A&A 664, A23 (2022)
https://doi.org/10.1051/0004-6361/202141964

Astronomy & Astrophysics

Large Interferometer For Exoplanets (LIFE)

III. Spectral resolution, wavelength range, and sensitivity requirements based on atmospheric retrieval analyses of an exo-Earth

B. S. Konrad[1,2], E. Alei[1,2], S. P. Quanz[1,2], D. Angerhausen[1,2,3], Ó. Carrión-González[4], J. J. Fortney[5], J. L. Grenfell[6], D. Kitzmann[7], P. Mollière[8], S. Rugheimer[9], F. Wunderlich[6], and the LIFE Collaboration*

[1] ETH Zurich, Institute for Particle Physics & Astrophysics, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland
e-mail: konradb@student.ethz.ch; sascha.quanz@phys.ethz.ch
[2] National Center of Competence in Research PlanetS, Gesellschaftsstrasse 6, 3012 Bern, Switzerland
[3] Blue Marble Space Institute of Science, Seattle, WA, USA
[4] Zentrum für Astronomie und Astrophysik, Technische Universität Berlin, Hardenbergstrasse 36, 10623 Berlin, Germany
[5] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[6] Department of Extrasolar Planets and Atmospheres (EPA), Institute of Planetary Research (PF), German Aerospace Center (DLR), Rutherfordstr. 2, 12489 Berlin, Germany
[7] University of Bern, Center for Space and Habitability, Gesellschaftsstrasse 6, 3012 Bern, Switzerland
[8] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[9] Dept of Physics, University of Oxford, Oxford OX1 3PU, UK

Received 5 August 2021 / Accepted 1 March 2022

ABSTRACT

Context. Temperate terrestrial exoplanets are likely to be common objects, but their discovery and characterization is very challenging because of the small intrinsic signal compared to that of their host star. Various concepts for optimized space missions to overcome these challenges are currently being studied. The Large Interferometer For Exoplanets (LIFE) initiative focuses on the development of a spacebased mid-infrared (MIR) nulling interferometer probing the thermal emission of a large sample of exoplanets.
Aims. This study derives the minimum requirements for the signal-to-noise ratio (*S/N*), the spectral resolution (*R*), and the wavelength coverage for the LIFE mission concept. Using an Earth-twin exoplanet as a reference case, we quantify how well planetary and atmospheric properties can be derived from its MIR thermal emission spectrum as a function of the wavelength range, *S/N*, and *R*.
Methods. We combined a cloud-free 1D atmospheric radiative transfer model, a noise model for observations with the LIFE interferometer, and the nested sampling algorithm for Bayesian parameter inference to retrieve planetary and atmospheric properties. We simulated observations of an Earth-twin exoplanet orbiting a G2V star at 10 pc from the Sun with different levels of exozodiacal dust emissions. We investigated a grid of wavelength ranges (3–20 µm, 4–18.5 µm, and 6–17 µm), *S/Ns* (5, 10, 15, and 20 determined at a wavelength of 11.2 µm), and *R*s (20, 35, 50, and 100).
Results. We find that H_2O, CO_2, and O_3 are detectable if $S/N \geq 10$ (uncertainty $\leq \pm 1.0$ dex). We find upper limits for N_2O (abundance $\lesssim 10^{-3}$). In conrtrast, CO, N_2, and O_2 are unconstrained. The lower limits for a CH_4 detection are $R = 50$ and $S/N = 10$. Our retrieval framework correctly determines the exoplanet's radius (uncertainty $\leq \pm 10\%$), surface temperature (uncertainty $\leq \pm 20$ K), and surface pressure (uncertainty $\leq \pm 0.5$ dex) in all cloud-free retrieval analyses. Based on our current assumptions, the observation time required to reach the specified *S/N* for an Earth-twin at 10 pc when conservatively assuming a total instrument throughput of 5% amounts to ≈6–7 weeks with four 2 m apertures.
Conclusions. We provide first order estimates for the minimum technical requirements for LIFE via the retrieval study of an Earth-twin exoplanet. We conclude that a minimum wavelength coverage of 4–18.5 µm, an *R* of 50, and an *S/N* of at least 10 is required. With the current assumptions, the atmospheric characterization of several Earth-like exoplanets at a distance of 10 pc and within a reasonable amount of observing time will require apertures ≥2 m.

Key words. methods: statistical – planets and satellites: terrestrial planets – planets and satellites: atmospheres

1. Introduction

Since the detection of 51 Peg b, the first planetary companion to a solar-type star (Mayor & Queloz 1995), exoplanet research has become one of the pillars of modern astrophysics. With more than 4000 exoplanets currently known[1], scientists have begun to uncover the vast diversity among exoplanet objects and extrasolar systems. A long-term goal is the discovery and the atmospheric characterization of a large sample of terrestrial exoplanets, with a specific focus on temperate objects.

In this context, a direct detection approach is essential in order to investigate the diversity of planetary atmospheres, assess the potential habitability of some objects, and look for biosignatures in their atmospheres. Different concepts for large exoplanet imaging space missions are currently being assessed, with LUVOIR (Peterson et al. 2017) and HabEx (Gaudi et al. 2020), which aim to directly measure the reflected spectrum of exoplanets in the visible (VIS) and near-infrared (NIR) range, being

* Webpage: www.life-space-mission.com
[1] https://exoplanetarchive.ipac.caltech.edu

prominent examples. The Large Interferometer For Exoplanets (LIFE) initiative[2] follows a complementary approach by focusing on the prospects of a large, space-based mid-infrared (MIR) nulling interferometer, which will observe the thermal emission spectrum and subsequently characterize the atmospheres of a large sample of (terrestrial) exoplanets (Quanz et al. 2018, 2021, 2022). The LIFE initiative aims to combine various efforts to push toward an eventual launch of such a large, space-based MIR nulling interferometer. The work we present here aims to constrain some of the instrument requirements for LIFE and it is the third paper in a series. The first paper of the series (Quanz et al. 2022) quantifies the exoplanet detection performance of LIFE and compares it with large single-aperture mission concepts for reflected light. The second paper (Dannert et al. 2022) introduces the LIFEsim instrument simulator and the necessary signal extraction algorithms.

The choice of the MIR wavelength range for LIFE pays dividends. More molecules have strong absorption bands in the MIR spectra of Earth-like planets, which allows us to better assess the atmospheric structure and composition (e.g., Des Marais et al. 2002; Line et al. 2019). Further, the infrared is less affected by the presence of hazes and clouds (see, e.g., Kitzmann et al. 2011; Rugheimer et al. 2013; Arney et al. 2018; Lavvas et al. 2019; Fauchez et al. 2019; Komacek et al. 2020; Wunderlich et al. 2021), which is a major challenge at visible wavelengths (see, e.g., Sing et al. 2016, for Jovian planets)[3]. Importantly, emission spectra allow us to constrain the planetary radius (e.g., Line et al. 2019), which is degenerate with the planetary albedo in reflected light measurements (Carrión-González et al. 2020).

Finally, the MIR range includes a rich portfolio of biosignatures (e.g., Catling et al. 2018). Biosignatures in an exoplanet context are gases or features that can be detected at interplanetary distances and that are produced by life. Among the main biosignature gases there are oxygen (O_2) and its photochemical product ozone (O_3), methane (CH_4), nitrous oxide (N_2O), chloromethane (CH_3Cl), phosphine (PH_3), and dimethyl sulfide (C_2H_6S, commonly known as DMS). Many of these gases can also be generated by abiotic processes and can therefore be false positives in the search for biosignatures (Schwieterman et al. 2018; Harman & Domagal-Goldman 2018). However, the presence of multiple biosignature gases in the atmosphere, along with a planetary context that points toward habitability, would increase the robustness of the detection of life on an exoplanet. The most widely known set of multiple biosignatures is the so-called triple fingerprint of carbon dioxide (CO_2), water vapor (H_2O), and O_3, which are well detectable in Earth's thermal emission spectrum and potentially detectable in terrestrial exoplanets (see, e.g., Selsis et al. 2002). The simultaneous presence of reducing and oxidizing species in an atmosphere, such as O_2 and O_3 in combination with CH_4, is a strong biosignature, with no currently known false positives. These species would not be both present in large quantities in an atmosphere over long timescales without disequilibrium processes driven by the presence of life (Lederberg 1965; Lovelock 1965). An exhaustive discussion on biosignatures can be found in Schwieterman et al. (2018) and Catling et al. (2018) introduce a Bayesian framework for assessing the confidence level of a biosignature detection.

While there is ample scientific justification for choosing the MIR wavelength regime for detailed (atmospheric) investigations of terrestrial exoplanets, deriving a concept for a space-mission, as pursued by the LIFE initiative, requires the derivation of fundamental mission and instrument requirements, including instrument sensitivity, wavelength range coverage, and spectral resolution. Earlier steps in this direction were presented in von Paris et al. (2013) for the former *Darwin* mission concept (Léger et al. 1996). In a more recent study, Feng et al. (2018) used a Bayesian atmospheric retrieval approach to quantify the power of reflected light observations, as foreseen by HabEx or LUVOIR, as a function of instrument parameters.

Here, we aim to provide the minimum requirements for the parameters listed above for the LIFE mission concept using an atmospheric retrieval framework. Such a framework allows us to derive quantitative estimates for the main atmospheric and planetary parameters from a simulated exoplanetary spectrum (see, e.g., Madhusudhan 2018; Deming et al. 2018; Barstow & Heng 2020, for recent reviews). Using a Bayesian approach, the space of input parameters (e.g., atmospheric abundances) is explored iteratively to assess which combination of values best fits the simulated observations. Repeating this for various combinations of signal-to-noise ratio (*S/N*) of the emission spectrum, wavelength range, and spectral resolution (R) allows us to understand, how well the simulated planet can be characterized. In the present study, we used a cloud-free modern Earth-twin exoplanet as our reference case. As Earth is the only planet known to host life, we are particularly interested in assessing if and how well some of its main atmospheric constituents can be detected for various combinations of instrument parameters. We are aware that by ignoring clouds, we are somewhat simplifying the problem. However, we emphasize here that the main aim of our analysis is to get first estimates for wavelength range coverage, R, and *S/N* requirements for LIFE. In subsequent work, we will investigate other types of exoplanets and atmospheric compositions, providing additional constraints on some of the requirements.

In Sect. 2, we introduce our retrieval framework. It combines a 1D atmospheric forward model based on the `petitRADTRANS` radiative transfer code (Mollière et al. 2019), the LIFEsim instrument simulator (Dannert et al. 2022), that adds astrophysical noise terms to the simulated Earth spectrum, and a Nested Sampling approach (Skilling 2006) for Bayesian parameter inference. We validate the retrieval framework in Sect. 3. Then, we perform a series of retrievals of the theoretical emission spectrum of a cloud-free Earth-twin exoplanet as it would be observed by LIFE, varying the wavelength range coverage, R, and *S/N*. We present our results in Sect. 4. In Sect. 5, we discuss our results and compare our study with other works. Finally, we summarize our main findings and conclusions in Sect. 6.

2. Methods

Our atmospheric retrieval framework aims to infer the atmospheric pressure-temperature (P–T) structure and composition from simulated mock observations of the MIR thermal emission spectrum of an Earth-twin planet. At its heart, the framework consists of two elements. First, we need a parametric model for the atmosphere that calculates the emergent light spectrum corresponding to a set of model parameters (Sect. 2.1). Second, a parameter estimation algorithm is required to optimize the model parameters, such that the spectrum produced by the atmospheric model best fits the simulated observational data

[2] `www.life-space-mission.com`

[3] However, clouds can also increase the reflectivity and signal for molecules such as molecular oxygen (O_2) that are well-mixed in the atmosphere of a terrestrial planet (Kawashima & Rugheimer 2019).

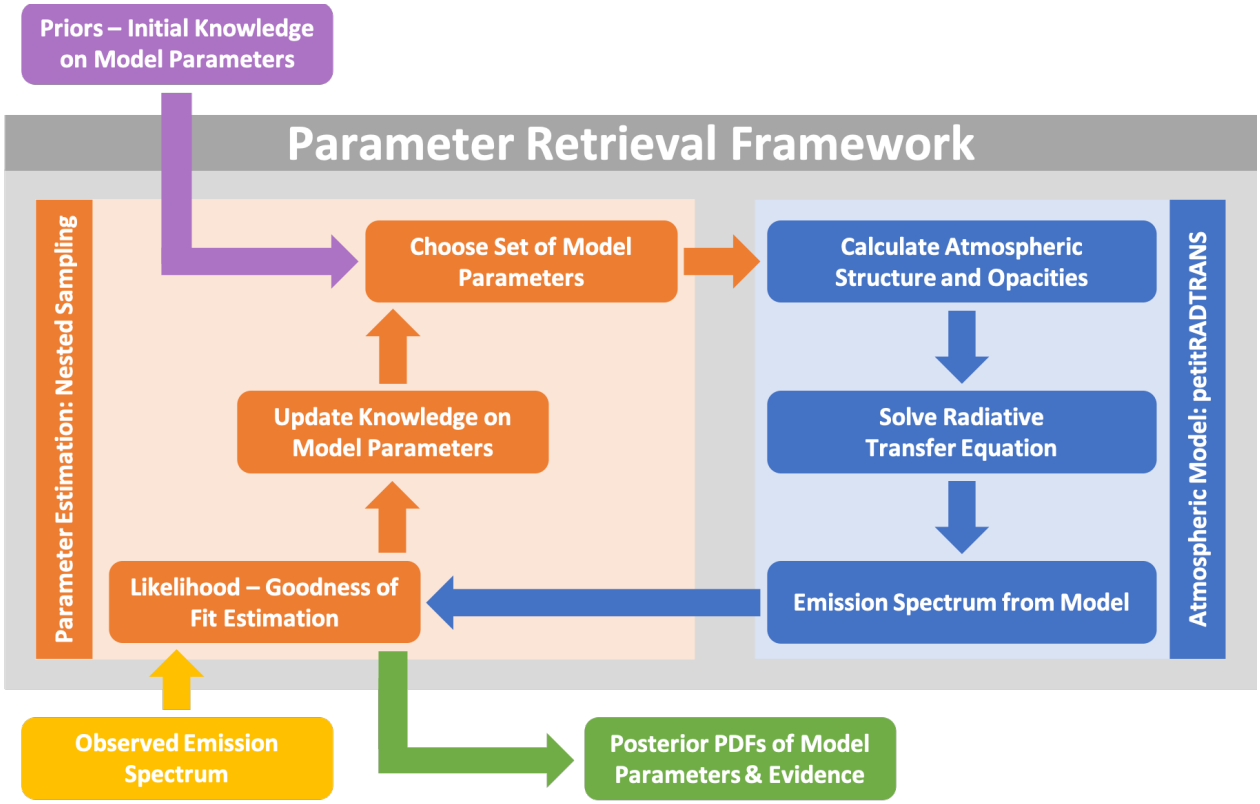


Fig. 1. Schematic illustrating our atmospheric retrieval framework.

(Sect. 2.2). These two elements are then combined to form our retrieval framework (Sect. 2.3). We provide an illustrative schematic summarizing our retrieval framework in Fig. 1.

2.1. Atmospheric model

As our atmospheric model, we used the 1D radiative transfer code `petitRADTRANS` (Mollière et al. 2019). To calculate the thermal emission spectrum of terrestrial exoplanets in the MIR wavelength range, `petitRADTRANS` passes a featureless black-body spectrum at the surface temperature through discrete atmospheric layers that interact with the radiation field. We characterized each layer by its temperature, pressure, and the opacity sources present.

2.1.1. Atmospheric *P–T* structure

Throughout this work, we parametrized the *P–T* structure of Earth's atmosphere using a fourth order polynomial:

$$T(P) = \sum_{i=0}^{4} a_i P^i. \tag{1}$$

Here, P denotes the atmospheric pressure and T the corresponding temperature. The parameters a_i are the parameters describing the polynomial *P–T* model. We chose this polynomial *P–T* model over other *P–T* profile parametrizations (e.g., models from Madhusudhan & Seager 2009; Guillot 2010; Mollière et al. 2019), since it provided a comparable description of the atmospheric *P–T* structure using fewer model parameters (see Appendix A for further information).

2.1.2. Opacity sources in `petitRADTRANS`

With `petitRADTRANS`, we can account for several different opacity sources that are potentially present in exoplanet atmospheres. In the computationally favorable low spectral resolution mode ($R = 1000$), the opacities originating from different atmospheric gases are calculated via the correlated-k method (Goody et al. 1989; Lacis & Oinas 1991; Fu & Liou 1992). Absorption lines from different molecules, collision-induced absorption (CIA), and atmospheric pressure broadening effects can be taken into account. Rayleigh and cloud scattering, as well as the scattering of direct radiation from the host star, can be considered. However, the scattering solution is achieved in an iterative way and the opacity handling is different (see Mollière et al. 2020, for more information). These changes increase the spectrum calculation time by roughly an order of magnitude when scattering is included in the computation.

In Fig. 2, we compare our scattering- and cloud-free Earth-twin MIR emission spectrum, for which we assumed uniform chemical abundances throughout the atmosphere, to other cloud-free and cloudy models (Daniel Kitzmann, priv. comm.; Rugheimer et al. 2015). These models considered scattering as well as nonuniform abundances. We see that scattering contributions and nonuniform abundances only have a minor impact on Earth's spectrum in the MIR and the differences between the cloud-free spectra are of similar magnitude as the most optimistic LIFESIM noise estimate. This justifies our approach of excluding scattering from our retrieval routine. Thus, we neglected effects linked to the incident stellar radiation, which reduces the number of retrieved parameters and thereby the computing time.

We also neglected scattering and absorption by clouds. As can be seen from Fig. 2, this is a significant simplification

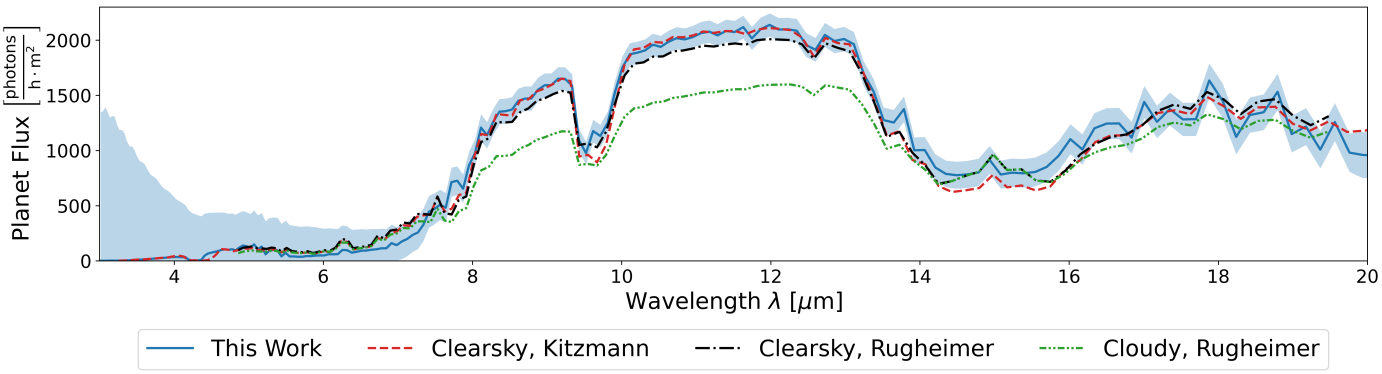


Fig. 2. Comparison of the Earth-twin MIR emission spectra calculated with various different models. We plot the photon flux received from an Earth-twin located 10 pc from the sun. The solid blue line is the MIR thermal emission calculated with `petitRADTRANS` using the settings discussed in Sect. 2.3.1. The blue-shaded region indicates the most optimistic LIFEsim uncertainty ($S/N = 20$) used in our retrievals. The red dashed line represents a cloud-free Earth model by Daniel Kitzmann (priv. comm.), which takes scattering into account. The green and black dashed-dotted lines are the cloudy (60% cloud coverage) and cloud-free modern Earth spectra from Rugheimer et al. (2015) that account for scattering.

of the problem. The presence of (opaque) clouds in an atmosphere will partially or fully obscure the view of the exoplanet's surface. Therefore, we expect systematic shifts in the surface temperature, pressure, and potentially also in the retrieved planetary radius if clouds are present in an atmosphere. Additionally, a partial cloud coverage combined with vertically nonconstant atmospheric abundances could lead to systematic offsets in the retrieved abundances. For example, if an atmospheric gas is only present below an optically thick cloud deck, it is likely not detectable via a MIR retrieval study. These considerations require further attention in future works. However, as we are interested in first estimates for specific instrument requirements, focusing on a cloud-free atmosphere is justifiable. Complete details on the implementation are given in Sect. 2.3.1.

Table 1. Jeffrey's scale (Jeffreys 1998).

$\log_{10}(K)$	Probability	Strength of evidence
<0	<0.5	Support for $\mathcal{M}_2$
0–0.5	0.5–0.75	Very weak support for $\mathcal{M}_1$
0.5–1	0.75–0.91	Substantial support for $\mathcal{M}_1$
1–2	0.91–0.99	Strong support for $\mathcal{M}_1$
>2	>0.99	Decisive support for $\mathcal{M}_1$

Notes. Scale for the interpretation of the Bayes' factor $K = \mathcal{Z}_{\mathcal{M}_1}(\mathcal{D})/\mathcal{Z}_{\mathcal{M}_2}(\mathcal{D})$. The scale is symmetrical, which means that negative values correspond to very weak, substantial, strong, or decisive support for $\mathcal{M}_2$.

2.2. Parameter estimation

Our retrieval study utilized a Bayesian parameter inference tool to sample the posterior probability distributions of the atmospheric forward model parameters. Bayesian parameter inference methods are based on Bayes' theorem, which provides a method for estimating model parameters based on experimental data (see, e.g., Trotta 2017; van de Schoot et al. 2021). Let us consider a model $\mathcal{M}$ described by a set of parameters $\Theta_\mathcal{M}$ and experimental data $\mathcal{D}$. Bayes' theorem states:

$$P(\Theta_\mathcal{M}|\mathcal{D},\mathcal{M}) = \frac{P(\mathcal{D}|\Theta_\mathcal{M},\mathcal{M})P(\Theta_\mathcal{M}|\mathcal{M})}{P(\mathcal{D}|\mathcal{M})}. \quad (2)$$

Here, $P(\Theta_\mathcal{M}|\mathcal{D},\mathcal{M})$ is the posterior probability (or "posterior") and represents the probability of different sets of model parameters $\Theta_\mathcal{M}$ under the constraint of the experimental data $\mathcal{D}$ and model $\mathcal{M}$. Further, $P(\mathcal{D}|\Theta_\mathcal{M},\mathcal{M})$ provides a probabilistic measure of how well a specific set of parameters $\Theta_\mathcal{M}$ for the model $\mathcal{M}$ describes the data $\mathcal{D}$. We calculated this probability via a log-likelihood function $\ln(\mathscr{L}(\Theta_\mathcal{M}))$:

$$\ln(\mathscr{L}(\Theta_\mathcal{M})) = \sum_{i=1}^{N}\left(-\frac{1}{2}\ln(2\pi\sigma_i^2) - \frac{1}{2}\frac{(\mathcal{D}_i - \mu_i(\Theta_\mathcal{M}))^2}{\sigma_i^2}\right). \quad (3)$$

Our log-likelihood function assumed that each of the N measured data points $\mathcal{D}_i$ behaves as a normally distributed quantity. The mean $\mu_i(\Theta_\mathcal{M})$ was assumed to be the value predicted by the model $\mathcal{M}$ using parameters $\Theta_\mathcal{M}$. The variable σ_i was the measurement error on the data point $\mathcal{D}_i$.

