017), which is inconsistent with Espinoza et al. (2019). However, given the weak evidence for Na absorption in the Gemini/GMOS spectrum, we cannot definitively resolve the tension with respect to Na detection between Sedaghati et al. (2017) and Espinoza et al. (2019). The Gemini/GMOS transmission spectrum is overall consistent with the sub-solar TiO scenario which is also reported by the previous studies from Espinoza et al. (2019) and Sedaghati et al., (2017, 2021) using Magellan/IMACS, VLT/FORS2, and ESPRESSO respectively.

Our work ultimately demonstrates that multi-epoch transmission spectra from exoplanet transiting variable stars cannot be simply co-added or combined to improve the precision on the final transmission spectrum before correcting them for stellar variability effects. The method to correct for the effect of multi-epoch stellar variability introduced in this paper becomes even more relevant for high precision observations from JWST (Zellem et al. 2017, Mayorga et al. 2021). The issue for the near-infrared JWST observations will be most sig-
significant for active stars with high level of stellar variability. However, JWST NIRISS observations, which will go up to \( \sim 600 \) nm in the blue optical, will have to correct for stellar variability effect before combining observations taken at different epochs.

4.8 Acknowledgements

Based on observations obtained at the Gemini Observatory (acquired through the Gemini Observatory Archive and Gemini Science Archive), which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), Ministério da Ciência, Tecnologia e Inovação (Brazil), and Korea Astronomy and Space Science Institute (Republic of Korea). Based in part on Gemini observations obtained from the National Optical Astronomy Observatory (NOAO) Prop. ID: 2012B-0398; PI: J.-M Désert. We are very grateful to the anonymous reviewer for their careful and thorough feedback which greatly improved this work. V.P. is grateful to Filipe Matos for helping with the raw data reduction of the LCOGT light curves. V.P. acknowledges stimulating discussions with Lorenzo Pino and Jacob Arcangeli on ultra-hot Jupiters. V.P. acknowledges help from Ben Montet on using the eleanor package. J.M.D acknowledges support from the Amsterdam Academic Alliance (AAA) Program, and the European Research Council (ERC) European Union’s Horizon 2020 research and innovation programme (grant agreement no. 679633; Exo-Atmos). This work is part of the research programme VIDI New Frontiers in Exoplanetary Climatology with project number 614.001.601, which is (partly) financed by the Dutch Research Council (NWO). This material is based upon work supported by the NWO TOP Grant Module 2 (Project Number 614.001.601). This material is based upon work supported by the National Science Foundation (NSF) under Grant No. AST-1413663. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. This research has made use of NASA’s Astrophysics Data System. The authors also acknowledge the significant cultural role and reverence the summit of Mauna Kea has within the indigenous Hawaiian community. This research has made use of Astropy,\(^1\) a community-developed core Python package for Astronomy Astropy Collaboration et al. (2013, 2018), NumPy Harris et al. (2020), matplotlib Hunter (2007), SciPy Virtanen et al. (2020) and IRAF Tody (1986) distributed by the NOAO, which is operated by AURA under a cooperative agreement with the NSF.

\(^1\)http://www.astropy.org
Appendices

4.A Photometric monitoring of WASP-19

4.A.1 TESS observations of WASP-19b

The TESS spacecraft observed WASP-19 (TIC 35516889, TOI 655) in Sector 9 (from 28 February 2019 to 26 March 2019) and Sector 36 (from 7 March 2021 to 2 April 2021) covering a total of 59 transits of WASP-19b (Figure 4.12). TESS photometric observations are obtained in the broadband wavelength range of 600 to 1000 nm which overlaps with the GMOS-R150 wavelength range of 500 to 900 nm covered by our WASP-19 observations. Hence, TESS photometric light curves obtained over multiple epochs can be used to probe the effect of stellar variability on the broadband transit depth of WASP-19b in the GMOS bandpass. Moreover, the large number of transits observed by TESS can be used to benchmark the transit parameters of the system.

For the analysis in this paper we obtained the Simple Aperture Photometry (SAP) TESS light curves for both sector 9 and 36 extracted using an optimal aperture size computed by the Science Processing Operations Center (SPOC) pipeline (Jenkins et al. 2016) and publicly available on the Mikulski Archive for Space Telescopes (MAST). As also previously noted by Wong et al. (2020) who used the Sector 9 data, the Presearch Data Conditioned (PDC) light curves made available by SPOC for WASP-19 introduce correlated features in the light curve which are not originally seen in the SAP light curve.

A recent update to the TESS SPOC pipeline as described in ? implemented better sky subtraction which overcomes the overestimation of sky background which could lead to overestimation of the measured transit depth. MAST provides reduced light curves for Sector 9 from both the old and the latest version of the TESS SPOC pipeline, and for Sector 36 only from the latest version of the pipeline. We find that measured flux in the Sector 9 SAP light curves from the old and new version of the pipelines differ by on average 8 % as shown in Figure 11 in the supplementary material. We also find that the light curves from the old TESS SPOC pipeline overestimate the transit depth of WASP-19b by 1500 ppm on average. Hence, we choose to use the SAP light curves from the latest version of the TESS SPOC pipeline for our analysis in this paper, which we describe in further detail.
Figure 4.12: TESS Simple Aperture Photometry (SAP) light curve of WASP-19 observed in Sector 9 (in the left panel) and Sector 36 (in the right panel) obtained from the latest version of the TESS SPOC pipeline from MAST. Both sectors have been corrected for dilution and bad quality exposures have been masked out (see Section 4.A.1.1 for more detail). We have also normalized the two orbits for each sector by their respective median fluxes. The transits are marked by red points. The stellar flux varies by \(~ 2\%\) peak to trough in both the sectors, with the Lomb-Scargle periodogram (inset) of the out of transit flux times series showing a peak at 11.4 days for Sector 9 and 10.13 days for Sector 36. The orange line shows the sinusoidal fit to the photometry corresponding to the peak of the Lomb Scargle periodogram for each sector.
4.A Photometric monitoring of WASP-19

4.A.1.1 Analysis of TESS Light Curves of WASP-19b

The range of TESS photometry of WASP-19 is 27 days for individual sectors which is just over two stellar rotational cycles ($P_{\text{rot}} \sim 10.5$ days) and is evident as spot modulated stellar flux variability of $\sim 2\%$ peak to trough. We first mask out the bad quality exposures using the one-hot encoded quality mask in the 'QUALITY' keyword in the header of the light curve files provided by SPOC (Jenkins et al. 2016). After masking out the bad quality exposures, one transit in sector 36 is masked out and hence we have 58 transits in total from both sectors for our final analysis. We then use the ‘CROWDSAP’ keyword from the header for each sector light curve file to get an estimate of the ratio of target flux to total flux in optimal aperture used for the SAP photometry. This value can be used to subtract the dilution from nearby sources. The dilution flux we subtracted this way was 10.73\% and 7.8\% of the median measured flux for Sector 9 and Sector 36 respectively.

We additionally clip any remaining outliers in SAP light curve at more than 3 $\sigma$ using a moving box average before fitting individual transit light curves. We sliced the SAP light curves for WASP-19 into 58 individual transit light curves manually with approximately 4 hours before and after transit, to provide enough out of transit baseline to fit the transit signal. We then fit each light curve with the combined transit and a GP noise model similar to that used for the GMOS transit light curve in Section 4.5.2 with time stamps of each exposure as a GP regressor. We fit for the orbital inclination ($i$), normalized orbital separation ($a/R_*$), central transit time ($T_0$), planet to star radius ratio ($R_p/R_*$), and a linear limb darkening parameter. We used wide uniform priors for $R_p/R_*$, mid-transit time ($T_0$), semi major axis ($a/R_*$), and inclination ($i$), and a Gaussian prior on the linear limb darkening coefficient with mean and variance fixed by the theoretically computed values from PyLDTk for the TESS bandpass (600 - 1000 nm).

The best fit TESS light curves are shown in Figures 13 and 14 in the supplementary material. The weighted average value of inclination measured from both the sectors of TESS light curves is about $5\sigma$ different from the most precise literature value reported by Espinoza et al. (2019). In order to ensure that variation of the measured $a/R_*$ and $i$ with each transit do not affect the measured transit depths, we also conducted light curve fits by fixing the inclination to the value measured by Espinoza et al. (2019). We find that the transit depths from both the cases, whether we fit for the inclination or fix it, are consistent with each other within 1 $\sigma$ as seen in Figure 12 in the supplementary material.

4.A.1.2 TESS Light Curves of the comparison stars

We also inspect the TESS light curves of the comparison stars observed by Gemini/GMOS in this paper. We obtained the TESS SPOC light curves for comparison star 2 (TIC 35516889) from MAST. Since the comparison star 1 (TIC 35516848) is faint, its light curves from the QLP pipeline (Huang et al. (2020)) available on MAST are significantly contaminated by the flux of WASP-19. Hence, we derived the light curves for the comparison star 1 by obtaining the full-frame images obtained from MAST and analysing them using the eleanor package.
Multi-epoch stellar variability effect in the transmission spectrum of WASP-19b

(Feinstein et al. (2019)). We show the light curves and the Lomb Scargle periodograms for both comparison stars in Figure 9 in the supplementary material. After independently normalizing the two TESS orbits within each sector to the median flux in the orbit, we find that the comparison star 1 shows $\pm 0.2\%$ and $\pm 0.6\%$ variability in Sector 9 and 36 respectively, while the comparison star 2 shows $\pm 0.1\%$ and $\pm 0.2\%$ variability in sector 9 and 36 respectively. From their Lomb Scargle periodograms, the light curves for comparison star 1 show periodicity at 11 and 10.74 days in sector 9 and 36 respectively as seen in Figure 11 in supplementary material. Similarly, comparison star 2 shows periodicity at 12.72 and 13.48 days for sector 9 and 36 respectively. For comparison star 2 this could represent the actual period of its variability as it is brighter than WASP-19. However, for the comparison star 1, which is fainter than both the other stars, we cannot rule out the possibility that the observed periodicity could stem from low level of contamination from either or both the other stars.

4.4.2 Ground based photometric monitoring of WASP-19 by LCOGT

We obtained broadband long-term monitoring photometric observations of WASP-19 in the Johnson Cousins/Bessell B and R band using the Las Cumbres Observatory Global Telescope (LCOGT) network of robotic telescopes (Brown et al. 2013) to monitor the photometric variability of WASP-19 due to stellar activity. The LCO network consists of 42 telescopes with mirrors with diameters of 40 cm, 1 m, and 2 m spread across Earth in latitude and longitude, providing full sky coverage. We used the 40 cm and 1 m telescopes with the SBIG 4k x 4k to obtain 303 B-band photometric observations from 2014 May 5 to 2014 July 13, and 234 B-band and 190 R-band photometric observations from 2016 March 10 to 2017 January 26.

We used the scientific data outputs from the BANZAI and ORAC pipelines of LCOGT on which we further perform WCS correction, centroid fitting, and simple aperture photometry using a custom pipeline that uses modules from the Astropy package (Astropy Collaboration et al. 2013, Astropy Collaboration et al. 2018). For performing differential photometry we first select the comparison stars to use according to their magnitude in the B band, choosing the ones with $\Delta B_{mag} \leq 1$ and within 5" from WASP-19. We then choose the combination of comparison stars that results in a minimum scatter in the light curve with respect to the median. We do not notice any significant correlation between the differential flux and airmass for both seasons and choose not to perform any airmass correction.

The extracted LCOGT photometry of WASP-19 shows peak to trough variation of $\sim 2\%$ during both seasons. We find period of flux variations as 10.47 and 10.29 days for the two seasons from their respective Lomb-Scargle periodograms as shown in Figure 4.13. We also fit the LCOGT photometry using a GP model with a quasi-periodic exponential sine-squared kernel implemented in george (Ambikasaran et al. 2015), which is also plotted in Figure 4.13. We find the ‘Period’ hyperparameter from the GP fit to be $10.91^{+0.1}_{-0.09}$ days which is consistent with the period of stellar flux variability measured from the Lomb-Scargle periodograms.
Figure 4.13: **Upper panel**: Photometric monitoring of WASP-19b obtained from the LCOGT network of telescopes in B band in two seasons (Upper left is Season 1, and upper right is Season 2). Both the seasons have been normalized with respect to their respective seasonal median. The red curve shows the best-fit obtained from the quasi-periodic Gaussian process regression model for both seasons combined which indicates quasi-periodic variations in stellar flux due to stellar rotation. Vertical lines mark the dates of our Gemini/GMOS observations (subscript indicating the observation number as specified in Table 4.2). The period estimated from the best fit quasi-periodic Gaussian process regression model is \( P_{\text{rot}} = 10.91 + 0.10 \) days which is close to that found from the Lomb Scargle periodograms of the corresponding seasons and to that from TESS photometry in this paper and previous ground based photometry (Espinoza et al. 2019). The horizontal dotted lines in the LS periodograms in the **lower panel** indicate the 0.1, 0.05, and 0.01 False Alarm Probability levels, and the vertical dashed red line mark the location of maximum normalized power.
4.B Illustration of the transmission spectral slope vs offset analysis

In Figure 4.14 we visually illustrate how we construct the transmission spectra mean level vs slope space (shown in Figure 4.6) to measure the relative effects of stellar variability on the transmission spectra obtained at different epochs.

4.C Transmission spectra tables

In this appendix we present the tables of transmission spectra obtained in Section 4.5.3. Table 4.8 shows the transmission spectra from the 8 transits obtained using the conventional method described in Section 4.5.3.1 and Table 4.9 shows the transmission spectra obtained using the new method in Section 4.5.3.2. The spectroscopic light curve fits from both the new and the conventional methods are shown in Figure 3 to 10 in the supplementary material. Table 4.11 shows the median combined transmission spectrum after applying the slope and offset correction in Section 4.6.2.

4.D Supplementary material
Figure 4.14: Illustration of the construction of transmission spectra mean level vs slope space presented in Figure 4.6. The black curve on the right panel shows platon transmission spectrum model without any stellar spot or faculae contribution and normalized to the HST/WFC3 spectrum from Huitson et al. (2013). In black and green points on the right panel are shown the transmission spectra for observations 3 and 7 respectively. Overplotted are the linear fits to both the spectra along with randomly sampled fits from the emcee posterior for the fit. We plot the mean of the two spectra from observation 3 and 7 vs their measured linear slope on the left panel (shown by the black arrows). We repeat the process for all the eight epochs to populate the spectral mean vs slope space as shown here as black points in the left panel which is the same as Figure 4.6.
Table 4.8: Wavelength dependent transit depths ($\delta(\lambda)$ in ppm) for the individual GMOS-R150 observations (marked in the columns) obtained using the conventional method described in more detail in Section 4.5.3.1.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>5200 - 5400</td>
<td>19582 ± 505</td>
<td>20831 ± 766</td>
<td>19588 ± 570</td>
<td>18685 ± 497</td>
<td>19576 ± 552</td>
<td>21251 ± 497</td>
<td>21143 ± 767</td>
<td>-</td>
</tr>
<tr>
<td>5400 - 5600</td>
<td>18694 ± 364</td>
<td>20821 ± 275</td>
<td>19443 ± 276</td>
<td>19269 ± 287</td>
<td>19924 ± 371</td>
<td>20522 ± 901</td>
<td>22418 ± 447</td>
<td>18876 ± 986</td>
</tr>
<tr>
<td>5600 - 5800</td>
<td>19781 ± 250</td>
<td>20913 ± 255</td>
<td>19232 ± 220</td>
<td>19022 ± 267</td>
<td>20197 ± 237</td>
<td>21330 ± 287</td>
<td>21091 ± 363</td>
<td>19064 ± 743</td>
</tr>
<tr>
<td>5800 - 6000</td>
<td>19975 ± 257</td>
<td>20712 ± 228</td>
<td>19386 ± 237</td>
<td>19620 ± 216</td>
<td>19839 ± 197</td>
<td>20587 ± 342</td>
<td>20123 ± 1049</td>
<td>17760 ± 765</td>
</tr>
<tr>
<td>6000 - 6200</td>
<td>19890 ± 278</td>
<td>21453 ± 349</td>
<td>19890 ± 191</td>
<td>19620 ± 199</td>
<td>20082 ± 242</td>
<td>20919 ± 324</td>
<td>21557 ± 377</td>
<td>18060 ± 975</td>
</tr>
<tr>
<td>6200 - 6400</td>
<td>20047 ± 240</td>
<td>21329 ± 213</td>
<td>19773 ± 169</td>
<td>19326 ± 203</td>
<td>20260 ± 270</td>
<td>20614 ± 511</td>
<td>20786 ± 508</td>
<td>16215 ± 935</td>
</tr>
<tr>
<td>6400 - 6600</td>
<td>19672 ± 452</td>
<td>20805 ± 211</td>
<td>19063 ± 1042</td>
<td>19465 ± 298</td>
<td>20344 ± 173</td>
<td>20785 ± 248</td>
<td>20972 ± 465</td>
<td>16657 ± 953</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>19611 ± 212</td>
<td>21082 ± 222</td>
<td>19470 ± 190</td>
<td>19652 ± 255</td>
<td>20241 ± 174</td>
<td>20709 ± 521</td>
<td>20631 ± 543</td>
<td>19570 ± 1445</td>
</tr>
<tr>
<td>6800 - 7000</td>
<td>19708 ± 271</td>
<td>21491 ± 309</td>
<td>19945 ± 196</td>
<td>19336 ± 799</td>
<td>19794 ± 207</td>
<td>20500 ± 611</td>
<td>21142 ± 428</td>
<td>18182 ± 683</td>
</tr>
<tr>
<td>7000 - 7200</td>
<td>19696 ± 212</td>
<td>20714 ± 1033</td>
<td>19391 ± 169</td>
<td>19161 ± 601</td>
<td>20241 ± 165</td>
<td>20312 ± 196</td>
<td>21215 ± 464</td>
<td>19424 ± 647</td>
</tr>
<tr>
<td>7200 - 7400</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19664 ± 543</td>
</tr>
<tr>
<td>8199 - 8401</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19921 ± 692</td>
</tr>
<tr>
<td>8401 - 8600</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19860 ± 836</td>
</tr>
<tr>
<td>8600 - 8800</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>18842 ± 711</td>
</tr>
<tr>
<td>8800 - 8999</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>19345 ± 900</td>
</tr>
</tbody>
</table>
Table 4.9: Wavelength dependent transit depths ($\delta(\lambda)$ in ppm) for the individual GMOS-R150 observations (marked in the columns) obtained using the new method introduced by P22 and described in Section 4.5.3.2.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>5200 - 5400</td>
<td>20414 ± 983</td>
<td>21446 ± 542</td>
<td>19418 ± 541</td>
<td>20228 ± 483</td>
<td>23945 ± 922</td>
<td>22099 ± 1079</td>
<td>22173 ± 547</td>
<td>22913 ± 1486</td>
</tr>
<tr>
<td>5400 - 5600</td>
<td>20927 ± 250</td>
<td>20901 ± 364</td>
<td>19451 ± 337</td>
<td>20057 ± 226</td>
<td>22365 ± 482</td>
<td>21217 ± 378</td>
<td>22294 ± 423</td>
<td>23280 ± 1287</td>
</tr>
<tr>
<td>5600 - 5800</td>
<td>21026 ± 335</td>
<td>20842 ± 297</td>
<td>19211 ± 240</td>
<td>20365 ± 316</td>
<td>22291 ± 391</td>
<td>21335 ± 371</td>
<td>22583 ± 596</td>
<td>23100 ± 957</td>
</tr>
<tr>
<td>5800 - 6000</td>
<td>21431 ± 87</td>
<td>20798 ± 205</td>
<td>19819 ± 187</td>
<td>20468 ± 249</td>
<td>22309 ± 440</td>
<td>21735 ± 313</td>
<td>22387 ± 409</td>
<td>22409 ± 714</td>
</tr>
<tr>
<td>6000 - 6200</td>
<td>20916 ± 401</td>
<td>21402 ± 205</td>
<td>19769 ± 167</td>
<td>20363 ± 263</td>
<td>22488 ± 179</td>
<td>21984 ± 414</td>
<td>22468 ± 395</td>
<td>21279 ± 769</td>
</tr>
<tr>
<td>6200 - 6400</td>
<td>20999 ± 218</td>
<td>21135 ± 200</td>
<td>19523 ± 235</td>
<td>21150 ± 176</td>
<td>22192 ± 309</td>
<td>21018 ± 301</td>
<td>22113 ± 305</td>
<td>21682 ± 460</td>
</tr>
<tr>
<td>6400 - 6600</td>
<td>21640 ± 815</td>
<td>20833 ± 286</td>
<td>20381 ± 382</td>
<td>21158 ± 208</td>
<td>22189 ± 177</td>
<td>21402 ± 161</td>
<td>23054 ± 578</td>
<td>21691 ± 283</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>20929 ± 297</td>
<td>21341 ± 236</td>
<td>20066 ± 357</td>
<td>20478 ± 268</td>
<td>22362 ± 357</td>
<td>21926 ± 389</td>
<td>22547 ± 656</td>
<td>21727 ± 258</td>
</tr>
<tr>
<td>6800 - 7000</td>
<td>21125 ± 227</td>
<td>21146 ± 206</td>
<td>19969 ± 330</td>
<td>20842 ± 222</td>
<td>23154 ± 410</td>
<td>21474 ± 229</td>
<td>22714 ± 438</td>
<td>21804 ± 132</td>
</tr>
<tr>
<td>7000 - 7200</td>
<td>21336 ± 294</td>
<td>21143 ± 516</td>
<td>20700 ± 783</td>
<td>21289 ± 301</td>
<td>22558 ± 558</td>
<td>21279 ± 499</td>
<td>22832 ± 661</td>
<td>22478 ± 699</td>
</tr>
<tr>
<td>7200 - 7400</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>22605 ± 1528</td>
</tr>
<tr>
<td>8199 - 8401</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>21162 ± 1799</td>
</tr>
<tr>
<td>8401 - 8600</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>21464 ± 1078</td>
</tr>
<tr>
<td>8600 - 8800</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>22006 ± 1199</td>
</tr>
<tr>
<td>8800 - 8999</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>22312 ± 1449</td>
</tr>
</tbody>
</table>
Figure 4.15: GMOS R150 transmission spectrum for each of the eight transits obtained using different methods to fit the spectroscopic light curves. The red, blue and green points are obtained by fitting the target spectroscopic light curves ($\lambda$LC) using the following GP regressor combinations respectively: common-mode and time, comparison $\lambda$LC and time, and comparison $\lambda$LC. The black points show the transmission spectrum obtained from the conventional method of fitting the target divided by the comparison star $\lambda$LC.
Table 4.10: Combined transmission spectrum (δ(λ) in ppm) obtained after applying the wavelength-dependent offset (slope) correction to the spectrum at each epoch measured using the new method, and then weighted median combining them, as described in more detail in Section 4.6.2.

