Detecting the proposed CH$_4$-CO$_2$ biosignature pair with the James Webb Space Telescope: TRAPPIST-1e and the effect of cloud/haze

Thomas Mikal-Evans$^{1,2}$,*

$^1$Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
$^2$MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

* tmevans@mpia.de

Accepted 2021 November 6. Received 2021 October 19; in original form 2021 July 10.

ABSTRACT

It is widely anticipated that the James Webb Space Telescope (JWST) will be transformative for exoplanet studies. It has even been suggested that JWST could provide the first opportunity to search for biosignatures in an alien atmosphere using transmission spectroscopy. This claim is investigated, specifically for the proposed anoxic biosignature pair CH$_4$-CO$_2$. The most favourable known target is adopted (TRAPPIST-1e), with an assumed atmospheric composition similar to the Archean Earth. Compared to previous studies, a more systematic investigation of the effect that cloud/haze-layers have on the detectability of CH$_4$ and CO$_2$ is performed. In addition to a clear atmosphere scenario, cloud/haze-layers are considered at eight pressure levels between 600 mbar and 1 mbar. These pressures cover a plausible range for H$_2$O cloud and photochemical haze, based on observations of solar system atmospheres and physical models of tidally-locked planets such as TRAPPIST-1e, although no assumptions regarding the cloud/haze-layer composition are made in this study. For the clear atmosphere and cloud/haze-layer pressures of 600-100 mbar, strong ($\sim 5\sigma$) detections of both CH$_4$ and CO$_2$ are found to be possible with approximately 5-10 co-added transits measured using the Near Infrared Spectrograph (NIRSpec) prism, assuming a dry stratosphere. However, approximately 30 co-added transits would be required to achieve the same result if a cloud/haze-layer is present at 10 mbar. A cloud/haze-layer at 1 mbar would prevent the detection of either molecule with the NIRSpec prism for observing programs up to 50 transits (>200 hours of JWST time), the maximum considered.

Key words: planets and satellites: general – planets and satellites: atmospheres – planets and satellites: terrestrial planets

1 INTRODUCTION

A number of gases or gas combinations have been proposed as atmospheric biosignatures (for a review see Schwieterman et al. 2018). For a biosphere similar to the modern Earth, this includes the simultaneous presence of O$_2$/O$_3$ and CH$_4$. Future space telescopes, such as LUVOIR (Roberge & Moustakas 2018) and the Origins Space Telescope (Battersby et al. 2018), will be designed with the explicit aim of detecting this biosignature pair in a sample of temperate rocky exoplanets. However, CH$_4$ may prove too challenging for a LUVOIR-like telescope to detect for an atmosphere similar to that of the modern Earth, due to the weakness of predicted features across the 0.2-2.0 $\mu$m wavelength range (Wang et al. 2018; Kawashima & Rugheimer 2019). In the meantime, the James Webb Space Telescope (JWST) could potentially detect CH$_4$ in the atmosphere of particularly favourable rocky exoplanets, but it most likely will not be able to detect O$_2$/O$_3$ (e.g. Wunderlich et al. 2019). Regarding the latter, Fauchez et al. (2020a) have recently highlighted a previously overlooked O$_2$–X collision-induced-absorption band at 6.4 $\mu$m, which appears to be the most promising O$_2$ feature for detection with JWST. In practice, the same authors find that even for the especially favourable target TRAPPIST-1e, a $5\sigma$ detection of this spectral band would require co-adding > 700 transits. To put such a program in context, each transit observation for TRAPPIST-1e would cost approximately 4.3 hours of JWST time. More significantly, there will be fewer than 200 observable transits over the maximum JWST lifetime of 10 years, given the approximately 6 day orbital period of TRAPPIST-1e and the fact that the TRAPPIST-1 system is only accessible to JWST for approximately 100 days per year (Gillon et al. 2020). Likewise, the strong O$_3$ band at $\sim 10$ $\mu$m appears impractical to detect with JWST, even assuming a highly favourable target with an atmosphere similar to that of the modern Earth. For example, Lin et al. (2021) have estimated that for TRAPPIST-1e, 100 transit observations (430 hours) with JWST would be insufficient to detect the 10 $\mu$m O$_3$ band at $3\sigma$ significance.

Biospheres differing from that of the modern Earth may pro-
duce distinctive gas combinations that are more readily detectable. Considering an anoxic atmosphere similar to that prevailing on the Archean Earth, Krissansen-Totton et al. (2018b) proposed the simultaneous detection of CH$_4$ and CO$_2$ as constituting a biosignature, which would be strengthened if combined with a non-detection of CO. Subsequently, Krissansen-Totton et al. (2018a) (KT2018a) investigated the prospects for detecting the CH$_4$-CO$_2$ pair in the atmosphere of TRAPPIST-1e, assuming the same atmospheric composition as the Archean Earth. KT2018a concluded that this could be achieved by co-adding approximately 10 transit observations made with the JWST Near Infrared Spectrograph (NIRSpec) prism, which covers the ∼1-5 μm wavelength range. Furthermore, KT2018a reported that the result was not significantly affected by the presence cloud that truncated the transmission spectrum for pressures above 10 mbar.

This paper presents an independent assessment of the prospects for detecting CH$_4$-CO$_2$ in the atmosphere of an Archean-like TRAPPIST-1e, inspired by the work of KT2018a. Of the currently known transiting rocky planets that may be habitable, TRAPPIST-1e and f have the highest expected signal-to-noise for transmission spectroscopy, according to the Exoplanet Atmosphere Observatory Table maintained by the Space Telescope Science Institute. General circulation model (GCM) simulations also indicate that TRAPPIST-1e is more likely than TRAPPIST-1f to support liquid water on its surface (e.g. Wolf 2017; Turbet et al. 2018; Fauchez et al. 2020b; Sergeev et al. 2020). For these reasons, it is reasonable to regard TRAPPIST-1e as the most favourable target available for conducting a biosignature search with JWST using transmission spectroscopy.