Next, $P(\Theta_\mathcal{M}|\mathcal{M})$ is the prior probability (or "prior") of the model parameters $\Theta_\mathcal{M}$ and represents the knowledge on the model parameters before taking the observational data into account. Finally, $P(\mathcal{D}|\mathcal{M})$ is frequently referred to as the Bayesian evidence $\mathcal{Z}_\mathcal{M}(\mathcal{D})$ and is a normalization constant, which ensures that the posterior is normalized to unity:

$$\mathcal{Z}_\mathcal{M}(\mathcal{D}) = P(\mathcal{D}|\mathcal{M}) = \int \mathscr{L}(\Theta_\mathcal{M})P(\Theta_\mathcal{M}|\mathcal{M})\mathrm{d}\Theta_\mathcal{M}. \quad (4)$$

Additionally, the evidence $\mathcal{Z}_\mathcal{M}(\mathcal{D})$ enables us to compare the performance of different models $\mathcal{M}_i$ to each other and to select the model that best describes the observed data $\mathcal{D}$. One can compare two models $\mathcal{M}_1$ and $\mathcal{M}_2$ by considering the Bayes' factor K:

$$K = \frac{P(\mathcal{M}_1|\mathcal{D})}{P(\mathcal{M}_2|\mathcal{D})} = \frac{\mathcal{Z}_{\mathcal{M}_1}(\mathcal{D})}{\mathcal{Z}_{\mathcal{M}_2}(\mathcal{D})}, \quad (5)$$

where the prior probabilities $\mathcal{P}(\mathcal{M}_i)$ for both models are assumed to be the same. An approach to interpreting the value of K is via the Jeffrey's scale (Jeffreys 1998), which is given in Table 1.

Due to the model comparison capabilities of the evidence $\mathcal{Z}_\mathcal{M}(\mathcal{D})$, we chose the nested sampling algorithm (Skilling 2006) over MCMC algorithms (e.g., the Metropolis-Hastings algorithm: Hastings 1970; Metropolis et al. 1953), since it provides a direct estimate for the Bayesian evidence $\mathcal{Z}_\mathcal{M}(\mathcal{D})$. Furthermore,

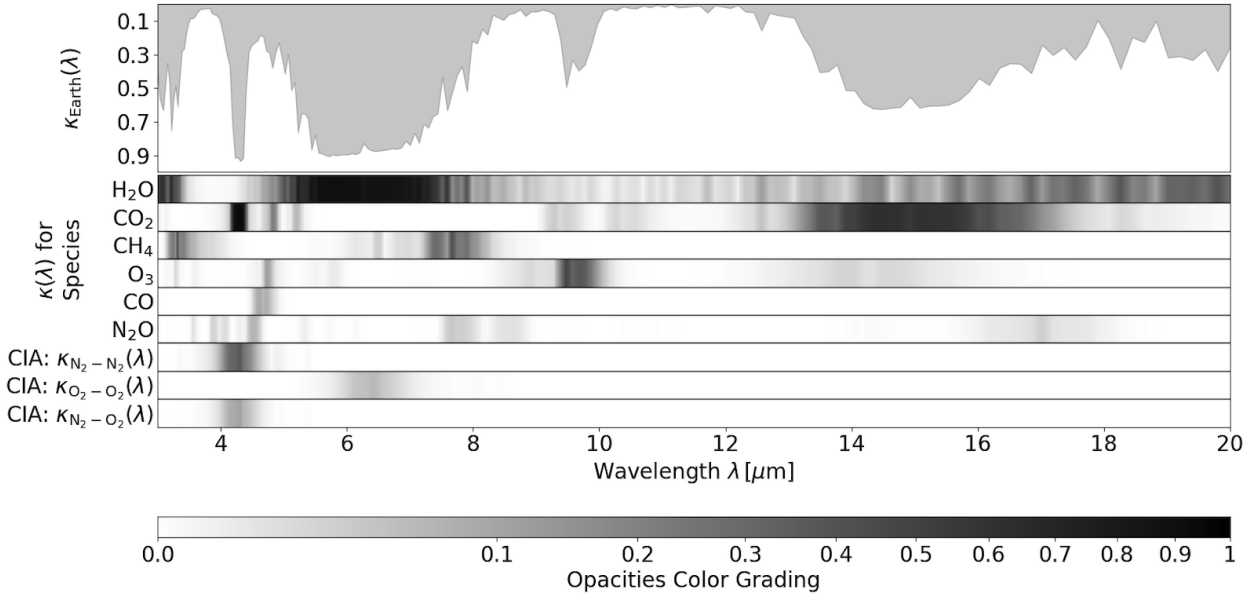


Fig. 3. Wavelength dependence of different opacity sources. *Upper panel*: the opacity of a cloudless Earth-twin atmosphere as a function of wavelength. Gray shading shows the amount of light blocked by the atmosphere. *Lower panel*: contribution of the different molecules to the opacity of the Earth-twin atmosphere as a function of wavelength. Dark regions indicate a high opacity, as is indicated by the colorbar.

Table 2. Line lists used throughout this study.

Species	Database	References
CO_2	ExoMol	Yurchenko et al. (2020)
O_3	HITRAN 2012	Mollière et al. (2019); Rothman et al. (2013)
CH_4	ExoMol	Chubb et al. (2021); Yurchenko et al. (2020)
CO	HITEMP	Rothman et al. (2010)
H_2O	HITEMP	Rothman et al. (2010)
N_2O	ExoMol	Chubb et al. (2021)

nested sampling is computationally less expensive and better at handling multimodal posterior distributions (Skilling 2006).

Specifically, we utilized the open-source `pyMultiNest` package (Buchner et al. 2014), which makes the nested sampling implementation `MultiNest` (Feroz et al. 2009) accessible to the Python language. `MultiNest` is based on the original nested sampling algorithm (Skilling 2006) and uses the importance nested sampling algorithm (Feroz et al. 2013) to obtain more accurate estimates of the Bayesian evidence $\mathcal{Z}_{\mathcal{M}}(\mathcal{D})$.

2.3. Retrieval setup

2.3.1. Forward model and noise terms

We generated a correlated-k `petitRADTRANS` spectrum ($R = 1000$) of an Earth-twin exoplanet using the parameter values listed in Table 3 (Input column) assuming an atmosphere consisting of 100 layers. We used the line lists from the ExoMol database (Tennyson et al. 2016) for CO_2, CH_4, and N_2O and from the HITRAN/HITEMP database (Rothman et al. 1995, 2010) for O_3, CO, and H_2O. We summarize the reference papers corresponding to the opacity line lists in Table 2. For N_2 and O_2, we considered CIA (as discussed in Schwieterman et al. 2015). All other atmospheric gases have distinct absorption features within the MIR range, even at low abundances (see Fig. 3). Furthermore, we assumed constant abundances of all molecules vertically throughout the atmosphere.

We then resampled the calculated MIR emission spectrum to the desired spectral resolution, keeping a constant resolution $R = \lambda/\Delta\lambda$ across the spectrum. This resulted in wavelength bins of variable width, where the width at short wavelengths was smaller than at long wavelengths. For the spectral resampling, we used the `SpectRes` tool (Carnall 2017), which allows for time efficient resolution reduction of MIR spectra whilst keeping the overall flux and energy conserved. We then used the radius R_{pl} and the distance from Earth d_{Earth} to scale the photon flux found with `petitRADTRANS` F_{pRT} to the flux LIFE would detect (F_{LIFE}):

$$F_{\mathrm{LIFE}} = F_{\mathrm{pRT}} \frac{R_{\mathrm{pl}}^2}{d_{\mathrm{Earth}}^2}. \tag{6}$$

Throughout this work, we defined the *S/N* of a spectrum as the *S/N* calculated in the reference bin at 11.2 µm. We chose this wavelength because it lies close to the peak flux and it does not coincide with strong absorption lines. The *S/N* at all other wavelengths was determined by the noise model, which relates the *S/N* for the reference bin to the *S/N* in all other wavelength bins.

For the retrieval validation we perform in Sect. 3, we only considered the photon noise of the planet spectrum. For the grid retrievals performed in Sect. 4.2, we obtained noise estimates via the LIFEsim tool (see Dannert et al. 2022). LIFEsim accounts for photon noise contributions from the planet's emission spectrum, stellar leakage, as well as local- and exozodiacal dust emission. In our simulations, we assumed that the noise did not impact the predicted flux values, but instead only added uncertainties to the simulated spectral points. As discussed in Feng et al. (2018), randomization of the individual spectral points based on the *S/N* allows us to simulate accurate observational instances. At the same time, retrieval studies on such instances result in biased results, since the random placement

Table 3. Parameters used in the retrievals, their input values, prior distributions, and the validation results.

Parameter	Description	Input	Prior	Validation
$\sqrt[4]{a_4}$	P–T Parameter (Degree 4)	1.14	$\mathcal{U}(0.5, 1.8)$	$1.12^{+0.05}_{-0.04}$
a_3	P–T parameter (Degree 3)	23.12	$\mathcal{U}(0, 100)$	$22.65^{+1.74}_{-1.54}$
a_2	P–T Parameter (Degree 2)	99.70	$\mathcal{U}(0, 500)$	$98.90^{+3.72}_{-3.44}$
a_1	P–T Parameter (Degree 1)	146.63	$\mathcal{U}(0, 500)$	$146.07^{+4.75}_{-4.84}$
a_0	P–T Parameter (Degree 0)	285.22	$\mathcal{U}(0, 1000)$	$285.01^{+3.16}_{-3.08}$
$\log_{10}(P_0$ [bar])	Surface pressure	0.006 [(1)]	$\mathcal{U}(-4, 3)$	$0.01^{+0.02}_{-0.02}$
$R_{\rm pl}$ $[R_\oplus]$	Planet radius	1.0	$\mathcal{G}(1.0, 0.2)$	$1.00^{+0.01}_{-0.01}$
$\log_{10}\left(M_{\rm pl}\ [M_\oplus]\right)$	Planet mass	0.0	$\mathcal{G}(0.0, 0.4)$	$-0.09^{+0.18}_{-0.22}$
$\log_{10}(N_2)$	N_2 Mass fraction	−0.107 [(1),(2)]	$\mathcal{U}(-3, 0)$	$-0.14^{+0.09}_{-0.11}$
$\log_{10}(O_2)$	O_2 Mass fraction	−0.679 [(1),(2)]	$\mathcal{U}(-3, 0)$	$-0.75^{+0.18}_{-0.17}$
$\log_{10}(H_2O)$	H_2O Mass fraction	−3.000	$\mathcal{U}(-15, 0)$	$-3.09^{+0.18}_{-0.22}$
$\log_{10}(CO_2)$	CO_2 Mass fraction	−3.387 [(1),(2)]	$\mathcal{U}(-15, 0)$	$-3.47^{+0.18}_{-0.22}$
$\log_{10}(CH_4)$	CH_4 Mass fraction	−5.770 [(1),(2)]	$\mathcal{U}(-15, 0)$	$-5.85^{+0.18}_{-0.23}$
$\log_{10}(O_3)$	O_3 Mass fraction	−6.523 [(2)]	$\mathcal{U}(-15, 0)$	$-6.61^{+0.18}_{-0.23}$
$\log_{10}(CO)$	CO Mass fraction	−6.903 [(2)]	$\mathcal{U}(-15, 0)$	$-7.02^{+0.23}_{-0.28}$
$\log_{10}(N_2O)$	N_2O Mass fraction	−6.495 [(2)]	$\mathcal{U}(-15, 0)$	$-6.60^{+0.24}_{-0.27}$

Notes. The last column shows the output value with 1σ uncertainties for each parameter obtained in the validation run (Sect. 3). Here, $\mathcal{U}(x, y)$ denotes a boxcar prior with a lower threshold x and upper threshold y; $\mathcal{G}(\mu, \sigma)$ represents a Gaussian prior with mean μ and standard deviation σ. For a_4 we choose a prior on $\sqrt[4]{a_4}$, which allows us to sample small values more densely, typical of a fourth order coefficient, and then take the fourth power to obtain a_4.

References. [(1)]NASA's planet factsheet: https://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html [(2)]Table A.1 in Seager et al. (2016)

of the small number of data points will impact the retrieval's performance. The ideal analysis would study many (at least 10) data realizations for each considered spectrum and assess instrument performance by considering the posteriors found for these different noise instances. However, the number of cases (96) we considered and the computation time per case (≈1 day on 90 CPUs) made such a study computationally unfeasible (≳30 months of cluster time). By not randomizing the individual spectral points, we eliminated the biases introduced by noise instances. However, we are aware that this approach likely leads to overly optimistic results. Namely, we expect an unrealistic centering of posteriors on the truths. Additionally, for molecules at the sensitivity limit (weak spectral features), we expect overly optimistic results (see, e.g., Feng et al. 2016). In Appendix C, we performed retrievals on randomized noise instances to study these effects in more detail.

2.3.2. Priors

We provide the assumed prior ranges for all model parameters in Table 3. For the polynomial P–T profile and the surface pressure P_0, we chose priors that cover a wide range of atmospheric structures. The prior range included Venus- or Mars-similar atmospheric structures (e.g., P_0 is considered over the range from 10^{-4} to 1000 bar, as indicated by $\mathcal{U}(-4, 3)$).

As demonstrated in Dannert et al. (2022), the detection of a planet during the search phase of the LIFE mission will already provide first estimates for the radius $R_{\rm pl}$ of the object. Specifically, for small, rocky planets within the habitable zone, the authors showed that a detection in the search phase would provide an estimate $R_{\rm est}$ for the true planet radius $R_{\rm true}$ with an accuracy of $R_{\rm est}/R_{\rm true} = 0.97^{\pm 0.18}$. For our simulations, we therefore assumed that a rough estimate of the radius is known and we chose a Gaussian prior with 20% uncertainty for this parameter. Next, we used the $R_{\rm pl}$ to derive a constraint on the planet's mass $M_{\rm pl}$ via a statistical mass-radius relation (see, e.g., Hatzes & Rauer 2015; Wolfgang et al. 2016; Zeng et al. 2016; Chen & Kipping 2016; Otegi et al. 2020). In our retrievals, we used `Forecaster`[4] (Chen & Kipping 2016), which allowed us to set a prior on the planetary mass $M_{\rm pl}$ based on the estimate for the radius $R_{\rm pl}$. The tool relies on the statistical analysis of 316 objects (Solar System objects and exoplanets) for which well-constrained mass and radius estimates are available. It produces accurate predictions for a large variety of different objects, spanning from dwarf planets to late-type stars.

For the trace gases, we assumed uniform priors between −15 and 0 in $\log_{10}$ mass fraction. For the bulk constituents N_2 and O_2, we used uniform priors between −3 and 0 in $\log_{10}$ mass fraction. This range gave us an increased sampling density in the high abundance regime, where we expected the sensitivity limit to be located for N_2 and O_2. Furthermore, we used N_2 as filling gas in our atmosphere, to ensure that: $\sum$(gas abundances) = 1.

3. Validation

Before running our retrieval framework for simulated LIFE data, we validated its accuracy and performance. For the retrieval validation, we retrieved a full resolution ($R = 1000$) Earth-twin MIR thermal emission spectrum covering the wavelength range 3–20 µm. We generated the validation input spectrum using `petitRADTRANS` and the parameter values provided in Table 3. We only considered photon noise from the spectrum itself and

[4] https://github.com/chenjj2/forecaster

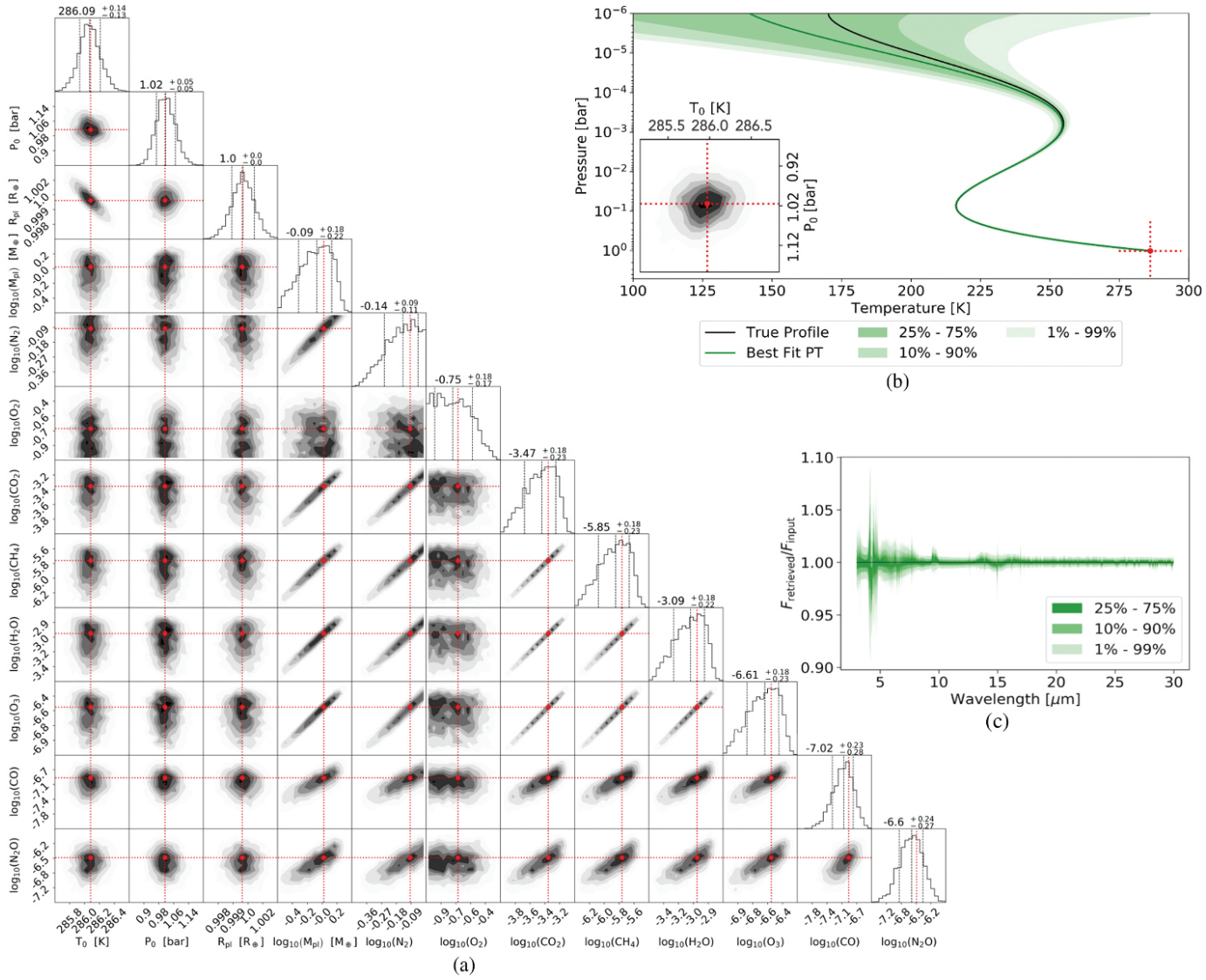


Fig. 4. Results from the validation run outlined in Sect. 3. (*a*) Corner plot for the posterior distributions of the planetary surface temperature T_0, surface pressure P_0, radius $R_{\rm pl}$, mass $M_{\rm pl}$, and retrieved abundances of different molecules. The red lines indicate the values used to generate the input spectrum. Additionally, we plot the retrieved median and the 16th and 84th percentile as dashed lines in every posterior plot. (*b*) Retrieved *P*–*T* profile. The shaded green regions show the uncertainties on the retrieved profile. In the bottom left corner of the *P*–*T* profile plot, we display P_0 and T_0. The red cross marks the input values. (*c*) The retrieved emission spectrum $F_{\rm retrieved}$ relative to the input emission spectrum for the retrieval $F_{\rm input}$.

chose an *S/N* of 50 at 11.2 µm. Furthermore, we assumed that the photon noise did not impact the simulated flux values, but instead only added uncertainties to the spectral points. In our validation retrieval, we ran the `pyMultiNest` package using 700 live points and a sampling efficiency of 0.3 as suggested for parameter retrieval by the documentation[5].

We summarize the results in Fig. 4 and Table 3 (last column). The corner plot in Fig. 4a suggests that the exoplanet's radius $R_{\rm pl}$ is retrieved to a very high precision with an uncertainty of roughly 0.001 $R_\oplus$, a significant improvement over the assumed prior distribution. Similarly, the retrieved posterior for the exoplanet mass $M_{\rm pl}$ is more strongly constrained than the assumed prior distribution, with the standard deviation of the posterior ($0.2 \cdot \log_{10}(M_\oplus)$) being significantly smaller than the standard deviation of the prior ($0.4 \cdot \log_{10}(M_\oplus)$). The centering of the $R_{\rm pl}$ and $M_{\rm pl}$ posteriors on the true values and the lack of significant correlation between the two posteriors implies that the surface gravity g_{pl} is estimated accurately. The surface pressure P_0 and surface temperature T_0 are both accurately retrieved to a very high precision, with an uncertainty of roughly 0.1 K for T_0 and 0.1 bar for P_0 (see Figs. 4a and 4b). Further, we observe a correlation between the planetary radius $R_{\rm pl}$ and the surface temperature T_0. This indicates that a higher T_0, which results in more emission per surface area, can be compensated by a smaller $R_{\rm pl}$, which results in a smaller emitting area.

From the retrieved posterior distribution of N_2, we see that our retrieval framework allows us to rule out low N_2 abundances in Earth's atmosphere via the N_2–N_2 CIA feature at 4 µm. However, the same N_2-N_2 CIA feature is too weak to rule out very high N_2 abundances. In contrast, the retrieval did not manage to find evidence for O_2 in Earth's atmosphere. However, the retrieval managed to limit the O_2 abundance to maximally 0.35 in mass fraction.