<table>
<thead>
<tr>
<th>Wavelength [Å]</th>
<th>δ(λ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5200 - 5400</td>
<td>21812 ± 236</td>
</tr>
<tr>
<td>5400 - 5600</td>
<td>21477 ± 120</td>
</tr>
<tr>
<td>5600 - 5800</td>
<td>21522 ± 124</td>
</tr>
<tr>
<td>5800 - 6000</td>
<td>21831 ± 67</td>
</tr>
<tr>
<td>6000 - 6200</td>
<td>21745 ± 89</td>
</tr>
<tr>
<td>6200 - 6400</td>
<td>21497 ± 87</td>
</tr>
<tr>
<td>6400 - 6600</td>
<td>21626 ± 88</td>
</tr>
<tr>
<td>6600 - 6800</td>
<td>21664 ± 110</td>
</tr>
<tr>
<td>6800 - 7000</td>
<td>21690 ± 79</td>
</tr>
<tr>
<td>7000 - 7200</td>
<td>21842 ± 159</td>
</tr>
</tbody>
</table>
Figure 4.16: BIC comparison for each wavelength bin for the spectroscopic light curve fits corresponding to the transmission spectra obtained using the various GP regressor combinations of the new method as shown in Figure 1. The blue points show the $\Delta BIC$ for the comparison ALC and time combination with respect to common-mode and time combination. The green points show the $\Delta BIC$ for comparison ALC with respect to the common-mode and time combination. We find that the common-mode and time combination is the best combination with the lowest BIC among all methods. When using the comparison ALC as one of the GP regressors, using time as well as an additional regressor (blue points here) gives a better fit in terms of the BIC as seen in this figure.
Figure 4.17: Spectroscopic light curves for observation 1 fit using the conventional method (three panels on this page) of fitting the common-mode corrected Target/Comparison λLC as described in Section 5.3.1, and the new method (three panels in the continued Figure on next page) of fitting the Target λLCs using the common-mode trend as a GP regressor as described in Section 5.3.2. The leftmost panel for each method shows the best fit to the light curves for each wavelength bin, the middle panel shows the detrended light curves with their best fit transit models, and the rightmost panel shows the corresponding residuals, their histograms, and RMS of the residuals.
Figure 4.17: Continued.
Figure 4.18: Same as Figure 4.24 for observation 2.
Multi-epoch stellar variability effect in the transmission spectrum of WASP-19b

Figure 4.18: Continued.
Figure 4.19: Same as Figure 4.24 for observation 3.
Figure 4.19: Continued.
Figure 4.20: Same as Figure 4.24 for observation 4.
Figure 4.20: Continued.
Figure 4.21: Same as Figure 4.24 for observation 5.
Figure 4.21: Continued.
Figure 4.22: Same as Figure 4.24 for observation 6.
Figure 4.22: Continued.
Figure 4.23: Same as Figure 4.24 for observation 7.
Figure 4.23: Continued.
Figure 4.24: Same as Figure 4.24 for observation 8.
Figure 4.24: Continued.
Figure 4.25: TESS sector 9 and 36 (left and right parts respectively of top, middle, and bottom panels) light curves of WASP-19, comparison star 1, and 2. Both sector 9 and 36 data for for WASP-19 were analysed by an update to the TESS SPOC pipeline plotted in black in the top panel along with the sector 9 data from the older version of TESS SPOC pipeline plotted in blue. Middle and bottom panels show the Sector 9 and 36 light curves and their Lomb Scargle (LS) periodograms (horizontal dashed lines marking the 10% false alarm probability levels and vertical lines marking the peak of the periodogram) for comparison star 1 and 2 using eleanor and the new version of TESS SPOC pipeline respectively. The comparison star 1 LS periodogram peaks at 11 days and 10.74 days for sector 9 and 36 respectively. The comparison star 2 LS periodogram peaks at 12.72 days and 13.48 days for sector 9 and 36 respectively. Further details of the analysis of TESS data are described in Appendix A1.
Figure 4.26: Best fit transit parameters of WASP-19 from the light curves obtained in the TESS sector 9 and 36 (left and right parts respectively of top, middle, and bottom panels). The black points for all the parameters are the best fit values for each transit obtained by fitting for all parameters including inclination, with the horizontal dashed black lines showing the weighted average value across all 58 transits. The red points are the best fit values for $a/R_*$, $R_p/R_*$, and $u_1$ when fixing the inclination to the value reported by the Magellan/IMACS observations, with the red dashed line showing their respective weighted average values. The horizontal red dashed line in the inclination panel marks the inclination measured by Magellan/IMACS observations.
Figure 4.27: Transit light curves for WASP-19b observed by TESS with their best fit transit model overplotted in red. The data are detrended SAP light curves secured in sector 9 and released on MAST through the latest TESS SPOC pipeline. The top two panels are the light curves from the first half of sector 9 and the bottom two panels are from the second half after the mid-sector change of TESS orbit. The right panels both top and bottom shows the residuals to the best fits for each transit along with the RMS of the residuals.
Figure 4.27: Continued.
Figure 4.28: Same as Figure 4.27 but for sector 36.
Figure 4.28: Continued.
Table 4.11: Best fit GP hyperparameters and their uncertainties for the white transit light curve fits. Similar to Table 4. in the paper, the column 'No.' specifies the transit observation, 'Method' specifies the method used to fit the white transit light curves, and 'GP regressors' specifies the GP regressor combination used. 'A' is the amplitude hyperparameter, and \( \eta_t, \eta_c, \eta_a \) are the time, comparison star white light curve length scale hyperparameters respectively.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>GP regressors</th>
<th>( \ln(A) )</th>
<th>( \ln(\eta_t) )</th>
<th>( \ln(\eta_c) )</th>
<th>( \ln(\eta_a) )</th>
<th>( \sigma_w [\text{ppm}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>New</td>
<td>Time, Comp</td>
<td>-4.67 \pm 0.92</td>
<td>24.22 \pm 2.41</td>
<td>-1.06 \pm 1.94</td>
<td>-          \pm 1.27</td>
<td>383 \pm 28</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Airmass</td>
<td>-14.02 \pm 0.44</td>
<td>-</td>
<td>-          \pm 0.92</td>
<td>-8.17 \pm 0.68</td>
<td>395 \pm 20</td>
</tr>
<tr>
<td>2</td>
<td>New</td>
<td>Time, Comp</td>
<td>-4.77 \pm 0.41</td>
<td>24.1 \pm 2.99</td>
<td>2.22 \pm 2.45</td>
<td>-</td>
<td>373 \pm 25</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Airmass</td>
<td>-11.82 \pm 1.22</td>
<td>-3.55 \pm 2.08</td>
<td>-          \pm 1.59</td>
<td>-</td>
<td>431 \pm 23</td>
</tr>
<tr>
<td>3</td>
<td>New</td>
<td>Comp</td>
<td>-4.53 \pm 0.81</td>
<td>-</td>
<td>24.37 \pm 1.97</td>
<td>-</td>
<td>345 \pm 24</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time</td>
<td>-12.77 \pm 0.81</td>
<td>-5.81 \pm 1.07</td>
<td>-          \pm 0.84</td>
<td>-</td>
<td>338 \pm 22</td>
</tr>
<tr>
<td>4</td>
<td>New</td>
<td>Time, Comp, Airmass</td>
<td>-6.19 \pm 2.39</td>
<td>23.3 \pm 3.16</td>
<td>430.8 \pm 363.91</td>
<td>-0.9 \pm 2.38</td>
<td>353 \pm 25</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time</td>
<td>-13.82 \pm 0.65</td>
<td>-7.64 \pm 0.89</td>
<td>-          \pm 1.06</td>
<td>-</td>
<td>375 \pm 25</td>
</tr>
<tr>
<td>5</td>
<td>New</td>
<td>Time, Comp</td>
<td>-5.96 \pm 2.30</td>
<td>23.51 \pm 3.09</td>
<td>-1.9 \pm 2.19</td>
<td>-</td>
<td>326 \pm 24</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time, Airmass</td>
<td>-14.23 \pm 0.71</td>
<td>-8.29 \pm 1.16</td>
<td>-          \pm 0.94</td>
<td>-</td>
<td>484.2 \pm 317.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>380 \pm 26</td>
</tr>
<tr>
<td>6</td>
<td>New</td>
<td>Comp</td>
<td>-3.34 \pm 2.32</td>
<td>-</td>
<td>25.53 \pm 2.45</td>
<td>-</td>
<td>402 \pm 25</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time</td>
<td>-14.21 \pm 1.33</td>
<td>-5.57 \pm 1.87</td>
<td>-          \pm 1.22</td>
<td>-</td>
<td>399 \pm 25</td>
</tr>
<tr>
<td>7</td>
<td>New</td>
<td>Time, Comp</td>
<td>-3.61 \pm 0.89</td>
<td>25.62 \pm 2.44</td>
<td>0.26 \pm 2.61</td>
<td>-</td>
<td>378 \pm 49</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time, Airmass</td>
<td>-13.82 \pm 0.62</td>
<td>123.54 \pm 364.52</td>
<td>-3.83 \pm 39.95</td>
<td>-</td>
<td>402 \pm 42</td>
</tr>
<tr>
<td>8</td>
<td>New</td>
<td>Time, Comp</td>
<td>-4.85 \pm 0.73</td>
<td>26.512 \pm 1.103</td>
<td>-0.38 \pm 1.16</td>
<td>-</td>
<td>753 \pm 58</td>
</tr>
<tr>
<td></td>
<td>Conventional</td>
<td>Time</td>
<td>-12.48 \pm 0.42</td>
<td>-8.08 \pm 1.09</td>
<td>-          \pm 0.61</td>
<td>-</td>
<td>1051.0 \pm 91</td>
</tr>
</tbody>
</table>
On the follow-up of small and rocky exoplanets orbiting bright stars using ground-based spectrophotometry

Vatsal Panwar, Jean-Michel Désert et al.

To be submitted to the Monthly Notices of the Royal Astronomical Society

Abstract

Small transiting exoplanets with radii \( \leq 2R_\oplus \) around very bright stars pose an observational challenge in case of both their detection and characterization. Theoretical planet formation models and detections from space-based transit missions like Kepler indicate that earth to super-Earth/sub-Neptune exoplanets are the most common outcomes of planet formation. Ongoing missions like TESS continue to detect such planets around bright stars. However, despite being so ubiquitous, follow-up studies of small planets around very bright stars of \( V_{\text{mag}} \leq 9.5 \), especially from ground-based observations, hitherto remains limited. This is mainly so because ground-based high-precision photometry of very bright stars is limited by non-availability of suitable comparison stars conventionally required for correction of transit light curve systematics. In this work we explore ways to overcome this limitation in the context of ground-based spectrophotometry observations which measure a transit signal in both time and wavelength dimensions. Recently developed systematics modelling techniques utilize the shared information about time-dependent systematics across the wavelength range to precisely measure the broadband transit signal. By applying these techniques to a transit of WASP-18b (stellar \( V_{\text{mag}} = 9.3 \)) we show that it is possible to use ground-based medium to large telescopes to reach transit depth precisions required for the follow-up of small and rocky exoplanets. This has implications for improving prospects of precise transit depth and mid-transit time measurements to constrain the planet radius and mass via TTVs. We discuss the relevance of our results on advancing the ground-based follow-up of small planets around very bright stars in the context of TESS and future transit survey mission PLATO.
Small exoplanets around bright stars from ground

5.1 Introduction

Since the discovery of the first transiting hot Jupiter by Charbonneau et al. (2001), one of the most technically challenging goals in exoplanet science has been the detection and characterization of small and rocky exoplanets orbiting sun like stars. Small and rocky exoplanets are the most common outcome of planet formation from theoretical perspective (Miguel et al. (2011)), with a majority of them statistically expected to be found around low-mass stars (Dressing & Charbonneau (2015)). Small exoplanets with earth to super-Earth sizes are exciting avenues for understanding the outcomes of physical mechanisms e.g. core-accretion (Bodenheimer & Pollack (1986)) and atmospheric mass loss (Lammer et al. (2003)) that sculpt the exoplanet population. Additionally, small exoplanets are of furthermore interest to detect and characterize when they reside in the habitable zones (Kasting et al. (1993)) of their hosts stars.

Measuring accurate mass and radius properties can shed light on statistical trends (Fulton et al. (2017)) in the exoplanet population that are likely connected to the planet formation and evolution histories Owen & Wu (2017); Luque & Pallé (2022). The initial and ongoing ground-based transit detection surveys (e.g. WASP and NGTS) have been most sensitive to detecting transiting gas giants around FGK stars. Long duration space-based transit surveys have been more successful at tackling the challenge of detecting small rocky planets around FGK stars. The most notable of these missions have been the Kepler (Borucki et al. (2010)) and the extended Kepler-K2 missions, which have confirmed detections of 811 and 134 exoplanets respectively with radii less than 1.6 R_⊕. Statistically, these candidates are most likely rocky and do not sustain a volatile gas rich envelope of H/He (Rogers (2015)). Kepler and K2 surveys have been most sensitive to detecting planets around optically faint stars in the brightness range V_\text{mag} = 12 to 16 (Thompson et al. (2018)). Specifically the K2 mission was restricted to fields around the ecliptic plane and was most sensitive to short period planets owing to only on average 60 days spent observing each campaign.

The ongoing TESS mission (Ricker et al. (2014)), which has been surveying a larger part of the sky, is more sensitive to transit searches around brighter host stars as compared to Kepler but still restricted to short period planets on account of spending about a month in each sector. 90% of planets TESS will detect are expected to be around stars in the brightness range of 8.2 to 13.1 TESS magnitude (Barclay et al. (2018)). TESS has already detected and confirmed 39 planets<1.6 R_⊕, and is expected to detect around 250 planets < 2 R_⊕ based on yield simulations by Barclay et al. (2018). Future missions like PLATO (Rauer et al. (2014)) are poised to expand this parameter space even more around V_\text{mag} range around 9 mag and brighter and in its long duration operation phase will also be sensitive to planets at longer orbital periods up to that of an earth analog (Heller et al. (2022)).

Ground-based photometry follow-up of space-based transit detections has been important for confirming space-based detections and further refining the properties of transiting planets. Precise ground-based transit follow up can be useful for confirming the transit in case of periods longer than a continuous coverage of the space-based facility (e.g. 27 days per

---

1As of September 2022, NASA Exoplanet Archive
sector for TESS), improving the planet radius measurement, constraining the ephemerides of the system for future follow-up, and in case of multi-planet system to constrain the transit timing variations (TTVs) to measure the planet masses (Holman & Murray (2005)). TTVs can be especially useful for mass determination of small and rocky planets as a majority of them are known to occur as part of multi-planet systems (Steffen et al. (2012); Zhu et al. (2018)).

Several advancements have been made recently in terms of instrumentation to achieve space like photometric precision from ground. One such way is to use beam shaping diffusers (Stefansson et al. (2017); Greklek-McKeon et al. (2022)) to stabilize the PSF and overcome the effects of inter-pixel variations, atmospheric seeing, and telescope systematics on the photometric precision. However, beam shaper assisted photometry often also relies on performing differential photometry, and non-availability of nearby suitable comparison stars – which is especially the case for very bright host stars – could be a limiting factor for this approach (e.g. Vissapragada et al. (2020)). Hence, alternative observational strategies are needed for ground-based high-precision transit follow-up of small exoplanets orbiting very bright stars.

In this paper, we introduce and apply a ground-based spectrophotometry approach for the transit follow-up of small and rocky exoplanets around bright stars with $V_{\text{mag}} \leq 9.5$. Multi-object spectrographs (MOS) on medium to large ground-based telescopes have been routinely used now for a decade to conduct high signal-to-noise spectrophotometry for studying the atmospheres of transiting exoplanets (e.g. Bean et al. (2010)). Recent advancements in transit light curve systematics methods have made it possible to extend this technique to very bright stars with possibly no comparison stars (Panwar et al. (2022b)). In this work we show how using the spectroscopic information in MOS time series spectra can be helpful in disentangling systematics from the transit signal in both time and wavelength dimension, which can be used to accurately and precisely measure the transit depths and mid-transit times of small and terrestrial planets around bright stars even without using the comparison star. We demonstrate our application on the GMOS transit observation of WASP-18b, a gas-giant around a bright ($V_{\text{mag}} = 9.3$) star.

This paper is distributed as follows. In Section 5.2 we describe the details of the GMOS observations of a transit of WASP-18b, and in Section 5.3 we describe the reduction of the data. In Section 5.4 we elaborate the various transit light curve analysis methods we apply to the WASP-18b transit with the goal of measuring precise broadband transit and mid-transit time of the planet. In particular, we apply a recently introduced 2D Gaussian Process (GP) approach by Gordon et al. (2020) to analyse spectroscopic light curves in both time and wavelength dimension together. In Section 5.5 we present our results for WASP-18b and base them to predict nominal transit parameter precisions that can be obtained for similar ground-based observations of small and rocky exoplanets in future. In Section 5.6 we present our summary and conclusions.
5.2 Observations

We observed a single transit of WASP-18b (Observation ID: GS-2012B-Q-6-57) using the Gemini Multi-Object Spectrograph (GMOS) on Gemini South from 29-10-2012 00:39 UT to 29-10-2012 4:51 UT. The transit was observed using a setup similar to other observations in our GMOS survey program described in Huitson et al. (2017). We used the R150 grism along with the blocking filter OG515_G0330 to block the light blueward of 515 nm. We also observed a comparison star Gaia DR3 4931352054788013312 which is 3.2’ away from WASP-18 and has a $G_{mag} = 12.75$ (3.6 mag fainter as compared to WASP-18 in $G_{mag}$). To avoid slit losses, the MOS mask we used had wide slits of 10 arcsec width and 30 arcsecond length for both the stars. We did not use an additional narrow slit for the sky, so there is no high-resolution sky spectrum available in the dataset. We took the observation prior to June 2014 and hence used the older e2v detectors with three amplifiers. The older e2v detectors on GMOS have been known to suffer from fringing redward of 710 nm (Huitson et al. 2017). The observations were taken in slow readout mode with $1 \times 1$ readout binning mode and gains varying from 1.89 to 2 e/ADU. The observations started at seeing values of $\sim 1$ to 1.9 arcseconds and improved to $\sim 0.8$ arcseconds towards the later part of the night. The exposure times stayed between 25 and 30 seconds. The resultant duty cycle was on average 33%.

5.3 Data reduction

We used a custom IDL pipeline designed for extracting the time-series of 1D stellar spectra from the series of raw GMOS 2D images. The pipeline was first introduced and described in detail in Huitson et al. (2017) and also used in Panwar et al. (2022b,a). We describe here the key features of spectral extraction and the steps taken specifically for the WASP-18b dataset analysed in this paper.