As elaborated on below, a point of difference setting this study apart from that of KT2018a is the methodology adopted for determining if a gas species is detected. Here, a Bayesian evidence framework is employed to quantify the confidence with which CH$_4$ and CO$_2$ could each be uniquely identified in the atmosphere of an Archean-like TRAPPIST-1e among a variety of plausible gas species that may be spectrally active. A second point of difference is that the present study investigates in greater detail how the presence of cloud and/or photochemical haze at the planetary day-night terminator might complicate detections. In addition, the present study adopts the latest system properties for TRAPPIST-1e from Agol et al. (2021).

2 METHODS

To investigate the detectability of CH$_4$ and CO$_2$ in the atmosphere of TRAPPIST-1e assuming an Archean Earth composition, a hypothetical observing program with the NIRSpec prism is considered. The NIRSpec prism mode is chosen due to the broad wavelength coverage it affords (0.8-5 μm), encompassing strong absorption features of both CH$_4$ and CO$_2$ as illustrated in Figure 1. Although it is conceivable that alternative instrument mode combinations may strictly be more efficient for detecting CH$_4$ and CO$_2$ (e.g. a hybrid NIRSpec program making use of both the prism and G395H grism), investigating such possibilities is beyond the scope of the present work. Furthermore, the broad wavelength coverage of the NIRSpec prism is likely to appeal to a wider range of scientific programs than the specific application considered here. The latter point is reinforced by the fact that of the approved Cycle 1 Guaranteed Time Observations (GTO) and General Observer (GO) programs that will be observing TRAPPIST-1 planets, those covering the 3-5 μm wavelength range will all be using the NIRSpec prism (GTO-1201, GTO-1331, GO-1981, GO-2420, GO-2589).

2.1 Observation cost and synthetic datasets

The publicly available code petitRADTRANS (Mollière et al. 2019) was used to generate model transmission spectra for TRAPPIST-1e. An isothermal upper atmosphere with a temperature of $T_{\text{top}} = 215$ K was assumed for pressures below 100 mbar, close to the mean stratospheric temperature of Earth. At higher pressures, a moist adiabatic temperature profile was adopted following Robinson & Catling (2012), with the surface temperature set to 300 K (Figure 2). Two atmospheric compositions were considered, differing only in the assumed H$_2$O vapor abundance. The first ‘moist’ scenario was identical to the Archean Earth composition used by KT2018a to enable direct comparison. For the spectrally active species, the log$_{10}$ mole fractions were: −2.301 dex for CH$_4$; −1.301 dex for CO$_2$; −8 dex for CO; and −2 dex for H$_2$O. However, the transmission spectrum is primarily sensitive to gases in the stratosphere. For the present-day Earth atmosphere, most H$_2$O is cold-trapped below the tropopause, resulting in typical stratospheric H$_2$O mole fractions of a few ppm (e.g. Chiu et al. 1997). Similarly low stratospheric H$_2$O abundances likely prevailed during the Archean (e.g. Arney et al. 2016). For this reason, a second ‘dry’ scenario was considered, in which all abundances were the same as for the moist scenario, except the H$_2$O abundance was conservatively set to zero. Although the dry stratosphere scenario is more likely to coincide with habitable surface conditions, Kopparapu et al. (2017) have identified cases in which habitable conditions are maintained with a moist stratosphere.

Under both the moist and dry scenarios, the remainder of the atmosphere was assumed to be composed of N$_2$, giving an atmosphere with mean molecular weight of approximately 29 atomic mass units. The surface pressure was set to $P_{\text{surf}} = 1$ bar and gases were assumed to be uniformly mixed throughout the atmosphere. A stellar radius of $R_*$ = 0.1192 $R_{\odot}$ was adopted (Agol et al. 2021). For the planet, a radius of $R_p = 0.920 R_{\oplus}$ and mass of $M_p = 0.692 M_{\oplus}$ were used (Agol et al. 2021), translating to a surface gravity of 8.0 m s$^{-2}$. Transmission spectra were generated for eight different assumptions of an opaque cloud-top pressure outlined below in Section 2.2, as well as an idealised ‘clear’ atmosphere without any cloud.

2.1.1 Allowing partial saturation of the detector

The publicly available pandexo code (Batalha et al. 2017) was used to determine the noise properties associated with a the NIRSpec prism transit observation for TRAPPIST-1e, adopting the partial saturation strategy described in Batalha et al. (2018). The resulting observing parameters for each transit were: SUB512 readout mode (i.e. 512 × 32 pixel subarray); NRSRAPID readout pattern; and 6 groups per integration with 6,515 integrations. This corresponds to a total observing time of 2.86 h, which includes 0.932 h for the transit, 0.932 h for the minimal out-of-transit baseline, and an additional 1 h of out-of-transit baseline to avoid the 1 h penalty imposed for phase-constrained observations. Including observatory overheads, this translates to a total JWST cost of 4.3 h per TRAPPIST-1e transit, according to the Space Telescope Science Institute Astronomer’s Proposal Tool (APT, version 2021.2).

Using the partial saturation strategy, 44 pixels were found to be
Biosignature detectability with JWST

Figure 1. Transmission spectra assuming a cloud-free atmosphere and Archean Earth composition. The moist atmosphere case is shown in panels (a) and (c) and the dry atmosphere case is shown in panels (b) and (d). For both cases, the top and bottom panels show, respectively, the same spectrum with the opacity of CH$_4$ and CO$_2$ switched off. Shading highlights the wavelengths at which each gas contributes uniquely to the absorption.