[5] https://johannesbuchner.github.io/PyMultiNest/pymultinest_run.html

The retrieved posterior distributions for the remaining molecules show a strong correlation with the $M_{\rm pl}$ posterior and consequently with the surface gravity g_{pl}. This correlation is evident in the corner plot of Fig. 4a since both $M_{\rm pl}$ and the molecular abundances of most retrieved trace gases exhibit a similarly shaped, non-Gaussian posterior distribution. This is a well known physical degeneracy and has been described in other studies (see, e.g., Mollière et al. 2015; Feng et al. 2018; Madhusudhan 2018; Quanz et al. 2021); it is not related to a numerical artifact. The degeneracy appears since the same spectral feature can be explained by different combinations of gravity and atmospheric composition. This degeneracy originates from the mass appearing in the form of the surface gravity in the hydrostatic equilibrium. Therein, the surface gravity is degenerate with the mean molecular weight of the atmospheric species. Since we derive the mean molecular weight from the abundances of the trace gases, this connects the planet's mass to the trace gas abundances present in its atmosphere. Further evidence for this degeneracy can be found in Fig. 4c. Despite the degeneracy between the retrieved mass and molecule abundances, the relative difference between the input spectrum and the spectra corresponding to the retrieved parameters is small.

The posteriors of CO and N_2O are broader, roughly Gaussian, and less correlated with $M_{\rm pl}$. This indicates that the constraint imposed by the retrieval framework on the abundances of these species is not solely limited by the degeneracy with $M_{\rm pl}$, but also by our retrieval's sensitivity for CO and N_2O.

A method of dealing with these strong correlations is to consider relative instead of absolute abundances. This allows us to minimize the impact of systematic uncertainties that affect all retrieved trace gas abundances in the same way at the cost of losing information on the absolute abundances. Relative abundances of trace gases are of interest to us since they provide a probe to whether an atmosphere is in chemical disequilibrium, which could potentially be upheld by life. For example, we could consider the abundance of CH_4 or N_2O relative to a strongly oxidizing species such as O_2 (or O_3, which is a photochemical product of O_2) as first suggested in Lovelock (1965) and Lippincott et al. (1967). These gases react rapidly with each other and therefore the simultaneous presence of both molecules is only possible if they are continually replenished at a high rate. On Earth, O_2 is constantly produced via photosynthesis and there is a continuous flux of CH_4 into the atmosphere due to biological methanogenesis and anthropogenic methane production. Similarly, the N_2O in Earth's atmosphere is continually replenished by a large range of microorganisms via incomplete denitrification. On Earth, these biological processes lead to CH_4 and N_2O abundances that are many orders of magnitude larger than the chemical equilibrium. Another interesting ratio of atmospheric gases to consider is the ratio between CO and CH_4. A large amount of CO accompanied by a lack of significant CH_4 could be interpreted as an antibiosignature as suggested in Zahnle et al. (2008). A more exhaustive discussion of potential biosignatures can be found in Schwieterman et al. (2018), for example.

The retrieved P–T structure is displayed in Fig. 4b. Our retrieval framework extracts the P–T structure of Earth's lower atmosphere accurately to very high precision. With decreasing pressure, the uncertainty on the retrieved P–T profile increases due to a lack of signatures from the low-pressure atmospheric layers ($\lesssim 10^{-4}$ bar) in the exoplanet's emission spectrum.

4. Results

In the following, we analyze the performance of the retrieval suite in characterizing an Earth-twin planet orbiting a Sun-like star at 10 pc from our Earth. We first estimate the fundamental detection limits of our retrieval suite for the trace gas abundances for different input spectra, as described in Sect. 4.1. We then introduce the full grid of retrievals that was run in Sect. 4.2. The results are presented in Sects. 4.3 and 4.4.

4.1. Detection limit analysis

We define the fundamental detection limits of our retrieval suite as the lowest possible molecular abundances that are retrievable for the atmospheric gases considered. For abundances below the detection limit, the corresponding spectral features are lost in the observational noise. Specifically, we estimate the detection limits of the trace gases considered in our retrievals at different values for R (20, 50, 1000), S/N (10, 20, 50), and for the largest and smallest assumed bandwidth (3–20 μm, 6–17 μm).

To do so, we generated spectra from the parameter values listed in Table 3, but set the abundance of one trace gas to 0 for each generated spectrum. We then used these spectra as retrieval input, assumed only photon noise of the planet to be present, and retrieved for the absent trace gas. We passed all other model parameters to the retrieval as known.

By construction, the retrieval routine should rule out any abundance of the missing gas down to the detection limit abundance, where the spectral features are no longer distinguishable from the noise. The posterior distribution should therefore be a flat distribution (all values are equally probable) for abundances smaller than the threshold value. For values greater than the detection limit, the probability should be close to null. The threshold value can therefore be interpreted as a detection limit for the gas abundance.

We approximated the retrieved posterior for the trace gas abundance as a logistic function (a soft step function):

$$f(x) = \frac{c}{1 + e^{a \cdot x + b}}. \tag{7}$$

Here, x is the abundance of the trace gas for which the detectability is being considered. The constants a, b, and c are unique for each posterior. In Table 4, we provide the abundances corresponding to the half maximum and the 16th and 84th percentile of the fitted logistic function for all tests we ran. The half maximum together with the percentiles provides an estimate for the detection limit of our retrieval suite.

The results for the $R = 1000$ test case predict that all atmospheric trace gases should be detectable in such a retrieval. This is in agreement with the validation results presented in Sect. 3. For the cases where $R \leq 50$, we predict H_2O, CO_2, and O_3 to be easily detectable in an Earth-like atmosphere, since the true abundances are more than one order of magnitude larger than the retrieval's estimated detection limit. For CH_4, the true abundance is comparable to the retrieval's detection limit. Thus, we expect the performance for CH_4 to depend strongly on the R and S/N of the input spectrum. The true CO and N_2O abundances lie at least one order of magnitude below the estimated detection limits and are therefore irretrievable. The upper limit of CO exhibits a strong dependence on the wavelength range considered because the only CO feature in the MIR is located at ~4.7 μm (see Fig. 3). Excluding the 3–6 μm wavelength range from the analysis makes the CO abundance impossible to constrain.

Table 4. Retrieval sensitivity analysis results.

R	S/N	Range [μm]	$\log_{10}(H_2O)$	$\log_{10}(CO_2)$	$\log_{10}(O_3)$	$\log_{10}(CH_4)$	$\log_{10}(CO)$	$\log_{10}(N_2O)$
20	10	3–20	$-6.13^{\pm 0.38}$	$-6.51^{\pm 0.33}$	$-7.22^{\pm 0.38}$	$-6.04^{\pm 0.51}$	$-4.14^{\pm 1.32}$	$-3.39^{\pm 1.23}$
		6–17	$-6.11^{\pm 0.54}$	$-6.61^{\pm 0.52}$	$-7.04^{\pm 0.30}$	$-5.91^{\pm 0.57}$	UC	$-3.09^{\pm 1.04}$
	20	3–20	$-6.68^{\pm 0.45}$	$-7.08^{\pm 0.48}$	$-7.55^{\pm 0.27}$	$-6.68^{\pm 0.53}$	$-5.28^{\pm 0.63}$	$-4.55^{\pm 1.05}$
		6–17	$-6.63^{\pm 0.40}$	$-7.05^{\pm 0.44}$	$-7.47^{\pm 0.37}$	$-6.67^{\pm 0.56}$	UC	$-4.30^{\pm 1.02}$
50	10	3–20	$-6.60^{\pm 0.49}$	$-7.08^{\pm 0.52}$	$-7.45^{\pm 0.31}$	$-6.49^{\pm 0.42}$	$-5.26^{\pm 0.67}$	$-4.33^{\pm 1.07}$
		6–17	$-6.47^{\pm 0.45}$	$-7.05^{\pm 0.46}$	$-7.46^{\pm 0.28}$	$-6.56^{\pm 0.55}$	UC	$-3.86^{\pm 0.98}$
	20	3–20	$-6.99^{\pm 0.34}$	$-7.48^{\pm 0.34}$	$-7.75^{\pm 0.22}$	$-7.09^{\pm 0.45}$	$-5.82^{\pm 0.36}$	$-5.28^{\pm 0.75}$
		6–17	$-6.98^{\pm 0.34}$	$-7.50^{\pm 0.41}$	$-7.84^{\pm 0.36}$	$-7.08^{\pm 0.33}$	UC	$-5.10^{\pm 0.87}$
1000	50	3–20	$-9.29^{\pm 0.25}$	$-9.27^{\pm 0.27}$	$-9.02^{\pm 0.27}$	$-8.81^{\pm 0.29}$	$-7.88^{\pm 0.29}$	$-8.35^{\pm 0.35}$
Terrestrial abundances:			−3.00	−3.39	−5.77	−6.52	−6.90	−6.50
LIFE retrieval expectation:			✓	✓	✓	~	×	×

Notes. All abundances are given in mass fractions in $\log_{10}$. We provide the threshold abundances for a detection at different R and S/N configurations, which we define as the half maximum of the retrieved soft step posterior distribution. Additionally, we provide the 16th and 84th percentile of the logistic function as denoted by the ±. A UC indicates that the retrieval does not manage to constrain the species of interest. In the last two rows, we display the assumed Earth-twin abundances and whether we expect such an abundance to be retrievable in low R and S/N cases, respectively: ✓ = retrievable; ~ = case dependant; × = not retrievable.

This test study has provided best case detection limits for the abundance in Earth-like atmospheres. However, retrieval of all parameters will lead to an overall increase in these detection limits. Additionally, adding additional astrophysical noise terms will also negatively impact the detection limits for the trace gases.

4.2. Retrieval grid

We chose to consider the following grid of wavelength ranges, R values, and S/N values in our final grid of retrieval studies:

- Wavelength ranges: 3–20 μm, 4–18.5 μm, and 6–17 μm.
- Spectral resolutions: $R = 20$, 35, 50, and 100.
- Signal-to-noise ratios: $S/N = 5$, 10, 15, and 20 fixed at the wavelength bin centered at 11.2 μm.

The short end of the wavelength range tests, whether the CO band at ~4.7 μm can be retrieved and whether including the ~3.3 μm band of CH_4 helps with the detection of this molecule. At the long wavelength end, the main question is how much of the extended H_2O feature should be included in the analysis. The choice for the spectral resolution range was motivated by earlier studies (e.g., Des Marais et al. 2002).

For the reference planet, we again assumed the thermal emission spectrum of an Earth-twin in orbit around a G2V star located at a distance of 10 pc from Earth. We considered two observational cases with slightly different noise properties and observational setups:

1. Nominal case: (1) the LIFE baselines (physical distance between the four mirrors) were optimized for the detection of habitable zone planets at a wavelength of $\lambda = 15$ μm (cf. Quanz et al. 2022); (2) the level of exozodiacal dust emission corresponded to three times the level of the local zodiacal light.
2. Optimized case: (1) the LIFE baselines were optimized for the detection of habitable zone planets at the short wavelength end; (2) the level of exozodiacal dust emission corresponded to 0.5 times the level of the local zodiacal light[6].

[6] According to the brightness distribution of exozodi disks used in recent LIFE detection yield estimate (Quanz et al. 2022), which is based on the results from the HOSTS survey (Ertel et al. 2020), ≲20% of the disks show such low emission.

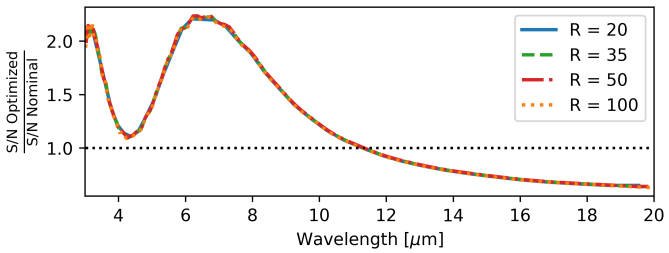


Fig. 5. Ratio between the wavelength-dependent S/N of the optimized case and the nominal case. This ratio is independent of the overall S/N and the R of the Spectrum.

Figure 5 visualizes the difference between the nominal and the optimized S/N case by plotting the ratio between the two S/N instances. We computed the full noise terms (including stellar leakage, local zodi and exozodi emission, and photon noise from the planet) using LIFEsim (see Dannert et al. 2022). In total, the grid comprised 96 retrieval analyses. Figure 6 visualizes the highest ($R = 100$, $S/N = 20$) and lowest ($R = 20$, $S/N = 5$) quality input spectra for the nominal case. For every grid point specified above, we ran one retrieval assuming the prior distributions listed in Table 3. Furthermore, we used the same `pyMultiNest` settings applied in the retrieval validation run (Sect. 3).

Taking the 3–20 μm nominal case input spectra as an example, we plot the posteriors of the planetary parameters (Fig. 7), as well as those for the absolute (Fig. 8) and the relative (Fig. 9) abundances. We use the retrieved posterior probability distributions to underline trends with respect to the wavelength range, R, and S/N. For the molecular abundances, we differentiate between four different classes of posterior distributions. First, a constrained (C-type) posterior signifies that the true atmospheric abundance lies above the retrieval's detection limit. A C-type posterior is well approximated by a Gaussian. Thus, both significantly higher and lower abundances are ruled out. Second, for a sensitivity limit (SL-type) posterior, the true atmospheric abundance is comparable to the detection limit of the retrieval. We observe a distinct peak in the posterior. However, lower abundances cannot be fully ruled out. A SL-type posterior is best described by the convolution of a logistic function with a Gaussian. Third, a upper limit (UL-type) posterior indicates

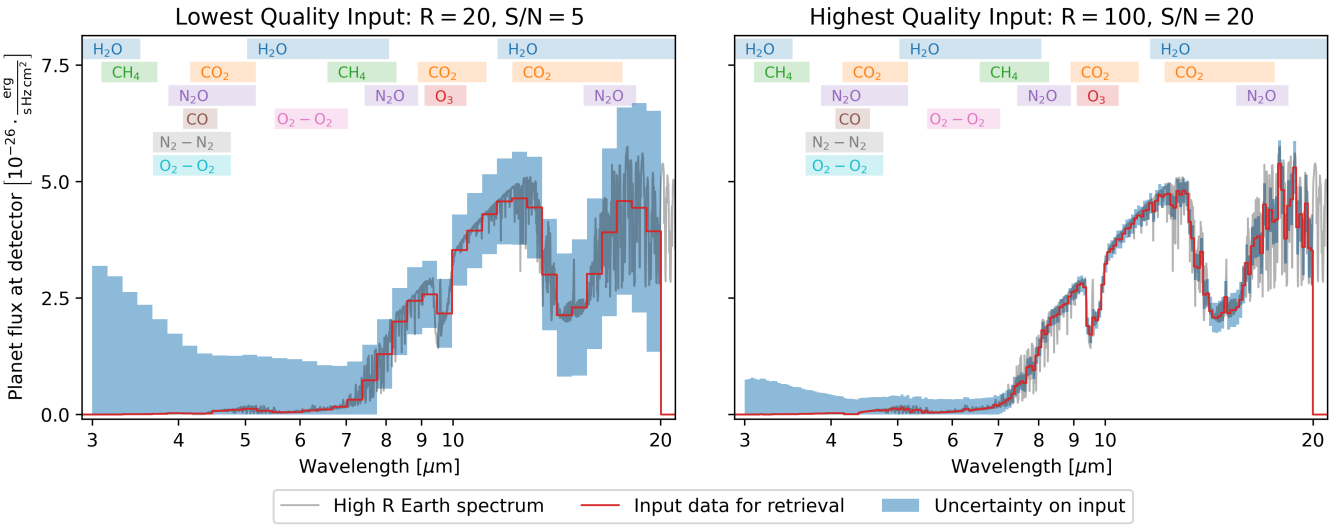


Fig. 6. Examples of input spectra used in the grid retrievals for the nominal case (*left*: the lowest quality input with $R = 20$, $S/N = 5$; *right*: the highest quality input with $R = 100$, $S/N = 20$). In gray, we provide the full resolution `petitRADTRANS` Earth spectrum. The red step function represents the wavelength-binning of the input data. Further, the blue shaded region represents the uncertainty for the corresponding bin. We also mark the absorption features of the considered atmospheric gases.

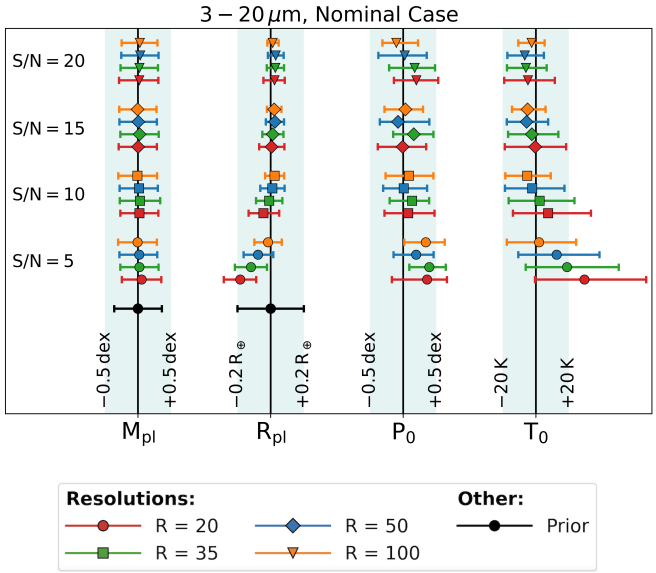


Fig. 7. Retrieved exoplanet parameters for the different grid points for the wavelength range 3–20 μm in the nominal case. Here, $M_{\rm pl}$ is the mass, $R_{\rm pl}$ the radius, P_0 the surface pressure, and T_0 the surface temperature of the exoplanet. The error bars denote the 68% confidence intervals. For $M_{\rm pl}$ and $R_{\rm pl}$, we also plot the assumed prior distributions. For T_0 and P_0, we assumed flat, broad priors. The vertical lines mark the true parameter values.

that the true atmospheric abundance lies below the retrieval's detection limit and could be zero. A UL-type posterior is best described by the logistic function introduced in Sect. 4.1. Abundances above the detection limit can be excluded, while all abundances below the limit are equally likely. Fourth, an unconstrained (UC-type) posterior signifies that no information about the atmospheric abundance can be retrieved. The posterior is best described by a constant function. For more detailed information about our posterior classification, see Appendix B.

Figure 10 summarizes the type of the retrieved posterior distribution for the molecular abundances for all considered *R*, *S/N*, wavelength ranges, and the two different cases. Tables containing the median as well as the 16th and 84th percentile of the retrieved posteriors for all retrieval runs are provided in the Appendix D (Tables D.1–D.6).

4.3. Retrieved planetary parameters

In Fig. 7, we provide the retrieval results for the exoplanet parameters $M_{\rm pl}$, $R_{\rm pl}$, P_0, and T_0 (3–20 μm input spectra in the nominal case). Our retrieval framework estimates all planetary parameters correctly.

For the planetary mass $M_{\rm pl}$, the retrieved posterior is centered on the true value and roughly corresponds to the assumed prior distribution. This result indicates that our retrieval framework cannot extract further information from the input spectrum. In contrast to $M_{\rm pl}$, we manage to strongly constrain the exoplanet's radius $R_{\rm pl}$ with respect to the assumed prior distribution. For all $S/N \geq 10$, we retrieve an accurate estimate for $R_{\rm pl}$ ($\Delta R_{\rm pl} \leq \pm 10\%$).

For all $S/N \geq 10$, our retrievals yield strong constraints for the surface pressure P_0. The retrieved value lies within maximally ± 0.5 dex (a factor of ≈ 3) depending on the *R* and *S/N* combination. Similarly, the surface temperature T_0 (calculated from the retrieved *P*–*T* profile parameters and P_0) is accurately estimated by our retrieval framework for all $S/N \geq 10$. The retrieved values are centered on the correct value for most cases and the $1\,\sigma$ uncertainties are mostly smaller than $\pm$ 20 K.

For an *S/N* of 5, the retrieved parameters T_0 and P_0 exhibit significant offsets with respect to the input value. Furthermore, we find similar offsets for $R_{\rm pl}$. These deviations indicate that an *S/N* of 5 is too low for the accurate characterization of an Earth-twin exoplanet's atmospheric structure, as we discuss further in Sect. 4.4. The same is true for all other wavelength ranges and both the nominal and the optimized cases considered in our work (see Appendix D).

4.4. Retrieved abundances

The retrieved molecular abundances of the atmospheric trace gases are provided in Fig. 8 (3–20 μm input spectra for the

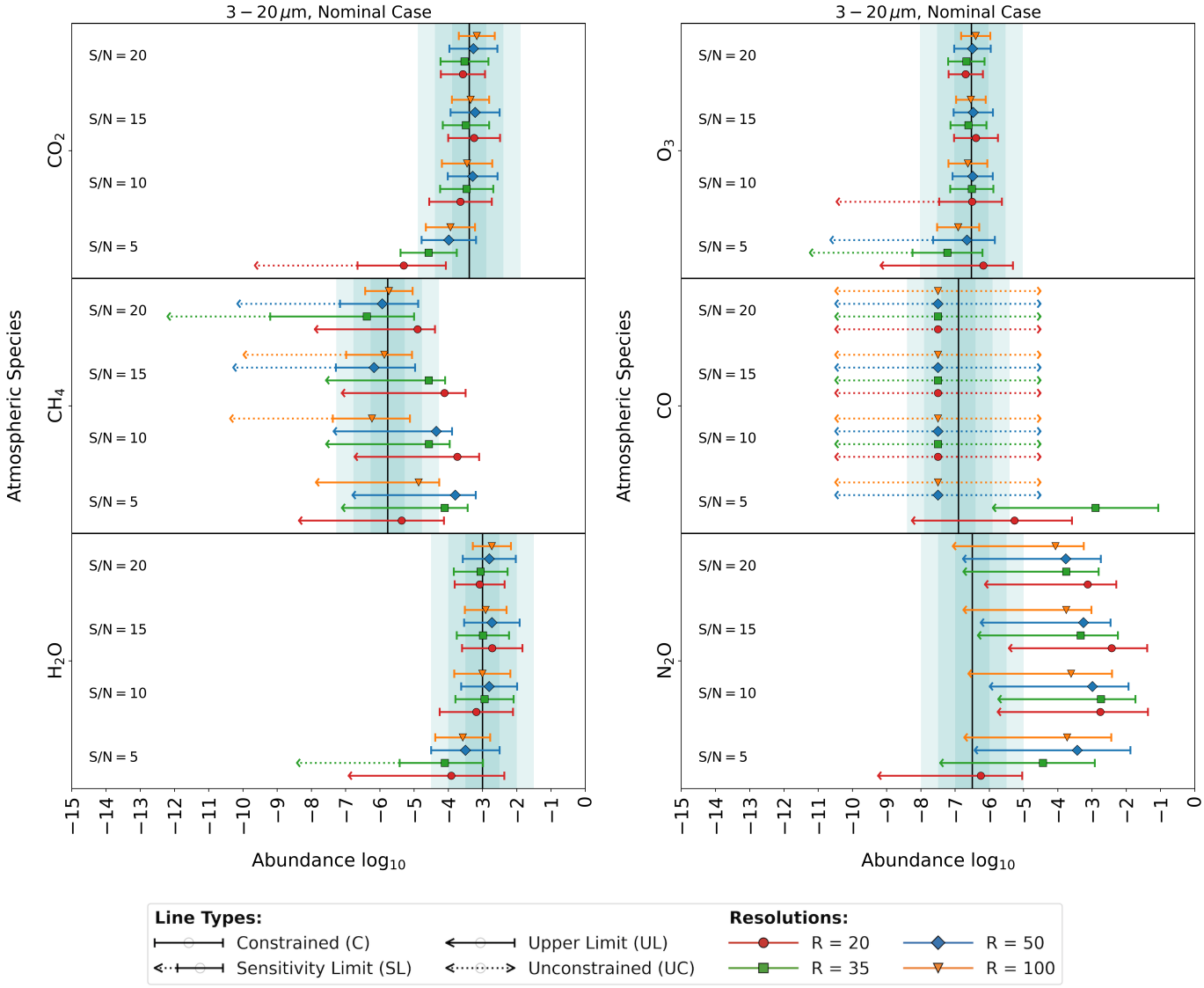


Fig. 8. Retrieved mass mixing ratios of the different trace gases present in Earth's atmosphere for an input spectrum wavelength range of 3–20 μm in the nominal case. The vertical lines mark the true abundances whereas the shaded regions mark the ±0.5 dex, ±1.0 dex, and ±1.5 dex regions.

nominal case). We observe that the abundances of CO_2, H_2O, and O_3 are accurately retrieved to a precision of $\lesssim \pm 1$ dex for all cases where the *S/N* is ≥ 10. In contrast, the molecule CH_4 only becomes retrievable for high combinations of *R* and *S/N* (for $S/N = 10$ at $R \geq 100$, for $S/N = 15$ at $R \geq 50$, and for $S/N = 20$ a $R \geq 35$). For other combinations of *R* and *S/N*, we only retrieve upper limits on the CH_4 abundance.