Similar to Stevenson et al. 2014, we performed tilt correction to the spectral traces for both the target and comparison stars by measuring the tilts in lines of the CuAr lamp spectrum taken on the same day with the same slit positions as the slits in the mask for science observations. We noticed up to 4 pixels difference in the spatial positions of the top and bottom of the lines in the lamp spectra. We did not perform flat field correction as it did not improve the scatter in the white and spectroscopic transit light curves significantly as compared to photon noise. We estimated the background by performing a linear fit to the cross dispersion background fluxes well above and below the spectral trace. Following the same steps as described in Huitson et al. (2017) we performed cosmic ray removal, and subtracted the background while performing optimal extraction (Horne 1986).

We performed wavelength calibration by using the CuAr lamp spectrum and also independently verified it by fitting the known positions of prominent stellar and telluric lines (e.g., Na I, Hα, and O₂ A band) in the time mean of all the spectra using an order 3 Chebyshev polynomial. The time mean comparison star spectrum deviated from the time averaged target star spectrum by at most 2 pixels. A constant offset (with respect to time) measured from the mean of spectra was applied to the spectra from all the exposures so that the target
and comparison star spectra are aligned. Through cross correlation of the spectral features we noticed that the target and comparison star spectra individually suffered from shifts of ±2 pixels through the night. We did not apply additional corrections for these shifts as they are significantly small as compared to the typical bin sizes of 10 nm that we use in this work, and also because we let the transit light curve systematics model described in Section 5.4 account for them.

5.4 Transit light curve analysis

We constructed light curves for both the target and comparison star in the wavelength range 530 to 710 nm in 10 nm wide bins. The white light curves for both stars were computed by integrating the flux in the wavelength range 530 to 710 nm for each exposure, excluding the flux redward of 710 nm due to high levels of fringing as seen in Figure 5.1. For both white and spectroscopic light curves (hereafter referred to as $\lambda$LCs), we normalized the flux in each exposure by the corresponding exposure time. We then normalize the white and each $\lambda$LCs with their respective median out of transit flux before using them for further analysis. The raw GMOS white and $\lambda$LCs for the target and comparison star are shown in Figure 5.2.

Our main goal is to measure the absolute broadband transit depth of WASP-18b in the visible bandpass of GMOS observations. To achieve this, we use two different models to correct for the transit light curve systematics as described in the following subsections. Both the models differ in two main ways. The first is that one of them fits the integrated white transit light curve to measure the broadband transit depth, while the other model measures the broadband transit depth from fits to $\lambda$LCs. The second difference is that one of them uses a 1D
Figure 5.2: Spectroscopic (ALC) and white transit light curves of WASP-18b observed using GMOS-S. The top left panel shows the ALC normalized to their median and vertically offset for clarity. The top right panel shows the difference between each ALC from the ALC for the bluest wavelength bin (529 nm) to demonstrate that each bin suffers from similar high-frequency systematics, along with a low-frequency airmass dependent trend that changes nearly monotonically going from blue to red wavelength bins. The bottom panel shows the white transit light curve for WASP-18b and the simultaneously observed comparison star. The vertical dashed lines mark the predicted start and end of the transit in all panels.
only time dependent GP model, while the other uses a 2D time and wavelength dependent model. We first describe the application of the two models to the WASP-18b light curves and compare their performance in Section 5.4.1

**5.4.1 White light curve: 1D GP with and without comparison star as a GP regressor**

We start with the simplest approach to measure the broadband transit depth from time series spectrophotometry observations, which is to fit the white transit light curve. We focus on fitting the target white transit light curve directly given that we have a significantly fainter comparison star. Panwar et al. (2022b,a) contains detailed discussions on how analysing the target transit light curves directly avoids additional sources of uncertainty arising from potential comparison star variability and from performing differential spectrophotometry using a suboptimal comparison star.

The transit is clearly visible by eye in the target transit light curves in Figure 5.2. In addition to the transit signal, it is also visibly evident that the light curves for both the the stars suffer from systematic trends at short (\(\sim\) minutes, due to the GMOS shutters) and long (\(\sim\) hours, due to change in airmass) timescales. Such trends can affect the accuracy and precision of the measured transit signal when fitting the white light curve. In particular, the WASP-18b light curve suffers from the well known odd-even effect due to the GMOS shutters (Jorgensen (2009); Stevenson et al. (2014)) which are of the order of 0.5 millimag or \(\sim 500\) ppm.

Hence, the systematics model we require must be able to effectively model the trends at a range of timescales and amplitudes. Gaussian Process (GP) models, as introduced by Gibson et al. (2012) and routinely used for ground-based MOS observations, provide a powerful approach to infer the nature of systematics. A GP model can be used to model and correct systematics with minimal assumptions about their parametric form and within a Bayesian approach. This model was further extended recently by Panwar et al. (2022b) who showed that it is possible to constrain the systematics directly in the target transit light curves by using the comparison star light curves as one of the GP regressors. We refer the reader to Panwar et al. (2022b) for the details about the formalism of the model. We elaborate here on the specifics of its application to the GMOS transit light curves of WASP-18b.

We use the Python package celerite (Foreman-Mackey et al. (2017)) to implement a Matérn 3/2 covariance kernel function and the transit model as the mean function of the GP. We combine the contributions from each GP regressor in quadrature as described in Panwar et al. (2022b) with a single amplitude hyperparameter and a length scale hyperparameter for each regressor. We also include a diagonal variance term to account for the white noise, and hence the resultant GP covariance matrix is

\[
K = T + \sigma_w \delta_{ij}
\]

(5.1)

where \(T\) is the time dependent covariance matrix computed using the Matérn 3/2 kernel function and \(\sigma_w \delta_{ij}\) is the white noise diagonal matrix.
Table 5.1: Summary of priors and fixed values for the model parameters used to fit the GMOS transit light curves of WASP-18b using the 1D and 2D GP model. The first part of the table shows the exoplanet transit model parameters and the second part shows the systematics model parameters. U is a uniform prior within the specified range, and $\mathcal{N}$ is a Gaussian prior with the mean and standard deviation respectively. In bold and marked with ‘(\lambda)’ show the wavelength dependent parameters and their priors for exoplanet and speccgp used in the case of 2D GP model specifically. $R_*$ is the stellar radius, $b$ is the impact parameter ($a/R_\ast \cos(i)$), $T_c$ is the predicted mid-transit time, $e$ and $\omega$ are the eccentricity and argument of periapsis respectively, $R_p/R_\ast$ is the planet to star radius ratio, $T_0$ is the mid-transit time, $u_1$ and $u_2$ are the quadratic limb-darkening coefficients, $A$, $\eta_p$, and $\sigma_w$ are the amplitude, length scale, and white noise hyperparameters of the 1D and 2D GP models. $\alpha$ is the vector of wavelength dependent scaling factors in the 2D GP model, with the value for the bluest bin fixed to 1. $a_1(\lambda)$ and $a_2(\lambda)$ are the coefficients for a wavelength dependent linear model of airmass used in addition to the 2D GP model to account for low-frequency wavelength dependent trends in the $\lambda$LCs. Note that the quadratic limb darkening coefficients specified are for the white light curve calculated using PyLDTk (Parviainen & Aigrain (2015)), and wavelength dependent coefficients similarly calculated by PyLDTk are fixed for each spectroscopic bin in the 2D GP model.

<table>
<thead>
<tr>
<th>Transit model</th>
<th>Parameter</th>
<th>Prior/Fixed value</th>
<th>Reference/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_\ast [R_\odot]$</td>
<td>1.26</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$P [d]$</td>
<td>0.9414526</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$e, \omega$</td>
<td>$\mathcal{N}$ (0.318, 0.018)</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>$\mathcal{U}$ (0, 1), $\mathcal{N}$ (0.09716, 0.001)</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$R_p/R_\ast$</td>
<td>$\mathcal{U}$ (T_c-0.002, T_c+0.002)</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$\Delta (R_p/R_\ast) (\lambda)$</td>
<td>$\mathcal{N}$ (0., 0.0005)</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$T_0[d]$</td>
<td>$\mathcal{N}$ (T_c, 0.001)</td>
<td>Shporer et al. (2019)</td>
<td></td>
</tr>
<tr>
<td>$u_1(\lambda), u_2(\lambda)$</td>
<td>0.435,0.164</td>
<td>PyLDTk</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systematics model</th>
<th>Parameter</th>
<th>Prior/Fixed value</th>
<th>Reference/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>In ($A$)</td>
<td>$\mathcal{U}$ (-100, 100)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>In ($\eta_p$)</td>
<td>$\mathcal{U}$ (-100, 100)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>In ($\sigma_w$)</td>
<td>$\mathcal{U}$ (-20., 10.)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>In ($\alpha(\lambda)$)</td>
<td>$\mathcal{U}$ (-20., 10.)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>$a_1(\lambda)$</td>
<td>$\mathcal{N}$ (-0.2, 0.98)</td>
<td>Fit to WLC baseline</td>
<td></td>
</tr>
<tr>
<td>$a_2(\lambda)$</td>
<td>$\mathcal{N}$ (1.2, 0.88)</td>
<td>Fit to WLC baseline</td>
<td></td>
</tr>
</tbody>
</table>
This is essentially a 1D model, and we use it to model the correlations in only the time dimension for a single white transit light curve. We fit the white transit light curve using two combinations of GP regressors: 1) only time, 2) time and comparison star white light curve.

We use the Python package exoplanet (Foreman-Mackey et al. (2021)) to model the transit of WASP-18b. exoplanet employs analytic closed-form solutions of exoplanetary transits with quadratic limb darkening as presented by Agol et al. (2020). Both exoplanet and celerite have backends compatible with the Python tensor library theano (now being developed as aesara1) and the probabilistic programming package PyMC3 (Salvatier et al. 2016). The ease of specifying priors and likelihood distributions in PyMC3, and its theano based speed optimization of tensor operations together are particularly suited for computationally intensive GP regression models we use for fitting the transit light curves. We specify the priors for all the transit model parameters and the GP hyperparameters in Table 5.1. For the transit model in particular, we are interested in measuring the planet to star radius ratio i.e., \( r_p/R_* \) and the mid-transit time (\( T_0 \)). Hence, when fitting for either \( r_p \) and \( T_0 \) we fix the other to the values inferred from literature. Doing so is representative of cases when the goal of an observation is to do a follow-up measurement of the transit and at least one of the two transit parameters – \( r_p \) or \( T_0 \) – is decently constrained using previous measurements.

We first derive the Maximum a-posteriori (MAP) solution by optimizing the GP likelihood. We then use the MAP solution as a starting point for PyMC3’s inbuilt NUTS MCMC sampler to sample the marginal likelihoods for each free parameter in the model. We use the PyMC3 ‘extra’ utilities2 released with exoplanet which are essential to let the sampler account for the covariances between the free model parameters. For both \( r_p \) and \( T_0 \), after tuning the model parameters including the GP hyperparameters in the first iteration, we run another iteration of PyMC3 NUTS MCMC sampler in which we fix all model parameters except the transit model parameter of interest. For both the GP regressor combinations, we show the MCMC posteriors hence obtained in Figure 5.5 and 5.6, and the corresponding best fit to the white target transit light curve in Figure 5.4.

5.4.2 \( \lambda \)LCs: 2D GP without using the comparison star as a GP regressor

The traditional approach to obtain the transmission spectrum of a planet is to measure the transit signals in the \( \lambda \)LCs. A common strategy to achieve this in the context of ground-based MOS data is to first fit the white transit light curve to 1) measure the best fit transit parameters, and 2) to obtain a common-mode trend that can be used to correct the systematics in each spectroscopic light curve. Multiple ground-based MOS studies have used the white light curve derived common-mode trend to correct for the common time-dependent systematics in \( \lambda \)LCs, either by dividing it out from each spectroscopic light curve (e.g., Gibson et al. (2013)) or using it as a GP regressor when fitting the \( \lambda \)LCs (e.g., Panwar et al. (2022b)). In both cases, the white light curve is fit first and subsequently the average information about

---

2https://exoplanet-docs.readthedocs.io/en/v0.2.3/tutorials/pymc3-extras/
the systematics and the absolute transit depth contained in the white light curve is used to constrain the systematics in \(\lambda LCs\).

An alternative and to our knowledge relatively unexplored approach to ground-based MOS spectrophotometric observations is the inverse of the steps taken for measuring the transmission spectra, which is to use the \(\lambda LCs\) themselves directly to infer transit properties like the absolute broadband transit depth and mid-transit time. The prime motivation of this approach is that the information on the timescales of time dependent systematics is largely common across the wavelength range. As shown in Figure 5.2, the wavelength dependence manifests as variation in the amplitude of systematics e.g. in low flux regions of the stellar spectrum, and due to the wavelength dependent atmospheric extinction with changing airmass. Recently, Ahrer et al. (2022) in their derivation of the optical transmission spectrum of WASP-94A b show that jointly fitting \(\lambda LCs\) with shared GP length scale hyperparameters can improve the precision in the transmission spectrum. However, this approach doesn’t explicitly account for the wavelength dependence of systematics and relies on fitting the white light curve a-priori to ascertain the best fit transit parameters. For our goal of directly using the target \(\lambda LCs\) to measure the absolute transit depth, a 2D time and wavelength dependent GP model is one potential solution, which we explore here.

The 2D celerite model specgp introduced by Gordon et al. (2020) is well suited to our application. specgp is an extension of the 1D celerite model to the two dimensions of time and wavelength with fast scaling for the special circumstances when only the amplitude of the correlated systematics change for the individual \(\lambda LCs\). The model can be applied directly to the data cube (time, wavelength, flux) of spectroscopic target light curves. For a dataset with \(N\) time points and \(M\) wavelength bins, the specgp GP covariance matrix, analogous to the 1D expression in Equation 5.1, is

\[
K = T \otimes R + \sigma_w \delta_{ij} \tag{5.2}
\]

where \(T\) is the \(N \times N\) covariance matrix along the time axis, \(R\) is the \(M \times M\) covariance matrix along the wavelength dimension, and \(\sigma_w \delta_{ij}\) is the white noise diagonal matrix. The operation \(\otimes\) signifies a Kronecker product between the time and wavelength covariance matrices, which essentially results in an \(NM \times NM\) block matrix formed by multiplying each element of \(T\) by the matrix \(R\). The \(T \otimes R\) matrix encodes the cross-covariances between each data point in the data cube of \(\lambda LCs\). The fastest scaling for the 2D GP model is achieved when \(R\) can be simply expressed as \(\alpha \alpha^T\) where \(\alpha\) is the vector of scaling factors for each wavelength bin \((\alpha_1, \alpha_2 \ldots \alpha_M)\) and \(\alpha^T\) is its transpose. Rather than specifying another kernel function for \(R\), we also adopt this approach of specify constructing \(R\) using the vector of scaling factors \(\alpha\) with each factor for a wavelength bin as a free parameter in the model.

We also additionally use an additive linear function of airmass added to the GP + transit model, with the linear coefficients varying for each wavelength bin to account for the wavelength dependent effect of airmass. We use one free parameter for the absolute \(R_p/R_*\) \((r_p)\) we aim to measure, and a vector of wavelength dependent offsets to it \((\Delta r_p)\) to account for the wavelength dependent changes to the basal \(r_p\) due to either systematics or the planetary
5.5 Results and Discussion

We use a Gaussian prior for each $\Delta r_p$ with the standard deviation set to 10 times the contribution expected from the planet scale height. For WASP-18b the contribution from 1 scale height is about 10 ppm or $\Delta r_p = 0.00005$, which is 6% of the precision expected from photon noise. Similar to the 1D GP model in Section 5.4.1, when fitting for $r_p$ we fix $T_0$ and vice versa, to independently assess the capability of the model in constraining the two quantities. All the other fixed and free parameters and their priors used are shown in Table 5.1.

We adapt the PyMC3 framework as for the 1D model in Section 5.4.1 to run the 2D GP model to fit the (Time $\times$ Wavelength $\times$ Flux) data cube formed using 19 $\lambda$LCs. Similar to the 1D GP model, we conduct the fits in two iterations. In the first iteration we run the PyMC3 sampler to fit the GP hyperparameters, airmass coefficients, and the free transit model parameters ($r_p$ or $T_0$ in our case). This iteration tunes the GP hyperparameters and the airmass trend parameters by estimating their marginal posteriors. In the next iteration, we fix the GP hyperparameters and the airmass trend parameters to their tuned median value estimated from their marginal posteriors in the previous iteration, and only fit for the transit model parameter ($r_p$ or $T_0$) we are interested in. We hence obtain the final posterior for the transit model parameter. We calculate the 50, 16 and 84 percentile values of the marginal posterior to estimate the median and $\pm 1 \sigma$ uncertainty values respectively of the transit model parameter. We repeat this process for two cases: 1) Using a uniform prior, and 2) using a Gaussian prior for the transit model parameter in the first iteration when tuning the hyperparameters. For both cases, in the second iteration we use a uniform prior for the transit model parameter. We show the marginal posteriors obtained for the transit parameters of interest $r_p$ and $T_0$ in Figure 5.6. The best fit transit parameter values from the 1D and 2D GP models are shown in Table 5.2.

In the next section, we compare our measurements of $r_p$ and $T_0$ for WASP-18b from the 1D and 2D GP models with each other and with the values measured from previous observations. We also discuss the implications of our measurements for the future follow-up of transits of rocky exoplanets around bright stars ($V_{mag} \leq 9$) from ground-based MOS observations.

5.5 Results and Discussion

5.5.1 Transit parameters measured from the 1D and the 2D GP models

We first compare the measurement of $r_p$ from the 1D GP and the 2D GP models. The left panel in Figure 5.5 shows the posterior distributions for $r_p$ from the 1D GP model — with and without using the comparison star as a GP regressor, and the right panel shows posteriors for the 2D GP model. We find that the 1D GP model itself when using a comparison star as a GP regressor performs better than when not using the comparison star as a GP regressor, giving nearly 3 times better precision for $r_p$ as compared to the latter. This is intuitively expected because using both time and comparison star light curve as a GP regressor in the 1D model are essential in correcting for both the high and low frequency systematics respectively in the WASP-18b white light curve.
Table 5.2: Summary of the best fit transit parameters and their ±1σ uncertainties for WASP-18b transit analysed from the 1D GP model (using only time and using time and comparison star light curve as GP regressors) and 2D GP model (using uniform (U) and Gaussian (N) priors for tuning the GP hyperparameters).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>1D GP</th>
<th>2D GP</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_p/R_⋆</td>
<td>Time</td>
<td>Time, Comp</td>
</tr>
<tr>
<td></td>
<td>0.1161±0.011</td>
<td>0.0970±0.0032</td>
</tr>
<tr>
<td>T_0 [BJD] - 2456229</td>
<td>0.600198 ± 7.24 × 10^{-4}</td>
<td>0.600553 ± 3.26 × 10^{-4}</td>
</tr>
</tbody>
</table>
Figure 5.3: Fits to the spectroscopic light curves (λLC) of WASP-18b using the 2D GP model described in Section 5.4.2. The left panel shows the λLC for each wavelength bin vertically offset for clarity and the best fit model (median of the posterior samples for the Transit + 2D GP systematics model) overplotted in black. The right panel shows the residuals for each λLC with respect to the best fit model shown in the left panel. Note that the last seven λLC show low levels of correlated noise not corrected by the 2D GP model as well as the first twelve bins. We discuss these as the caveats of the 2D GP model in Section 5.5.2.
204 Small exoplanets around bright stars from ground

Figure 5.4: Fits to the white transit light curve of WASP-18b using the 1D GP model described in Section 5.4.1. The left and right panels show the fits to the WASP-18b transit light curve when using only time, and using time and comparison star light curve as GP regressors respectively. The detrended transit light curve and the corresponding model is vertically offset for clarity. For the 1D GP model, using both time and comparison star light curve as regressor improves the residual RMS by an order of magnitude as compared to when not using the comparison star.

Figure 5.5: Posterior probability densities for $R_p/R_*$ ($r_p$) of WASP-18b measured from a single Gemini/GMOS transit light curve using the 1D GP model (left panel, see section 5.4.1) and 2D GP model (left panel, see section 5.4.2). In the left panel, the yellow and green posteriors are from the 1D GP model fit to the white light curve using only time, and using time and comparison star light curve as GP regressors respectively. In the right panel, the blue and red posteriors are from the 2D GP model fit to the $\lambda$LCs using uniform and Gaussian priors on $R_p/R_*$ when tuning the systematics model parameters. In both panels the dashed and dotted vertical lines mark the $\pm 1\sigma$ of $R_p/R_*$ measured from TESS by Shporer et al. (2019) and Pearson (2019). The shaded regions in both panels mark the $\pm 1\sigma$ bound for the posterior of the corresponding colour. Note that the horizontal scale for both panels is different for clarity. As specified in the figure legend, it can be seen that the 2D GP model for both prior cases yields about 4 times better precision on $R_p/R_*$ as compared to the best precision from the 1D GP model.
In comparison, we observe that the 2D GP model performs the best in terms of the precision. The 2D GP model yields $r_p = 0.0941 \pm 0.00078$ and $0.0962 \pm 0.00078$ when using uniform and Gaussian priors on $r_p$ respectively for tuning the hyperparameters. Both cases of uniform and Gaussian priors give identical precisions, and the respective median $r_p$ are consistent within $2\sigma$. The 2D GP model precision on $r_p$ is 4 times better precision as compared to the best precision from the 1D GP model. This translates to a transit depth precision $\Delta r_p^2 = 150$ ppm from the 2D GP model, which is within $\sim 5\%$ above the transit depth precision expected purely from photon noise ($\sim 140$ ppm) for a white light curve in the wavelength range of 530 to 710 nm. The 2D GP model achieves this by modelling the correlations in both time and wavelength dimension. By taking advantage of the shared information about systematics across the wavelength bins, the 2D GP model constrains both high and low frequency systematics without needing to use the comparison star light curve as an additional regressor. The ability of the 2D GP model to attain a remarkable precision from a single transit of a planet orbiting a bright star and without needing to use the comparison star has important implications for the ground-based follow-up of rocky exoplanets. We discuss these in more detail in Section 5.5.3.