Figure 2. (a) Moist atmosphere transmission spectra generated for different cloud-top assumptions with major absorption bands labelled. (b) The PT profile adopted in the atmospheric models that were used to generate all transmission spectra, with isothermal stratosphere and adiabatic troposphere.

2.2 Cloud-top altitudes

Evidence for condensible cloud and photochemically-produced haze has been uncovered in many of the exoplanet transmission spectra measured to date for gas giants and sub-Neptunes (e.g. Kreidberg et al. 2014; Sing et al. 2016). An optically thick layer of cloud or haze can substantially alter the detection prospects for absorption bands in the transmission spectrum, as can be appreciated by inspecting Figure 2. Indeed, only a trace amount of cloud or haze is required to be a significant opacity source in a transmission spectrum due to the slant viewing geometry (Fortney 2005). Note that unlike for an Earth-like planet transiting a Sun-like star (Bétrémieux & Kaltenegger 2014), refraction has a negligible effect on the transmission spectra of planets orbiting within the much closer-in habitable zones of M dwarfs (Robinson et al. 2017). As

saturated by the final (sixth) read of each integration. Noise properties for 5 and 39 pixels of the saturated pixels were calculated from the second and third non-destructive detector reads, respectively, by which point they had not saturated. This required applying a correction factor to account for the reduced duty cycle of the pixels saturating before the final read, as described in Batalla et al. (2018). To illustrate the advantage of a partial saturation read-out strategy, Figure 3 compares the predicted noise properties for the resulting transmission spectrum against those that would be obtained with a conservative read-out strategy that does not allow saturation. Example synthetic datasets generated using the pandexo noise properties are shown for the different cloud-top models in Figure 4.
such, cloud/haze should be the primary source of atmospheric obscuration for TRAPPIST-1e.

For a temperate planet such as TRAPPIST-1e, the most plausible condensible species is H$_2$O. On average, approximately 70% of the Earth is covered by H$_2$O cloud at any given time, with typical cloud-tops becoming optically thick at pressures in the ~500-700 mbar range when observed from close to zenith (e.g. see Figures 4 and 7 of King et al. 2013). This translates to cloud-top altitudes of approximately 3-6 km. However, H$_2$O clouds could be mixed to higher altitudes in the atmospheres of close-in rocky planets such as TRAPPIST-1e, which have almost certainly been locked into synchronous orbits by tidal forces (e.g. Barnes 2017). Under this scenario, numerous 3D GCMs have found that typical cloud-top altitudes at the day-night terminator are likely to be around 15 km, corresponding to pressures of approximately 100 mbar (Fauchez et al. 2019; Komacek et al. 2020; Suissa et al. 2020; Pidhorodetska et al. 2020). The only habitable zone exoplanet with a measured transmission spectrum precise enough to reveal spectral features – K2-18b, a sub-Neptune orbiting an M dwarf – appears to be cloudy, with a cloud-top pressure estimated between approximately 10-100 mbar, close to where the temperature profile crosses the H$_2$O condensation curve (Benneke et al. 2019). Although, it should be noted that the latter interpretation remains open to question, as Barclay et al. (2021) have recently argued that stellar inhomogeneities cannot yet be ruled out as an alternative explanation for the existing K2-18b data. For other transmission spectra published to date for planets in habitable zones, the data have only been precise enough to rule out cloud-free, H$_2$/He-dominated atmospheres (de Wit et al. 2018) or produced inconclusive results (Edwards et al. 2021).

In addition to condensable clouds, an optically thick layer of hydrocarbon haze could form in the stratosphere of TRAPPIST-1e at pressures below 100 mbar, driven by ultraviolet photochemistry. Hazes of this kind are thought to have been present in the atmosphere of Archean Earth (e.g. Arney et al. 2016), given the favourably high abundance of CH$_4$, which could be broken up and recombined into longer hydrocarbon chains (Atreya & Romani 1985). In the atmosphere of Titan, formation of an analogous photochemical haze is favoured by a reducing atmosphere with high CH$_4$ and N$_2$ abundances (Hörst 2017). Using a stellar occultation measurement, Robinson et al. (2014) demonstrated that the transmission spectrum for Titan’s atmosphere is insensitive to pressures above approximately 0.1-10 mbar across the 0.9-5 µm wavelength range due to this haze.

For the present study, experiments were performed for eight different cloud-top pressures spanning the range of plausible scenarios outlined above: 600 mbar (4 km), 250 mbar (11 km), 120 mbar (15 km), 100 mbar (16 km), 50 mbar (21 km), 10 mbar (31 km), 5 mbar (36 km), and 1 mbar (48 km). In particular, the 600 mbar case represents a conservative Earth-like H$_2$O cloud scenario. The 100 mbar and 120 mbar cases reflect more plausible H$_2$O cloud scenarios for a slowly-rotating planet such as TRAPPIST-1e. The 250 mbar cloud/haze-layer is included as an intermediate case between the 600 mbar and 100-120 mbar physically-motivated scenarios, to provide additional coverage of the parameter space. The 10 mbar case can be considered an optimistic Titan-like atmosphere, as it neglects the steep rise in opacity towards shorter wavelengths. Studying the atmosphere of TRAPPIST-1e specifically, Fauchez et al. (2019) modeled the production of photochemical haze assuming an Archean Earth composition and found it could become optically thick at effective altitudes spanning approximately 45-20 km across the 1-5 µm wavelength range (see their Figure 9). These altitudes correspond to pressures of approximately 1-100 mbar, somewhat deeper than is observed for Titan. It is therefore also possible that an Archean haze would be intermediate between the 100 mbar and 1 mbar cloud-top cases considered here, although haze altitudes can be sensitive to numerous factors (Kawahshima & Ikoma 2019). In any case, the present study does not presume a specific origin for the 100 mbar, 50 mbar, 10 mbar, 5 mbar, and 1 mbar cloud-tops, but presents them as representative cases covering a range of plausible scenarios.