The CO abundance is not constrained by our retrievals for any $S/N \geq 10$. At an *S/N* of 5 we retrieve an upper limit on the abundance of CO. This upper limit is a result of the poor overall retrieval performance in the $S/N = 5$ case as we motivate below. Similarly, we do not retrieve the N_2O abundance, but find an upper limit for the maximal possible abundance. The position of this upper limit decreases significantly with increasing *R* and *S/N* of the input spectrum. Further, the input spectra do not contain sufficient information to constrain the abundances of the bulk constituents N_2 and O_2 for all considered *R* and *S/N* combinations (not shown). We retrieve UC-type posteriors and cannot provide any constraint for N_2 and O_2 in our retrievals.

Finally, by adding up the retrieved abundances, we find that there is at least one additional atmospheric gas that has no MIR signature probable by LIFE, constitutes ≈ 99% of Earth's atmospheric mass, and has no significant absorption feature in the MIR. We can exclude H and He dominated atmospheres due to the small retrieved radius and CO_2, CH_4, or H_2O dominated atmospheres due to our retrieval results. These findings already put strong constraints on the bulk atmospheric composition.

Overall, the findings we presented agree well with the predictions obtained in the sensitivity analysis (Sect. 4.1). This indicates that the simplified assumptions we relied on for the retrieval sensitivity analysis are reasonable.

Input spectra with an *S/N* of 5 do not contain sufficient information to yield accurate retrieval predictions for the absolute abundances of the considered trace gases. This is in accordance with our findings for the planetary parameters. For CO_2, H_2O, and O_3, we tend to underestimate the true atmospheric abundances. Similarly, the upper limits on the abundances of N_2O retrieved at $S/N = 5$ are lower than at $S/N \geq 10$. For CO we retrieve upper limits, which are no longer found at higher *S/N*. The underestimation of abundances at $S/N = 5$ is coupled with the overestimation of the surface pressure P_0 and temperature T_0, which are again compensated by an underestimation of the

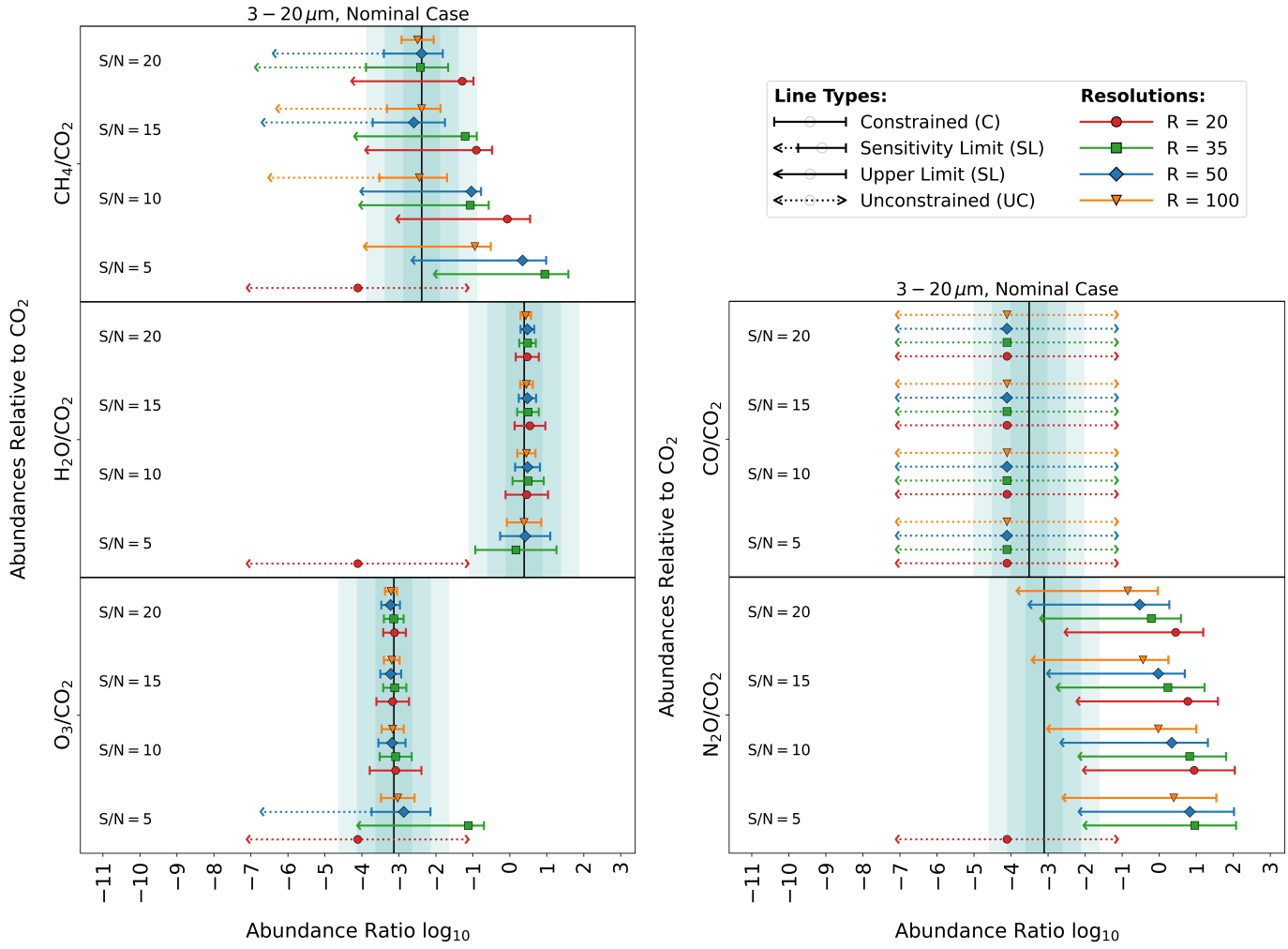


Fig. 9. As for Fig. 8 but for retrieved abundances relative to CO_2.

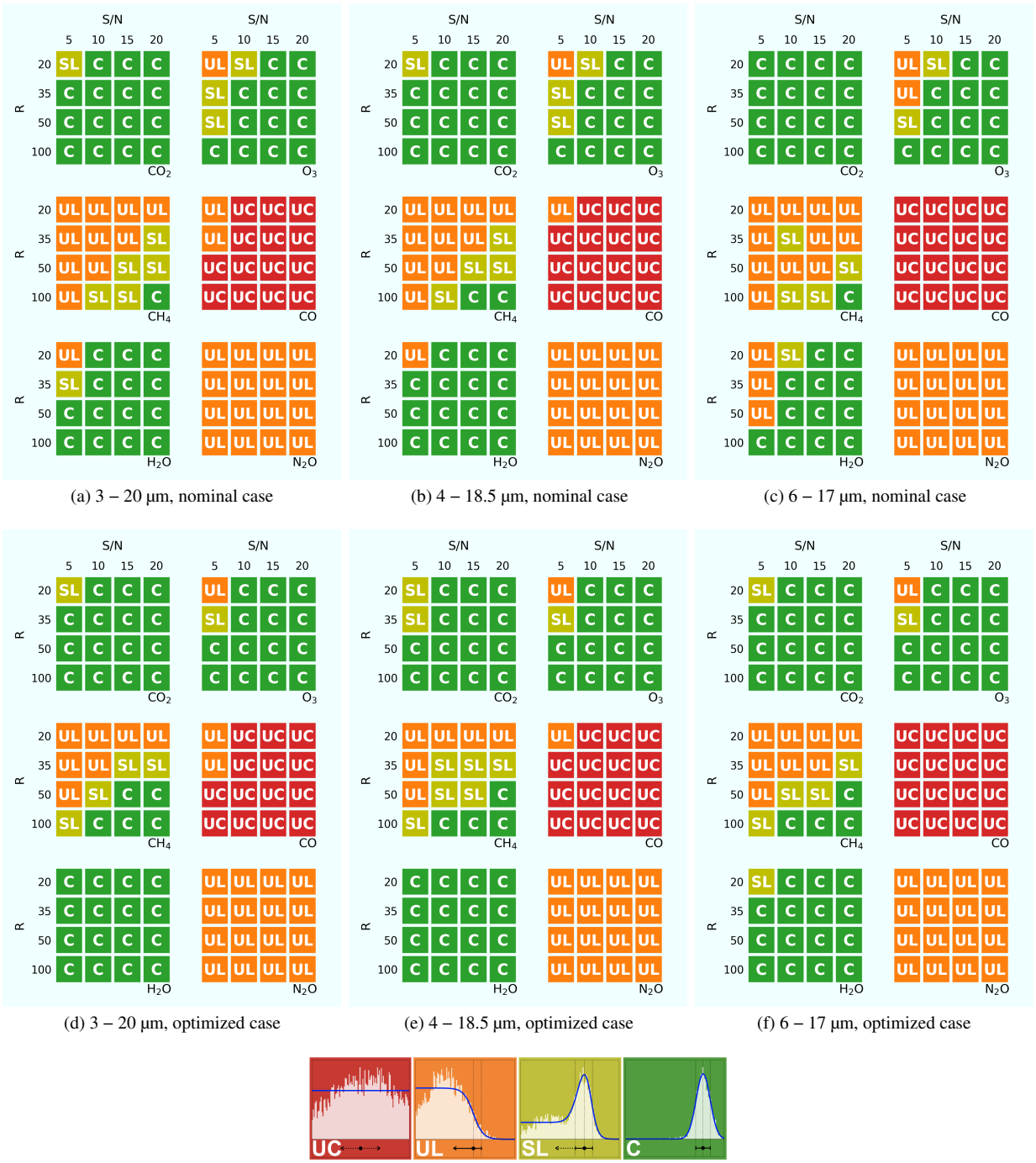
radius R_{pl}. A higher P_0 leads to a higher atmospheric mass and therefore to more absorbing material between the planet surface and the observer, which will lead to deeper absorption features at constant molecular abundances; the same line in a MIR emission spectrum would then be produced by a smaller abundance of the atmospheric species. Hence, in this case the retrieved abundances lie below the true ones. By considering the relative abundances we can reduce these offsets (Fig. 9 shows the trace gas posteriors relative to the CO_2 posterior for the nominal case, 3–20 μm input spectrum). This indicates that the offsets share the same properties for all atmospheric gases, indicating that they are caused by degeneracies between parameters (pressure-abundance and gravity-abundance degeneracies). These degeneracies are larger in the small *S/N* and *R* cases, since the constraints posed by the input spectrum are smaller. We further observe that the true values still lie within the posterior range (note that the plots only show the 16th, 50th, and 84th percentiles). If the degeneracies are asymmetric (e.g., lower surface pressures and temperatures are easier to exclude than high ones), this would result in the observed asymmetric positioning of the posterior distribution around the truth, causing the observed offsets. The offsets diminish at higher *R* and *S/N* because the retrieval input provides stronger constraints and thus manages to reduce/break these degeneracies.

For $S/N \geq 10$, the relative abundances of H_2O, O_3, and CH_4 (if retrieved) are centered on the true values and the corresponding uncertainties are significantly smaller than for the absolute abundances ($\leq \pm 0.5$ dex). The reduction in uncertainty is due to the elimination of the gravity-abundance degeneracy, since this degeneracy affects all retrieved abundances comparably. Likewise, at $S/N = 5$, the offsets we observed for the retrieved absolute abundances are strongly diminished when considering relative abundances. This allows us to find accurate estimates for the relative abundances of CO_2, H_2O, and O_3 despite the inaccurately retrieved absolute abundances.

These findings demonstrate that considering relative abundances can significantly diminish the effects of degeneracies between trace gas abundances and other atmospheric parameters. This occurs at the cost of losing information on the absolute abundances. However, the relative abundances of trace gases can still contain vital information on planetary conditions and provide potential biosignatures (see Sect. 3 and, e.g., Lovelock 1965; Lippincott et al. 1967; Meadows et al. 2018).

5. Discussion

After providing an overview of the results we obtained for the nominal case (with an input spectrum covering 3–20 μm), we now compare the retrieval results for different wavelength ranges in the nominal and optimized case. Thereby, we investigate the minimal requirements that the LIFE mission needs to fulfill to



(a) 3 – 20 µm, nominal case

(b) 4 – 18.5 µm, nominal case

(c) 6 – 17 µm, nominal case

(d) 3 – 20 µm, optimized case

(e) 4 – 18.5 µm, optimized case

(f) 6 – 17 µm, optimized case

Fig. 10. Wavelength-dependent posterior types retrieved for the different trace gases in the nominal case (*a*)–(*c*) and the optimized case (*d*)–(*f*). The lowermost panel gives the color coding for the different posterior types. The abbreviations used for the different posteriors are introduced in Sect. 4.2. (*a, d*): 3–20 µm, (*b, e*): 4–18.5 µm, (*c, f*): 6–17 µm.

characterize the atmospheric structure of an Earth-twin exoplanet and detect biosignature gases in the emission spectrum. In Sects. 5.1 and 5.2 we derive requirements for the wavelength range, the R, and the S/N. We then quantify the integration times required to reach the desired S/N for a given R, conservatively assuming a total instrument throughput of 5% (cf. Dannert et al. 2022) and discuss the implications for the LIFE design in Sect. 5.3. In Sect. 5.4, we compare our work to similar retrieval studies in the literature to underline the unique scientific potential of LIFE for atmospheric characterization. The limitations inherent to our work and possible future investigations follow in Sect. 5.5.

5.1. Wavelength range requirement

Figure 10 provides a concise summary of the retrieved posterior types for the trace gas abundances for varying wavelength range, R, and S/N for both the nominal and optimized case. Here, we are primarily interested in large systematic differences in the overall retrieval performance between the different wavelength ranges. Small differences (e.g., in the nominal case for H_2O at $R = 35$, $S/N = 5$ or for CH_4 at $R = 35$, $S/N = 10$) are expected to disappear when averaging over multiple retrieval runs. In the following discussion, we focus on the results obtained for the nominal case (first row in Fig. 10).

We observe that for the trace gases CO_2, O_3, CO, and N_2O, considering a broader wavelength range input spectrum will result in slightly smaller uncertainties on the retrieved parameters. This is expected, since more information is passed to the retrieval framework. However, there is no significant difference in performance between the considered wavelength ranges.

For the trace gases H_2O and CH_4 the 6–17 µm case exhibits a considerably lower sensitivity than the 3–20 µm and 4–18.5 µm cases. An upper limit at 17 µm excludes the strong H_2O absorption features at >17 µm (see Fig. 3). The H_2O lines between 8 and 17 µm are either weak or overlap with stronger absorption features of other molecules. Therefore, at low *R* and *S/N*, these features cannot constrain the H_2O abundance satisfactorily. Below ≈ 8 µm, we observe a strong H_2O band that overlaps with the CH_4 feature at 7.7 µm. In addition, at these wavelengths the *S/N* in the simulated spectra decreases drastically, further reducing the constraints on H_2O. Thus, when considering a wavelength range from 6–17 µm, both the H_2O and CH_4 abundance are estimated via their overlapping absorption feature at ≈ 8 µm. At low *R* and *S/N* this leads to larger uncertainties. If we consider a larger wavelength range, the long-wavelength tail of the emission spectrum provides more robust constraints on the H_2O abundance, which directly leads to an improvement for the CH_4 estimates as well. We therefore cannot recommend limiting the wavelength to 6–17 µm for LIFE due to its negative impact on the sensitivity for H_2O and CH_4.

On the other hand, choosing the largest wavelength range considered in our study (3–20 µm) provides only a negligible improvement over the 4–18.5 µm wavelength range. From this, we conclude that LIFE should opt for a long wavelength cutoff of at least 18.5 µm. For the short wavelength limit, we suggest a value of 4 µm. This boundary would include the CO feature at 4.67 µm, as well as the N_2 CIA line at 4.3 µm. In the Earth-twin retrievals we present in this work, the abundances of CO and N_2 are not constrainable for any of the considered cases that include these spectral features. However, making these spectral features accessible to LIFE would enable us to potentially constrain the abundances of these important molecules in non-Earth-twin atmospheres. More tests to explore this idea are foreseen for future retrieval studies.

For the retrievals performed for the optimized case, we reach similar conclusions and our recommended wavelength range remains unchanged. However, the interfering effects described above are less pronounced due to an improved *S/N* at short wavelengths (see Fig. 10 and the tables in Appendix D).

5.2. R and S/N requirements

For most trace gases, we observe no significant difference between the nominal and optimized case for all $S/N \geq 10$ as can be seen from Fig. 10. For an *S/N* of 5, the optimized case yields better results for the retrieved abundances. However, due to the generally poor performance at this noise level (as previously seen), an *S/N* of 5 is not sufficient to characterize the atmospheric structure and composition of an Earth-twin exoplanet satisfactorily.

Generally, we observe that CO_2, H_2O, and O_3 are easily retrievable (within ± 1 dex) for an Earth-like atmosphere for all $S/N \geq 10$. This finding is in accordance with the results presented in Cockell et al. (2009). However, other studies suggest that for a clear O_3 detection a higher *R* or *S/N* are necessary (see, e.g., von Paris et al. 2013; Léger et al. 2019). In contrast, CO is not recoverable for any of the considered cases due to the large astrophysical noise at short wavelengths, which indicates that detecting CO from the MIR emission spectrum of an Earth-twin exoplanet around a Sun-like star is extremely challenging and would require very high *S/N* and higher spectral resolution. Similarly, the N_2O abundance present in Earth's atmosphere is too low to be detected in all cases considered. However, we retrieve upper abundance limits, which indicates that high atmospheric concentrations of N_2O ($\gtrsim 10^{-3}$) would likely be detectable.

In contrast, the retrieval results for CH_4 depend strongly on the considered *R* and *S/N* for both the nominal and optimized case (see Fig. 10). Generally, our retrieval results for CH_4 improve as we consider higher *R* and *S/N*. For both baseline configurations, we retrieve a threshold above which the retrieval framework manages to accurately estimate the CH_4 abundance. This threshold manifests itself as a diagonal in the *R*–*S/N* space, which indicates that, for instance, a lower *S/N* can be compensated with a higher *R* without impacting the retrieval results significantly. The main difference is that *R* is intrinsic to the LIFE design (dependent on the spectrograph specifications), whereas the *S/N* depends on the design (aperture size) and the integration time.

Generally, we find that in the optimized case the accuracy of the retrieval for CH_4 is improved. First, LIFE is configured to optimize the signal at short wavelengths, which directly leads to an improvement in the *S/N* of the ~7.7 µm CH_4 feature. Second, the reduced noise contribution from exozodiacal dust also improves the *S/N* at short wavelengths. The resulting *S/N* enhancement at short wavelengths can be seen in Fig. 5. These two factors lead to an increase in the retrieval's overall performance.

We observe that, in the optimized case for the 3–20 µm range, CH_4 is detectable (at least an SL-type posterior) in an Earth-twin atmosphere if $S/N \geq 15$ for $R = 35$, $S/N \geq 10$ for $R = 50$, or $S/N \geq 5$ for $R = 100$. We further find that $R = 20$ is too low to allow for a meaningful constraint on the CH_4 abundance for all considered cases. But if for technical reasons one would prefer to keep *R* as low as possible, we are left with $R = 35$ and $R = 50$. The higher resolution case potentially allows for the detection of a C-type posterior when going to higher *S/N*. In the nominal case, the *S/N* has to be increased to obtain comparable results at the same *R* ($S/N = 20$ for $R = 35$, or $S/N = 15$ for $R = 50$).

Since the detection of CH_4 depends strongly on the combination of *R* and *S/N*, it is important to evaluate the significance of the CH_4 detection in the cases $R = \{35, 50\}$ and $S/N = \{10, 15, 20\}$. For every *R*–*S/N* pair, we ran retrievals with and without CH_4 being included in the forward model. We then compared the log-evidences corresponding to the retrievals for both scenarios via the Bayes' factor *K*. The results are summarized in Table 5.

The Bayes' factor *K* is given by the ratio between the Bayesian evidence of the model that considers CH_4 and the evidence of the methane-free model. It provides a measure for which model is better at describing the observed data (see Table 1). A positive $\log_{10}(K)$ indicates that the model including CH_4 (model evidence is denoted $\mathcal{Z}_{CH_4}$) describes the observed emission spectrum better than the model without CH_4 (model evidence is denoted $\mathcal{Z}_{-CH_4}$). In contrast, negative values favor the CH_4-free atmospheric model.

For all cases where $R = 35$, the Bayes' factor $\log_{10}(K) \approx 0.0$ (within 1-σ) indicating that there is no difference between the models with and without CH_4. Therefore, despite retrieving SL-type posteriors for $S/N \geq 15$, the retrieval framework does not provide unambiguous evidence that CH_4 is indeed present in the observed atmosphere. This finding underlines the nature of

Table 5. Log-evidence for retrievals with and without CH_4.