There is a $\sim 100$ nm overlap between the GMOS and TESS bandpass as seen in figure 5.1. Hence, the value of $r_p$ measured by TESS can serve as a benchmark for assessing the accuracy of the models we use to measure $r_p$ from the GMOS transit. The TESS $r_p$ value averaged over 40 transits by two different studies, Shporer et al. (2019) and Pearson (2019), are marked in Figure 5.5. We find that the 1D GP model with its relatively large uncertainties on $r_p$, is consistent with the TESS values within $2\sigma$. The 2D GP model deviates by $4\sigma$ and $1.5\sigma$ from the Shporer et al. (2019) value when using uniform and Gaussian priors for $r_p$ respectively. As compared to the Pearson (2019) value, the GMOS value deviates by $3\sigma$ and less than $1\sigma$ when using uniform and Gaussian priors for $r_p$ respectively. We conclude that using a Gaussian prior for $r_p$ when tuning the systematics model parameters can help in improving the accuracy of measurement while retaining its precision.

In summary, we do not observe more than $4\sigma$ deviation between the value measured GMOS and TESS value. We speculate that the reason behind this deviation could be because of differing treatments of limb-darkening in our work and the TESS analyses. The limb-darkening differences could especially be important because the TESS bandpass covers a significant part of the red optical stellar spectrum beyond 710 nm which is inaccessible to the GMOS measurement due to high level of fringing. Future high-precision transit observations at multiple epochs, from space or ground, could confirm if there are evidences of transit depth variations for the planet. It should also be noted that the TESS value is obtained from averaging across 40 transits, while the GMOS measurement in this work is from a single transit. Assuming a $\sqrt{N}$ improvement in precision, the per transit precision on $r_p$ from TESS can be estimated to be $\pm 0.0006$, which is about 15% better as compared to the precision from a single GMOS transit using the 2D GP model. The GMOS precision from the 2D GP model is nearly at photon noise. Hence, ways to improve the precision would involve improving the photon noise e.g. by using new detectors, as we discuss in more detail in Section 5.5.2.
Figure 5.6: Posterior probability densities for the mid-transit time ($T_0$) of WASP-18b measured from a single Gemini/GMOS transit light curve using the 1D GP model (left panel, see section 5.4.1) and 2D GP model (left panel, see section 5.4.2). The posterior distribution colours correspond to the same cases as that in Figure 5.5. The shaded regions in both panels mark the $\pm 1\sigma$ bound for the posterior of the corresponding colour. The horizontal scale for both panels is the same. As specified in the legend, it can be seen that the 2D GP model, for both prior cases, yields better precision on $T_0$ as compared to the 1D GP models.

Similar to the results for $r_p$, we find that the 2D GP model also performs best in case of $t_0$. The posterior distributions for $T_0$ are shown in Figure 5.6. We observe that the 1D GP model with time and comparison star as GP regressors yields a precision of 28 seconds on $T_0$, which is about 2 times better as compared to when not using the comparison star. In comparison, the 2D GP model gives a precision of 21 seconds, which is 25% better than the 1D GP model. Early TESS study of WASP-18b by Shporer et al. (2019) ruled out significant TTVs. However, more recent analyses by Pearson (2019) show significant evidence for TTVs at the order of $\pm 25$. The $T_0$ precision we obtain is of the same order and could provide marginal additional constraint to the TTV model of the planet. Improved precision from future observations from newer detectors as we discuss in Section 5.5.2 could provide better $T_0$ precisions and better constraint to the TTVs.

The 2D GP model also gives median $T_0$ consistent well within $1\sigma$ for when using uniform or Gaussian priors on $T_0$ for tuning the GP hyperparameters. We discuss the implication of such precision on the mid-transit times for constraining the TTVs for small planets in Section 5.5.3 and 5.5.4.

5.5.2 Caveats of the GMOS measurements

We now discuss some caveats associated with our GMOS measurement of transit parameters of WASP-18b and describe potential ways to resolve them. The first of them is the observational setup of the GMOS observations which were taken in 2012 with the old e2v detectors. As seen in Figure 5.1, the e2v detectors suffer from decreasing efficiency and fringing beyond 700 nm which was significantly improved upon by the new Hamamatsu detectors (Scharwächter et al. (2018)). The Hamamatsu detectors provide an increase in the usable wavelength range by $\sim 50\%$ and improvement in quantum efficiency by 20 to 60% in the red
optical. As a result of the combination of increased wavelength coverage and increased detector efficiency, we expect that future measurements for WASP-18b from the new detector could yield nearly 2 times better photometric precision on the light curves. This could result in a $\sqrt{2}$ improvement on the transit depth precision. Additionally, the transit of WASP-18b analysed in this paper was observed in slow readout and high gain mode of GMOS, resulting in a duty cycle of 33%. For targets as bright or brighter than WASP-18, fast-readout and low-gain mode can improve the duty cycle and hence also improved photometric precision on the light curves based on photon noise.

Another caveat stems from the limitations of the 2D GP model as we apply to the WASP-18b data in this paper. We use the simplest form of the 2D GP model as implemented by specgp where the wavelength dependent change in the GP model for M wavelength bins is governed by the scaling vector $\mathbf{\alpha} (\alpha_1, \alpha_2, ... \alpha_M)$ where each $\alpha_i$ is a free parameter. As discussed in Gordon et al. (2020), this form of specgp, with the wavelength dependent covariance matrix $\alpha \alpha^T$, also scales fastest in terms of computational time. However, as shown in Figure 5.3, we notice that this form of the model works relatively less efficiently for the 7 reddest bins for which there is still low-frequency correlated noise present in the residuals. Note that the wavelength dependent low-frequency trend in our 2D GP model is only taken care of by the wavelength dependent airmass trend, and the single timescale hyperparameter ($\eta_t$) governs the high-frequency systematics scaled for each wavelength bin by $\alpha$. We found that the PyMC3 NUTS sampler for such a model converged fastest. A future implementation of specgp with faster scaling for a more general wavelength dependent covariance matrix constructed using a wavelength dependent length scale hyperparameter ($\eta_\lambda$) could improve the performance across the whole bandpass.

Interestingly, as seen in Figure 5.2, the 7 reddest $\lambda$LCs also visually show larger deviation in the behaviour of time dependent systematics as compared to the $\lambda$LCs for the bluer bins. The origin of this deviating behaviour for the reddest $\lambda$LCs are unclear, but we speculate that it could be due to lower levels of fringing present even in the 650 to 700 nm range for the e2v detectors. This won’t be an issue for observations taken using newer Hamamatsu detectors.

### 5.5.3 Predictions for ground-based follow-up of rocky exoplanets

We find that the 2D GP model measures the transit depth of WASP-18b at a precision very close to that expected from photon noise despite not using the comparison star. This opens up the unprecedented possibilities for ground-based follow-up of transits of small planets orbiting bright stars with no suitable nearby stars in the field of view. To quantify the expected performance of a similar observational setup as that for WASP-18b in this paper for future follow-up observations of small rocky planets, we use the results from Section 5.5.1 as a starting point.
The transit depth precision of WASP-18b ($\Delta \delta_{W_{18b}} = 150$ ppm) attained by GMOS observations is close to that expect from just photon noise, hence we scale that to estimate the the expected transit depth precision for a planet $P$ ($\Delta \delta_P$), which is given by

$$\Delta \delta_P = \left( \frac{\sqrt{N_{W_{18b}}}}{\sqrt{N_P}} \right) \left( \frac{\Delta \delta_{W_{18b}}}{10^{V_{W_{18b}} - V_P}} \right)$$  \hspace{1cm} (5.3)$$

where $N_{W_{18b}}$ is the number of exposures during transit for WASP-18b, $N_P$ is the number of exposures in transit for the planet $P$, $V_{W_{18b}}$ is the $V_{mag}$ of WASP-18, and $V_P$ is the $V_{mag}$ of the host star of planet $P$. $N_{W_{18b}}$ is 95 from the WASP-18b GMOS transit light curve we analyse in this work. We use the Integration Time Calculator (ITC) for GMOS-S\(^1\) to calculate the optimal exposure times for $V_{mag}$ ranging from 5 to 9.5 in steps of 0.5 mag. We set up the ITC for a source with given $V_{mag}$ and a typical Gaussian profile with FWHM of 1.5 arcsecond (including seeing). For a plate scale of 0.073 arcsec/pixel for e2v detectors, this corresponds to PSF FWHM of about 22 pixels, which is the median FWHM of the cross-dispersion PSF for the WASP-18b GMOS observation. The plate scale for the Hamamatsu detectors is slightly different (0.080 arcsec/pixel).

We set the grating to R150 with a G2V input spectrum, central wavelength set to 760 nm, blocking filter OG515, and 5 arcsecond wide slit let as the focal plane unit. Note that this is the highest slit size available on GMOS-S ITC which is smaller than the typical custom slit widths of 10 arcseconds used for exoplanet observations from MOS. However, for arcsecond level of seeing typical at Mauna Kea or Cerro Pachon, wider slits are essential for avoiding slit losses. While we present calculations for the widest slit size of 5 arcseconds available on GMOS-ITC, for observation planning we recommend adjusting exposure times according to the slit sizes chosen. We select the Hamamatsu detector array with $1 \times 1$ spatial and spectral binning and a readout of only the central 80" spectrum to reduce the readout overhead. We select the high gain and fast read mode of the detector and set the calculation to airmass of 1.5. By varying the exposure time that we obtain counts below 85% saturation limit, we obtain the optimal exposure times for $V_{mag}$ ranging from 5 to 9.5 in steps of 0.5 mag.

We then use the Equation 5.3 to calculate the expected transit depth precisions from photon noise for a grid of $V_P$ and transit durations ranging from 0.5 to 5 hours. We restrict our calculation to maximum transit duration 5 hours because for a typical ground-based observation that would be around the maximum transit duration that will also allow a decent before and after transit baseline. The transit duration controls the $N_P$, for a given optimal exposure time obtained from GMOS-S ITC. A caveat in our scaling calculation is that the WASP-18b observations were taken using older e2v CCD, while the GMOS-S ITC configuration we have used to calculate the scaled precisions is for the new Hamamatsu detectors which only show better efficiency than the e2v CCD. Hence, our scaled precisions represent an upper limit on the expected transit depth uncertainties from photon noise because the same WASP-18b transit we analyse in this paper, if repeated with the Hamamatsu CCD, would have higher...

\(^1\)http://www.gemini.edu/instrumentation/gmos/exposure-time-estimation
photon noise SNR and duty cycle. We show the predicted transit depth precisions hence computed in Figure 5.7 and Table 5.3.

From Figure 5.7, we find that for planets with stellar host brightness in the range of $V_{\text{mag}}$ 5 to 9.5 and transit duration 0.5 to 5 hours, the expected transit depth uncertainties from just photon noise and a single transit ranges from $\sim 10$ to 400 ppm. As expected, planets with brighter host stars and longer transit durations should yield the smallest transit depth uncertainties. From the overplotted 22 confirmed small planets ($R \leq 2 R_\oplus$, Rogers (2015)), with $P_{\text{orb}} \leq 20$ days and stellar $V_{\text{mag}} \leq 9.5$, we find that the transits of 9 planets can be detected at more than $3\sigma$ using GMOS with just a single transit. Observing two transits extends this number to 11 planets. We summarize these predictions in Table 5.3. Note from Table 5.3 that as expected for a majority of bright stars there are no suitable comparison stars within 5.5' (FoV of GMOS) with $\Delta V_{\text{mag}} \leq 1$. For some targets for which we do have such a comparison star available, it is likely a variable star which makes multi-epoch observations with it unfeasible (see Panwar et al. (2022a)).

A caveat in our calculations here is that smaller exposure times for very bright stars e.g. 55 Cnc will also increase the scintillation noise contribution in the observations (see e.g., Osborn et al. (2015)). Additional modifications to the observational setup, e.g. a ‘spatial scan’ of the star similar to that done for the program GN-2015A-DD-1\(^1\), could help in reducing the amplitude of scintillation noise. An alternative would be to use a spectrometer on a smaller 4 or 2 metre class telescope which can allow for higher exposure times while staying below saturation and also keeping the scintillation noise amplitude low. The NTT/EFOSC2 on the 4-metre William Herschel Telescope which has been used for transmission spectroscopy observations (Kirk et al. (2018); Ahrer et al. (2022)) is one such instrument that could provide the right sweet-spot for wide slit time series spectroscopy of very bright stars while maintaining high duty cycle and keeping the scintillation noise contribution low. Additionally, using high-speed CCDs like ULTRACAM (Dhillon et al. (2007)) can also help in keeping the duty cycle high for very bright stars. ULTRACAM has already been used for transit spectrophotometry in the past (e.g. Föhring et al. (2013); Kirk et al. (2016)). It uses frame transfer CCDs with readout overhead as small as 24 ms which is perfectly suited for taking multiple short exposures on bright targets before saturating the CCD. Moreover, ULTRACAM can also simultaneously observe in three different wavelength bands. Hence, transit observations taken using ULTRACAM would be excellent cases for applying the 2D GP approach and measuring precise transit properties for small planets orbiting very bright stars.

The observability of multiple transits to improve the transit depth precision would depend on the planet ephemeris uncertainties from previous measurements by the detection facilities e.g. TESS. However, given that $\sim 20$ second precision is achievable on $T_0$ as we show from GMOS observation of WASP-18b, repeated attempts at transit follow-up would eventually help in significantly constraining the TTVs. Repeated transit follow-up would yield multiple complementary outcomes that can help in improving the characterization of small transiting exoplanets. Firstly, precise recent mid-transit times would safeguard the planet ephemeris for potential follow-up using other facilities like JWST. Secondly, improved TTV constraints

\(^1\)https://archive.gemini.edu/programinfo/GN-2015A-DD-1
Figure 5.7: Expected transit depth uncertainties (colormap) from observations taken using GMOS-S R150 grism for exoplanets with transit duration $\leq 5$ hours and orbiting bright stars with $V_{\text{mag}} \leq 9.5$. The expected transit depth uncertainties have been obtained by scaling the precision obtained for WASP-18b using the 2D GP model and by obtaining the optimal exposure time from the GMOS-S ETC. Overplotted in white circles are confirmed exoplanets with radii less than $2\,R_{\oplus}$, orbital period $\leq 20$ days, and host star $V_{\text{mag}} \leq 9.5$. The names and the ratio of their measured transit depths to the scaled precision (in terms of $N\sigma$) are annotated. The planets marked with a red box are the ones whose transit can be measured at more than $3\sigma$ from a single transit using GMOS and simply based on scaled photon noise precisions.
Table 5.3: Predicted photon noise transit depth uncertainties scaled from the WASP-18b transit depth measured in this work for confirmed exoplanets shown in Figure 5.7, and displayed in the table here sorted by their expected detection significance from a GMOS single transit. These are all the confirmed exoplanets orbiting bright stars ($V_{mag} \leq 9.5$), radii less than 2 R$_{\oplus}$, and orbital period $\leq 20$ days. The columns show the planet name, planet radius (in terms of earth radius), planetary orbital period in days (P[d]), equilibrium temperature in K ($T_{eq}[K]$), transit duration in hours ($T_{dur}[h]$), measured transit depth in ppm ($\delta[ppm]$), scaled uncertainties from WASP-18b transit in ppm ($\Delta\delta[ppm]$), number of sigma detection expected based on $\Delta\delta(\delta/\Delta\delta)$, and the $V_{mag}$ of the comparison star within 5.5’ (field of view (FoV) of GMOS) of the target star closest to the $V_{mag}$ of the target star. The magnitudes were derived using an astroquery (Ginsburg et al. 2019) search on SIMBAD. In the column planet name, a marks the stars for which no stellar object was found in GMOS FoV, b denotes that the comparison star is a rotating variable, and c denotes that the comparison star is an RR Lyrae variable.

<table>
<thead>
<tr>
<th>Planet</th>
<th>Radius (R$_{\oplus}$)</th>
<th>P [d]</th>
<th>$T_{eq}$[K]</th>
<th>$T_{dur}$[h]</th>
<th>$\delta$[ppm]</th>
<th>$V_{mag}$</th>
<th>$\Delta\delta$[ppm]</th>
<th>$\delta/\Delta\delta$</th>
<th>$V_{mag,comp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD 219134 c</td>
<td>1.511</td>
<td>6.7646</td>
<td>782</td>
<td>1.66</td>
<td>318</td>
<td>5.57</td>
<td>5</td>
<td>61.8$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 136352 b</td>
<td>1.664</td>
<td>11.578</td>
<td>905</td>
<td>3.94</td>
<td>208</td>
<td>5.65</td>
<td>3</td>
<td>59.96$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 219134 b</td>
<td>1.602</td>
<td>3.0929</td>
<td>1015</td>
<td>0.94</td>
<td>358</td>
<td>5.57</td>
<td>6</td>
<td>52.49$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>55 Cnc e</td>
<td>1.875</td>
<td>0.7365</td>
<td>1958</td>
<td>1.56</td>
<td>375</td>
<td>5.95</td>
<td>8</td>
<td>46.03$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HR 858 c</td>
<td>1.939</td>
<td>5.9729</td>
<td>1326</td>
<td>2.9</td>
<td>184</td>
<td>6.38</td>
<td>9</td>
<td>20.15$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 158259 b</td>
<td>1.292</td>
<td>2.178</td>
<td>1478</td>
<td>2.04</td>
<td>70</td>
<td>6.48</td>
<td>11</td>
<td>6.15$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 21749 c</td>
<td>0.892</td>
<td>7.7899</td>
<td>701</td>
<td>2.52</td>
<td>205</td>
<td>8.08</td>
<td>42</td>
<td>4.78$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 213885 b</td>
<td>1.745</td>
<td>1.008</td>
<td>2128</td>
<td>1.5</td>
<td>200</td>
<td>7.95</td>
<td>52</td>
<td>3.82$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 15337 b</td>
<td>1.64</td>
<td>4.7562</td>
<td>1001</td>
<td>2.49</td>
<td>344</td>
<td>9.1</td>
<td>109</td>
<td>3.15$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 108236 b</td>
<td>1.615</td>
<td>3.796</td>
<td>1099</td>
<td>2.38</td>
<td>285</td>
<td>9.25</td>
<td>119</td>
<td>2.39$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 3167 b</td>
<td>1.6</td>
<td>0.9596</td>
<td>1759</td>
<td>1.65</td>
<td>294</td>
<td>8.97</td>
<td>126</td>
<td>2.33$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>TOI-431 b</td>
<td>1.28</td>
<td>0.49</td>
<td>1862</td>
<td>1.01</td>
<td>291</td>
<td>9.12</td>
<td>172</td>
<td>1.69$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-21 b</td>
<td>1.639</td>
<td>2.7858</td>
<td>2025</td>
<td>3.59</td>
<td>62</td>
<td>8.25</td>
<td>38</td>
<td>1.6$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>TOI-1860 b</td>
<td>1.31</td>
<td>1.0662</td>
<td>1885</td>
<td>1.14</td>
<td>130</td>
<td>8.4</td>
<td>92</td>
<td>1.4$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 93963 A b</td>
<td>1.35</td>
<td>1.0391</td>
<td>2042</td>
<td>1.81</td>
<td>141</td>
<td>9.18</td>
<td>132</td>
<td>1.07$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>HD 80653 b</td>
<td>1.613</td>
<td>0.7196</td>
<td>2301</td>
<td>1.8</td>
<td>159</td>
<td>9.45</td>
<td>184</td>
<td>0.86$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-444 f</td>
<td>0.741</td>
<td>9.7405</td>
<td>673</td>
<td>1.69</td>
<td>75</td>
<td>8.87</td>
<td>119</td>
<td>0.63$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-444 e</td>
<td>0.546</td>
<td>7.7435</td>
<td>727</td>
<td>2.79</td>
<td>51</td>
<td>8.87</td>
<td>92</td>
<td>0.55$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-444 d</td>
<td>0.53</td>
<td>6.1894</td>
<td>783</td>
<td>2.21</td>
<td>47</td>
<td>8.87</td>
<td>104</td>
<td>0.46$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-444 c</td>
<td>0.497</td>
<td>4.5459</td>
<td>868</td>
<td>1.74</td>
<td>44</td>
<td>8.87</td>
<td>117</td>
<td>0.37$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-444 b</td>
<td>0.403</td>
<td>3.6001</td>
<td>938</td>
<td>2.05</td>
<td>27</td>
<td>8.87</td>
<td>108</td>
<td>0.25$\sigma$</td>
<td>–</td>
</tr>
<tr>
<td>Kepler-408 b</td>
<td>0.82</td>
<td>2.465</td>
<td>1561</td>
<td>1.23</td>
<td>30</td>
<td>8.83</td>
<td>137</td>
<td>0.22$\sigma$</td>
<td>–</td>
</tr>
</tbody>
</table>
would be crucial for constraining the mass of small planets, the majority of which are known to reside in multi-planet systems. This is especially critical for measuring the mass of small planets around active stars which impede precise RV mass measurement (e.g. V1298 Tau system David et al. (2019), Maggio et al. (2022)). Thirdly, repeated transit follow-up would be also helpful in measuring the impact of stellar activity on the measured transit depth and transit spectra at multiple epochs (Panwar et al. (2022a)). Note that the calculations we present here are a general and idealized prediction scaling from the WASP-18b results and only accounting for the brightness of host star and transit duration of the planet. The actual observational setup for the follow-up of a planet may need to be adjustments to exposure time depending on the nature of the field of view, observing conditions, and use of alternative GMOS gratings.