### 2.3 Bayesian retrieval analyses

Ideally, an atmospheric biosignature detection would be made by convincingly demonstrating that a measured signal could only be plausibly explained by the specific gas/gases that constitute the biosignature. A distinction should be made between this fundamental first step and the subsequent process of constraining the abundances of the biosignature gases. In published studies investigating the detectability of gas species in temperate rocky exoplanets, it has been common practice to effectively skip the first step and assume the atmospheric composition is known a priori. For example, a synthetic dataset is generated assuming some atmospheric composition, then a retrieval analysis or similar is performed with those same gases included and no others (e.g. Barstow & Irwin 2016; Batalha et al. 2018; Krissansen-Totton et al. 2018a; Lustig-Yaeger et al. 2019). In practice, however, it is preferable to remain agnostic about the atmospheric composition and avoid the possibility of a false positive, in which a combination of unrelated gases could conspire to produce a signal similar to that of the expected biosignature.

To achieve this, the Bayesian framework described in Benneke & Seager (2013) is employed. This involves computing the Bayesian evidence for a ‘full’ model including the suite of numerous possible gas species, then systematically computing the Bayesian evidence for models with one gas species removed at a time. By comparing the Bayesian evidences of the latter models to that of the full model,

---

3 U.S. Standard Atmosphere, 1976.
it is possible to compute the associated Bayes factors and thus determine the level of evidence for each individual gas species. Notably, Kaltenegger et al. (2020) applied this approach in assessing the detectability of gas species in a temperate terrestrial planet orbiting a white dwarf. However, unlike the present study, those authors did not explicitly consider the effect of cloud on detectability, which as shown below, can be significant for transmission spectroscopy. Lin et al. (2021) also employed the same Bayesian approach to assess detectability of various molecules in the atmosphere of TRAPPIST-1e assuming prebiotic and modern Earth compositions, but only considered cloud-top altitudes of 6 km and 12 km. Another recent application was presented by Tremblay et al. (2020) as part of a trade-space study for a hypothetical infrared observatory covering various wavelength ranges between 1-30 \( \mu \)m. Those authors took TRAPPIST-1e as the target, but assumed a modern Earth-like atmospheric composition rather than the Archean-like composition considered here and only considered a single cloud-layer scenario with a cloud-top pressure of 560 mbar.

For the full model considered in the present analysis, the spectrally active gas species \( \text{H}_2\text{O}, \text{CH}_4, \text{CO}, \text{CO}_2, \text{NH}_3, \text{H}_2\text{S}, \text{HCN}, \text{and} \text{C}_2\text{H}_2 \) were included as free parameters. The \( \log_{10} \) mass fractions of these gases were allowed to vary uniformly between \([-10, 0]\) dex. For each model evaluation, the \( \text{N}_2 \) abundance was set to ensure the mass fractions summed to unity. The other free parameters were: the

Figure 4. Examples of synthetic datasets assuming a partial saturation observing strategy for different cloud-top scenarios assuming moist stratosphere models (solid black lines). Corresponding dry stratosphere models are also shown (dashed pink lines). The assumed number of transit observations is labelled in the lower left corner of each axis. Plotted datasets had the median \( \chi^2 \) from 1000 random noise realisations, representing a statistically typical outcome, and have been binned in wavelength to \( R = 100 \) for visual clarity. In practice, retrieval analyses were performed on the models binned to the plotted spectral resolution (\( R = 100 \)) without adding random noise perturbations, but still accounting for the measurement uncertainty as described in the text. Yellow shading indicates the measurement uncertainty at the resolution of the plotted data for the specified number of transits. Grey shading indicates the corresponding measurement uncertainties for a single transit observation. Also shown are illustrative models with \( \text{CO} \) as the only active gaseous absorber at inflated mole fractions of -1.1 dex (dot-dashed green lines) and -2.5 dex (solid green lines), highlighting the difficulty of placing strong upper limits on the \( \text{CO} \) abundance using the NIRSpec prism in the case of an Archean-like atmosphere for TRAPPIST-1e.
Table 1. Median and 1σ credible ranges for CO2 and CH4 mole fractions, with corresponding 2σ and 3σ upper bounds for the CO mole fraction, assuming different cloud cases and number N of co-added transits observed with the NIRSpec prism.

| Cloud | N  | CH4 (dex) | CO2 (dex) | CO upper (dex) 2σ | CO upper (dex) 3σ |
|-------|----|-----------|-----------|-------------------|-------------------|
| Clear | 10 | -0.90±0.52 | -1.95±0.47 | < -2.53 | < -1.11 |
|       | 600 mbar | -0.93±0.54 | -1.98±0.50 | < -2.47 | < -1.08 |
|       | 250 mbar | -0.93±0.54 | -1.94±0.47 | < -2.65 | < -1.06 |
|       | 120 mbar | -0.97±0.57 | -1.99±0.53 | < -3.00 | < -0.93 |
|       | 100 mbar | -0.96±0.58 | -1.97±0.54 | < -2.20 | < -0.64 |
|       | 10 mbar | -1.46±0.88 | -2.32±0.67 | < -1.12 | < -0.17 |
|       | 20 | -1.16±0.68 | -2.16±0.57 | < -1.68 | < -0.33 |
|       | 30 | -1.09±0.67 | -2.16±0.47 | < -2.01 | < -0.70 |