	S/N		10	15	20
Nominal Case	$R = 35$	$\ln(\mathcal{Z}_{CH_4})$	$-23.6^{\pm 0.2}$	$-27.0^{\pm 0.2}$	$-28.8^{\pm 0.2}$
		$\ln(\mathcal{Z}_{-CH_4})$	$-23.7^{\pm 0.2}$	$-27.2^{\pm 0.2}$	$-28.8^{\pm 0.2}$
		$\log_{10}(K)$	$0.0^{\pm 0.1}$	$0.1^{\pm 0.1}$	$0.0^{\pm 0.1}$
	$R = 50$	$\ln(\mathcal{Z}_{CH_4})$	$-25.1^{\pm 0.2}$	$-27.8^{\pm 0.2}$	$-29.8^{\pm 0.2}$
		$\ln(\mathcal{Z}_{-CH_4})$	$-25.4^{\pm 0.2}$	$-28.4^{\pm 0.2}$	$-30.6^{\pm 0.2}$
		$\log_{10}(K)$	$0.1^{\pm 0.1}$	$0.3^{\pm 0.1}$	$0.3^{\pm 0.1}$
Optimized Case	$R = 35$	$\ln(\mathcal{Z}_{CH_4})$	$-24.8^{\pm 0.2}$	$-27.9^{\pm 0.2}$	$-30.2^{\pm 0.2}$
		$\ln(\mathcal{Z}_{-CH_4})$	$-24.9^{\pm 0.2}$	$-28.2^{\pm 0.2}$	$-30.1^{\pm 0.2}$
		$\log_{10}(K)$	$0.0^{\pm 0.1}$	$0.1^{\pm 0.1}$	$0.0^{\pm 0.1}$
	$R = 50$	$\ln(\mathcal{Z}_{CH_4})$	$-26.7^{\pm 0.2}$	$-29.6^{\pm 0.2}$	$-31.5^{\pm 0.2}$
		$\ln(\mathcal{Z}_{-CH_4})$	$-28.8^{\pm 0.2}$	$-31.7^{\pm 0.2}$	$-35.1^{\pm 0.2}$
		$\log_{10}(K)$	$0.8^{\pm 0.1}$	$0.9^{\pm 0.1}$	$1.6^{\pm 0.1}$

Notes. With CH_4: $\ln(\mathcal{Z}_{CH_4})$; Without CH_4: $\ln(\mathcal{Z}_{-CH_4})$. The results are for input spectra covering the 3–20 µm wavelength range. The last row for a given R lists $\log_{10}(K)$ of the Bayes' factor K, which we calculate via: $\log_{10}(K) = \left(\ln\left(\mathcal{Z}_{CH_4}\right) - \ln\left(\mathcal{Z}_{-CH_4}\right)\right) / \ln(10)$ (derived from Eq. (5)). We compare the performance of the models with and without CH_4 to each other using Jeffrey's scale (Jeffreys 1998) (see Table 1). The $\pm$ indicates the 68% confidence interval $\log_{10}(K)$.

Table 6. Required observation time in days.

		Observation Time [days]					
		Nominal Case			Optimized Case		
R	S/N	1 m	2 m	3.5 m	1 m	2 m	3.5 m
35	15	771	74	15	922	64	9
50	10	490	47	10	589	41	6

Notes. Observation times for an Earth-twin at 10 pc for the two combinations of R and S/N discussed in Sect. 5.2 and three different aperture diameters (1 m, 2 m, and 3.5 m). We conservatively assume a total instrument throughput of 5% (cf. Dannert et al. 2022).

SL-type posteriors, where the long tail toward low abundances indicates that the retrieval can do without CH_4.

For $R = 50$, we generally observe larger $\log_{10}(K)$. In the nominal case, we find weak support for the model including CH_4. In the optimized case, we find substantial to strong evidence for the presence of CH_4 as is indicated by $0.5 < \log_{10}(K) < 2.0$. These findings suggest that LIFE requires a minimal R of 50 to confidently rule out the CH_4-free atmospheric model in retrievals of Earth-twin MIR emission spectra.

5.3. Observation time estimates

It is crucial to derive estimates for the integration time required to reach a certain S/N. However, the integration time does not only depend on R and S/N, but also on the aperture diameter of the LIFE collector spacecraft and the instrument throughput. Hence, our analyses can provide first order requirements for the aperture size. We summarize the integration time estimates for the above-mentioned combinations of R and S/N in Table 6. These estimates are based on the conservative assumption of a total instrument throughput of 5% (cf. Dannert et al. 2022). We give a more exhaustive list containing observation time estimates for all retrieved input spectra in Appendix E.

We find that with four 1 m apertures LIFE will not be able to characterize terrestrial exoplanet atmospheres at a distance of 10 pc because unrealistic integration times > 1 yr would be required. While Bryson et al. (2021) used the results from NASA's *Kepler* mission to estimate with 95% confidence that the nearest terrestrial exoplanet orbiting with the habitable zone around a G or K dwarfs is only ≈6 pc away, they also estimate that there are only approximately four such objects within 10 pc. Hence, a 4×1 m configuration will not allow us to probe a somewhat sizable sample of temperate terrestrial exoplanets.

For the other two setups (4×2 m and 4×3.5 m), the observation times are more feasible and would allow for the characterization of up to a few tens of terrestrial exoplanet atmospheres within LIFE's characterization phase. Specifically, for both the nominal and optimized case and assuming the 4×2 m setup, having $R = 50$ and $S/N = 10$ would require less time (47 days for the nominal case, 41 days for the baseline optimized case) compared to $R = 35$ and $S/N = 15$ (74 and 64 days, respectively). This again underlines that a resolution of $R = 50$ is more suitable. The saved integration time could be used to characterize the atmospheres of additional exoplanets or to measure higher S/N spectra for the most promising objects. Similar conclusions hold for the 4×3.5 m case. However, the required observation times required would be significantly smaller.

5.4. Comparison to other studies

In order to demonstrate the validity and understand the full implications of our retrieval results, we compare the results to previous studies in the literature. In Sect. 5.4.1, we compare our work to previously published work on MIR thermal emission retrievals. In Sect. 5.4.2, we consider findings from a comparable retrieval study for the NIR-VIS wavelength range and demonstrate the unique scientific potential of MIR observations with LIFE.

5.4.1. MIR thermal emission studies

In Quanz et al. (2018), a similar retrieval study was performed for an Earth-twin exoplanet. The study assumed $R = 100$, $S/N = 20$, covered a larger wavelength range (3–30 µm), and considered only photon noise of the planet on the input spectrum, neglecting any additional noise terms. We compare their findings to our results for the 3–20 µm wavelength range, $R = 100$, $S/N = 20$ for the nominal case.

Our study reaches a comparable accuracy (± 0.5 dex) for the atmospheric trace gases CO_2, H_2O, O_3, and CH_4. This is achieved despite our usage of the more realistic LIFEsim noise model, which features additional noise sources that lead to considerably larger errors especially at short wavelengths ($\lambda \lesssim 8$ µm), where the LIFEsim noise is dominated by the contributions from stellar leakage. The performance of our retrieval suite is likely a result of our prior assumption for the exoplanet mass, which was not made in the Quanz et al. (2018) study. The width of the retrieved abundance posteriors is limited by the exoplanet's mass posterior due to the degeneracy between trace gas abundances and the surface gravity. The Gaussian prior we assumed for the exoplanet's mass limits the range of allowed masses. Thereby, also the surface gravity was constrained, limiting the range of possible abundances. However, we note that our mass and radius priors were informed by empirical measurements.

In contrast to Quanz et al. (2018), we do not succeed in constraining the CO and N_2O abundances in our retrievals. Both atmospheric trace gases have their main absorption features at wavelengths $\lambda \lesssim 8$ µm (see Fig. 3), where the observational

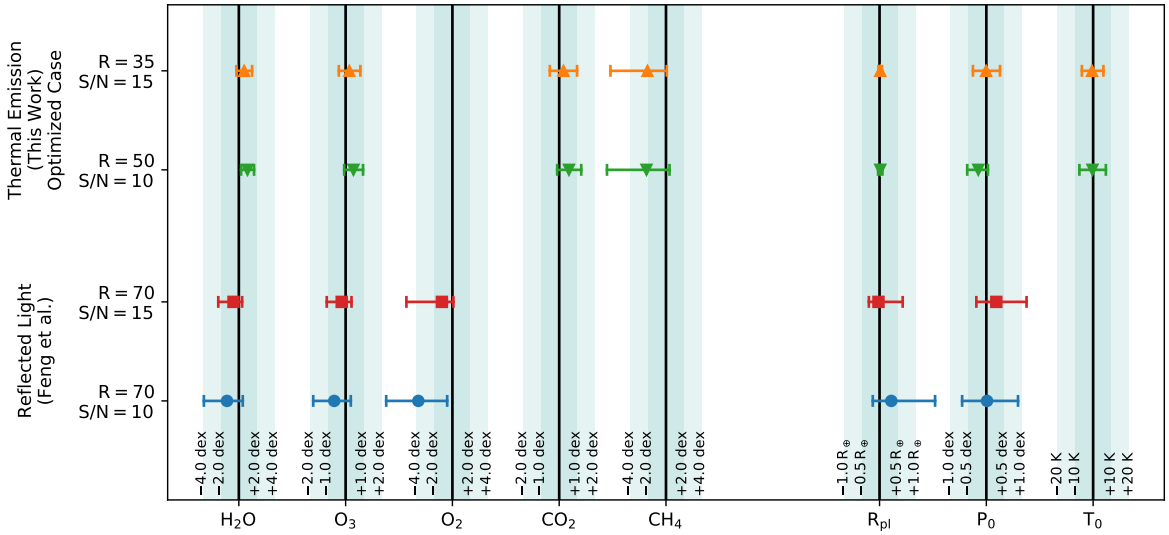


Fig. 11. Performance comparison between different retrieval studies. The error bars correspond to the 68% confidence intervals of the retrieved posterior distributions. The emitted light study from this work for ($R = 50$, $S/N = 10$ (green downward triangle) and $R = 35$, $S/N = 15$ (orange upward triangle), 3–20 μm in the optimized case); the reflected light study by Feng et al. (2018) ($R = 70$, $S/N = 10$ (blue dot), $R = 70$, $S/N = 15$ (red square), 0.4–1.0 μm).

LIFEsim noise dominates over the absorption features of interest. Furthermore, the retrievals performed in Quanz et al. (2018) find significantly stronger constraints for the shape of the P–T profile and the planetary parameters P_0 and T_0. This is likely a combined result of the more heavily constrained P–T profile model and the more optimistic noise estimates they used in their retrieval analysis.

Finally, the Earth-twin's radius is retrieved to an extremely high precision in both studies. This underlines the scientific potential of observations probing the thermal emission of planets in the MIR wavelength range. Determining the radius in NIR/VIS reflected light spectra is not robust (e.g., Feng et al. 2018) due to a degeneracy between the planet's albedo and its radius (a large surface area and low albedo can lead to the same flux as a smaller area and higher albedo).

A similar study, albeit for lower resolutions ($R = 5$, 20) and a less complete noise model, has been performed by von Paris et al. (2013) for the former *Darwin* mission concept. They found, that at $R = 20$, a MIR nulling interferometer would be capable of constraining the surface conditions (P_0 to ± 0.5 dex, T_0 to ± 10 K) and planetary parameters (R_{pl} to $\pm 10\%$, M_{pl} to ± 0.3 dex) of a cloud-free exoplanet, which agrees well with the results presented here. Further, they manage to retrieve comparable constraints for the abundances of CO_2 and O_3 at 1σ confidence levels. However, they also discuss, that for a 5σ detection of CO_2 and O_3, higher resolutions are mandatory. Similar results are presented in Léger et al. (2019), where a resolution limit of $R = 40$ is derived for a robust detection of CO_2, H_2O, and O_3 in the atmospheres of Earth-similar planets around M- and K-type stars.

5.4.2. NIR-VIS reflected light studies

In Feng et al. (2018), results from a similar retrieval study in the reflected light are presented (the input spectra cover the wavelength range 0.4–1.0 μm). These results are similar to the findings presented in Brandt & Spiegel (2014). We consider their results for the $R = 70$, $S/N = (10, 15)$ cases and compare them with our findings for the $R = 50$, $S/N = 10$ and $R = 35$, $S/N = 15$ optimized case. In Fig. 11, we show the results for all parameters of interest that are retrievable with at least one of the two approaches.

This comparison demonstrates that NIR-VIS and MIR wavelength studies yield partially complementary results. The MIR range enables us to search for signatures of the important trace gases CO_2 and CH_4. Both these gases are not accessible through reflected NIR-VIS light observations at Earth's mixing ratios. Additionally, studying the MIR thermal emission spectra of exoplanets enables us to pose significant constraints on the important planetary parameters T_0, P_0, and R_{pl}. These three parameters are not easily accessible via reflected light observations in the NIR-VIS wavelength range (Quanz et al. 2018; Feng et al. 2018; Carrión-González et al. 2020).

However, studies in the NIR-VIS wavelength range provide a direct probe for the O_2 abundance, whereas MIR observations can only probe the O_3 abundance, a photochemical byproduct of O_2 in our atmosphere. Additionally, the NIR-VIS wavelength range may allow us to characterize the surface composition of an exoplanet (e.g., Brandt & Spiegel 2014) via the wavelength-dependent surface scattering albedo. Such NIR-VIS observations could potentially allow for a detection of liquid surface water via the ocean glint as suggested in Robinson et al. (2010) or a detection of the vegetation red edge, which is an increased reflectivity in the NIR due to photosynthetic life and therefore a surface biosignature (see, e.g., Seager et al. 2005; Schwieterman 2018). A planet accessible to both techniques would be a prime target for atmospheric characterization.

5.5. Limitations and future work

Even though we have achieved our main goal of deriving first order quantitative requirements for LIFE, the results have to be interpreted with care since there are fundamental limitations inherent to our approach. First, we generated the input spectra for our retrievals using a 1D radiative transfer model and a fully mixed atmosphere, which is a simplification of reality. Additionally, we did not account for the effect of partial, full, or varying cloud coverage on the MIR Earth-twin emission spectrum. Retrieval analyses based on more complex forward models

are foreseen in the future to investigate the effects of these simplifying assumptions.

Next, we did not retrieve additional molecules that are not present in the input spectrum. However, when analyzing spectra from observations, we do not know what species are present in the atmosphere. Not retrieving for additional molecules that are not present might lead to overly confident estimates for LIFE's technical requirements. We provide a first test for the robustness of our results with respect to additional species in Appendix C and plan further investigations for future work.

Further, our input spectra were static and represented the average emission spectrum of the Earth-twin. However, the real emission spectrum varies over time (day vs. night and summer vs. winter) and it also depends on the viewing geometry (pole-on vs. equator-on, e.g., Mettler et al. 2020).

Additionally, the presence of a moon can have an influence on the integrated thermal emission spectrum of the planet-moon system. This effect can be particularly important, if the moon is as large as the Earth's moon and features day side temperatures higher than that of the planet (Robinson 2011). The quantitative impact a moon has on the retrieval results will be investigated in future work.

Also, we used `petitRADTRANS` both to generate the simulated input spectra and as atmospheric forward model for the retrieval framework. As demonstrated in Barstow et al. (2020), retrieving the same input spectrum with different forward models can lead to inconsistencies between the retrieved parameter values due to differences in the forward models. Similar problems will likely arise when retrieving model parameters from an experimentally measured spectrum, since the forward model does not capture the full physics (and/or chemistry) of the observed atmosphere. Our results could therefore be overly-optimistic. Estimating the magnitude of this bias is the subject of future work.

Furthermore, we added the LIFEsim noise as uncertainty to the theoretically simulated theoretical flux without randomizing the value of the individual spectral points. This may lead to over-optimistic retrieval results, in particular for the small *S/N* cases. This is especially true for CH_4, whose Earth's abundance is close to the detection limit of LIFE. A discussion on the potential impact of this simplification is provided in Appendix C.

Moreover, LIFEsim currently only features dominant astrophysical noise terms (Dannert et al. 2022). However, systematic instrumental effects will also impact the observations, even though, ideally, the instrument will only contribute to the noise, but not dominate the noise budget. Still, the required integration times should be considered a lower limit until the optical, thermal, and detector designs of LIFE have further matured and LIFEsim is updated accordingly.

Finally, the study of an Earth-twin exoplanet is clearly a simplification. The known diversity of terrestrial exoplanets demonstrates that future studies will have to look beyond the Earth-twin case and consider a wider range of worlds. In order to obtain more rigorous constraints on the requirements for LIFE, we will perform similar retrieval studies on a variety of different exoplanet types.

6. Summary and conclusions

For this study, we considered an Earth-twin exoplanet orbiting a solar-type star at 10 pc. Using an atmospheric retrieval framework, we derived the minimal requirements for the spectral resolution R, the wavelength coverage, and the *S/N* that need to be met by a space-based MIR nulling interferometer such as LIFE to characterize the atmospheres of such an exoplanet.

In our atmospheric model, we described the atmospheric P–T structure of an Earth-twin using a fourth order polynomial with surface-pressure P_0 and surface temperature T_0. We assumed constant, modern Earth abundances of N_2, O_2, CO_2, H_2O, O_3, CH_4, and N_2O throughout the vertical extent of the atmosphere; clouds were not included. We generated the thermal emission spectrum corresponding to this Earth-twin atmosphere using the 1D radiative transfer model `petitRADTRANS` (Mollière et al. 2019), assuming an exoplanet mass $M_{\rm pl} = 1\,M_\oplus$ and a radius $R_{\rm pl} = 1R_\oplus$. Further, we used the LIFEsim tool (Dannert et al. 2022) to estimate the wavelength-dependent observational noise (including noise from stellar leakage and from local- and exozodiacal dust emission) for different wavelength ranges and combinations of R and *S/N*.

We created a Bayesian retrieval framework coupling `petitRADTRANS` and `pyMultiNest` (Buchner et al. 2014) (python access of the `FORTRAN MultiNest` (Feroz et al. 2009) implementation of the Nested Sampling algorithm (Skilling 2006)) to extract information about the atmospheric structure and composition of the Earth-twin exoplanet from the input spectra. These retrievals were performed considering different wavelength ranges (3–20 µm, 4–18.5 µm, 6–17 µm), spectral resolutions R (20, 35, 50, 100), and signal-to-noise ratios *S/N* (5, 10, 15, 20) at a wavelength of 11.2 µm (the corresponding noise at other wavelengths was derived via LIFEsim).

The performed retrieval analyses suggest that MIR observations with LIFE at an $S/N \geq 10$ can robustly constrain the radius (uncertainty $\leq \pm 10\%$), surface pressure P_0 (uncertainty $\leq \pm 0.5$ dex), and surface temperature T_0 (uncertainty $\leq \pm 20$ K) of an Earth-twin exoplanet. These parameters cannot be probed accurately via reflected light observations of an Earth-twin at NIR-VIS wavelengths. Furthermore, we predict CO_2, H_2O, and O_3 to be detectable by LIFE (error $\leq \pm 1.0$ dex) given an input spectrum with an $S/N \geq 10$ for all the considered wavelength ranges and R. In contrast, the potential biosignature N_2O and the potential antibiosignature CO are not detectable for any of the considered LIFE configurations. For N_2O we find an upper limit on the abundance, indicating that high abundances ($\gtrsim 10^{-3}$ in mass fraction) are potentially detectable in MIR observations. For CO, we do not retrieve any constraint.

Concerning the potential biosignature CH_4, our retrieval results strongly depend on the properties of the input spectrum. If we aim to detect CH_4 at Earth-like abundances ($\approx 10^{-6}$ in mass fraction) in exoplanet atmospheres, we estimate a minimal requirement of $R = 50$ and $S/N = 10$. Furthermore, we observe a performance drop for CH_4 for the 6–17 µm wavelength range, which is a result of cutting off the H_2O absorption bands at wavelengths ≥17 µm. Including H_2O bands at wavelengths >17 µm improves the accuracy of the H_2O abundance estimate, which in turn helps to disentangle contributions from species overlapping at shorter wavelengths. Between the 3–20 µm and 4–18.5 µm wavelength range, we do not observe significant differences in the performance of the retrieval framework. Therefore, a wavelength coverage of at least 4–18.5 µm is desirable.

Turning the *S/N* requirement into estimates for integration times (conservatively assuming a total instrument throughput of 5%, cf. Dannert et al. 2022) and considering recent estimates for the number of temperate terrestrial exoplanets around Solar-type stars within 10 pc, we find that LIFE should feature at least 4×2 m apertures to be able to investigate a somewhat sizeable sample of these objects. With 4×1 m apertures, the integration

time to study an Earth-twin located at 10 pc will be prohibitively long (>1 yr).

By comparing our results with those obtained in similar studies focusing on reflected light observations of Earth-twin exoplanets (e.g., Feng et al. 2018), we find that both approaches complement each other. MIR emission spectra of terrestrial exoplanets can provide access to surface conditions and accurately constrain the radii of the objects, parameters that are challenging for reflected light observations. In addition, we have shown that MIR spectra will allow us to probe for the simultaneous presence of O_3 and CH_4, a strong combinatory biosignature with no currently known false positives, underlining again the large and unique opportunity that MIR observations provide.

Overall, our results suggest that pursuing a concept for a space-based MIR nulling interferometer such as LIFE in addition to the proposed NIR-VIS mission concepts, currently under consideration by NASA, will be a key element for the future of exoplanet characterization. In particular, the combination of results from both approaches would vastly expand our knowledge about worlds outside the Solar System.

Acknowledgements. This work has been carried out within the framework of the National Center of Competence in Research PlanetS supported by the Swiss National Science Foundation. S.P.Q. and E.A. acknowledge the financial support from the SNSF. This work benefited from the 2019 Exoplanet Summer Program in the Other Worlds Laboratory (OWL) at the University of California, Santa Cruz, a program funded by the Heising-Simons Foundation. Further, we would like to thank Michael Line for useful discussions and analyses from which the project benefited. P.M. acknowledges support from the European Research Council under the European Union's Horizon 2020 research and innovation program under grant agreement no. 832428. J.L.G. thanks ISSI Team 464 for useful discussions. Author contributions. B.S.K. carried out the analyses, created the figures, and wrote the bulk part of the manuscript. S.P.Q. initiated the project. S.P.Q. and E.A. guided the project and wrote part of the manuscript. All authors discussed the results and commented on the manuscript.