5.5.4 Implications for current and future facilities

The fact that it is possible to obtain transit depth precisions required to follow up the niche of small planets around bright stars from ground-based spectrophotometry has implications for the current transit survey like TESS and in the near future for PLATO. The limited duration of TESS spent per sector (~27 days) means that several small planets with orbital periods within 20 days require additional follow-up for refining their transit parameters and constraining their ephemerides. Ground-based spectrophotometry over one or multiple epochs, as we show in this work, can provide flexible and less expensive (as compared to space based telescopes) means to do so. In case of TTVs for multi-planet systems, constraining mid-transit times precisely can be helpful in measuring the planetary masses. This is furthermore relevant in the context of multi-planet systems around bright and active stars e.g. V1298 Tau system (David et al. (2019); Feinstein et al. (2019)) for which RV measurements are impeded by the high levels of stellar activity contamination.

The PLATO mission, slated to launch in December 2026, will primarily focus on finding terrestrial planets in the habitable zones of sun like stars. The goal of PLATO is to observe at least 15000 F5 to K7 type stars with $V_{mag} \leq 11$ and at least 1000 stars of the same spectral type range with $V_{mag} \leq 8.5$ (Montalto et al. (2021)). The expected precision for these two samples of stars is 50 ppm in 1 hour. One of the science goals of PLATO as defined in the PLATO Definition Study report is to detect at least a sample of 100 exoplanets precisely characterized for their transit properties and radii (better than 3%)1. From yield simulations by Heller et al. (2022), it is expected that PLATO could detect about 8 to 34 earth size planets (radii between 0.5 $R_\oplus$ and 1.5 $R_\oplus$) in their habitable zones around F5 to K7 main-sequence stars. In the context of planets detected around $V_{mag} \leq 8.5$, ground-based spectrophotometry as we demonstrate in this work will be a useful complement for improving constraints on the the transit parameters and safeguarding the ephemeris for future follow-up measurements. Moreover, the 2D GP approach we apply in this work could also be applicable to the photometry conducted by the two PLATO fast cameras in two filters for planets with $V_{mag}$ between 4 and 8.

1PLATO Definition Study report:https://sci.esa.int/web/plato/-/59252-plato-definition-study-report-red-book
5.6 Conclusion

In this work we have explored the feasibility of following up the observationally formidable (from ground) sample of small and rocky exoplanets orbiting bright stars ($V_{mag} \leq 9.5$). We analyse a single GMOS transit observation of WASP-18b, a hot-Jupiter transiting a bright star ($V_{mag} = 9.3$), using 1D and 2D GP models to correct the systematics in the white and spectroscopic target transit light curves. We find that the 2D GP model, implemented using specgp, provides the best precision on the transit depth and mid-transit times despite not using comparison star light curves at any stage. Specifically, the 2D GP model is able to retrieve the transit depth within $\sim 5\%$ of the photon noise, and yields mid-transit time with a $1\sigma$ precision of 21 seconds. Based on these results, we use the GMOS-ITC to predict that ground-based spectrophotometry from instruments like GMOS can be used to follow-up 9 out of 22 known exoplanets around stars with $V_{mag} \leq 9.5$, radii $\leq 2R_\oplus$ within 20 days orbital period and 5 hours transit duration. A majority of these bright stars do not have suitable comparison stars within 6", which is where the 2D GP model could aid in correcting for the transit light curve systematics for them. Ultimately, our results show that high-precision ground-based spectrophotometry could be highly complementary in the follow-up of terrestrial exoplanets discovered around bright stars by TESS and near future by PLATO.

5.7 Acknowledgements

Based on observations obtained at the Gemini Observatory (acquired through the Gemini Observatory Archive and Gemini Science Archive), which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil). Based in part on Gemini observations obtained from the National Optical Astronomy Observatory (NOAO) Prop. ID: 2012B-0398; PI: J.-M Désert. J.M.D acknowledges support from the Amsterdam Academic Alliance (AAA) Program, and the European Research Council (ERC) European Union's Horizon 2020 research and innovation programme (grant agreement no. 679633; Exo-Atmos). This work is part of the research programme VIDI New Frontiers in Exoplanetary Climatology with project number 614.001.601, which is (partly) financed by the Dutch Research Council (NWO). This material is based upon work supported by the NWO TOP Grant Module 2 (Project Number 614.001.601). This material is based upon work supported by the National Science Foundation (NSF) under Grant No. AST-1413663. This research has made use of NASA’s Astrophysics Data System. The authors also acknowledge the significant cultural role and reverence the summit of Mauna Kea has within the indigenous Hawaiian community. This research has made use of Astropy,\(^1\) a community-developed core Python package for Astronomy Astropy Collaboration et al. (2013, 2018), NumPy Harris et al. (2020), matplotlib Hunter (2007), SciPy Virtanen et al.

\(^1\)http://www.astropy.org
Small exoplanets around bright stars from ground

(2020) and IRAF Tody (1986) distributed by the NOAO, which is operated by AURA under a cooperative agreement with the NSF.
Chapter 6

Common trends in close-in gas giant exoplanet atmospheres from MOS surveys

*Using a library of tools for modelling spectrophotometric transit light curves with Gaussian Processes (GHALIB)*

Vatsal Panwar, Jean-Michel Désert et al.

*To be submitted to the Monthly Notices of the Royal Astronomical Society*

**Abstract**

Transit spectrophotometry has long been the prevalent observational strategy for characterizing close-in exoplanets and their atmospheres. Both ground and space based observations commonly suffer from systematics of non-astrophysical origin along both time and wavelength axes, either due to the instrument or the earth's atmosphere. This often distorts the underlying planetary signal and has proven to be a major hurdle in measuring the planetary transit and atmospheric properties at the required precision and accuracy. We introduce ghali, a library of tools tailored for a typical analysis of transit spectrophotometry observations from ground or space based facilities. ghali brings together state-of-the-art codes used for modelling exoplanet transits, data-driven systematics model using Gaussian Processes, and efficient likelihood sampling for parameter estimation. The workflow of the code is nested in a flexible modelling framework suitable for easy model manipulation via e.g. changing combination of regressors, and changing free and fixed model parameters. We benchmark the application of ghali by using it to analyse ground-based transit spectrophotometry observations of ten gas giants from Gemini/GMOS and Keck/MOSFIRE, totalling
to ∼220 hours of cumulative observations. Our results demonstrate the capability of ghalib to analyse a diverse set of observations taken using different instruments and in varying observing setups and conditions.

6.1 Introduction

With their large and diverse population, characterization of exoplanets presents the opportunity to study the outcomes of planet formation and evolution from a statistical perspective. Elemental abundance ratio measurement of solar system planets and their atmospheres over the last several decades have helped to constrain theoretical planet formation models. Analogous bulk density measurement from precise mass and radius, and atmospheric measurements from transit spectroscopy for a large sample of exoplanetary systems can provide further constraints and observational evidence of the effect of system properties like stellar metallicity (Gonzalez (1997); Meléndez et al. (2009)), stellar activity evolution (Poppenhaeger et al. (2021)), and planetary architecture (e.g. system obliquities, Bourrier et al. (2022)). Early efforts, primarily restricted to using space based facilities like HST and Spitzer due to the required high transit/eclipse depth precision (∼ 100 ppm), detected the atmospheric signatures of hot Jupiters e.g. HD 209458b in both transmission (Charbonneau et al. (2002), Vidal-Madjar et al. (2003))) and emission (Deming et al. (2005)). Intensive space based surveys since then have been conducted for transiting exoplanets using HST/STIS and HST/WFC3 (Sing et al. (2016); Crossfield & Kreidberg (2017); Gao et al. (2020); Mansfield et al. (2021)), and Spitzer/IRAC (Baxter et al. (2020), Baxter et al. (2021)) which reveal a complex connection between the nature of exoplanets and their birth and evolution histories.

The last decade has also seen the medium to large ground-based telescopes become the workhorses of exoplanet characterization, especially their atmospheres. At the forefront of ground-based time series spectroscopic observations of exoplanet transits have been mainly two techniques — low-resolution multi-object spectroscopy (MOS) (e.g. Bean et al. (2010), Panwar et al. (2022b)) and high-resolution Doppler spectroscopy (e.g. Snellen et al. (2010); Birkby (2018)). Both independently probe different pressures levels and thermal regimes of an exoplanet’s atmosphere and have been extensively used independent of each other to detect and constrain abundances of prominent molecular absorbers. Particularly in the context of low-resolution spectroscopy, the visible wavelength range (400 to 900 nm) of the transmission spectrum of hydrogen dominated atmospheres is sensitive to absorption features due to alkali species (Na, Li, K), Rayleigh scattering or scattering due to cloud condensates, and molecular absorbers like TiO and VO (Hubeny et al. (2003)). Ground-based MOS observations aiming to measure these atmospheric signatures have yielded low-resolution (10-20 nm wide wavelength bins) transit spectra in the visible (e.g. Espinoza et al. (2019)) and near-infrared (e.g. Bean et al. (2011)) for archetypes of individual subclasses of exoplanets across the population. These range from cool super-Earths like GJ-1132b (Diamond-Lowe et al. (2018)) around M-dwarfs to ultra-hot inflated Jupiters like WASP-103b (Wilson et al. (2020); Kirk et al. (2021)).
A common bottleneck apparent from the prolific number of studies in both ground and space-based spectroscopy of exoplanets is the ability to disentangle transit light curve systematics from planetary signals. These systematics can sometimes originate at the level of the detector e.g. intra-pixel gain variations in Spitzer IRAC (Stevenson et al. (2012)), charge-trapping in HST/WFC3 (Deming et al. (2013)), fringing in case of GMOS. They can also originate at a higher level in the instrument that can affect the optical path of the light, e.g., due to pointing or thermal variations in Spitzer IRAC and HST/WFC3, instrument shutter blades and cassegrain rotator in case of GMOS, and imperfect ADC in VLT/FORS2. For both ground and space-based instruments, these various effects eventually manifest together in form features in the transit light curves correlated in time and wavelength, and at timescales and amplitudes similar to the planetary signal. Some examples are the hook pattern seen in HST/WFC3 light curves (e.g. Wakeford et al. (2016)), high-frequency systematics correlated with the PSF centroid in Spitzer IRAC light curves (e.g. Désert et al. (2011)), and odd-even offsets due to the shutters and abrupt kink in the light curve due to the Cassegrain rotator position angle (CRPA) changing sign close to the zenith in GMOS light curves (e.g. Stevenson et al. (2014)). For ground-based observations, telluric absorption adds another layer of non-astrophysical variations in the light curves. In summary, exoplanet transit light curves from both and space need to correct complex systematics at various timescales while measuring the transit signals, and a comprehensive set of tools are needed for this purpose. In this paper we present ghalib, a library of such tools tailored for modelling spectroscopic transit light curves.

For ground-based observations, a survey approach to measuring transit light curves can be advantageous in several aspects. Specifically in the context of low-resolution spectroscopy, using the same instrument and observational setup for different targets can be helpful in identifying and compensating for the sources of uncertainties and systematics that degrade the accuracy and precision of such measurements. This includes telluric and instrumental systematics that are expected to behave similarly for the same instrument and observing site. An example of this is the odd-even effect aforementioned for GMOS light curves, which has been seen in several observations taken using the instrument (e.g. Stevenson et al. (2014); Panwar et al. (2022b,a)). Moreover, repeated measurements for the same target over multiple epochs can enable correction of the effect of stellar variability on the transmission spectrum (e.g., Zellem et al. (2017); Espinoza et al. (2019); Panwar et al. (2022a)). In summary, a comprehensive survey can ensure the reliability of ground-based optical measurements and reinforce their joint interpretation with space-based infrared observations (e.g. Nikolov et al. (2018); Nikolov et al. (2022)).

To date MOS based measurements of exoplanet spectra have been obtained for specific targets and using diverse instruments with variations in observational setups and techniques used to correct for transit light curve systematics. A comparative study using the same ground-based instrument for an ensemble of exoplanets spanning a broad range of system parameters (e.g., planet surface gravity and stellar insolation) and all observed and analysed in a homogeneous manner is yet to be done. With this motivation we apply ghalib to a survey of transiting gas-giants observed in visible band using Gemini/GMOS and in near-infrared
K band using Keck/MOSFIRE. In this work, we present the outcomes of ghalib in the form of exoplanet transmission and emission spectra, along with refined transit parameters. Application of the same framework to a large set of light curves provides insights particularly useful for current and future ground-based observations and their analysis strategies.

This paper is distributed as follows. In Section 6.2 we provide an overview of ghalib elaborating on its features, workflow, and the type of datasets it is applicable to. In Section 6.3, we describe the observations from our ground-based survey of transiting gas-giants to which we apply ghalib. In Section 6.4 we describe the data reduction steps to obtain the spectral time series from 2D raw frames. In Section 6.5 we discuss the application of ghalib to the survey observations. In Section 6.6 we present our results and discuss key outcomes of the application of ghalib to the survey of MOS observations. In Section 6.6.3 we provide our conclusions and comment on the future improvements to the framework of our code.

6.2 Overview of ghalib

We present the Python package ghalib\(^1\) which implements a versatile framework for analysing spectroscopic time series observations of exoplanet transits. ghalib at its core is a compilation of tools commonly used in exoplanet science assembled in a framework suited for analysing exoplanet transit light curves, especially those observed in time and wavelength dimension. We illustrate the important aspects of ghalib in a flow-diagram in Figure 6.1.

Overall ghalib contains modules that individually parse the data, model the transit light curve based on calculated limb darkening coefficients, model the the systematics using Gaussian Processes (GP), compute maximum a-posteriori (MAP) solution and perform posterior sampling based on user specified priors.

A typical dataset should have at minimum the following components: time series of 1D stellar spectra of all the stars observed simultaneously (in case of MOS observations), wavelength solution, time stamps of each spectrum, and a set of time series of variables corresponding to each exposure like airmass, cassegrain rotator position angle (CRPA), PSF width of the stellar spectral traces etc. ghalib parses the data into white light curve, spectroscopic light curves (in user specified wavelength bins), and a set of regressors for the GP model.

Transit and GP modelling is nested in a modelling protocol adapted and generalized for a combined transit and GP model from the model class introduced in the GP package george\(^2\). This modelling class provides an easy and flexible way for fixing or setting free the transit model and GP kernel parameters whenever desired for an instance of model. This gives exceptional control to the user for manipulating model parameters at any point in a workflow of a typical light curve analysis.

For transit modelling we provide a choice of two models : 1) numerical integration transit modelling using the code batman (Kreidberg et al. (2014)), 2) analytic transit model using the code starry (Luger et al. (2019)). The limb-darkening coefficients for given stellar prop-

\(^1\)Under development in GitHub repository at https://github.com/vatsalpanwar/ghalib
\(^2\)https://george.readthedocs.io/en/latest/user/modeling/
6.2 Overview of ghalib

Figure 6.1: A flow diagram describing the main modules of ghalib. The code takes the data in the form of time series of 1D stellar spectra for both target and available comparison stars, along with time series of observation metadata which mainly include the FWHM and position of stellar spectral traces and the cassegrain rotator position angle (CRPA). It then extracts the white and spectroscopic light curves for each star. Based on the user specifications it calculates the model using a choice of tools to calculate the transit model, GP systematics model, and limb darkening coefficients. Sampling of the posterior distribution is done using emcee or dynesty, accounting for the user specified priors. The main outputs, among other useful diagnostics, are the best fit transit parameters and transmission or emission spectrum of the planet.
erties (stellar effective temperature, surface gravity, and metallicity) are calculated using the package PyLDTk (Parviainen & Aigrain (2015)). For GP modelling, we provide a choice of using george (Ambikasaran et al. (2015)) and celerite (Foreman-Mackey et al. (2017)) for fast and scalable GP computation.

We emphasize that the current implementation of ghalib is designed to fit a single light curve i.e. a white light curve or a single spectroscopic light curve in an instance. We intend to extend ghalib to also include joint fitting of multiple light curves in future. In the current design, ghalib computes the posterior probability by combining the GP likelihood for a light curve and the prior computed based on the user specifications for priors for each free parameter in the model. We provide a priors class that can be easily edited to include additional priors. For computing the MAP solution we use the optimize module of the SciPy package. The MAP solution can then be used as a starting point for MCMC sampling using either emcee (Foreman-Mackey et al. (2013)) or nested sampling using dynesty (Speagle (2020)). In case of dynesty, the current implementation of ghalib provides prior transforms for only uniform and Gaussian priors.

While the development of ghalib occurred in the context of low-resolution ground-based MOS observations, especially for the results presented in Panwar et al. (2022b), we emphasize that with little tweaking it can easily be applied to spectroscopic transits at higher resolutions and those observed from space based facilities like HST and JWST. The prime utility of ghalib is that it provides a generalized and flexible framework for modelling transit light curves and their systematics using various model configurations specified by covariance kernel types and GP regressor combinations. Additionally, it also computes criteria like Bayesian Information Criterion (BIC) and Bayesian evidence useful for Bayesian model comparison between each model configuration.

We demonstrate application ghalib to a collection of spectrophotometric observations of transiting gas-giants taken using Gemini/GMOS and Keck/MOSFIRE in the subsequent sections.

### 6.3 Observations from a ground-based survey of transiting gas-giants

#### 6.3.1 Gemini/GMOS observations

We observed transits of a sample of gas-giants using GMOS-N and GMOS-S as part of two large Gemini programs (Proposal IDs: 2012B-0398, PI: Désert; 2017A-LP-6, PI: Huitson; 2015B-DD-3, PI: Dragomir) with the goal of performing comparative exoplanetology. These were the first of their kind survey programs implemented from 2012 to 2017 to characterize a set of confirmed well known exoplanets. Note that these programs were implemented well before TESS era and hence do not include many of the newly discovered TESS planets which would be feasible to follow-up using ground-based MOS observations. The planets targeted in our sample along with the details of the observations are mentioned in Table 6.1. As shown in Figure 6.2, the planets in our sample span a wide range of equilibrium temperature, planetary mass, and planetary radius.
Table 6.1: Details of the GMOS and MOSFIRE transit observations newly analysed in this work. Observation details for XO-2Nb, HAT-P-26b, and WASP-19b can be found in Chapters 2, 3, and 4 respectively. In order the rows show the planet name, the observation number and the grating used, program ID, date of observation, magnitude of the target, separation of the comparison star used in the analyses (θ in arc minutes), and difference between the magnitude of the target and comparison stars.