| Cloud | N  | CH4 (dex) | CO2 (dex) | CO upper (dex) 2σ | CO upper (dex) 3σ |
|-------|----|-----------|-----------|-------------------|-------------------|
| Clear | 10 | -1.08±0.60 | -2.05±0.50 | < -2.51 | < -1.27 |
|       | 600 mbar | -1.07±0.61 | -2.09±0.51 | < -2.63 | < -1.21 |
|       | 250 mbar | -1.12±0.67 | -2.09±0.52 | < -2.68 | < -1.38 |
|       | 120 mbar | -1.26±0.73 | -2.13±0.60 | < -2.94 | < -1.57 |
|       | 100 mbar | -1.15±0.70 | -2.09±0.55 | < -2.53 | < -0.79 |
|       | 10 mbar | -1.50±0.89 | -2.37±0.78 | < -1.22 | < -0.26 |
|       | 20 | -1.10±0.68 | -2.23±0.54 | < -1.99 | < -0.66 |
|       | 30 | -1.07±0.61 | -2.14±0.43 | < -2.17 | < -1.20 |

planetary mass \((M_p)\) with Gaussian prior \((M_p = 0.692 ± 0.022 M_\oplus)\) (Agol et al. 2021); a reference level planetary radius \((R_p)\) with a uniform prior between \([0.8, 1.1]\) as adopted by KT2018a and encompassing the posterior distribution \(R_p = 0.929^{+0.013}_{-0.014} \) \(R_\oplus\) reported by Agol et al. (2021); the \(\log_{10}\) reference pressure \((\log_{10} P_0)\) corresponding to the reference radius, with a uniform prior between \([-8, 2]\) dex bar; the \(\log_{10}\) pressure of an opaque cloud deck \((\log_{10} P_C)\), also allowed to vary between \([-8, 2]\) dex bar; and the atmospheric temperature \((T_{\text{atm}})\), assumed to be isothermal with a uniform prior between \([100, 700]\) Kelvin. The stellar radius \((R_\star)\) was held fixed to the value quoted in Section 2.1.

The evidence for the full atmosphere model was also compared to that of a simple ‘flat’ model, which assumed the planet lacks an atmosphere. For the flat model, the transmission spectrum is featureless, with the opaque planet radius \((R_p)\) being the only free parameter. Although the full atmosphere model could also account for a featureless transmission spectrum by, for instance, placing a cloud deck at low pressures or decreasing abundances of spectrally active species to negligible levels, the flat model represents a physically plausible scenario with a smaller prior volume (i.e. fewer free parameters). In principle, the latter means that the flat model would be conservatively favoured by the Bayesian evidence, unless the data itself contains relatively strong evidence in favour of the full atmosphere model. Hence, a strong preference for the full model over a simple flat model is an additional prerequisite to be satisfied for claiming a biosignature detection.

To avoid sampling bias, random perturbations were not added to the synthetic datasets prior to performing the retrieval analyses, following Feng et al. (2018). Instead, the models described in Section 2.1 were simply binned to the data resolution and assumed to have the associated measurement uncertainties returned by pandex (Figure 4). Retrieval analyses were then performed using petitRADTRANS to generate model transmission spectra and PyMultiNest (Buchner et al. 2014) to evaluate the Bayesian evidences using nested sampling with 500 live points per retrieval. The ‘n-sigma’ \((n\sigma)\) detection significances for CH4 and CO2 were computed from the relevant Bayes factors, using the method outlined in Benneke & Seager (2013). This process was repeated for different numbers of co-added transit observations between \(N = 5\) and \(N = 50\) in increments of 5, assuming the measurement uncertainties binned down optimally as \(1/\sqrt{N}\). However, observing programs of \(N > 50\) transits were not considered, primarily because this would equate to more than 200 hours of JWST time devoted to a single target using only one observing mode. Such a program would likely pose significant scheduling challenges, given that JWST can only observe the TRAPPIST-1 system for approximately 100 days per year (Gillon et al. 2020). The scheduling challenge could be especially pronounced if the community decides to undertake additional transit observations of TRAPPIST-1e with instrument modes beside the NIRSpec prism (e.g. see Guzmán-Mesa et al. 2020). However, if \(N > 50\) transit observations with the NIRSpec prism are approved, the present analysis could be extended as appropriate in a future study.

3 RESULTS

The detection significances for each cloud case versus the number of co-added transits are shown in Figure 5, assuming a dry stratosphere with the partial-saturation observing strategy described in Section 2.1.1. A direct comparison of the detection significances are shown for the dry and moist stratosphere scenarios in Figure 6. As expected, a moist stratosphere generally increases the required observing time to reach a given detection significance, due to H2O spectral bands overlapping those of CH4 and CO2. However, the increased observing time typically corresponds to just 1-3 additional transits, depending on the molecule, cloud-top pressure, and detection significance (Figure 6).

Under the assumption of a dry stratosphere, strong \((5\sigma)\) detections of both CH4 and CO2 are made for \(N = 5-10\) transits for cloud/haze-layer pressures down to 100 mb (Figure 5). For the specific case of a cloud-free atmosphere with moist stratosphere, \(N \approx 10\) transits are found to be required for strong detections of both gases (Figure 6), in agreement with KT2018a. However, the 10 mbar cloud/haze-layer case, \(N \approx 30\) transits are found to be necessary for ruling out the flat scenario described in Section 2.3 and make \(5\sigma\) detections of both CH4 and CO2 (Figures 5 and 6). This contrasts with KT2018a, who reported that the detectability of CH4 and CO2 was insensitive to the effect of a cloud layer if they were to truncate the transmission spectrum at the 10 mbar pressure level. For the 5 mbar cloud/haze-layer, identifying CH4 becomes substantially more challenging, with \(N \approx 35\) transits required for a \(> 4\sigma\) detection and \(N > 45\) transits required for a \(5\sigma\) detection (Figure 5). A cloud/haze-layer at 1 mbar would likely prevent \(> 3\sigma\)
detections of spectral features even with 200 hr of JWST time, the maximum considered (Figure 5).