References

Arney, G. N., Domagal-Goldman, S. D., & Meadows, V. S. 2018, AsBio, 18, 311
Barstow, J. K., & Heng, K. 2020, SSR, 216, 82
Barstow, J. K., Changeat, Q., Garland, R., et al. 2020, MNRAS, 493, 4884-909
Brandt, T. D., & Spiegel, D. S. 2014, Proc. Natl. Acad. Sci. U.S.A., 111, 13278
Bryson, S., Kunimoto, M., Kopparapu, R. K., et al. 2021, AJ, 161, 36
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Carnall, A. C. 2017, ArXiv e-prints [arXiv:1705.05165]
Carrión-González, Ó., García Muñoz, A., Cabrera, J., et al. 2020, A&A, 640, A136
Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., et al. 2018, Astrobiology, 18, 709
Chen, J., & Kipping, D. 2016, ApJ, 834, 17
Chubb, K. L., Rocchetto, M., Yurchenko, S. N., et al. 2021, A&A, 646, A21
Cockell, C., Léger, A., Fridlund, M., et al. 2009, Astrobiology, 9, 1, pMID: 19203238
Dannert, F., Ottiger, M., Quanz, S. P., et al. 2022, A&A, 664, A22 (Paper II)
Deming, D., Louie, D., & Sheets, H. 2018, PASP, 131, 013001
Des Marais, D. J., Harwit, M. O., Jucks, K. W., et al. 2002, Astrobiology, 2, 153
Ertel, S., Defrère, D., Hinz, P., et al. 2020, AJ, 159, 177
Fauchez, T. J., Turbet, M., Villanueva, G. L., et al. 2019, ApJ, 887, 194
Feng, Y. K., Line, M. R., Fortney, J. J., et al. 2016, ApJ, 829, 52
Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, AJ, 155, 200
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601-14
Feroz, F., Hobson, M., Cameron, E., & Pettitt, A. 2013, Open J. Astrophys., 2
Fu, Q., & Liou, K. N. 1992, J. Atmos. Sci., 49, 2139
Gaudi, B. S., Seager, S., Mennesson, B., et al. 2020, ArXiv e-prints [arXiv:2001.06683]
Goody, R., West, R., Chen, L., & Crisp, D. 1989, J. Quant. Spectr. Rad. Transf., 42, 539
Guillot, T. 2010, A&A, 520, A27
Harman, C. E., & Domagal-Goldman, S. 2018, Biosignature False Positives, eds. H. J. Deeg, & J. A. Belmonte, 71
Hastings, W. K. 1970, Biometrika, 57, 97
Hatzes, A. P., & Rauer, H. 2015, ApJ, 810, L25
Jeffreys, H. 1998, The Theory of Probability (Oxford Classic Texts in the Physical Sciences (OUP Oxford), 432
Kawashima, Y., & Rugheimer, S. 2019, ApJ, 157, 213
Kitzmann, D., Patzer, A. B. C., von Paris, P., Godolt, M., & Rauer, H. 2011, A&A, 531, A62
Komacek, T. D., Fauchez, T. J., Wolf, E. T., & Abbot, D. S. 2020, ApJ, 888, L20
Lacis, A. A., & Oinas, V. 1991, J. Geophys. Res.: Atmos., 96, 9027
Lavvas, P., Koskinen, T., Steinrueck, M. E., Muñoz, A. G., & Showman, A. P. 2019, ApJ, 878, 118
Lederberg, J. 1965, Nature, 207, 9
Léger, A., Mariotti, J. M., Mennesson, B., et al. 1996, Icarus, 123, 249
Léger, A., Defrère, D., García Muñoz, A., et al. 2019, Astrobiology, 19, 797
Line, M., Quanz, S. P., Schwieterman, E. W., et al. 2019, BAAS, 51, 271
Lippincott, E. R., Eck, R. V., Dayhoff, M. O., & Sagan, C. 1967, ApJ, 147, 753
Lovelock, J. E. 1965, Nature, 207, 568
Madhusudhan, N. 2018, Handbook of Exoplanets, 2153
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Mayor, M., & Queloz, D. 1995, Nature, 378, 355
Meadows, V. S., Reinhard, C. T., Arney, G. N., et al. 2018, Astrobiology, 18, 630
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087
Mettler, J.-N., Quanz, S. P., & Helled, R. 2020, AJ, 160, 246
Mollière, P., Boekel, R. v., Dullemond, C., Henning, T., & Mordasini, C. 2015, ApJ, 813, 47
Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67
Mollière, P., Stolker, T., Lacour, S., et al. 2020, A&A, 640, A131
Mueller-Wodarg, I. C. F., Strobel, D. F., Moses, J. I., et al. 2008, Neutral Atmospheres, eds. A. F. Nagy, A. Balogh, T. E. Cravens, M. Mendillo, & I. Mueller-Wodarg (New York: Springer New York), 191
Otegi, J. F., Bouchy, F., & Helled, R. 2020, A&A, 634, A43
Peterson, B. M., Fischer, D., & LUVOIR Science and Technology Definition Team. 2017, AAS, 229, 405.04
Quanz, S. P., Kammerer, J., Defrère, D., et al. 2018, SPIE Conf. Ser., 10701, 107011I
Quanz, S. P., Absil, O., Angerhausen, D., et al. 2021, ArXiv e-prints [arXiv:1908.01316]
Quanz, S. P., Ottiger, M., Fontanet, E., et al. 2022, A&A, 664, A21 (Paper I)
Robinson, T. D. 2011, ApJ, 741, 51
Robinson, T. D., Meadows, V. S., & Crisp, D. 2010, ApJ, 721, L67
Rothman, L. S., Wattson, R. B., Gamache, R., Schroeder, J. W., & McCann, A. 1995, in Atmospheric Propagation and Remote Sensing IV, ed. J. C. Dainty, 2471, International Society for Optics and Photonics (SPIE), 105, 111
Rothman, L., Gordon, I., Barber, R., et al. 2010, J. Quant. Spectr. Rad. Transf., 111, 2139, xVIth Symposium on High Resolution Molecular Spectroscopy (HighRus-2009)
Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, J. Quant. Spec. Rad. Transf., 130, 4
Rugheimer, S., Kaltenegger, L., Zsom, A., Segura, A., & Sasselov, D. 2013, Astrobiology, 13, 251
Rugheimer, S., Segura, A., Kaltenegger, L., & Sasselov, D. 2015, ApJ, 806, 137
Schwieterman, E. W. 2018, Handbook of Exoplanets, 3173
Schwieterman, E. W., Kiang, N. Y., Parenteau, M. N., et al. 2018, Astrobiology, 18, 663
Schwieterman, E. W., Robinson, T. D., Meadows, V. S., Misra, A., & Domagal-Goldman, S. 2015, ApJ, 810, 57
Seager, S., Turner, E., Schafer, J., & Ford, E. 2005, Astrobiology, 5, 372
Seager, S., Bains, W., & Petkowski, J. 2016, Astrobiology, 16, 465
Selsis, F., Despois, D., & Parisot, J. P. 2002, A&A, 388, 985
Sing, D. K., Fortney, J. J., Nikolov, N., et al. 2016, Nature, 529, 59
Skilling, J. 2006, Bayesian Anal., 1, 833
Tennyson, J., Yurchenko, S. N., Al-Refaie, A. F., et al. 2016, J. Mol. Spectro., 327, 73
Trotta, R. 2017, Bayesian Methods in Cosmology
van de Schoot, R., Depaoli, S., King, R., et al. 2021, Nat. Rev. Methods Primers, 1, 3
von Paris, P., Hedelt, P., Selsis, F., Schreier, F., & Trautmann, T. 2013, A&A, 551, A120
Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, ApJ, 825, 19
Wunderlich, F., Scheucher, M., Grenfell, J. L., et al. 2021, A&A, 647, A48
Yurchenko, S. N., Mellor, T. M., Freedman, R. S., & Tennyson, J. 2020, MNRAS, 496, 5282
Zahnle, K., Haberle, R. M., Catling, D. C., & Kasting, J. F. 2008, J. Geophys. Res.: Planets, 113
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127

Appendix A: Selection of *P–T* profile model

In the following, we analyze the performance of different parametric models in describing the atmospheric *P–T* structure of different terrestrial planets in our Solar System. In a second step, we assess their applicability in our retrieval framework.

Appendix A.1. Considered P–T models

For our *P–T* model selection, we considered the following four *P–T* models:

1. Polynomial *P–T* parametrization: The *P–T* structure of an atmosphere is modeled via a nth order polynomial:

$$T(P) = \sum_{i=0}^{n} a_i P^i. \tag{A.1}$$

The $n+1$ constants a_i are the model parameters. In our *P–T* model selection, we considered polynomials up to the ninth order.

2. *P–T* parametrization proposed by Madhusudhan & Seager (2009): This is a *P–T* model for terrestrial planets. It is based on the *P–T* profiles of atmospheres of rocky Solar System objects and 1D self-consistent exoplanet *P–T* profiles generated via model atmosphere calculations. It is defined by the following equations:

$$P_0 < P < P_1: \quad P = P_0 e^{\alpha_1 (T - T_0)^{\beta_1}}, \tag{A.2}$$

$$P_1 < P < P_3: \quad P = P_2 e^{\alpha_2 (T - T_2)^{\beta_2}}, \tag{A.3}$$

$$P > P_3: \quad T = T_3. \tag{A.4}$$

The middle pressure layer (Eq. A.3) allows for thermal inversion, the deepest layer (Eq. A.4) is set to be isothermal due to its large optical depth. Parameters β_1 and β_2 are set to 0.5. Further, P_0 is the pressure at the top of the atmosphere, which we fixed to 10^{-6} bar. The remaining tunable parameters are: T_0, α_1, α_2, P_1, P_2, P_3, T_2, and T_3. By requiring continuity between the three layers, the number of model parameters can be reduced to six by setting:

$$T_3 = T_2 + \left(\frac{\log(P_3/P_2)}{\alpha_2}\right)^2, \tag{A.5}$$

$$T_2 = T_0 + \left(\frac{\log(P_1/P_0)}{\alpha_1}\right)^2 - \left(\frac{\log(P_3/P_2)}{\alpha_2}\right)^2. \tag{A.6}$$

3. *P–T* parametrization proposed by Guillot (2010): This model calculates the *P–T* structure of an atmosphere assuming a radiative equilibrium for each atmospheric layer. The model is described by the following equation:

$$T_{Guillot}(P) = \frac{3T_{int}^4}{4}\left(\frac{2}{3} + \frac{\kappa_{IR}}{g_{pl}}P\right) + \frac{3T_{equ}^4}{4}\left[\frac{2}{3} + \frac{1}{\gamma\sqrt{3}} + \left(\frac{\gamma}{\sqrt{3}} - \frac{1}{\gamma\sqrt{3}}\right)e^{-\gamma\sqrt{3}\frac{\kappa_{IR}}{g_{pl}}P}\right]. \tag{A.7}$$

The factor $\left(\kappa_{IR}P/g_{pl}\right)$ is the atmosphere's optical depth τ. The model assumes that the IR opacity can be approximated by a constant κ_{IR}. Further, the opacity in the visible, κ_{VIS}, is assumed to be constant and linked to κ_{IR} via γ: $\kappa_{VIS} = \gamma\kappa_{IR}$. The planet's surface gravity is denoted as g_{pl}, T_{int} is the planet's internal temperature (remaining heat from planet's gravitational collapse or energy from radioactive decay of elements). Finally, T_{equ} is the planet's equilibrium temperature, which is the temperature the planet would have, if it were a black-body heated only by the radiation coming from its host star:

$$T_{equ} = T_* \sqrt{\frac{R_*}{2d}}\,(1 - A_B)^{1/4}. \tag{A.8}$$

Here, T_* is the host star's effective temperature, R_* its radius, d the separation between the planet and its host, and A_B the planet's Bond albedo. Overall, the Guillot (2010) *P–T* model requires four model parameters (T_{int}, T_{equ}, κ_{IR}, and γ; g_{pl} is already a parameter of the retrieval).

4. *P–T* parametrization proposed by Mollière et al. (2019): The Guillot (2010) model does not allow for a nonisothermal structure in the upper atmosphere. The small modification proposed by Mollière et al. (2019) allows for a nonisothermal upper atmosphere:

$$T(P) = T_{\text{Guillot}}(P) \cdot \left(1 - \frac{\alpha}{1 + P/P_{\text{Trans}}}\right), \tag{A.9}$$

Where $T_{\text{Guillot}}(P)$ is the Guillot (2010) *P–T* model (Eq. (A.7)). This modification adds two parameters, α and P_{Trans}, to the Guillot (2010) model, resulting in six model parameters.

Table A.1: Priors used in the *P–T* profile retrievals.

Model	Model Parameter	Prior
Polynomial	a_i	$\mathcal{U}(-1000, 1000)$
Madhusudhan & Seager	$\log_{10}(P_i)$	$\mathcal{U}(-10, 4)$
	T_0	$\mathcal{U}(0, 1000)$
	α_1, α_2	$\mathcal{U}(0, 100)$
Guillot	T_{equ}, T_{int}	$\mathcal{U}(0, 1000)$
	κ_{IR}	$\mathcal{U}(0, 1)$
	γ	$\mathcal{U}(0, 100)$
Addition for Modified Guillot	$\log_{10}(P_{Trans})$	$\mathcal{U}(-10, 4)$
	α	$\mathcal{U}(-1, 1)$

Notes. $\mathcal{U}(x, y)$: boxcar prior with lower limit x and upper limit y.

Appendix A.2. Choice of P–T model via retrievals

We used a Bayesian retrieval framework to assess the performance of the different *P–T* models in describing the *P–T* structure of the Solar System planets Venus, Earth, and Mars. We included Venus and Mars to prevent the selected *P–T* model from being biased toward Earth-like *P–T* profiles. In the retrievals, we used the *P–T* models introduced in Appendix A.1 as forward models. Further, we utilized the `MultiNest` algorithm (Feroz et al. 2009) for parameter estimation, via `pyMultiNest` (Buchner et al. 2014). We chose 400 live points and a sampling efficiency of 0.8 as suggested for model comparison by the `pyMultiNest` documentation. The prior distributions we assumed for the different model parameters are summarized in Table A.1.

The *P–T* model we choose for our retrieval study should provide an optimal combination of two properties. First, the model

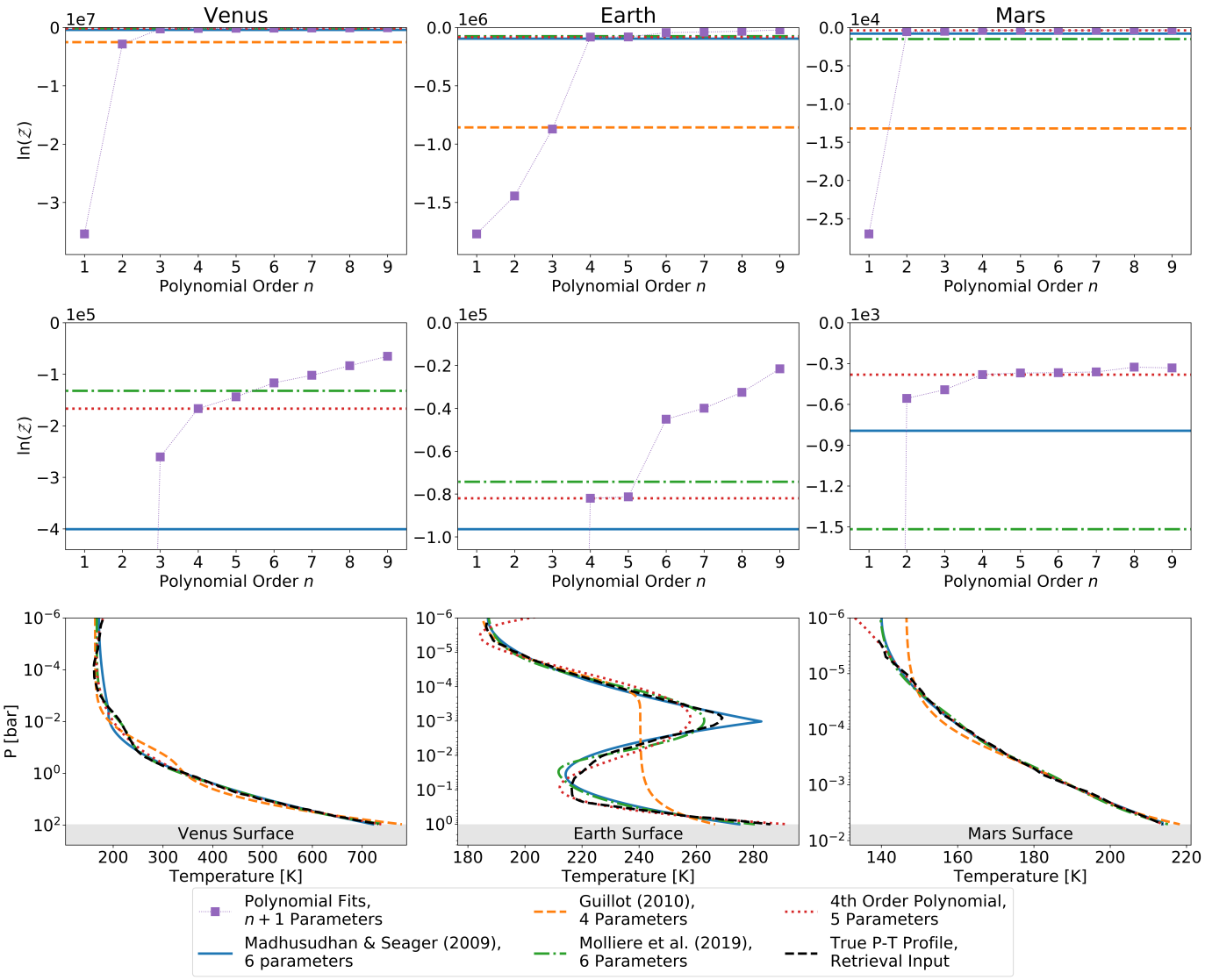


Fig. A.1: Results from the *P*–*T* model retrievals. The first row displays the log-evidence ln($\mathcal{Z}$) of the different models. The second row shows a zoomed in view of the high log-evidence region of the first row to allow for better comparison. The bottom row displays the retrieved best fit *P*–*T* profiles for the fourth order polynomial, the Madhusudhan & Seager (2009) model, and the two Guillot models (Guillot 2010; Mollière et al. 2019). The data for the true *P*–*T* profiles is taken from Fig. 1 in Mueller-Wodarg et al. (2008).

evidence $\mathcal{Z}_{\mathcal{M}}$ should be maximal, since a large $\mathcal{Z}_{\mathcal{M}}$ implies a good fit of the *P*–*T* model (see Sect. 2.2). Second, the number of model parameters should be minimal, since additional model parameters increase the computational cost of the retrieval significantly.

In the first two rows of Fig. A.1, we plot the log-evidence ln($\mathcal{Z}$) corresponding to the retrieval results for the considered polynomial *P*–*T* models. Additionally, we indicate the ln($\mathcal{Z}$) corresponding to the nonpolynomial *P*–*T* models as horizontal lines. For all considered atmospheres, we observe a continuous increase in ln($\mathcal{Z}$) with rising order of the polynomial *P*–*T* model. This indicates that higher order polynomials provide a better fit to the atmosphere's *P*–*T* profile. However, we observe a distinct flattening in the increase of ln($\mathcal{Z}$) with increasing polynomial order for considered atmospheres. This flattening occurs at a third order polynomial for Venus, a fourth order polynomial for Earth, and a second order polynomial for Mars. Increasing the polynomial order beyond these thresholds does not lead to a significant improvement in the polynomial fit. The last row in Fig. A.1 indicates that the Guillot (2010) model fails to fit the inversion in Earth's *P*–*T* profile. This further manifests itself in the corresponding ln($\mathcal{Z}$), which is an order of magnitude smaller than the ln($\mathcal{Z}$) associated with most of the other models considered (order greater than three for the polynomial model).

The Madhusudhan & Seager (2009) profile accurately models the *P*-*T* structure of all considered atmospheres. However, when comparing ln($\mathcal{Z}$) to the other models via the Bayes' factor, we find that for Venus and Earth $\log_{10}(K)$ exceeds two. This suggests that both polynomials of order greater than three and the Mollière et al. (2019) model provide a significantly better *P*-*T* fit (see Table 1). For the Martian atmosphere, the Madhusudhan & Seager (2009) model outperforms the Mollière et al. (2019) model, but still underperforms compared to all polynomials of order greater than one.

The Mollière et al. (2019) accurately fits all three considered *P*–*T* structures. For Venus and Earth, the model provides a better fit than all polynomials of order smaller than seven. However, compared to the fourth order polynomial, one additional param-

eter is required. For Earth, the main difference between the Mollière et al. (2019) model and the fourth polynomial occurs at the inversions at 10^{-1} and 10^{-3} bar, where both models struggle to describe the true *P-T* structure accurately. For Mars, all polynomials of order greater than one outperform the modified Guillot model, as is indicated by the significantly larger $\ln(\mathcal{Z})$.

For our purposes, we choose the fourth order polynomial *P-T* model, since it adds fewer parameters to our retrievals. Despite relying on fewer parameters, it manages to yield a comparable fit and describes all three considered *P-T* structures satisfactorily. The saved parameter allows us to retrieve for one additional atmospheric parameter of interest.

Appendix B: Posterior classification

For the presentation of the grid retrieval results, we chose to classify the retrieved posterior distributions for the different parameters into four main classes based on their visual appearance. Figure B.1 gives illustrative examples for each of the four different posterior shapes.

The unconstrained (UC-type) posterior does not pose a strong constraint on the parameter of interest for the assumed prior distribution. In this case, we described the posterior probability as a constant value throughout the entire prior range (Fig. B.1a).

The upper limit (UL-type) posterior provides an upper limit for the retrieved parameter. For the abundance of a molecule, this implies that the retrieval rules out any concentration above the found threshold value. Below this threshold, all abundances are equally likely and cannot be ruled out. We described such a posterior via the Logistic function:

$$f(x) = \frac{c}{1 + e^{a \cdot x + b}}. \tag{B.1}$$

The constants a, b, and c are unique for each posterior. In the data analysis, we marked the half-maximum as well as the 16th percentile of the logistic function (Fig. B.1b).

The sensitivity limit (SL-type) posterior is similar to the UL-type posterior. The retrieval is capable of excluding the high, but not the low molecular abundances. However, the SL-type posterior exhibits a distinct peak at the boundary between these two regimes, which roughly corresponds to the true abundance. We convolved the Logistic function (Eq. (B.1)) with a Gaussian distribution to describe this posterior type:

$$f(x) = \frac{c + d \cdot e^{-(x+\mu)^2/\sigma^2} / \left(\sqrt{2\pi}\sigma\right)}{1 + e^{a \cdot x + b}}. \tag{B.2}$$

The constants a, b, c, d, μ, and σ are unique for each posterior. Here, μ corresponds to the mean, σ to the standard deviation of the Gaussian. In our analysis, we marked the maximum value of the distribution and the position of the half maxima to the left and the right of it (Fig. B.1c).

Finally, the constrained (C-type) posterior strongly constrains the prior range and was approximated via a Gaussian distribution:

$$f(x) = \frac{d \cdot e^{-(x+\mu)^2/\sigma^2}}{\sqrt{2\pi}\sigma}. \tag{B.3}$$

The constants d, μ, and σ are unique for each posterior. Here, μ corresponds to the mean and σ to the standard deviation of the Gaussian. In the analysis, we marked the 50th, 16th, and 84th percentile, which corresponds to the 1 σ range (Fig. B.1d).