<table>
<thead>
<tr>
<th>Planet No., Grating</th>
<th>Program ID</th>
<th>Date</th>
<th>V\text{mag}</th>
<th>θ['']</th>
<th>ΔV\text{mag}</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAT-P-7b 1,R150</td>
<td>GN-2012B-Q-8-60</td>
<td>25/07/2012</td>
<td>10.48</td>
<td>2</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>GN-2012B-Q-8-69</td>
<td>05/08/2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3,R150</td>
<td>GN-2013B-Q-14-14</td>
<td>06/08/2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASP-12b 1,R600</td>
<td>GN-2011B-C-1</td>
<td>25/01/2012</td>
<td>11.57</td>
<td>3</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>2,R600</td>
<td>GN-2011B-C-1</td>
<td>26/01/2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3,B600</td>
<td>GN-2015B-LP-6-94</td>
<td>10/01/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,B600</td>
<td>GN-2016A-LP-3-28</td>
<td>02/02/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5,B600</td>
<td>GN-2016A-LP-3-52</td>
<td>14/02/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6,B600</td>
<td>GN-2016A-LP-6-38</td>
<td>22/12/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7,B600</td>
<td>GN-2016A-LP-6-80</td>
<td>02/01/2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASP-18b 1,R150</td>
<td>GS-2012B-Q-6-57</td>
<td>28/10/2012</td>
<td>9.3</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td>WASP-4b 1,B600</td>
<td>GS-2011B-Q-45-5</td>
<td>10/10/2011</td>
<td>12.48</td>
<td>1.3</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>GS-2012B-Q-6-30</td>
<td>16/10/2012</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3,R150</td>
<td>GS-2013B-Q-44-16</td>
<td>11/10/2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,R150</td>
<td>GS-2014B-Q-45-12</td>
<td>24/09/2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAT-P-1b 1,R150</td>
<td>GN-2012B-Q-67-61</td>
<td>12/11/2012</td>
<td>9.87</td>
<td>0.2</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>2,B600</td>
<td>GN-2015B-LP-3-29</td>
<td>02/11/2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASP-80b 1,R150</td>
<td>GN-2013B-Q-14-18</td>
<td>04/08/2013</td>
<td>11.9</td>
<td>0.6</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>GN-2014A-Q-32-24</td>
<td>22/06/2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3,B600</td>
<td>GN-2015B-LP-3-21</td>
<td>10/09/2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,B600</td>
<td>GN-2016B-LP-6-29</td>
<td>28/07/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5,B600</td>
<td>GN-2016B-LP-6-70</td>
<td>06/09/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GJ-3470b 1,R150</td>
<td>GN-2014B-Q-12-88</td>
<td>16/11/2014</td>
<td>12.3</td>
<td>2.9</td>
<td>-2</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>GS-2014B-Q-9-93</td>
<td>02/01/2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3,B600</td>
<td>GN-2015B-DD-3-31</td>
<td>04/03/2016</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4,B600</td>
<td>GN-2015B-DD-3-39</td>
<td>03/04/2016</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\begin{tabular}{llll}
K\text{mag} & ΔK\text{mag} & \hline
WASP-12b MOSFIRE Ks & C182 & 28/02/2013 & 10.18 & 3 & -1
\end{tabular}
Figure 6.2: Mass-radius diagram of transiting exoplanets observed using GMOS and MOSFIRE and analysed using ghalib in this work. The colour of each point representing a planet shows the planetary equilibrium temperature ($T_{eq}$, while the size of the circle around each point is proportional to the scale height of the planet. It can be seen that the planets in our sample span a wide range of radii, masses, and equilibrium temperatures.)
All the planets were observed during their primary transit ensuring a before and after transit baseline of duration approximately equal to the transit duration. The observational setup for all the observations was similar to that described in past publications from the two programs (Huitson et al. (2017); Todorov et al. (2019); Panwar et al. (2022b,a)). For both GMOS-N and GMOS-S, we used the MOS mode of the instrument to obtain time series spectra of the target star during the transit, and at the most two comparison stars in the field of view that were closest in brightness as compared to the target star. We used 10" wide and 30" long slits for each star in the mask to minimize slit loss and to make sure that there was enough sampling of the background. The exposure times were set to ensure that the counts for the target star remained in the linear regime and well below the saturation limit.

Both GMOS-S and GMOS-N were upgraded in the 14B and 17A semesters respectively from the older e2v detectors to the new Hamamatsu detectors. The Hamamatsu detectors overall provide better quantum efficiency and improved response and minimal fringing at the red end of the spectrum (Gimeno et al. (2016); Scharwächter et al. (2018)). Hence, some targets in our sample were observed using both detectors over the course of our survey program.

6.3.2 Keck/MOSFIRE observations

We observed primary transits and secondary eclipses of a sample gas-giants in the Ks band (≈1.9 to 2.4 μm) using the newly installed Keck/MOSFIRE instrument in 2012 as part of the program C182M (PI : Désert). In this work we choose to analyse the secondary eclipse of the ultra-hot Jupiter WASP-12b, which was also intensively followed up using GMOS as described in Section 6.3.1. Ultra-hot Jupiters have emerged as an exciting subclass of exoplanets which provides the opportunity to study planetary physics in rare and extreme regimes (Arcangeli et al. (2018); Lothringer et al. (2018); Lothringer & Barman (2019)). Similar to the GMOS observations, the secondary eclipse of WASP-12b using MOSFIRE was observed in the MOS mode of the instrument with a mask providing slits for the target and 2 other comparison stars. Appropriate background subtraction is can be more consequential in terms of the extracted flux in the near-infrared as compared to the visible bands. Hence, the observations were taken using an ABBA pattern. However, as we discuss in Section 6.4, ABBA may not be the best strategy for cases when the wings of the PSF overlap between the consequent A and B position. Further details of the MOSFIRE observations are given in Table 6.1.

6.4 Data reduction

We use the custom IDL pipeline presented in Huitson et al. (2017) to extract the 1D spectra from the 2D raw frames for both GMOS and MOSFIRE data. The pipeline initially developed for reducing GMOS data was adapted for MOSFIRE data. As described in more detail in Huitson et al. (2017), the pipeline performs flat field and bias correction, cosmic ray removal, and spectral tilt correction. The spectral tilt is particularly significant for MOSFIRE data. The pipeline then estimates the background and performs optimal extraction (Horne (1986)) to extract the 1D spectra for each stellar trace for each exposure. For MOSFIRE observation of
WASP-12b specifically, we find that performing classical background subtraction instead of using the consequent ABBA frames is more suitable. This is so because the nodding amplitude was relatively small which meant that overlap between the wings of PSF of adjacent A and B frames deteriorated background subtraction.

For the GMOS data, we use CuAr lamp spectrum obtained on the night of the observation for wavelength calibration. For the MOSFIRE data, we use the CO\(_2\) telluric absorption features in the 2 to 2.1 \(\mu m\) range to calibrate the spectra by fitting an order 3 Chebyshev polynomial to the position of the features. We apply constant offsets with respect to prominent stellar spectral features to align the comparison star spectra to that of the target star, and additionally interpolate the comparison star spectra to the wavelength solution of the target star. We let the GP model for fitting the spectroscopic light curves (\(\lambda LCs\)) account for any residual shifts and stretches of the stellar spectra through a transit observation, as described in Panwar et al. (2022b,a).

Figure 6.3 shows examples of extracted GMOS (from B600 and R150 grisms) and MOSFIRE stellar spectra to illustrate the wavelength coverage of the observations in our survey. For GMOS observations taken using older e2v detectors, we restrict the analysis up to 700 nm avoiding the regions of significant fringing. For MOSFIRE data, we restrict the analysis to 2.1 to 2.35 \(\mu m\), avoiding the regions of significant telluric absorption due to CO\(_2\) and H\(_2\)O.

We construct the white light curves for all observations by summing the flux in the wavelength range of interest for each exposure in an observation. For GMOS we construct \(\lambda LCs\) by integrating in 10 to 20 nm wide bins, while for MOSFIRE we use 25 nm wide bins. In the next section we describe the application of ghalib to the analysis of GMOS and MOSFIRE observations.
6.5 Application of ghalib to GMOS and MOSFIRE observations

We demonstrate one particular workflow of ghalib which we apply to the GMOS and MOSFIRE observations described in Section 6.3. The code can be setup to be used in many variations as desired by the nature of light curves. In this work we apply one single workflow for all the observations. This is the same setup also used in Panwar et al. (2022b,a), and we describe some specifics for this work in brief below:

1) We first find the MAP solution to fit the target star white light curves using comparison star white light curve time as a GP regressors. For GMOS transit light curves specifically, we fit for the planet to star radius ratio \( \frac{R_P}{R_\star} \), mid-transit time \( T_0 \), inclination \( i \), semi-major axis \( a/R_\star \), and linear limb darkening coefficient \( u_1 \). We use Gaussian priors for \( i \), \( a/R_\star \) based on previously measured values in the literature, and \( u_1 \) based on PyLDTk. We use uniform priors for \( \frac{R_P}{R_\star} \) and \( T_0 \).

For MOSFIRE observation of WASP-12b, we perform the analysis of A and B nod position light curves independently, and fix all the eclipse model parameters except planet to star flux ratio \( \frac{F_P}{F_\star} \).

2) We then use the MAP solution as a starting point for sampling the posterior using dynesty, with the default tolerance of \( \Delta \ln Z = 0.5 \) as the stopping criterion. The best fit white-light curve parameters, their \(+1\sigma\) and \(-1\sigma\) values are obtained from their respective marginal posteriors by estimating the 50th, 84th, and 16th percentile values.

3) We then derive a common-mode trend by subtracting the best fit transit model obtained in the previous step from the observed white light curve.

4) Fixing all transit parameters except the wavelength dependent \( \frac{R_P}{R_\star} \) (or \( \frac{F_P}{F_\star} \)) and \( u_1 \), we use the common-mode trend and time as GP regressors to fit each \( \lambda \)LC independently, repeating the steps for the white light curve in 1) and 2) for each \( \lambda \)LC. We use a Gaussian prior for \( u_1 \) around the PyLDTk value calculated for the bin, approximating a top-hat transmission function for each bin. This yields the wavelength dependent transit or eclipse depths for each bin.

We find that unless the comparison star is ideal in terms of similar brightness and spectral type as compared to the target star and relatively close by (e.g. for XO-2Nb in Chapter 2), this approach can be most generally applied to all observations including even those in which case the comparison star light curve shows undesirable behaviour that prove more detrimental than helpful when doing conventional differential spectrophotometry. The ‘quality’ of comparison star light curves for each target is can be identified in terms of brightness difference and distance from the target star as specified in Table 6.1. In general, as intuitively expected we find the closer the comparison stars are in terms of both brightness and distance to the target star the better is their performance when using them to correct for systematics in the target transit light curves.

In the next section, we present the outcomes of our application of ghalib to the GMOS and MOSFIRE observations. In cases except when the comparison star shows ideal correction via differential photometry (e.g. XO-2 Nb), we apply ghalib to fit the target light curves – both
white and spectroscopic – directly as described. We highlight the performance of the code in terms of accuracy and precisions of measurements across our observation sample, and also discuss its use for future ground-based spectrophotometry observations of transiting exoplanets.

### 6.6 Results and Discussion

#### 6.6.1 Transit parameters from the white light curves

We summarize the results in terms of the measured best fit transit parameters and their $\pm 1\sigma$ uncertainties from the white transit light curve fits for GMOS and MOSFIRE each target in our sample in Table 6.2. We show the individual white light curves and their fits in Appendix 6.A. From the the best fit transit parameters in Table 6.2 averaged across all observations for each target, we find that [ghalib](https://github.com/ghalib) is able to retrieve transit parameters consistent with the literature values in many cases improving upon them. For multiple observations of the same target, we obtain transit parameters largely consistent within $3\sigma$. Note that the targets in our sample range have brightness in the range of $V_{mag}$ 9 to 12 (see Table 6.1). On average across the sample, we obtain a residual RMS of the order of $\sim 300$ ppm in the GMOS transit light curves which is about 1 to 2 times the photon noise (depending on the planet and the observation). Note that we use the same [ghalib](https://github.com/ghalib) setup to fit the transit light curves for all observations which is using time and comparison star light curve as GP regressors. We note that each target and its observation could warrant further exploration of suitable GP regressors combinations which could improve the results. A more thorough approach would be using all combinations of relevant GP regressors (e.g. as done in [Panwar et al. (2022b)](https://arxiv.org/)) and perform Bayesian model comparison by using a model selection criterion like Bayesian evidence which is calculated by dynesty. However, we restrict ourselves to a single combination for all observations for simplicity of demonstration of [ghalib](https://github.com/ghalib) which is the main goal of this work. Our results with the same methodology shown in Table 6.2 nevertheless represent the typical outcomes expected from MOS observations of transiting gas-giants around stars in the brightness range of our sample.

We constrain mid-transit times of transits at average precision of 10 to 20 seconds. This level of transit timing precision can be crucial for constraining TTVs. An example is the study by [Bouma et al. (2019)](https://example.com) who use the mid-transit times measured for WASP-4b from the GMOS transit observations in our sample to constrain the long term TTV trend for the planet.

#### 6.6.2 Transmission and emission spectra across the sample

In Figure 6.4 we show the GMOS transmission spectra of all the planets obtained using [ghalib](https://github.com/ghalib) and combined by multiplying the transit depth posterior distributions obtained by [ghalib](https://github.com/ghalib) for each wavelength bin across all observations for a planet. The MOSFIRE emission spectrum of WASP-12b, along with the previous emission spectrum observed using HST and ground-based broadband secondary eclipse observations, is shown in Figure 6.5.
From an overall examination of the sample we find a diverse nature of transmission spectra ranging from featureless to those containing clear features due to either the planet's atmosphere or stellar photospheric heterogeneity or a combination of both. Some planets in our sample around late K and M dwarfs like WASP-80b and GJ-3470b indicate contributions from stellar contamination at the blue end of the spectrum. This effect has a larger slope and amplitude at the blue end of spectrum which means that simply accounting for offsets or a linear slope between the spectrum from each epoch, as we do for WASP-19b in Panwar et al. (2022a) may not be enough to correct for the effect. A thorough and detailed interpretation of individual transmission spectrum is beyond the scope of this work as we mainly aim to demonstrate the performance of the ghalib framework on a broad sample of observations. Overall we find that that the per 20 nm wavelength bins we obtain transit depth precisions at 1 to 2 times the photon noise.

### 6.6.3 Conclusions and future perspective

In this work we have presented ghalib, a library of tools that can be used for analysing transit spectrophotometry observations of exoplanets. ghalib implements a generalized framework by compiling an array of tools for modelling the transit, constraining the light curve systematics using Gaussian Processes with arbitrary combination of GP regressors, and calculating limb-darkening coefficients based on the host star. It also provides a choice of popular posterior sampling codes like emcee and dynesty for determining the uncertainties on the model parameters through a Bayesian approach. The overall analysis workflow is structured on a modelling framework that allows easy manipulation of fitting for, fixing, or employing priors for all free parameters including the transit and GP model hyperparameters.

We benchmark the performance of ghalib and demonstrate its versatility for different instruments by using it to analyse observations of a sample of ten transiting gas giants observed using ground-based multi object spectrographs GMOS and MOSFIRE. Cumulatively for the whole sample, we apply ghalib to ~220 hours of transit spectrophotometry observations. Our results demonstrate the capability of the framework to measure precise transit parameters and planetary spectra while accounting for time and wavelength dependent systematics arising from a variety of sources.

The package ghalib is under active development for a release to the community. We intend to make further extensions to it to improve its performance in three main directions. One of them is to also allow PyMC3 (Salvatier et al. (2016)) based approach which can be used to employ more complex Bayesian systematics models. Another direction is to add the implementation of the 2D GP model like specgp used in Chapter 5 to model both time and wavelength dependent systematics simultaneously for all the wavelength bins. This approach will account for cross-covariance along the wavelength dimension, which is lacking in the current approach of ghalib that employs a separate GP for each wavelength bin. A third direction is to improve the scalability of the code for larger datasets with large number of data points in time and wavelength dimension. One solution for this is to use GP libraries
Figure 6.4: *ghalib* transmission spectra of all the planets observed using Gemini/GMOS along with the overplotted equilibrium chemistry and solar abundance forward transmission spectrum models for each planet computed using *platon* (Zhang et al. 2019, 2020). Note that the models have been offset in transit depth for comparison with the data and are not fit to the data. The right Y axis of each panel shows the relative scale height for increments in transit depth shown on the respective left Y axis. The colour of markers represents the equilibrium temperature of the planets. Filled circles where shown are for the observations taken using the blue GMOS grating (B600), whereas the empty circles are for the red grating (R150 or R400).
Figure 6.5: Emission spectrum of the ultra-hot Jupiter WASP-12b in the Ks band (2.1 to 2.35 µm) extracted from a single secondary eclipse observation of WASP-12b using Keck/MOSFIRE. The Y axis shows the planet to star flux ratio or the eclipse depth. Also, overplotted in blue are the previous eclipse depth measurements in other bandpasses using broadband photometry: z' band (0.9 µm) from ground at APO (López-Morales et al. (2010)) and in Ks band using CFHT-WIRCam (Croll et al. (2011)), and from spectrophotometry using HST/WFC3 (1 to 1.6 µm, Stevenson et al. (2014)). In green, we show a forward emission spectrum model for isothermal temperature profile of the planet at 3000 K (essentially a blackbody spectrum with features originating due to the stellar spectrum.)
based on JAX e.g. tinygp$^1$. Such libraries can harness the power of GPUs to perform fast covariance matrix inversions which would be essential for feasibly applying them to higher resolution (especially in wavelength dimension) spectrophotometry using e.g. JWST or even ground-based medium to high resolution spectrographs.

### 6.7 Acknowledgements

Based on observations obtained at the Gemini Observatory (acquired through the Gemini Observatory Archive and Gemini Science Archive), which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA), under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil). Based in part on Gemini observations obtained from the National Optical Astronomy Observatory (NOAO) Prop. ID: 2012B-0398; PI: J.-M Désert. J.M.D acknowledges support from the Amsterdam Academic Alliance (AAA) Program, and the European Research Council (ERC) European Union’s Horizon 2020 research and innovation programme (grant agreement no. 679633; Exo-Atmos). This work is part of the research programme VIDI New Frontiers in Exoplanetary Climatology with project number 614.001.601, which is (partly) financed by the Dutch Research Council (NWO). This material is based upon work supported by the NWO TOP Grant Module 2 (Project Number 614.001.601). This material is based upon work supported by the National Science Foundation (NSF) under Grant No. AST-1413663. This research has made use of NASA’s Astrophysics Data System. The authors also acknowledge the significant cultural role and reverence the summit of Mauna Kea has within the indigenous Hawaiian community. This research has made use of Astropy,$^2$ a community-developed core Python package for Astronomy Astropy Collaboration et al. (2013, 2018), NumPy Harris et al. (2020), matplotlib Hunter (2007), SciPy Virtanen et al. (2020) and IRAF Tody (1986) distributed by the NOAO, which is operated by AURA under a cooperative agreement with the NSF.

---


$^2$[http://www.astropy.org](http://www.astropy.org)
Appendices

6.A  ghalib fits to white transit light curves from GMOS and MOSFIRE

We show in Figures 6.6 to 6.13 the ghalib fits to the white light curves from GMOS and MOSFIRE newly analysed in this work. Note that the white light curves for other planets in our sample, namely XO-2Nb, HAT-P-26b, and WASP-19b are shown in previous works (see Chapter 2, 3, and 4 respectively).

6.B  Measured transit properties from ghalib

In the Table 6.2 we summarize the results from the application of ghalib as described in Section 6.5. The Table 6.2 and its continued parts show the best fit transit parameters and their ±1σ uncertainties, along with the weighted averaged transit parameters for each grating.
Figure 6.6: White target transit light curves for HAT-P-7b and their fits obtained using time and comparison star light curve as GP regressors. The observation number and the GMOS grating used are mentioned in the title of each panel. In each panel, the black circles show the transit light curve of HAT-P-7b and the overplotted red line shows the best fit transit + GP model. The detrended transit light curve and its best fit transit model overplotted are shown in green circles and blue line respectively. In the lower part of each panel are shown the residuals between the observed light curve and the best fit transit + GP model. Also annotated is the RMS of the residuals in ppm.
Figure 6.7: Same as Figure 6.6 but for the planet WASP-12b.
Table 6.2: Best fit transit parameters and their $\pm 1\sigma$ uncertainties obtained from the fits to white transit light curves of eight GMOS and MOSFIRE transit and secondary eclipse observations newly analysed in this work using ghaliB. The columns in order show the planet name, grating or nod position used in case of MOSFIRE, planet to star radius or flux ratio ($R_p/R_*$ or $F_p/F_*$), mid-transit time ($T_0[BJD_{TDB}]$), semi-major axis ($a/R_*$), inclination ($i$ [°]), linear limb-darkening coefficient ($u_1$), GP white noise parameter ($\sigma_w$), and the root-mean-square (RMS) of the residuals between the data and the best fit model. For each grating we also show the weighted averaged transit parameters from all observations.