Posterior constraints for the retrieved mole fractions of CH₄, CO₂, and CO are reported in Table 1 for the dry and moist stratosphere scenarios. Note that these mole fractions have been converted from the mass fractions that were used to compute the model transmission spectra during sampling (Section 2.3). In general, the CH₄ mole fraction constraints are ~ 0.1-0.2 dex tighter for the dry stratosphere scenario compared to the moist stratosphere scenario, due to the overlap between CH₄ and H₂O spectral bands (Table 1). Since the strongest CO₂ spectral features are less degenerate with H₂O (Figure 2), the CO₂ mole fraction constraints are essentially identical for the dry and moist stratosphere scenarios. Covariance plots for a subset of model parameters are shown in Figure A1 for the clear atmosphere case with a dry stratosphere, assuming N = 10 co-added transits. For all cases listed in Table 1 except for the 10 mbar cloud/haze-layer, the mole fractions of both CO₂ and CH₄ have 1σ uncertainties between ~ 0.5-1 dex for N = 10 co-added transits. For the 10 mbar cloud/haze-layer, similar uncertainties are achieved for N ≳ 20 co-added transits, and Figure A2 shows how the posterior constraints are affected by the assumption of N = 10, N = 20, and N = 30 co-added transits.

4 DISCUSSION

Assuming the atmosphere of TRAPPIST-1e has a composition similar to the Archean Earth, the results of this study indicate that strong (5σ) detections of both CH₄ and CO₂ could potentially be achieved by co-adding N = 5-10 transit observations made with the NIRSpec prism (Figure 5). This is in broad agreement with the results of KT2018a, who reported that N ≈ 10 transits would be required to detect both molecules. However, unlike KT2018a, the presence of an optically thick cloud/haze layer at a pressure of 10 mbar is found to significantly increase the observing time required to N ≳ 30 transits to achieve detections at the same significance level. A cloud/haze layer at 1 mbar is found to prevent detections of both CH₄ and CO₂ even with 200 hr observing time (Figures 5 and 6).

If CH₄ and CO₂ are detected together in the atmosphere of TRAPPIST-1e, KT2018a and KT2018b suggest that the interpretation of these gases as an Archean biosignature would be strengthened by constraining the CO mole fraction to below about 100 ppm. The present study finds that this likely cannot be achieved with the NIRSpec prism. Table 1 reports 2σ and 3σ credible upper bounds for the CO mole fraction, which are at best around ~ 2.6 dex (2,500 ppm) and ~ 1.6 dex (25,000 ppm), respectively. The green lines in Figure 4 show the strength of CO absorption at these mole fractions, highlighting the challenge of using the NIRSpec prism to place strong upper limits on the CO mole fraction, at least for the case of an Archean-like atmosphere. However, complementary high-resolution spectroscopy measurements made with large ground-based telescopes may prove better suited for placing tighter upper limits on the CO mole fraction (e.g. Snellen et al. 2010; Brogi et al. 2014; Giacobbe et al. 2021).

In another relevant study, Fauchez et al. (2019) (F2019) modeled the atmosphere of TRAPPIST-1e using a 3D general circulation model coupled to a 1D photochemistry model, assuming an Archean Earth composition. F2019 investigated the detectability of CH₄ and CO₂ using a relatively simple approach for evaluating detection uncertainties, in which measurement uncertainties were directly compared to the amplitude of individual spectral features. Specifically, only the CH₄ band at 1.2 μm and CO₂ band at 4.3 μm were considered (Figure 1). This is different to the Bayesian methodology described in Section 2.3, which effectively accounts for all spectral features of a given species when considering if it can be detected. The updated Agol et al. (2021) system parameters for TRAPPIST-1e adopted in the present study (Section 2.1) were also not available at the time F2019 study. As with KT2018a, F2019 instead adopted the system parameters of Grimm et al. (2018) (i.e. Mₚ = 0.772 M⊕, Rₚ = 0.91 R⊕), corresponding to a surface gravity 14% higher than assumed for the present study. Nonetheless, the results of F2019 appear to be broadly consistent with those of the present study. In particular, at wavelengths close to the 1.2 μm CH₄ band, the F2019 model predicts an optically thick haze extending to altitudes of approximately 45 km, corresponding to pressures of ~ 1 mbar. For the 1 mbar cloud/haze-layer case considered in the present study, F2019 find that CH₄ would be undetectable. In addition, the F2019 model predicts a decreasing haze opacity with increasing wavelength, lowering to an effective altitude of approximately 17 km (~ 100 mbar in pressure) at wavelengths close to the 4.3 μm CO₂ band. F2019 find that a 3σ detection of CO₂ would require N = 8 co-added transits, which is reasonably close to the N = 6 transits estimated by the present study (Figure 6). However, for a 5σ detection of CO₂, F2019 report that N = 23 co-added transits would be required, which is significantly higher than the N = 10-14 co-added transits estimated by the present study, allowing for uncertainty in the moisture content of the stratosphere (Figure 6). This latter discrepancy could be due to a number of factors, such as the different methods used to evaluate detection significances, differences in the model transmission spectra (such as those arising due to the different assumed system properties), and F2019 perhaps not allowing for partial-saturation. To check the sensitivity of the results to these latter assumptions, additional calculations were performed using the Grimm et al. (2018) parameters to generate the synthetic transmission spectra and without allowing for partial saturation (Figure 3). With these settings, the present study estimates that a 5σ detection of CO₂ for the 100 mbar cloud/haze-layer would require N ≳ 35 co-added transits, which is a factor of 2-3 higher than the fiducial estimate of N = 10-14 co-added transits cited above, indicating that one or more such factors could potentially account for the different results obtained in this study and that of F2019. In any case, both F2019 and the present study conclude that cloud/haze at pressures below 100 mbar would significantly hinder the detection CO₂ and CH₄.