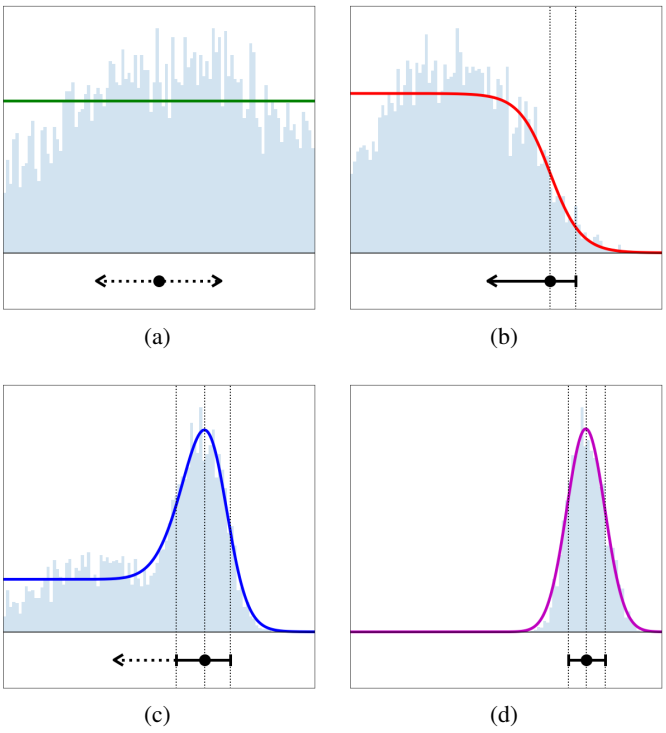


Fig. B.1: Classes of retrieved posterior distributions. The histogram gives the posterior distribution as found by the retrieval framework. The colored lines represent the best fit model for the posterior distribution, the black symbols below show how such a posterior will be represented in our data analysis plots. (*a*): unconstrained posterior (UC-type), (*b*): upper limit posterior (UL-type), (*c*): sensitivity limit posterior (SL-type), (*d*): constrained posterior (C-type).

We determined the best fit model for each of the retrieved posterior distributions by fitting all four models. Thereafter, we used the log-likelihood function (Eq. (3), Sect. 2.2) to determine which model best described the retrieved posterior distribution. For the SL-type posterior, we required the maximum value of the fitted function was at least 1.3 times and maximally 10 times larger than the continuum probability at low abundances. If the peak probability was less than 1.3 times the continuum value, we assumed a UL-type posterior. If the peak abundance was at least 10 times more probable than the low abundance continuum, we chose a C-type posterior instead. These cutoff values were chosen empirically to match what we would have obtained in a manual classification.

Appendix C: Impact of randomized noise and additional species on retrieval results

As discussed in Sect. 2.3.1, we relied on two major assumptions in order to render our retrieval study computationally feasible and bias-free. First, we did not randomize the placement of the individual spectral points. Instead, we ran retrievals for the non-randomized spectral points and treated the LIFEsim noise as uncertainty on the spectrum. Second, our retrieval model did not allow for any additional gas species in the atmosphere. However, in real observations, we do not know what species are present. Therefore, it is important to investigate how robust these retrievals are with respect to false positive detections of molecules.

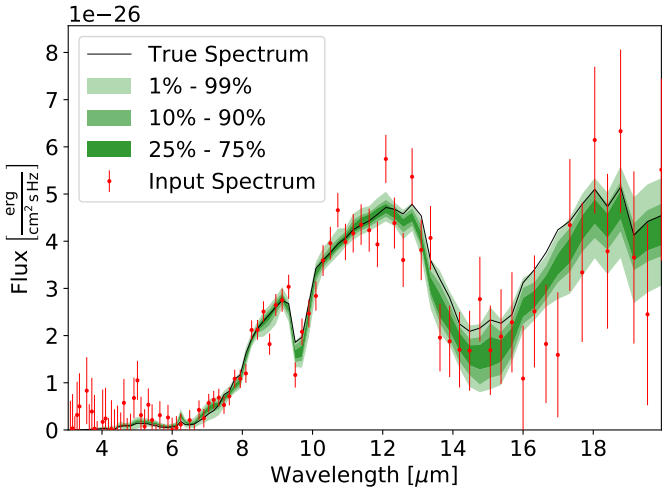


Fig. C.1: Fitted MIR emission spectra to a randomized input spectrum. The nonrandomized spectrum is marked by the black line, the randomized input for the retrieval are represented by red data points and error bars. The green shaded regions mark percentiles of the MIR spectra corresponding to the retrieved posteriors of the model parameters.

There could be a finite probability of false positive detections, especially if combined with the randomized placement of the spectral points.

In this section, we investigate how randomizing spectral points and adding more species to the forward model impacts the performance of our retrieval routine at the proposed instrument requirements for LIFE ($R = 50$, $S/N = 10$, optimized noise scenario). To this purpose, we generated 20 different noise realizations of the spectrum at these requirements. We randomized the placement of each individual spectral point by randomly sampling a Gaussian distribution, where the standard deviation is determined by the wavelength dependent LIFESIM noise. Next, we ran a retrieval for every noise realization using the `pyMultiNest` settings outlined in Sect. 3. In addition to the molecules present in the input spectrum, we also retrieved for traces of NH_3, PH_3, and SO_2. The resulting posteriors were then analyzed and classified using the method outlined in Appendix B.

We summarize our results in Figs. C.1, C.2, and C.3. Figure C.1 compares the MIR spectra corresponding to the posteriors retrieved from a single noise realization to the true spectrum. In Fig. C.2, we summarize all retrieved posterior types for the considered atmospheric gases. The leftmost column provides the retrieval findings for the nonrandomized input. Columns 1 to 20 list the results for the randomized input spectra. The rightmost column gives the percentage of retrievals on the randomized input that yielded the same posterior type as the nonrandomized retrieval. The last row provides the log-evidences $\ln(\mathcal{Z})$ for each retrieval run. In Fig. C.3, we show all retrieved posterior distributions as histograms and overlay them. Further, we plot the bin-wise mean and median of the posteriors obtained in the randomized retrievals and the posteriors found in the nonrandomized retrieval.

Appendix C.1. Truth versus spectra from noisy retrievals

As can be seen from Fig. C.1, our retrieval framework shows robust behavior even for randomized input spectra. The spectra corresponding to the retrieved posteriors roughly follow the true spectrum and generally do not significantly overfit to the noise. This observation holds for all noise realizations and is reinforced by the finding that the $\ln(\mathcal{Z})$ corresponding to the randomized retrievals are significantly smaller than in the nonrandomized case (see Fig. C.2). When overfitting, the retrieval models noise induced features accurately and thus the difference between the fitted spectrum and the input spectrum is small. This results in a $\ln(\mathcal{Z})$ value similar to the one obtained in retrievals of nonrandomized spectra. In our case, we are not fitting to the noise induced features. Thus, there is a larger difference between the fitted and input spectrum than in the nonrandomized case, which results in a smaller $\ln(\mathcal{Z})$.

Despite the good overall fit, a systematic offset between the true and the retrieved spectra starts to emerge above 10 μm. This offset becomes more pronounced at the longer wavelengths and results from the decreasing density of spectral points and increasing LIFESIM noise in this wavelength range. Between 10 and 12 μm, the molecules we consider have no strong absorption features and thus this range probes the planetary surface conditions. Offsets in this range will manifest themselves as shifts in the retrieved posteriors of P_0, T_0, and $R_{\rm pl}$. Above 12 μm the shape of the MIR spectrum is dominated first by CO_2 (up to $\approx$ 17 μm) and then by H_2O features. Offsets from the truth in this range correspond to shifts in the retrieved abundance posteriors of these molecules. The described offsets are unique to each of the considered noise realizations and are diminished when averaging the retrieval results obtained from multiple noise realizations (see Sect. C.3 for a detailed discussion).

Appendix C.2. Robustness of the retrieved abundance posterior types

In this section, we take a closer look at the retrieved abundance posterior types to understand how robust different types are with respect to the noise randomization of the input spectrum. There are two underlying questions. First, we analyze how randomization affects the results for molecules that are present in the atmosphere. Second, we discuss if randomization can trigger false positive detections of molecules not present in the atmosphere.

From Fig. C.2, we see that the C-type posterior is robust under randomization. This means that, if retrieving the nonrandomized spectrum yields a C-type posterior, retrievals on noise realizations will mostly also yield C-type posteriors. This also shows in the high accuracy percentage. For our study, this observation implies that, at the determined LIFE requirements, we can expect to be capable of detecting CO_2, H_2O, and O_3 in an Earth-twin atmosphere. We observe a similarly strong robustness for the UC-type posterior. From this, we can conclude that retrieval behavior for N_2, O_2, and CO is accurately predicted via the nonrandomized retrievals presented in our study.

Also, the UL-type posterior shows robust behavior with respect to noise randomization. Here we have to differentiate between two different scenarios: First, N_2O, which is present in the input atmosphere and second, NH_3, PH_3, and SO_2, which are not present.

For N_2O, the high robustness of the UL-type posterior reassures our finding that N_2O is likely not detectable in an Earth-twin at the proposed LIFE requirements. However, we do observe rare exceptions where an SL- or even a C-type posterior is retrieved. These exceptions correspond to cases where the randomization of the input spectrum results in a perceived amplification of the N_2O absorption feature. This results in a detection of N_2O. Naturally, the occurrence of a N_2O detection is characterized by a strong overestimation of the retrieved N_2O

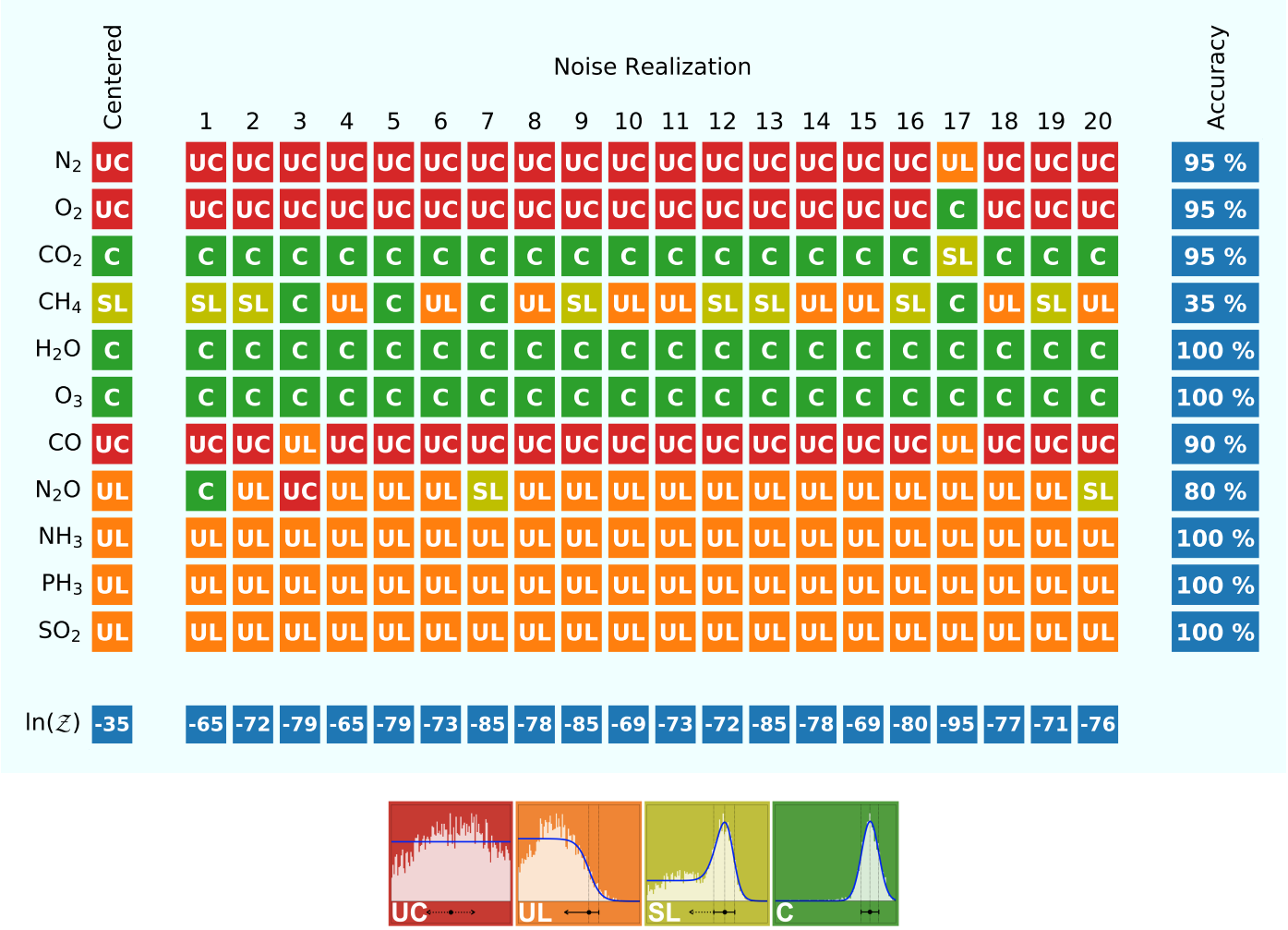


Fig. C.2: Retrieved posterior types for atmospheric gases for the nonrandomized case (first column) and the 20 different noise realizations (columns 1 to 20). For posterior classification, we used the method outlined in Appendix B. The last row lists the Bayesian log-evidence $\ln(\mathcal{Z})$, the last column provides the retrieval accuracy (percentage of noise realizations resulting in the same posterior type as the centered case).

abundance. This can be seen from the retrieved posteriors of N_2O displayed in Fig. C.3.

For NH_3, PH_3, and SO_2, we retrieve UL-type posteriors in all cases, which yields upper limits on the abundances of these molecules. We emphasize that retrieving a UL-type posterior is not a false positive detection. A UL-type posterior indicates that molecular abundances below the corresponding upper limit do not lead to an observable signature in the spectrum and therefore cannot be ruled out. The invariance of the retrieved UL-type posterior with respect to the different noise realizations indicates, that at LIFE requirements, we are robust to false positive detections of these molecules. We are aware that this analysis does not generally rule out false positive detections of arbitrary molecules. However, the fact that there was no false positive suggests that we are likely robust with respect to a large variety of false positives.

The most interesting case in this study is CH_4. When not randomizing the input spectrum, we retrieve an SL-type posterior, which indicates that CH_4 lies at the sensitivity limit (for the considered input data). When using the randomized spectra as input, the retrieved posterior types for CH_4 vary between UL-, SL-, and C-type. Therefore, SL-type posteriors are not robust with respect to spectrum randomization. The observed variance in the retrieved posterior type originates because the randomization of input spectra can lead to an amplification or a reduction of the CH_4 feature (similar as for N_2O). Since the true CH_4 abundance is close to the sensitivity limit, already small differences in the input spectrum can lead to differences in the retrieved posterior type. However, in contrast to N_2O, the peaks of the retrieved SL- and U-type posteriors lie roughly within ± 1 dex of the true CH_4 abundance and do not have a tendency to overestimate it (see Fig. C.3). Still, the fact that the posterior type for CH_4 depends on the noise realization leads us to the question of how we should interpret an SL-type posterior in a nonrandomized retrieval.

Appendix C.3. Interpretation of retrievals results for unrandomized spectra

As we discussed in the previous section, SL-type posteriors do not show robust behavior with respect to noise randomization. This observation raises the question of how to correctly interpret results obtained in retrievals of unrandomized input spectra. In Fig. C.3, we illustrate one possible answer to this question, which we further motivate in this section.

When considering the posteriors from the individual retrieval runs on randomized input spectra (gray shaded areas), we

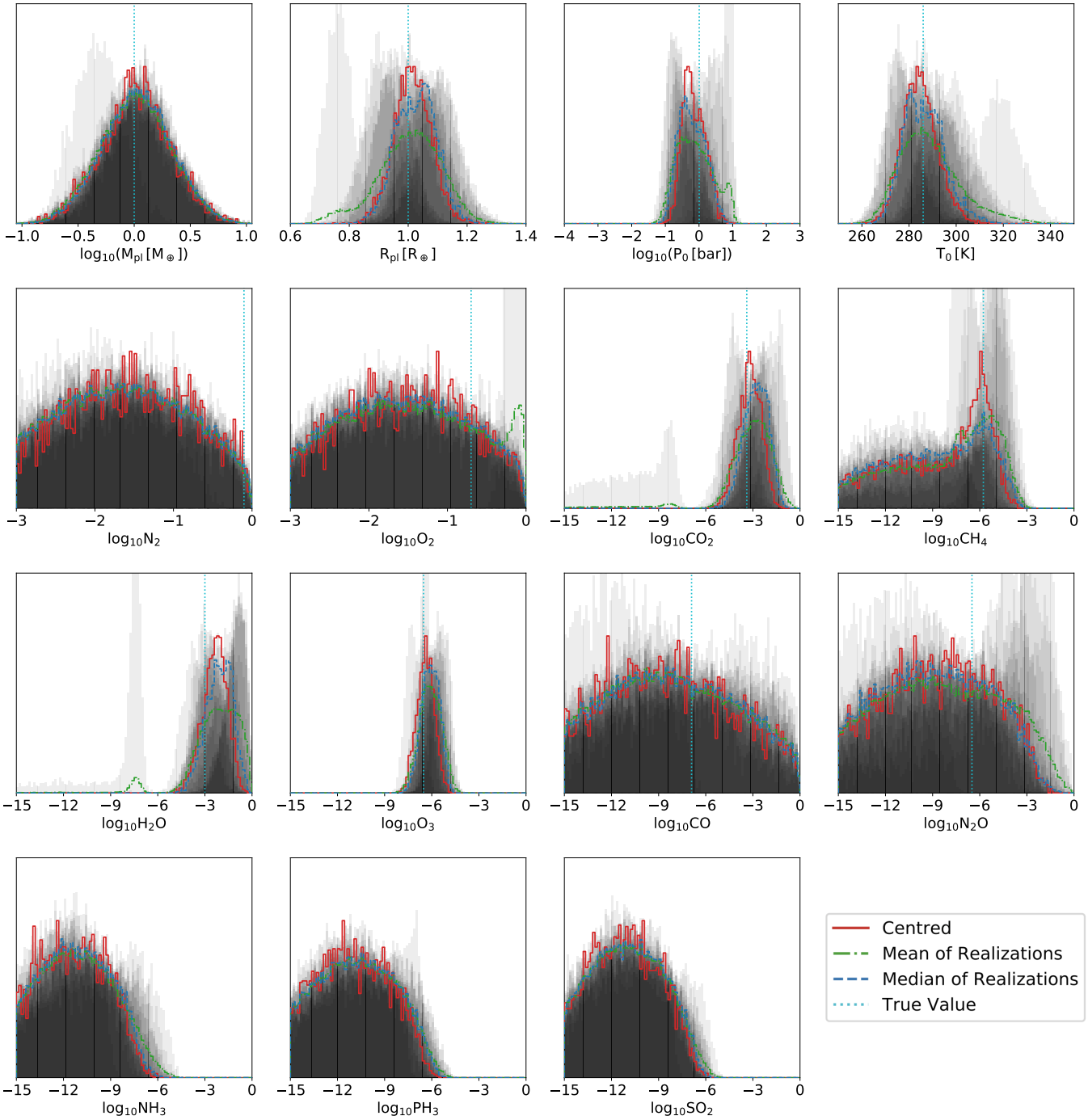


Fig. C.3: Retrieved normalized posterior distributions of model parameters for the 20 noise realizations. The posteriors of the individual noise realizations are plotted in light gray. The more posteriors overlap, the darker the gray becomes. The green dash-dotted line marks the mean, the blue dashed line the median of the posteriors for the different noise realizations. The posteriors found in the retrieval on the nonrandomized spectrum are plotted in solid red. The dotted line marks the true value for the Earth-twin atmosphere. Absence of the dotted line indicates that the species were not present in the atmosphere used to generate the input spectrum.

observe that in some cases there are non-neglectable differences between the different results. For the UC-type posteriors (N_2, O_2, and CO) we observe only small differences between the individual retrieval runs. This underlines the robustness of this posterior type with respect to noise randomization. Similarly, for NH_3, PH_3, and SO_2, we only observe little variance in the retrieved UL-type posteriors, which emphasizes the retrieval's robustness with respect to false positive detections. Small differences between the different posteriors correspond to small biases in the retrieved upper abundance limits of these molecules. Similar conclusions can be drawn for N_2O, where most of the retrieved posteriors are of the UL-type. The few deviances from the norm

correspond to the previously described cases where an SL-type posterior that overestimates the abundance is retrieved.

The interpretation of posteriors is more subtle when considering C-type (M_{pl}, R_{pl}, P_0, T_0, CO_2, H_2O, and O_3) or SL-type (CH_4) posteriors. Here we observe a considerable variance in both the shape and the position of the individual posteriors. This observed variance in the posteriors is a bias that is evoked by the noise randomization of the input spectra. We further find that the variation occurs around the true value and its magnitude is comparable to the standard deviation of the posteriors obtained in the nonrandomized retrievals. Furthermore, we observe that the regions where many posteriors overlap (the dark regions) are similar in shape and position to the posteriors found in the unrandomized retrieval. In other words, the regions of parameter space that most retrievals agree on corresponds to the retrieval result obtained in the unrandomized retrieval.

Given the observations outlined above, the straightforward interpretation for retrieval results from unrandomized spectra is the following: the retrieved posterior distributions provide an estimate for the average retrieval behavior. Thus, if we take the average over the results from retrievals on different noise instances, it will converge toward the posterior distributions found in a retrieval on the unrandomized spectrum. To further motivate this interpretation, we plot the bin-wise mean and median of the posteriors retrieved from the 20 noise realizations in Fig. C.3. We find that already for 20 noise realizations, both mean and median are similar in shape and position to the posteriors from the unrandomized retrieval. For all cases, the median resembles the unrandomized case more closely because it is less sensitive to outliers.

In conclusion, this study suggests that retrieving on an unrandomized input spectrum eliminates biases that arise from noise randomization and will provide reliable estimates for the average behavior of the retrieval on randomized spectra. Generally, UC-, UL-, and C-type posteriors show robust behavior with respect to noise randomization. In contrast, SL-type posteriors are less robust and vary between C-, SL-, and UL-type in retrievals of randomized spectra. This variance in posterior type is closely linked to the fact that SL-type posteriors signify that an abundance is at the sensitivity limit of the retrieval for the studied spectral input. cloud-fee

Appendix D: Retrieved parameter values

The retrieved parameter values are listed in Tables:
- Table D.1: 3–20 μm, nominal case,
- Table D.2: 3–20 μm, optimized case,
- Table D.3: 4–18.5 μm, nominal case,
- Table D.4: 4–18.5 μm, optimized case,
- Table D.5: 6–17 μm, nominal case,
- Table D.6: 6–17 μm, optimized case.

Table D.1: Retrieval results for the 3–20 μm wavelength range for the nominal case.