<table>
<thead>
<tr>
<th>Planet</th>
<th>No./Grating</th>
<th>$R_p/R_*$</th>
<th>$T_0[BJD_{TDB}]$</th>
<th>$a/R_*$</th>
<th>$i$ [°]</th>
<th>$u_1$</th>
<th>$\sigma_w$ [ppm]</th>
<th>RMS [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAT-P-7b</td>
<td>1,R150</td>
<td>0.073$^{+0.077}_{-0.089}$</td>
<td>2456133.892662$^{+0.000416}_{-0.000425}$</td>
<td>4.17$^{+0.02}_{-0.02}$</td>
<td>83.14$^{+0.02}_{-0.02}$</td>
<td>0.42$^{+0.08}_{-0.08}$</td>
<td>532$^{+14}_{-13}$</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>0.0802$^{+0.0039}_{-0.0069}$</td>
<td>2456144.91674$^{+0.00007}_{-0.00082}$</td>
<td>4.11$^{+0.03}_{-0.03}$</td>
<td>83.14$^{+0.02}_{-0.02}$</td>
<td>0.52$^{+0.09}_{-0.09}$</td>
<td>645$^{+16}_{-16}$</td>
<td>619</td>
</tr>
<tr>
<td></td>
<td>3,R150</td>
<td>0.0796$^{+0.0043}_{-0.0044}$</td>
<td>2456510.90182$^{+0.000398}_{-0.000401}$</td>
<td>4.16$^{+0.02}_{-0.02}$</td>
<td>83.14$^{+0.02}_{-0.02}$</td>
<td>0.51$^{+0.08}_{-0.08}$</td>
<td>573$^{+14}_{-13}$</td>
<td>559</td>
</tr>
<tr>
<td>Comb R150</td>
<td>0.0766$^{+0.0027}_{-0.0027}$</td>
<td>4.155$^{+0.013}_{-0.013}$</td>
<td>83.143$^{+0.011}_{-0.011}$</td>
<td>0.478$^{+0.047}_{-0.047}$</td>
<td>576$^{+8}_{-8}$</td>
<td>565</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASP-12b</td>
<td>1,R600</td>
<td>0.1171$^{+0.0001}_{-0.0001}$</td>
<td>2455951.835755$^{+9.3c_{-0.05}}$</td>
<td>3.02$^{+0.02}_{-0.01}$</td>
<td>82.67$^{+0.34}_{-0.19}$</td>
<td>0.43$^{+0.0}_{-0.0}$</td>
<td>165$^{+12}_{-11}$</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>2,R600</td>
<td>0.1229$^{+0.0012}_{-0.0013}$</td>
<td>2455952.927616$^{+0.000106}_{-0.000109}$</td>
<td>3.0$^{+0.01}_{-0.01}$</td>
<td>82.32$^{+0.23}_{-0.19}$</td>
<td>0.43$^{+0.0}_{-0.0}$</td>
<td>144$^{+11}_{-10}$</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>3,B600</td>
<td>0.12$^{+0.0031}_{-0.0033}$</td>
<td>2457397.966768$^{+0.0000271}_{-0.0000279}$</td>
<td>3.07$^{+0.02}_{-0.02}$</td>
<td>83.35$^{+0.44}_{-0.42}$</td>
<td>0.62$^{+0.0}_{-0.0}$</td>
<td>158$^{+16}_{-15}$</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>4,B600</td>
<td>0.1222$^{+0.0016}_{-0.0032}$</td>
<td>2457420.887115$^{+0.0001043}_{-0.0000878}$</td>
<td>3.05$^{+0.02}_{-0.02}$</td>
<td>83.32$^{+0.49}_{-0.45}$</td>
<td>0.62$^{+0.0}_{-0.0}$</td>
<td>192$^{+28}_{-28}$</td>
<td>155</td>
</tr>
<tr>
<td></td>
<td>5,B600</td>
<td>0.124$^{+0.0043}_{-0.0044}$</td>
<td>2457432.89302$^{+0.0000391}_{-0.0000384}$</td>
<td>3.05$^{+0.02}_{-0.02}$</td>
<td>83.39$^{+0.48}_{-0.47}$</td>
<td>0.62$^{+0.0}_{-0.0}$</td>
<td>139$^{+16}_{-14}$</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>6,B600</td>
<td>0.1203$^{+0.0035}_{-0.0046}$</td>
<td>2457745.03817$^{+0.0000322}_{-0.0000384}$</td>
<td>3.07$^{+0.02}_{-0.02}$</td>
<td>83.64$^{+0.46}_{-0.45}$</td>
<td>0.62$^{+0.0}_{-0.0}$</td>
<td>209$^{+22}_{-17}$</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>7,B600</td>
<td>0.1214$^{+0.0033}_{-0.0033}$</td>
<td>2457755.952769$^{+0.0000206}_{-0.0000209}$</td>
<td>3.07$^{+0.02}_{-0.02}$</td>
<td>83.7$^{+0.43}_{-0.42}$</td>
<td>0.62$^{+0.0}_{-0.0}$</td>
<td>125$^{+13}_{-13}$</td>
<td>84</td>
</tr>
<tr>
<td>Comb R600</td>
<td>0.1193$^{+0.0008}_{-0.0008}$</td>
<td>3.008$^{+0.001}_{-0.001}$</td>
<td>82.433$^{+0.181}_{-0.181}$</td>
<td>0.433$^{+0.001}_{-0.001}$</td>
<td>154$^{+7}_{-7}$</td>
<td>136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comb B600</td>
<td>0.1215$^{+0.0014}_{-0.0014}$</td>
<td>3.062$^{+0.001}_{-0.001}$</td>
<td>83.484$^{+0.201}_{-0.201}$</td>
<td>0.623$^{+0.001}_{-0.001}$</td>
<td>152$^{+7}_{-7}$</td>
<td>124</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$F_p/F_*$ [ppm]</th>
<th>$\sigma_w$ [ppm]</th>
<th>RMS [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASP-12b A</td>
<td>3117$^{+1492}_{-1492}$</td>
<td>2832$^{+1510}_{-1510}$</td>
</tr>
<tr>
<td>B</td>
<td>3170$^{+1075}_{-1075}$</td>
<td>1980$^{+1000}_{-1000}$</td>
</tr>
</tbody>
</table>
### Table 6.2: Continued Table 6.2.

<table>
<thead>
<tr>
<th>Planet</th>
<th>No., Grating</th>
<th>$R_p/R_*$</th>
<th>$T_0[BJD_{TDB}]$</th>
<th>$a/R_*$</th>
<th>$i [^\circ]$</th>
<th>$u_1$</th>
<th>$\sigma_w$ [ppm]</th>
<th>RMS [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASP-18b</td>
<td>1,R150</td>
<td>0.0985$^{+0.0023}_{-0.0022}$</td>
<td>2456229.600331$^{+0.000161}_{-0.000161}$</td>
<td>3.57$^{+0.02}_{-0.02}$</td>
<td>84.81$^{+0.25}_{-0.25}$</td>
<td>0.52$^{+0.06}_{-0.05}$</td>
<td>456$^{+23}_{-21}$</td>
<td>436</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASP-4b</td>
<td>1,B600</td>
<td>0.1524$^{+0.0006}_{-0.0008}$</td>
<td>2455844.663142$^{+0.000115}_{-0.000128}$</td>
<td>5.46$^{+0.02}_{-0.02}$</td>
<td>88.73$^{+0.48}_{-0.38}$</td>
<td>0.7$^{+0.01}_{-0.01}$</td>
<td>277$^{+53}_{-39}$</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>2,R150</td>
<td>0.1545$^{+0.0007}_{-0.0006}$</td>
<td>2456216.691768$^{+8.3e-05}_{-8.4e-05}$</td>
<td>5.45$^{+0.02}_{-0.02}$</td>
<td>88.83$^{+0.35}_{-0.35}$</td>
<td>0.59$^{+0.02}_{-0.02}$</td>
<td>297$^{+22}_{-20}$</td>
<td>271</td>
</tr>
<tr>
<td></td>
<td>3,R150</td>
<td>0.1572$^{+0.0014}_{-0.0014}$</td>
<td>2456576.675814$^{5.4e-05}_{-5.3e-05}$</td>
<td>5.47$^{+0.02}_{-0.02}$</td>
<td>88.14$^{+0.26}_{-0.23}$</td>
<td>0.53$^{+0.03}_{-0.03}$</td>
<td>221$^{+12}_{-11}$</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>4,R150</td>
<td>0.1592$^{+0.0041}_{-0.0042}$</td>
<td>2456924.615855$^{+0.000151}_{-0.000154}$</td>
<td>5.47$^{+0.02}_{-0.02}$</td>
<td>88.37$^{+0.48}_{-0.37}$</td>
<td>0.49$^{+0.06}_{-0.06}$</td>
<td>362$^{+31}_{-28}$</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comb R150</td>
<td></td>
<td>0.155$^{+0.0006}_{-0.0006}$</td>
<td></td>
<td>5.461$^{+0.01}_{-0.01}$</td>
<td>88.328$^{+0.189}_{-0.189}$</td>
<td>0.568$^{+0.014}_{-0.014}$</td>
<td>253$^{+9}_{-9}$</td>
<td>257</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAT-P-1b</td>
<td>1,R150</td>
<td>0.1134$^{+0.004}_{-0.0042}$</td>
<td></td>
<td>9.87$^{+0.05}_{-0.05}$</td>
<td>85.64$^{+0.04}_{-0.04}$</td>
<td>0.51$^{+0.07}_{-0.07}$</td>
<td>304$^{+8}_{-8}$</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td>2,B600</td>
<td>0.1193$^{+0.0029}_{-0.0029}$</td>
<td></td>
<td>9.83$^{+0.05}_{-0.05}$</td>
<td>85.65$^{+0.04}_{-0.04}$</td>
<td>0.7$^{+0.05}_{-0.05}$</td>
<td>248$^{+11}_{-10}$</td>
<td>219</td>
</tr>
<tr>
<td>Planet No.,Grating</td>
<td>R_P/R_*</td>
<td>T(b)[BJD_{TDB}]</td>
<td>a/R_*</td>
<td>i [°]</td>
<td>u_1</td>
<td>σ_w [ppm]</td>
<td>RMS [ppm]</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>-----------------</td>
<td>-------</td>
<td>-------</td>
<td>-----</td>
<td>------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>WASP-80b 1,R150</td>
<td>0.1727±0.0011</td>
<td>2456508.900412±6.7e-05</td>
<td>12.68±0.05</td>
<td>88.97±0.08</td>
<td>0.57±0.02</td>
<td>684±23</td>
<td>668</td>
<td></td>
</tr>
<tr>
<td>2,R150</td>
<td>0.1731±0.0035</td>
<td>2456831.024564±6.6e-05</td>
<td>12.66±0.05</td>
<td>89.0±0.09</td>
<td>0.58±0.05</td>
<td>543±19</td>
<td>514</td>
<td></td>
</tr>
<tr>
<td>3,B600</td>
<td>0.1687±0.0035</td>
<td>2457275.863228±0.000128</td>
<td>12.63±0.07</td>
<td>89.03±0.1</td>
<td>0.76±0.05</td>
<td>633±59</td>
<td>555</td>
<td></td>
</tr>
<tr>
<td>4,B600</td>
<td>0.1732±0.0021</td>
<td>2457597.987441±0.000133</td>
<td>12.63±0.07</td>
<td>89.0±0.1</td>
<td>0.72±0.05</td>
<td>834±144</td>
<td>645</td>
<td></td>
</tr>
<tr>
<td>5,B600</td>
<td>0.1658±0.0043</td>
<td>2457637.869969±0.000133</td>
<td>12.65±0.07</td>
<td>89.01±0.1</td>
<td>0.74±0.05</td>
<td>436±34</td>
<td>366</td>
<td></td>
</tr>
<tr>
<td>Comb R150</td>
<td>0.1727±0.0011</td>
<td>2456978.027±0.00362</td>
<td>12.67±0.036</td>
<td>88.985±0.06</td>
<td>0.572±0.019</td>
<td>602±14</td>
<td>591</td>
<td></td>
</tr>
<tr>
<td>Comb B600</td>
<td>0.1715±0.0018</td>
<td>2457024.739092±0.00362</td>
<td>12.638±0.04</td>
<td>89.012±0.059</td>
<td>0.74±0.03</td>
<td>511±29</td>
<td>522</td>
<td></td>
</tr>
<tr>
<td>GJ-3470b 1,R150</td>
<td>0.0801±0.0035</td>
<td>2456978.027±0.00362</td>
<td>14.07±0.34</td>
<td>88.81±0.46</td>
<td>0.58±0.09</td>
<td>1120±38</td>
<td>1076</td>
<td></td>
</tr>
<tr>
<td>2,R150</td>
<td>0.0719±0.005</td>
<td>2457451.831439±0.00362</td>
<td>13.93±0.39</td>
<td>88.65±0.37</td>
<td>0.8±0.05</td>
<td>376±23</td>
<td>362</td>
<td></td>
</tr>
<tr>
<td>4,B600</td>
<td>0.0716±0.0035</td>
<td>2457451.831439±0.00362</td>
<td>13.98±0.32</td>
<td>88.9±0.5</td>
<td>0.8±0.09</td>
<td>228±14</td>
<td>187</td>
<td></td>
</tr>
<tr>
<td>5,B600</td>
<td>0.0781±0.0107</td>
<td>2457481.860803±0.00362</td>
<td>14.095±0.255</td>
<td>88.78±0.282</td>
<td>0.581±0.061</td>
<td>1040±27</td>
<td>990</td>
<td></td>
</tr>
<tr>
<td>Comb R150</td>
<td>0.0757±0.0035</td>
<td>2457704.739092±0.00362</td>
<td>13.961±0.258</td>
<td>88.76±0.305</td>
<td>0.803±0.064</td>
<td>271±13</td>
<td>275</td>
<td></td>
</tr>
</tbody>
</table>
Figure 6.8: Same as Figure 6.6 but for the planet WASP-18b.
Figure 6.9: Same as Figure 6.6 but for the planet WASP-4b.
6.B Measured transit properties from ghalib

Figure 6.10: Same as Figure 6.6 but for the planet HAT-P-1b.
Figure 6.11: Same as Figure 6.6 but for the planet WASP-80b.
6.B Measured transit properties from 

Figure 6.12: Same as Figure 6.6 but for the planet GJ-3470b.

Figure 6.13: Same as Figure 6.6 but for the the Keck/MOSFIRE eclipse observation of WASP-12b. The WASP-12b eclipse was observed using ABBA nod pattern and the left and right panels here show the white eclipse light curve from the A and B nod positions respectively.
Bibliography

Adibekyan, V. 2019, Geosciences, 9, 105
Allen, C. W. 1963, Astrophysical quantities
Barclay, T., Pepper, J., & Quintana, E. V. 2018, A revised exoplanet yield from the Transiting Exoplanet Survey Satellite (TESS)
244    Bibliography

Casasayas-Barris, N., Pallé, E., Yan, F., et al. 2020, Astronomy and Astrophysics, 635
Chen, G., Palle, E., Parviainen, H., Murgas, F., & Yan, F. 2021
Crossfield, I. J., Barman, T., Hansen, B. M., & Howard, A. W. 2013, Astronomy and Astrophysics, 559, A33
Foreman-Mackey, D., Agol, E., Angus, R., & Ambikasaran, S. 2017, ArXiv
Bibliography


Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90


Bibliography


Kreidberg, L. 2015, PASP, 127, 1161


Lodders, K., Palme, H., & Gail, H. P. 2009, Landolt Börnstein, 4B, 712


May, E. M., Zhao, M., Haidar, M., Rauscher, E., & Monnier, J. D. 2018, MOPSS I: Flat optical spectra for the hot jupiters WASP-4 b and WASP-52b
Mayor, M., Queloz, D., Marcy, G., et al. 1995, IAU Circ., 6251, 1
Moehler, S., Freudling, W., Möller, P., et al. 2010, PASP, 122, 93
Patel, J. A. & Espinoza, N. 2022, AJ, 163, 228


Rackham, B. V., Apai, D., & Giampapa, M. S. 2017b, arXiv, 853, 122


Rasmussen, C. E. & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning


Bibliography

Shimansky, V. V., Bikmaev, I. F., Galeev, A. I., et al. 2003, Astronomy Reports, 47, 750
Takeda, Y. & Takada-Hidai, M. 2012, PASJ, 64, 42
Yan, F. & Henning, T. 2018, Nature Astronomy, 2, 714


Here we list bibliographic information of the papers included in this thesis. The relative contribution of every co-author is represented by their place in the author list.

**Chapter 2:** Enhanced metallicity of an exoplanet atmosphere from precise alkali lines measurements: Ground-based low-resolution detection of Na and K in the atmosphere of a cloud-free gasgiant

Vatsal Panwar, Jean-Michel Désert, Michael R. Line, , Jacob L. Bean, Timothy M. Brown
To be submitted to the Astrophysical Journal

**Chapter 3:** A new method to measure the spectra of transiting exoplanet atmospheres using multi-object spectroscopy

Vatsal Panwar, Jean-Michel Désert, Kamen O. Todorov, Jacob L. Bean, Kevin B. Stevenson, C. M. Huitson, Jonathan J. Fortney, Marcel Bergmann

**Chapter 4:** A new method to correct for host star variability in multi-epoch observations of exoplanet transmission spectra

Vatsal Panwar, Jean-Michel Désert, Kamen O. Todorov, Jacob L. Bean, Kevin B. Stevenson, C. M. Huitson, Jonathan J. Fortney, Marcel Bergmann

**Chapter 5:** On the follow-up of small and rocky exoplanets orbiting bright stars using ground-based spectrophotometry

Vatsal Panwar, Jean-Michel Désert et al.
To be submitted to the Monthly Notices of the Royal Astronomical Society
Chapter 6: Common trends in close-in gas giant exoplanet atmospheres from MOS surveys

Using a library of tools for modelling spectrophotometric transit light curves with Gaussian Processes (GHALIB)

Vatsal Panwar, Jean-Michel Désert et al.

To be submitted to the Monthly Notices of the Royal Astronomical Society
Summary

Transiting exoplanets, in all of their diverse flavours, provide a unique opportunity to study planet formation and evolution. Discoveries of planetary systems in physical and chemical regimes vastly different to the solar system, development of new instruments tailored for the accuracy and precisions required for exoplanet measurements, and advancements in observational and data analysis methods have been the prime movers of exoplanet science. Ground-based spectroscopic observations have been at the forefront of exoplanet detection and characterization efforts. This thesis presents results from a multitude of optical time series spectroscopic observations of transiting exoplanets using low-resolution multi-object spectrographs (MOS) on large ground-based telescopes. A majority of these results were obtained from a multi-year observational program using the Gemini Multi Object Spectrograph (GMOS) on both the Gemini North and Gemini South telescopes.

The traditional approach of differential spectrophotometry i.e., using a simultaneously observed nearby comparison star to remove systematics in the target transit light curves, performs best when the comparison star is of similar brightness and spectral type. We begin with such an ideal example in Chapter 2 where we present the GMOS transmission spectrum of the hot Jupiter XO-2Nb in the 420 to 920 nm range. The host star XO-2N has an ideal comparison star in the form of its visual binary companion XO-2S, with identical brightness and spectral type. Using differential spectrophotometry, we obtain average precision of 60 ppm per 10-20 nm bins in the combined transmission spectrum from 7 GMOS transits. The transmission spectrum shows pressure broadened absorption features due to Na and K consistent with a cloud-free atmosphere. We retrieve a super-solar metallicity ([M/H] = 0.8 ± 0.38) and a super solar [Na/K] = 1.26 ± 0.5 for the planetary atmosphere. Super solar metallicity could indicate that XO-2Nb formed in a region of disk rich in refractory material. The latter migration of XO-2Nb could have caused late accretion of refractory rich material onto the host star, potentially explaining the high refractory content of XO-2N as compared to XO-2S. Depletion of K in XO-2Nb’s atmosphere remains difficult to explain through the known sinks of K.

In Chapter 3 we consider a more general case of MOS observations where there is a non-linear relationship between the target and comparison star light curves. We introduce a new Gaussian Process (GP) regression based method that infers a generalized non-linear map between the target transit light curve systematics and a set of GP regressors including the comparison star light curve. We demonstrate the application of this method on GMOS observations of the warm Neptune HAT-P-26b. As compared to conventional methods, we
obtain an improvement on the transit light curve fits in terms of residual scatter and broadband transit depth precisions. We also show that it is possible to retrieve decent precisions on the transmission spectra without directly using the comparison star light curves to perform differential spectrophotometry. This result overcomes the commonly faced hurdle of non-availability of suitable comparison stars for especially bright targets and enables their future observations. In summary, our method is particularly applicable for cases where either a comparison star is not available or shows widely different systematics which can introduce additional systematics through conventional methods. We obtain a flat transmission spectra for HAT-P-26b consistent with previous measurements in the visible range.

In Chapter 4 we explore the impact of stellar activity on multi-epoch transit observations, and demonstrate a new empirical method to disentangle it from the planetary atmospheric signal. The subject of our study is WASP-19b, an ultra-hot Jupiter orbiting an active solar-type star. We use the method introduced in Chapter 3 to obtain the transmission spectra of WASP-19b from 8 GMOS transit observations. Using solely the target spectroscopic light curves through this method prevents additional uncertainties from a significantly faint comparison star. We observe a tentative trend between the slopes and offsets of transmission spectrum from each epoch, suggesting the impact of stellar variability. The range of relative offsets in the GMOS spectra is comparable with that calculated from previously observed occulted spots and faculae, and with that observed in TESS transits. We empirically remove the best fit slope and offsets from each spectrum before constructing the combined transmission spectrum. We rule out solar TiO atmosphere at high significance based on the GMOS transmission spectra of WASP-19b, indicating potential condensation of TiO near the terminator. The sub-solar TiO scenario is consistent with previous ground-based measurements. Due to the lack of coverage in the blue optical, we do not confirm or rule out the presence of scattering due to hazes.