Since 5σ detections of both CH₄ and CO₂ would require co-adding N = 30 transits for the 10 mbar cloud/haze-layer case (Figures 5 and 6), it is worth considering the corresponding measurement precisions that this would require of the NIRSpec prism. These precisions are shown in Figure 7 for N = 10, N = 20, and N = 30 co-added transits, assuming the partial-saturation strategy described in Section 2.1.1. Also indicated is the measurement precision that has been demonstrated by co-adding transit observations made with the Hubble Space Telescope (HST) Wide Field Camera 3 (WF3C) spectrograph. For example, measurement precisions of ~ 15-20 ppm have been achieved across the 1.12-1.64 μm wavelength range with WF3C at R = 25-80 by Line et al. (2016), Mikal-Evans et al. (2021), and Guilluy et al. (2021). In addition, these latter studies have demonstrated convincing transit-to-transit repeatability and resulting atmospheric spectra that are well explained by physically plausible models, which increases confidence in the quoted uncertainties being realistic. For N = 30 co-added transits, Figure 7 shows that the highest precisions anticipated for NIRSpec will be comparable to these best precisions demonstrated with WF3C. It therefore seems reasonable to expect that NIRSpec could be capa-
Figure 5. Detection significance versus number of coadded transits for CO$_2$ (top panel) and CH$_4$ (middle panel) assuming a dry stratosphere. The bottom panel shows the significance with which the full atmosphere model is favoured over a flat transmission spectrum. For each panel, the top horizontal axis gives the JWST time required, assuming a time charge of 4.3 hr per transit. Observing program categories of small ($< 25$ hr), medium (25-70 hr), and Large ($> 70$ hr) are indicated. Note that the 1 mbar cloud scenario falls below the displayed axis range in the bottom panel. Inspired by a similar figure in Kaltenegger et al. (2020).
Figure 6. Comparison of the detection significance versus observing time for the dry stratosphere (dashed lines) and moist stratosphere (solid lines) scenarios, with the same axis format as Figure 5. Each row corresponds to a different cloud/haze assumption, as labelled along the right vertical axes. Blue lines indicate the detection significance of CO$_2$ (left column) and CH$_4$ (right column). A moist stratosphere generally increases the required observing time for CO$_2$ and CH$_4$ detections at a given significance, due to H$_2$O spectral bands overlapping those of CO$_2$ and CH$_4$. Brown lines indicate the significance with which the full atmosphere model is favoured over a flat transmission spectrum. The latter are identical for both columns of each row and are displayed to allow direct comparison with each molecule individually.
ble of achieving these measurement precisions, given that NIRSpec will employ similar HgCdTe detectors to those of WFC3 (Greene et al. 2016). Indeed, the ultimate systematics noise floor of NIRSpec could even prove to be lower than that of WFC3, owing to the stable L2 orbit of JWST compared to the low-Earth orbit of HST. For example, the L2 orbit will allow continuous-stare observations to be made, which should help mitigate the charge-trapping systematic that affects WFC3 (Zhou et al. 2017). Ultimately, however, the on-sky performance of JWST is yet to be determined and the assumption of noise that bins down as \(\sqrt{N}\) might prove too optimistic, somewhat analogously to early predictions of transit survey yields that neglected the role of systematic noise sources (e.g. Pont et al. 2006).

The results presented here also do not account for the possible effect of stellar variability. This could be especially relevant for the TRAPPIST-1 system, given that molecular absorption bands in star spots are possibly a significant source of contamination for the planetary transmission spectrum (Zhang et al. 2018, cf. Wakeford et al. 2019). Another point to note is that each result shown in Figures 5 and 6 effectively corresponds to the convergent mean of many repeated \(N\)-transit observing programs. In practice, only a single \(N\)-transit program would be conducted, representing a random draw from a distribution of \(N\)-transit programs. Exploring the dispersion of the latter would require repeating the present analysis many times, each with random measurement perturbations, which were not included here.

Finally, it should be stressed that many atmospheric compositions are possible for TRAPPIST-1e, a number of which have been investigated elsewhere (e.g. Barstow & Irwin 2016; Morley et al. 2017; Lustig-Yaeger et al. 2019). Only the very specific scenario of an Archean-like composition was considered here. However, the results obtained support the basic finding of KT2018a that, under this specific scenario, the detection of both CH\(_4\) and CO\(_2\) may be achievable with \(N \leq 10\) transit observations made with the NIRSpec prism, corresponding to less than 50 hours of JWST time. Nonetheless, numerous uncertainties remain that cannot be resolved until such an observing program is executed. In particular, this includes the unknown presence or absence of a significant cloud/haze layer at pressures below 100 mbar, which as noted above would either increase the amount of observing time required or prevent the detection of spectral features altogether.