In Chapter 5, we explore the potential of ground-based MOS observations for the follow-up transit observations of small planets (radius \( \leq 2R_{\oplus} \)) around bright stars (\( V_{\text{mag}} \leq 9.5 \)). Precise ground-based transit follow-up of such planets is a challenge. In addition to the small amplitude of transit signal, this is also due to the reliance of ground-based observations on availability of suitable comparison stars and the lack thereof for bright stars. We apply a 2D GP method to measure the broadband transit properties from a single GMOS transit of WASP-18b. This approach leverages the shared information of transit light curve systematics across the spectroscopic channels, and retrieves the transit signal at precisions close to photon noise without using the comparison star. Based on this result, we show that it is possible to use instruments like GMOS on medium to large telescopes to conduct precise follow-up of small and rocky exoplanets around bright stars. This is relevant in the context of current and future transit detection surveys TESS and PLATO which are expected to yield an abundance of small and rocky exoplanets orbiting bright stars. Precise transit depth and mid-transit measurements of such planets from repeated ground-based observations will aid in refining their size and mass (via transit timing variations) measurement. This will make them amenable for bulk composition studies and atmospheric characterization.
Finally, in Chapter 6 we present ghalib – a library of tools to model transit light curve systematics in MOS observations. We show the application of ghalib to MOS observations of a sample of ten gas-giant exoplanets spanning a wide range of bulk densities and equilibrium temperatures. We explore the commonalities in the nature of telluric and instrumental systematics across \( \sim 220 \) hours of MOS observations in our sample. We show that the approach of ghalib is capable of disentangling the systematics from white and spectroscopic transit light curves from observations taken in a wide range of observing conditions and instrument setups. We elaborate on the essential components of ghalib and discuss future improvements to it in the context of more computationally challenging datasets.

In summary, we introduced new data analysis methods to overcome inherent limitations to ground-based transit spectrophotometry technique and improve it for future applications. In particular, we focussed on tackling the complex non-linear telluric, instrumental, and astrophysical systematics affecting ground-based observations. Our results enable observations of a swath of scientifically promising exoplanets orbiting bright stars which have been difficult using conventional techniques. As we transition to detecting and characterizing earth analogues with the new age of instruments like JWST, PLATO, and extremely large telescopes, it will be important to have a deeper understanding of various sources of uncertainties in time series exoplanet observations. The overarching goal of this thesis is to inspire new data analysis strategies for realizing the full potential of our most advanced observing facilities.
Exoplaneten, in al hun verschillende vormen, die voor hun moederster passeren, bieden een unieke kans om de vorming en evolutie van planeten te bestuderen. De ontdekkingen van planetenstelsels in fysische en chemische regimes die sterk verschillen van die in het zonnestelsel, de ontwikkeling van nieuwe instrumenten die zijn toegesneden op de nauwkeurigheid en precisie die voor metingen aan exoplaneten nodig zijn, en de vooruitgang op het gebied van observatie- en gegevensanalysemethoden zijn de belangrijkste motoren geweest voor de wetenschap van exoplaneten. Spectroscopische waarnemingen op aarde hebben een vooraanstaande rol gespeeld bij de opsporing en karakterisering van exoplaneten. Dit proefschrift presenteert de resultaten van een groot aantal optische, spectroscopische waarnemingen van de lichtkrommen van transiterende exoplaneten met behulp van multi-object spectrografen (MOS) met lage resolutie gemonteerd op grote telescoopen op aarde. De meeste van deze resultaten zijn verkregen met een meerjarig observatieprogramma met behulp van de Gemini Multi Object Spectrograph (GMOS) op zowel de Gemini-noord- als de Gemini-zuid-telescoop.

De traditionele benadering van differentiële spectrofotometrie, d.w.z. het gelijktijdig waarnemen van een nabijgelegen vergelijkingster om systematiek in de lichtkrommen van het doelwit te verwijderen, werkt het best wanneer de vergelijkingster een vergelijkbare helderheid en een vergelijkbaar spectraaltype heeft. We beginnen met zo’n ideaal voorbeeld in Hoofdstuk 2, waar we het GMOS-transmissiespectrum van de hete Jupiter XO-2Nb in het 420-920 nm-bereik presenteren. De moederster XO-2N heeft een ideale vergelijkingster in de vorm van zijn visuele dubbelster XO-2S, met identieke helderheid en spectraaltype. Met behulp van differentiële spectrofotometrie verkrijgen we een gemiddelde precisie van 60 ppm per bins van 10-20 nm in het gecombineerde transmissiespectrum van 7 GMOS-transits. Het transmissiespectrum vertoont drukverbrede absorptielijnen van Na en K die passen bij een wolkenloze atmosfeer. We vinden een metaalgehalte hoger dan dat van de zon ([M/H] = 0,8±0,38) en een [Na/K] = 1,26±0,5, hoger dan dat van de zon, voor de planeetatmosfeer. Het hoge metaalgehalte zou erop kunnen wijzen dat XO-2Nb is gevormd in een gebied van de schijf dat rijk is aan vuurvast materiaal. De late migratie van XO-2Nb zou de late aanwas van vuurvast materiaal op de moederster veroorzaakt kunnen hebben, wat het hoge vuurvaste gehalte van XO-2N in vergelijking met XO-2S zou kunnen verklaren. Het gebrek aan K in de atmosfeer van XO-2Nb blijft moeilijk te verklaren door de bekende afvoerkanalen van K.
In Hoofdstuk 3 beschouwden we een algemener geval van MOS-waarnemingen waarbij er een niet-lineaire relatie bestaat tussen de lichtkrommen van het doelwit en de vergelijkster. We introduceerden een nieuwe, op Gaussische Processen (GP) gebaseerde regressiemethode die een gegeneraliseerde niet-lineaire omzetting samenstelt van de systematiek van de lichtkromme van het doelwit naar een reeks GP regressoren waaronder de lichtkromme van de vergelijkster. Wij demonstrem de toepassing van deze methode op GMOS-waarnemingen van de warme Neptunus HAT-P-26b. In vergelijking met conventionele methoden verkrijgen we een verbetering van de fits van de lichtkromme in termen van de verspreiding van residuen en de nauwkeurigheid van de breedbandige transitdiepte. We laten ook zien dat het mogelijk is om een behoorlijke nauwkeurigheid van de transmissiespectra te verkrijgen zonder de lichtkrommen van de vergelijksters te gebruiken voor differentiële spectrofotometrie. Met dit resultaat kan de vaak voorkomende hinder nis van het niet beschikbaar zijn van geschikte vergelijksters voor bijzonder heldere doelen, overkomen worden en dit maakt toekomstige waarnemingen daarvan mogelijk. Kortom, onze methode is bijzonder geschikt voor gevallen waarin een vergelijkster niet beschikbaar is of sterf afwijkende systematiek vertoont waardoor conventionele methoden extra systematiek kunnen introduceren. We verkrijgen een vlak transmissiespectrum voor HAT-P-26b dat consistent is met eerdere metingen in het zichtbare gebied.

In Hoofdstuk 4 onderzoeken we de invloed van stellaire activiteit op transitwaarnemingen gedaan voor meerdere transits, en demonstreren we een nieuwe empirische methode om deze activiteit te scheiden van het atmosferische signaal van de planeet. Het onderwerp van onze studie is WASP-19b, een ultrahete Jupiter die rond een actieve, zonachtige ster draait. We gebruikten de methode uit Hoofdstuk 3 om de transmissiespectra van WASP-19b te verkrijgen uit 8 GMOS-transitwaarnemingen. Door deze methodede spectroscopische lichtkrommen van enkel het doel te gebruiken, worden extra onzekerheden door een aanzienlijk zwakke vergelijkster voorkomen. We vinden een voorzichtige trend waar tussen de hellingen en offsets van het transmissiespectrum van elke transit, wat suggereert dat de stellaire variabiliteit hier invloed op heeft. Het bereik van de relatieve offsets in de GMOS-spectra is vergelijkbaar met dat van eerdere waargenomen bedekte vlekken en faculae, en met dat van TESS-transits. We verwijderden empirisch de best passende helling en offsets uit elk spectrum voordat we het gecombineerde transmissiespectrum samenstellen. Op basis van de GMOS-transmissiespectra van WASP-19b sluiten we een solaire TiO-atmosfeer uit met grote significantie, wat wijst op mogelijke condensatie van TiO nabij de dag-nachtgrens. Het sub-solaire TiO-scenario komt overeen met eerdere metingen vanaf de aarde. Wegens een gebrek aan dekking in het blauw-optische spectrum kunnen we de aanwezigheid van verstrooing door fijnstof niet bevestigen of uitsluiten.

In Hoofdstuk 5 onderzochten we het potentieel van MOS-waarnemingen vanaf de aarde voor de follow-up van transitwaarnemingen van kleine planeten (straal ≤ 2R⊕) rond heldere sterren (Vmag ≤ 9, 5). Het nauwkeurig volgen van zulke planeten op de grond is een uitdaging. Naast de kleine amplitude van het transit-signal is dit ook te wijten aan het feit dat waarnemingen vanaf de aarde afhankelijk zijn van de beschikbaarheid van geschikte vergelijksters en dat het hieraan ontbreekt voor heldere sterren. Wij passen een 2D GP meth
Samenvatting

Samenvatting 267

Ode toe om de breedbandige transiteigenschappen van een enkele GMOS transit van WASP-18b te meten. Deze aanpak maakt gebruik van informatie van de lichtkromme-systematiek die gedeeld wordt door de spectroscopische kanalen, en levert het transit-signaal op met een nauwkeurigheid die dicht bij de fotonruis ligt, zonder gebruik te maken van de vergelijkingster. Op basis van dit resultaat laten wij zien dat het mogelijk is om instrumenten als GMOS op middelgrote tot grote telescopen te gebruiken voor een nauwkeurige follow-up van kleine en rotsachtige exoplaneten rond heldere sterren. Dit is relevant in het kader van de huidige en toekomstige transitdetectieonderzoeken TESS en PLATO, die naar verwachting een groot aantal kleine en rotsachtige exoplaneten bij heldere sterren zullen opleveren. Nauwkeurige metingen van de transittiepte en het transittijdstip van dergelijke planeten helpen om hun grootte en massa (via transit tijdstip variaties) beter te meten. Dit zal hen geschikt maken voor onderzoek naar de bulksamenstelling en atmosferische karakterisering.

Tenslotte presenteren we in Hoofdstuk 6 ghalib - een bibliotheek van hulpmiddelen om lichtkromme-systematiek in MOS-waarnemingen te modelleren. We tonen de toepassing van ghalib op MOS-waarnemingen van een steekproef van tien gasreuzen die een breed spectrum van dichtheden en evenwichtstemperaturen bestrijken. We onderzoeken de overeenkomensten in de aard van tellurische en instrumentele systematiek over 220 uur aan MOS-waarnemingen in onze steekproef. We laten zien dat de aanpak van ghalib in staat is de systematiek van witte en spectroscopische transitlichtkrommen van waarnemingen in een breed scala aan waarnemingsomstandigheden en instrumentele opstellingen te scheiden. We lichten de essentiële componenten van ghalib toe en bespreken toekomstige verbeteringen in de context van meer computationeel uitdagende datasets.

Kortom, we hebben nieuwe methoden voor gegevensanalyse geïntroduceerd om de inherente beperkingen van de spectrofotometrietechniek voor transit op de grond te overwinnen en deze te verbeteren voor toekomstige toepassingen. We hebben ons met name gericht op het aanpakken van de complexe niet-lineaire tellurische, instrumentele en astrofysische systematiek die van invloed is op waarnemingen vanaf de grond. Onze resultaten maken waarnemingen mogelijk van een aantal wetenschappelijk veelbelovende exoplaneten die rond heldere sterren draaien, die met conventionele technieken moeilijk te vinden waren. Naarmate we in het nieuwe tijdperk van instrumenten als JWST, PLATO en extreem grote telescopen overgaan op het detecteren en karakteriseren van aardse planeten, zal het belangrijk zijn om meer inzicht te krijgen in de verschillende bronnen van onzekerheden in tijdreekswaarnemingen van exoplaneten. Het overkoepelende doel van dit proefschrift is het inspireren van nieuwe data-analyse strategieën om het volledige potentieel van onze meest geavanceerde observatie faciliteiten te realiseren.
One tree is fed by waters of many rains and thunderstorms. The analogy for this PhD thesis goes way beyond the paper you are reading this on.

For those who know me well might remember the anecdote I often recount about the anxiously long Amsterdam to Delhi flight after my PhD interview, right when I was expecting that a PhD offer may or may not land in my inbox. Spoiler alert, it did. Sometime during those 8 hours at 36000 feet, I flew through the portal to the next five years of PhD life. Thank you Jean-Michel for this giving me the opportunity, and guiding me through the ups and downs of this eventful journey. You supported me and inculcated in me scientific integrity and the strength and endurance needed to venture into uncertain and risky research questions, and encouraged me in the importance of recognizing and celebrating team work. I am grateful to you for inspiring empathy for the well-being of self and peers even in the face of metaphorical storms one inevitably encounters in academia.

At its core, my excitement for science stems way back from my first teachers. I am indebted to my first teachers — my Mummy who gets the unique credit for making sure I don't mix up my 'd' and 'b' alphabet, and my Papa for the countless evenings patiently teaching me Mathematics and making sure I understand why numbers can be negative. I am grateful to my high school Science and Maths teachers Matin sir, Nighat teacher, and Tagde sir for instilling in me the foundational aptitude for scientific thinking.

To Physics lecturers at BITS Pilani - Rishikesh, Tapomoy, and RR Mishra sir — thank you for reinforcing my interest in the subject and setting an example of teaching. Kaushar ma'am, any amount of gratitude I thank you for introducing me to Astronomy would be understated. From the first Astronomy project with you, followed by the Summer research project in NCU Taiwan you arranged for me, and our many chats about science and life, you have been a constant mentor to me. I am grateful to Wing at NCU for advising me on my first foray into planetary science. To Isabelle at LAM, I am thankful for responding to a random cold email from an undergrad in India who just wanted to work on 'something about exoplanets'. You organized the first ever telescope observing trip for me at Observatoire de Haute-Provence, the birthplace of exoplanet science. It is no wonder after that I had no doubts about choosing to work on exoplanets for my PhD.

The streak of inspiring mentors continued in the form of the colleagues I had during my PhD. Thank you, Chris, for setting up the foundations of the Gemini survey, the observations from which form a large part of this thesis. Kamen, Jacob, Claire, and Lorenzo, thank you for be-
ing patient ears to my many silly science questions and pondering through the course of my PhD. I am sorry Kamen for making you part with some of your precious dark chocolate on more than one occasion. Claire, I couldn’t have asked for a more supporting elder academic sister and friend since I first met you during my PhD interview. Thank you for continuing to being an inspiring force. Jacob, your calmness in some of the most crunch times is a bar I strive to achieve. I am sure a good fraction of that calmness permeated into me during our shared time in the office. Lorenzo, I have looked up to you for your hard work and attention to scientific detail. Thank you for being a constantly inspiring collaborator and someone I would never shy away from writing a very last minute observing proposal with. Thank you Bob, Saugata, Hinna, Niloo, Gabbie, Ben, and Eleanor for being amazing colleagues and fellow exoplanet enthusiasts at API. Thank you Jayne, Ralph, and Carsten for your mentorship and support throughout my PhD. Nathalie, thank you for all the conversations checking up with me during my last year job-hunting and finishing up my thesis. Milena, Susan, and Renee, thank you for the countless admin help and troubleshooting. Thank you, Esther, for organizing the numerous outreach events and guiding me through the joys of doing public outreach.

Life in Amsterdam was punctuated by the most awesome bunch of friends I made during my PhD. I am grateful to all of you friends and bouldering companions at API and IBED for being part of my PhD life. Kaustubh, Gullo, Amruta, Smriti, Alicia, Macla, and Alice for being the inspiring PhD elders and friends in my early days in Amsterdam. Dante, Oli, Mark, Stefanie, Yves for keeping the spirits in the office high. Eva, Sarah, Kelly, and Dimitris, for being the awesome cohort of PhDs I had the fortune to start with. Vlad, for being the ever present social glue of API. Martin and Difeng, for the many card game evenings at Oerknal. Ben, for your travel stories, organizing long bike rides, and for never letting anyone forget my public flossing fails. Frank, for demonstrating that one just needs an orange coloured jacket to survive the Dutch winters. Kenzie, for being an awesome office mate, backup human to Sisi, and an ear to my most random gossips and factoids. David M., for trying your best to make me believe that you are not British. Ines, for your contagious excitement about dance and Karaoke. Maja and Dado, for the dinner get-togethers at your place with the tastiest food and the dangerously delicious Rakia. Morgan and Avery, for the bird walks at Diemerbos and all the games nights and barbecues.

Johannes, for those lockdown mornings at home drinking Moka and evenings cooking some insane dishes together (200 dolmades in one day was it?), teaching me how to make a good pasta, and for the furniture. Vatsal, for the good times at BITS conveniently dodging any name confusions, and for all the memes. You should really check if being my namesake might also allow you to claim authorship of this thesis. And all the other Young Uncles, for a decade of friendship. For tolerating my irrational rants about topics galore, you all, especially Kaul, Jha, and Sharma can claim co-authorship of this thesis.

Before I forget, there’s a list that deserves mention, an assortment of things that kept me going through the doldrums of the pandemic. Netherlands, for being the first place outside India I called home, for the people friendly urban design and the dense network of fiet-sknooppunts. Mutti passata, for your perfect tomatoey-ness and your versatility of use in
both pasta and curries. The green grocer in IJburg, for your supply of scrumptious veggies. Il Sogno for the pizza. Mezrab, for the memorable storytelling evenings and soup. Hummus bistro da for the falafel. Newslaundry, for leading the charge in subscriber based journalism, Ravish Kumar *sirf bolne ke liye*, and Amit Varma for the unmissable weekly podcasts I had the company of during my walks.

Deniz and Alex - the *beratnas* - we should really finish watching Expanse some day. Thank you both checking up on me and making sure we kept each other sane during the weird times when we had to count the number of people we could legally hang out with. Oh, what would life have been without those endless Windmill Columbuses, Oerknal Duvels and La Chouffes, and Flevopark BBQs with you two! Deniz, you were the first person I talked to in Amsterdam, and I cherish the times we spent as flatmates. I reserve the credit for charming you over to life in IJburg. Looking forward to many future Dutch holidays with you and Sinem. Thank you, in particular, for the unending chats about GPs, ambitious optimization algorithms, and funky YouTube things. Alex, as opposed to what many people think, I was definitely very sober back in 2018 when I asked you to be my paranymph. And I swear my mild Anglophilia had nothing to do with it. Thank you for all the water-cooler gossips and morning coffees, and being the utterly gracious chum and fellow office mate. I failed to attract you to IJburg, but I live in hope that I can tempt you to come retire with me in the UK and partake in the samosas I will inevitably start selling from that food truck. Alicia, thank you for your supply of delicious baked goodies, especially those that you brought to our bouldering sessions at monk! And thank you both Alex and Alicia for your adorably naughty feline babies. In the near future, I hope our cats can become half as good friends as us.

My family showered me with their unfailing care and support throughout the course of this journey and any words of gratitude for them would fall short. Thanks to my grandparents for being constant inspiration for my path in life from my earliest mischiefs as an infant to all the milestones that you never fail to be proud of. Nikku, Limca, Shalu, Golu, Ranu, Charu, Manu, for being a great bunch of cousins all those Bhilwara summers. Situ and Mitu didi, and Singh uncle for making sure my favorite songs as a toddler was on when I wanted it. Papa, you had the extraordinary intuition in figuring pretty early on that science is where I should be headed, and your hard work in nurturing me to that goal from the very start. Mummy, you were the pillar of strength to all of my fears and aspirations. I love you both so much for going beyond your means in building me. Kannu, thank you for keeping me rooted to my embarrassing pop-culture fandoms, and for maintaining the doctoral balance in the family. Elspeth, my love, my in-house mushroom expert, and my fellow lentil and poha connoisseur, thank you for all your unwavering love. At the moments when I needed it most, you reminded me what matters the most is to enjoy simple joys of life, a croissant on a lazy Saturday morning, the long bike rides and hike followed by a tasty snack, all the chocolate and crisps, and the fun of foraging delicious natural treats. And finally Sisi, thanks for adopting me and Elspeth.

So long, and thanks for all the lentils.