5 CONCLUSION

If TRAPPIST-1e has an atmospheric composition similar to that of the Archean Earth, strong (5\(\sigma\)) detections for both CH\(_4\) and CO\(_2\) are possible for \(N = 5\)–10 transit observations under the assumption of well-behaved instrumental noise and neglecting the effect of stellar variability. This result holds when cloud/haze-layers are included down to pressures of 100 mbar (i.e. up to altitudes of \(\sim 16\) km), encompassing the range of H\(_2\)O cloud-top pressures predicted by GCM simulations of tidally-locked planets such as TRAPPIST-1e. However, if a cloud/haze-layer is present at a pressure level of 10 mbar, the required observing time increases significantly to \(N \approx 30\) transits for 5\(\sigma\) detections of both gases, assuming the NIRSpec prism will be able to achieve similar measurement precisions to the best demonstrated by HST WFC3. A cloud/haze-layer at the 1 mbar pressure level would prevent the detection of both CH\(_4\) and CO\(_2\) even for a 200 hr observing program devoted exclusively to TRAPPIST-1e.

DATA AVAILABILITY

No new data were generated or analysed in support of this research.

REFERENCES

Agol, E., Dorn, C., Grimm, S. L., et al. 2021, Planet. Sci. J., 2, 1
Arney, G., Domagal-Goldman, S. D., Meadows, V. S., et al. 2016, Astrobiology, 16, 873
Atreya, S. K., & Romani, P. N. 1985, Photochemistry and clouds of Jupiter, Saturn and Uranus., ed. G. E. Hunt, 17–68
Barclay, T., Kostov, V. B., Colón, K. D., et al. 2021, arXiv e-prints, arXiv:2109.14608
Barnes, R. 2017, Celestial Mechanics and Dynamical Astronomy, 129, 509
Barstow, J. K., & Irwin, P. G. J. 2016, MNRS, 461, L92
Batalha, N. E., Lewis, N. K., Line, M. R., Valenti, J., & Stevenson, K. 2018, ApJ, 856, L34
Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, PASP, 129, 064501
Battersby, C., Armus, L., Bergin, E., et al. 2018, Nature Astronomy, 2, 596
Benneke, B., & Seager, S. 2013, ApJ, 778, 153
Benneke, B., Wong, I., Piaulet, C., et al. 2019, ApJ, 887, L14
Bétrémieux, Y., & Kaltenegger, L. 2014, ApJ, 791, 7
Brogi, M., de Kok, R. J., Birkby, J. L., Schwarz, H., & Snellen, I. A. G. 2014, A&A, 565, A124
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Chiao, E. W., McCormick, M. P., & Chu, W. P. 1997, Journal of Geophysical Research: Atmospheres, 102, 19105
de Wit, J., Wakeford, H. R., Lewis, N. K., et al. 2018, Nature Astronomy, 2, 214
Edwards, B., Changeat, Q., Mori, M., et al. 2021, AJ, 161, 44
Fauchez, T. J., Turbet, M., Villanueva, G. L., et al. 2019, ApJ, 887, 194
Fauchez, T. J., Villanueva, G. L., Schwieterman, E. W., et al. 2020a, Nature Astronomy, 4, 372
Fauchez, T. J., Turbet, M., Wolf, E. T., et al. 2020b, Geoscientific Model Development, 13, 707
Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, AJ, 155, 200
Fortney, J. J. 2005, MNRAS, 364, 649

Figure 7. Noise properties calculated with pandeox for the measured transmission spectrum of TRAPPIST-1e assuming a partial saturation observing strategy and different numbers of co-added transits. The grey region indicates precisions that have been demonstrated with HST WFC3 at the same spectral resolving power (\(R = 80\)) by co-adding multiple, repeatable transit observations.
APPENDIX A: POSTERIOR DISTRIBUTIONS

For a subset of the most relevant parameters, Figure A1 shows the posterior covariances and marginalised distributions that were obtained for the dry stratosphere scenario described in Section 2.1 assuming a clear atmosphere and $N = 10$ co-added transits. Figure A2 shows the same, but for the dry stratosphere scenario with a cloud/haze-layer at a pressure of 10 mbar and $N = 10$, $N = 20$, and $N = 30$ co-added transits. These figures were produced using the publicly available chainconsumer Python software (Hinton 2016).
Dry stratosphere, clear (no cloud)

10 x transits:

\[
\log_{10} P_0 = -1.91^{+0.31}_{-0.35} \text{ (bar dex)}
\]
\[
R_p = 0.925^{+0.031}_{-0.033} \text{ (} R_\oplus \text{)}
\]
\[
M_p = 0.69^{+0.06}_{-0.07} \text{ (} M_\oplus \text{)}
\]
\[
T = 227^{+23}_{-21} \text{ (Kelvin)}
\]
\[
\log_{10} \text{CH}_4 = -1.95^{+0.39}_{-0.67} \text{ (dex)}
\]
\[
\log_{10} \text{CO}_2 = -0.90^{+0.27}_{-0.76} \text{ (dex)}
\]
\[
\log_{10} \text{CO} = -6.67^{+2.28}_{-2.21} \text{ (dex)}
\]

Figure A1. Posterior distributions obtained for a subset of model parameters from the dry stratosphere retrieval analysis assuming a clear atmosphere and \( N = 10 \) co-added transits. Panels along the diagonal show marginalised posterior distributions with blue shading indicating the 1σ credible ranges. Off-diagonal panels show parameter covariances with contours corresponding to the 1σ, 2σ, and 3σ credible bounds. Horizontal and vertical dashed black lines show the values for each parameter that were used to generate the synthetic datasets prior to performing the retrieval analyses. Text in the upper right half of the figure reports the posterior distribution medians and 1σ credible bounds.
Figure A2. Similar to Figure A1, but showing posterior distributions obtained for a subset of model parameters for the dry stratosphere scenario with a cloud/haze-layer at a pressure of 10 mbar. Contours show the 1σ credible ranges obtained assuming N = 10 (green), 20 (blue), and 30 (orange) co-added transits. Text in the upper right half of the figure reports the posterior medians and 1σ credible bounds, with the same colour-coding as the plotted posterior distributions.