		R=20				*R=35*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.3^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$
a_3	23.12	$34.3^{+29.5}_{-19.4}$	$15.8^{+13.0}_{-8.0}$	$13.2^{+9.3}_{-6.6}$	$13.3^{+11.2}_{-6.7}$	$26.7^{+21.3}_{-14.2}$	$14.7^{+11.2}_{-7.2}$	$14.2^{+11.6}_{-7.0}$	$13.7^{+10.1}_{-6.2}$
a_2	99.70	$163.8^{+114.4}_{-80.0}$	$65.0^{+49.7}_{-33.3}$	$56.0^{+38.9}_{-30.0}$	$52.4^{+37.2}_{-25.7}$	$102.3^{+63.0}_{-48.4}$	$57.2^{+35.5}_{-27.4}$	$57.4^{+33.5}_{-24.9}$	$57.8^{+29.9}_{-21.7}$
a_1	146.63	$147.5^{+125.6}_{-76.7}$	$104.1^{+71.1}_{-44.5}$	$99.2^{+68.4}_{-38.5}$	$80.3^{+62.4}_{-31.3}$	$108.0^{+68.6}_{-50.5}$	$90.3^{+47.1}_{-28.8}$	$87.6^{+46.5}_{-31.1}$	$88.3^{+56.0}_{-32.5}$
a_0	285.22	$234.2^{+94.6}_{-93.5}$	$281.4^{+45.3}_{-42.3}$	$288.1^{+45.7}_{-34.3}$	$266.5^{+34.6}_{-26.2}$	$240.1^{+42.8}_{-54.8}$	$276.3^{+34.1}_{-29.7}$	$269.8^{+33.0}_{-27.2}$	$266.6^{+43.3}_{-26.1}$
$L(P_0\,[\mathrm{bar}])$	0.006	$0.4^{+0.3}_{-0.5}$	$0.1^{+0.4}_{-0.4}$	$-0.0^{+0.4}_{-0.4}$	$0.2^{+0.3}_{-0.3}$	$0.4^{+0.2}_{-0.3}$	$0.1^{+0.3}_{-0.3}$	$0.2^{+0.3}_{-0.3}$	$0.2^{+0.3}_{-0.4}$
$R\ [R_\oplus]$	1.0	$0.82^{+0.1}_{-0.1}$	$0.96^{+0.09}_{-0.09}$	$1.01^{+0.08}_{-0.08}$	$1.02^{+0.06}_{-0.07}$	$0.88^{+0.10}_{-0.10}$	$0.99^{+0.08}_{-0.08}$	$1.01^{+0.06}_{-0.06}$	$1.03^{+0.05}_{-0.05}$
$L(M\ [M_\oplus])$	0.0	$0.1^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-7.3^{+4.1}_{-4.4}$	$-7.8^{+4.6}_{-4.2}$	$-7.6^{+4.6}_{-4.5}$	$-7.6^{+4.5}_{-4.5}$	$-7.5^{+4.4}_{-4.3}$	$-7.4^{+4.4}_{-4.4}$	$-7.7^{+4.6}_{-4.4}$	$-7.8^{+4.7}_{-4.4}$
$L(O_2)$	-0.679	$-7.0^{+4.2}_{-4.6}$	$-7.1^{+4.4}_{-4.6}$	$-7.4^{+4.5}_{-4.5}$	$-7.5^{+4.6}_{-4.7}$	$-7.7^{+4.5}_{-4.3}$	$-7.4^{+4.3}_{-4.6}$	$-7.6^{+4.5}_{-4.5}$	$-7.5^{+4.7}_{-4.8}$
$L(H_2O)$	-3.000	$-9.4^{+3.8}_{-3.4}$	$-3.2^{+1.1}_{-1.1}$	$-2.7^{+0.9}_{-0.8}$	$-3.1^{+0.7}_{-0.7}$	$-5.3^{+1.7}_{-5.2}$	$-2.9^{+0.8}_{-0.8}$	$-3.0^{+0.7}_{-0.7}$	$-3.0^{+0.8}_{-0.7}$
$L(CO_2)$	-3.387	$-6.6^{+1.9}_{-4.8}$	$-3.6^{+0.9}_{-0.8}$	$-3.2^{+0.7}_{-0.7}$	$-3.6^{+0.7}_{-0.6}$	$-4.6^{+0.8}_{-1.0}$	$-3.4^{+0.8}_{-0.7}$	$-3.5^{+0.7}_{-0.6}$	$-3.5^{+0.7}_{-0.6}$
$L(CH_4)$	-5.770	$-9.9^{+3.1}_{-3.0}$	$-8.7^{+3.2}_{-3.8}$	$-9.2^{+3.4}_{-3.6}$	$-9.6^{+3.3}_{-3.2}$	$-9.2^{+3.4}_{-3.3}$	$-9.7^{+3.6}_{-3.2}$	$-9.3^{+3.2}_{-3.3}$	$-9.0^{+2.8}_{-3.8}$
$L(O_3)$	-6.523	$-10.3^{+2.7}_{-2.8}$	$-6.8^{+0.9}_{-2.2}$	$-6.4^{+0.6}_{-0.6}$	$-6.7^{+0.5}_{-0.5}$	$-9.6^{+2.5}_{-3.3}$	$-6.5^{+0.6}_{-0.6}$	$-6.6^{+0.5}_{-0.5}$	$-6.7^{+0.6}_{-0.5}$
$L(CO)$	-6.903	$-9.8^{+3.3}_{-3.0}$	$-7.8^{+4.3}_{-4.3}$	$-7.5^{+4.3}_{-4.5}$	$-7.7^{+4.5}_{-4.3}$	$-8.7^{+4.0}_{-3.7}$	$-7.7^{+4.3}_{-4.2}$	$-7.7^{+4.4}_{-4.4}$	$-7.7^{+4.3}_{-4.5}$
$L(N_2O)$	-6.495	$-10.3^{+2.8}_{-2.6}$	$-8.5^{+3.8}_{-3.9}$	$-8.1^{+3.7}_{-3.9}$	$-8.5^{+3.6}_{-4.0}$	$-9.4^{+3.5}_{-3.2}$	$-8.3^{+3.6}_{-3.9}$	$-8.7^{+3.6}_{-3.9}$	$-8.9^{+3.4}_{-3.7}$

		R=50				*R=100*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.3}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.3}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$22.2^{+29.6}_{-10.6}$	$14.4^{+11.8}_{-7.4}$	$18.0^{+20.9}_{-8.4}$	$15.3^{+11.2}_{-6.3}$	$12.9^{+10.0}_{-6.6}$	$14.1^{+11.4}_{-6.2}$	$16.2^{+12.1}_{-6.6}$	$19.3^{+16.8}_{-8.4}$
a_2	99.70	$94.0^{+69.0}_{-41.9}$	$61.6^{+43.5}_{-29.0}$	$76.6^{+69.7}_{-33.1}$	$71.8^{+34.7}_{-26.9}$	$47.8^{+26.9}_{-21.8}$	$62.6^{+35.2}_{-24.5}$	$73.5^{+34.7}_{-22.7}$	$88.6^{+56.3}_{-30.0}$
a_1	146.63	$120.6^{+72.6}_{-47.8}$	$102.0^{+70.2}_{-38.2}$	$134.5^{+83.4}_{-69.7}$	$114.0^{+81.2}_{-42.8}$	$70.3^{+33.5}_{-24.8}$	$101.3^{+64.2}_{-42.1}$	$115.4^{+68.2}_{-36.3}$	$154.1^{+68.0}_{-62.4}$
a_0	285.22	$265.1^{+47.0}_{-41.7}$	$285.1^{+39.9}_{-34.3}$	$294.1^{+49.0}_{-44.8}$	$280.4^{+61.1}_{-32.2}$	$256.0^{+30.6}_{-24.4}$	$273.9^{+45.1}_{-30.8}$	$279.2^{+47.7}_{-26.3}$	$298.8^{+41.7}_{-38.8}$
$L(P_0\,[\mathrm{bar}])$	0.006	$0.2^{+0.3}_{-0.3}$	$0.0^{+0.4}_{-0.3}$	$-0.1^{+0.5}_{-0.3}$	$0.0^{+0.3}_{-0.4}$	$0.3^{+0.3}_{-0.3}$	$0.1^{+0.4}_{-0.4}$	$0.0^{+0.3}_{-0.3}$	$-0.1^{+0.3}_{-0.2}$
$R\ [R_\oplus]$	1.0	$0.92^{+0.09}_{-0.09}$	$1.01^{+0.07}_{-0.07}$	$1.03^{+0.06}_{-0.05}$	$1.03^{+0.05}_{-0.05}$	$0.98^{+0.08}_{-0.08}$	$1.02^{+0.06}_{-0.06}$	$1.02^{+0.04}_{-0.04}$	$1.01^{+0.04}_{-0.03}$
$L(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-6.5^{+4.0}_{-4.6}$	$-7.6^{+4.5}_{-4.5}$	$-7.3^{+4.4}_{-4.7}$	$-7.6^{+4.7}_{-4.5}$	$-7.9^{+4.5}_{-4.1}$	$-7.9^{+4.8}_{-4.5}$	$-7.4^{+4.4}_{-4.6}$	$-7.4^{+4.6}_{-4.7}$
$L(O_2)$	-0.679	$-7.8^{+4.5}_{-4.3}$	$-7.2^{+4.4}_{-4.5}$	$-7.1^{+4.5}_{-4.7}$	$-7.5^{+4.7}_{-4.5}$	$-7.4^{+4.2}_{-4.4}$	$-7.8^{+5.0}_{-4.3}$	$-7.3^{+4.5}_{-4.7}$	$-7.5^{+4.6}_{-4.6}$
$L(H_2O)$	-3.000	$-3.7^{+1.1}_{-2.1}$	$-2.8^{+0.8}_{-0.8}$	$-2.7^{+0.7}_{-0.8}$	$-2.8^{+0.8}_{-0.7}$	$-3.6^{+0.9}_{-0.8}$	$-3.0^{+0.8}_{-0.8}$	$-2.9^{+0.6}_{-0.6}$	$-2.7^{+0.5}_{-0.6}$
$L(CO_2)$	-3.387	$-4.0^{+0.8}_{-0.8}$	$-3.3^{+0.7}_{-0.7}$	$-3.2^{+0.7}_{-0.7}$	$-3.3^{+0.7}_{-0.7}$	$-3.9^{+0.8}_{-0.7}$	$-3.4^{+0.7}_{-0.7}$	$-3.3^{+0.6}_{-0.5}$	$-3.2^{+0.5}_{-0.5}$
$L(CH_4)$	-5.770	$-8.5^{+3.1}_{-3.8}$	$-9.0^{+3.1}_{-3.6}$	$-8.8^{+2.9}_{-3.7}$	$-7.7^{+2.1}_{-4.4}$	$-9.5^{+3.2}_{-3.3}$	$-8.7^{+2.7}_{-3.7}$	$-6.8^{+1.3}_{-4.2}$	$-5.9^{+0.7}_{-1.4}$
$L(O_3)$	-6.523	$-7.7^{+1.4}_{-4.1}$	$-6.5^{+0.6}_{-0.6}$	$-6.5^{+0.6}_{-0.6}$	$-6.5^{+0.5}_{-0.5}$	$-7.0^{+0.7}_{-0.7}$	$-6.6^{+0.6}_{-0.5}$	$-6.5^{+0.4}_{-0.4}$	$-6.4^{+0.4}_{-0.4}$
$L(CO)$	-6.903	$-7.8^{+4.0}_{-4.0}$	$-7.8^{+4.5}_{-4.3}$	$-7.7^{+4.4}_{-4.5}$	$-7.9^{+4.4}_{-4.3}$	$-8.1^{+4.5}_{-4.0}$	$-7.7^{+4.7}_{-4.5}$	$-7.8^{+4.6}_{-4.4}$	$-8.0^{+4.6}_{-4.4}$
$L(N_2O)$	-6.495	$-9.0^{+3.8}_{-3.4}$	$-8.5^{+3.7}_{-3.8}$	$-8.7^{+3.5}_{-3.9}$	$-8.9^{+3.4}_{-3.7}$	$-9.1^{+3.5}_{-3.4}$	$-9.0^{+3.6}_{-3.7}$	$-9.0^{+3.4}_{-3.6}$	$-9.3^{+3.5}_{-3.7}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Table D.2: Retrieval results for the 3–20 μm wavelength range for the optimized case.

		R=20				*R*=35			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.2}_{-0.3}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$
a_3	23.12	$25.5^{+20.0}_{-12.3}$	$14.2^{+9.4}_{-6.7}$	$10.7^{+8.0}_{-5.3}$	$13.2^{+11.7}_{-6.4}$	$16.4^{+14.3}_{-8.2}$	$10.6^{+7.1}_{-5.1}$	$11.5^{+10.3}_{-5.4}$	$11.8^{+8.2}_{-5.0}$
a_2	99.70	$108.3^{+59.0}_{-41.4}$	$60.4^{+36.3}_{-27.5}$	$38.9^{+30.2}_{-20.8}$	$61.3^{+51.6}_{-32.6}$	$64.1^{+41.2}_{-30.6}$	$43.3^{+33.8}_{-22.8}$	$47.0^{+36.5}_{-21.8}$	$50.0^{+26.1}_{-20.2}$
a_1	146.63	$128.8^{+67.5}_{-60.6}$	$105.4^{+51.8}_{-37.0}$	$71.4^{+42.2}_{-25.9}$	$130.8^{+97.0}_{-60.7}$	$86.5^{+56.6}_{-43.8}$	$91.4^{+53.7}_{-31.0}$	$89.4^{+59.0}_{-32.5}$	$89.1^{+42.1}_{-30.6}$
a_0	285.22	$201.5^{+73.6}_{-73.0}$	$266.4^{+42.4}_{-33.2}$	$266.6^{+33.0}_{-27.8}$	$323.4^{+66.8}_{-45.3}$	$227.9^{+46.2}_{-59.1}$	$294.7^{+40.7}_{-37.9}$	$285.8^{+43.6}_{-32.5}$	$277.9^{+29.5}_{-24.8}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.5^{+0.3}_{-0.4}$	$0.2^{+0.3}_{-0.4}$	$0.2^{+0.4}_{-0.4}$	$-0.3^{+0.4}_{-0.3}$	$0.6^{+0.3}_{-0.4}$	$-0.1^{+0.4}_{-0.4}$	$0.0^{+0.4}_{-0.4}$	$0.1^{+0.3}_{-0.3}$
$R\ [R_\oplus]$	1.0	$0.9^{+0.1}_{-0.1}$	$0.99^{+0.09}_{-0.09}$	$1.01^{+0.07}_{-0.06}$	$1.02^{+0.06}_{-0.05}$	$0.9^{+0.1}_{-0.1}$	$1.01^{+0.07}_{-0.07}$	$1.02^{+0.05}_{-0.05}$	$1.02^{+0.04}_{-0.04}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-7.6^{+4.3}_{-4.2}$	$-7.9^{+4.5}_{-4.2}$	$-7.7^{+4.4}_{-4.5}$	$-7.8^{+4.7}_{-4.3}$	$-7.6^{+4.2}_{-4.3}$	$-7.2^{+4.2}_{-4.6}$	$-7.5^{+4.6}_{-4.5}$	$-7.5^{+4.6}_{-4.4}$
$L(\mathrm{O_2})$	-0.679	$-7.0^{+4.8}_{-4.6}$	$-7.3^{+4.4}_{-4.6}$	$-7.1^{+4.3}_{-4.3}$	$-7.9^{+4.5}_{-4.3}$	$-7.1^{+4.2}_{-4.5}$	$-7.8^{+4.6}_{-4.4}$	$-7.4^{+4.4}_{-4.5}$	$-7.5^{+4.6}_{-4.5}$
$L(\mathrm{H_2O})$	-3.000	$-4.2^{+1.1}_{-3.0}$	$-2.9^{+0.9}_{-0.8}$	$-2.9^{+0.8}_{-0.8}$	$-1.7^{+0.7}_{-0.9}$	$-3.8^{+1.0}_{-0.9}$	$-2.1^{+0.9}_{-1.0}$	$-2.4^{+0.9}_{-0.9}$	$-2.7^{+0.7}_{-0.6}$
$L(\mathrm{CO_2})$	-3.387	$-6.6^{+2.0}_{-4.6}$	$-3.9^{+0.9}_{-0.9}$	$-3.7^{+0.8}_{-0.8}$	$-2.6^{+0.6}_{-0.8}$	$-5.1^{+1.2}_{-2.4}$	$-3.1^{+0.9}_{-0.9}$	$-3.1^{+0.8}_{-0.8}$	$-3.3^{+0.6}_{-0.6}$
$L(\mathrm{CH_4})$	-5.770	$-9.8^{+2.8}_{-3.0}$	$-9.9^{+3.1}_{-3.1}$	$-9.5^{+3.0}_{-3.4}$	$-7.9^{+2.6}_{-4.1}$	$-10.2^{+2.8}_{-2.8}$	$-9.4^{+3.1}_{-3.3}$	$-8.0^{+2.2}_{-4.2}$	$-6.5^{+0.9}_{-4.0}$
$L(\mathrm{O_3})$	-6.523	$-9.7^{+2.3}_{-3.0}$	$-6.7^{+0.7}_{-0.7}$	$-6.6^{+0.6}_{-0.6}$	$-5.9^{+0.6}_{-0.6}$	$-7.5^{+0.8}_{-2.7}$	$-6.2^{+0.7}_{-0.7}$	$-6.3^{+0.6}_{-0.6}$	$-6.5^{+0.5}_{-0.4}$
$L(\mathrm{CO})$	-6.903	$-9.0^{+3.9}_{-3.6}$	$-7.5^{+4.4}_{-4.5}$	$-7.8^{+4.5}_{-4.5}$	$-7.9^{+4.5}_{-4.3}$	$-8.3^{+3.9}_{-3.9}$	$-7.5^{+4.4}_{-4.5}$	$-7.7^{+4.4}_{-4.4}$	$-8.1^{+4.3}_{-4.4}$
$L(\mathrm{N_2O})$	-6.495	$-9.8^{+3.3}_{-3.0}$	$-8.9^{+3.7}_{-3.6}$	$-9.0^{+3.6}_{-3.6}$	$-8.3^{+3.4}_{-3.9}$	$-8.8^{+3.4}_{-3.5}$	$-8.8^{+3.8}_{-3.7}$	$-9.4^{+3.8}_{-3.3}$	$-9.4^{+3.3}_{-3.4}$

		R=50				*R*=100			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.3^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$
a_3	23.12	$30.4^{+16.3}_{-13.0}$	$14.1^{+11.0}_{-6.7}$	$11.7^{+9.5}_{-5.1}$	$13.9^{+9.7}_{-6.2}$	$10.7^{+7.6}_{-5.1}$	$14.7^{+10.1}_{-6.4}$	$19.1^{+18.2}_{-8.4}$	$17.0^{+10.6}_{-6.7}$
a_2	99.70	$112.7^{+47.3}_{-41.3}$	$64.4^{+41.8}_{-30.6}$	$43.8^{+25.2}_{-18.1}$	$53.2^{+28.0}_{-19.9}$	$40.3^{+31.5}_{-21.0}$	$67.9^{+38.4}_{-26.6}$	$86.6^{+59.2}_{-37.8}$	$78.0^{+41.7}_{-26.9}$
a_1	146.63	$141.1^{+81.8}_{-54.4}$	$126.8^{+72.6}_{-44.3}$	$69.1^{+32.7}_{-20.7}$	$79.3^{+37.2}_{-27.2}$	$79.7^{+55.1}_{-33.1}$	$128.1^{+72.9}_{-46.1}$	$148.4^{+81.3}_{-72.5}$	$132.2^{+61.8}_{-56.2}$
a_0	285.22	$255.5^{+48.8}_{-41.8}$	$311.1^{+52.0}_{-31.6}$	$260.9^{+23.3}_{-19.7}$	$260.8^{+22.6}_{-19.2}$	$273.8^{+44.5}_{-32.1}$	$306.9^{+56.2}_{-39.7}$	$298.4^{+54.4}_{-42.5}$	$286.7^{+39.6}_{-34.3}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.2^{+0.2}_{-0.3}$	$-0.2^{+0.3}_{-0.3}$	$0.3^{+0.2}_{-0.3}$	$0.3^{+0.3}_{-0.2}$	$0.1^{+0.4}_{-0.4}$	$-0.2^{+0.4}_{-0.3}$	$-0.1^{+0.4}_{-0.3}$	$-0.0^{+0.4}_{-0.2}$
$R\ [R_\oplus]$	1.0	$0.98^{+0.09}_{-0.09}$	$1.02^{+0.06}_{-0.06}$	$1.01^{+0.05}_{-0.04}$	$1.01^{+0.03}_{-0.03}$	$1.01^{+0.08}_{-0.08}$	$1.01^{+0.05}_{-0.04}$	$1.0^{+0.03}_{-0.03}$	$1.0^{+0.02}_{-0.02}$
$L\,(M\ [M_\oplus])$	0.0	$0.1^{+0.2}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-7.8^{+4.2}_{-4.0}$	$-7.7^{+4.7}_{-4.3}$	$-7.8^{+4.6}_{-4.4}$	$-7.1^{+4.2}_{-4.5}$	$-7.3^{+4.5}_{-4.8}$	$-7.4^{+4.6}_{-4.6}$	$-7.5^{+4.6}_{-4.7}$	$-7.1^{+4.5}_{-4.9}$
$L(\mathrm{O_2})$	-0.679	$-7.1^{+4.2}_{-4.2}$	$-7.6^{+4.5}_{-4.6}$	$-7.5^{+4.5}_{-4.6}$	$-7.4^{+4.6}_{-4.7}$	$-7.2^{+4.3}_{-4.6}$	$-7.1^{+4.4}_{-4.7}$	$-7.7^{+4.8}_{-4.5}$	$-7.9^{+5.0}_{-4.2}$
$L(\mathrm{H_2O})$	-3.000	$-3.2^{+0.8}_{-0.7}$	$-2.0^{+0.7}_{-0.7}$	$-3.2^{+0.7}_{-0.6}$	$-3.3^{+0.6}_{-0.6}$	$-2.8^{+1.1}_{-0.9}$	$-2.4^{+0.8}_{-0.8}$	$-2.8^{+0.6}_{-0.7}$	$-2.9^{+0.5}_{-0.6}$
$L(\mathrm{CO_2})$	-3.387	$-4.2^{+0.8}_{-1.0}$	$-2.9^{+0.7}_{-0.7}$	$-3.7^{+0.6}_{-0.5}$	$-3.7^{+0.5}_{-0.5}$	$-3.5^{+1.0}_{-0.9}$	$-2.9^{+0.7}_{-0.7}$	$-3.2^{+0.6}_{-0.6}$	$-3.3^{+0.5}_{-0.5}$
$L(\mathrm{CH_4})$	-5.770	$-9.5^{+2.9}_{-3.0}$	$-8.0^{+2.6}_{-4.4}$	$-6.6^{+0.7}_{-3.3}$	$-6.2^{+0.5}_{-0.6}$	$-8.9^{+2.8}_{-3.8}$	$-5.7^{+0.9}_{-1.1}$	$-5.7^{+0.6}_{-0.7}$	$-5.8^{+0.5}_{-0.6}$
$L(\mathrm{O_3})$	-6.523	$-7.0^{+0.6}_{-0.9}$	$-6.1^{+0.5}_{-0.5}$	$-6.7^{+0.4}_{-0.4}$	$-6.7^{+0.4}_{-0.4}$	$-6.6^{+0.8}_{-0.6}$	$-6.2^{+0.6}_{-0.5}$	$-6.4^{+0.5}_{-0.5}$	$-6.5^{+0.4}_{-0.4}$
$L(\mathrm{CO})$	-6.903	$-7.8^{+4.3}_{-4.1}$	$-8.0^{+4.6}_{-4.4}$	$-8.0^{+4.5}_{-4.2}$	$-8.2^{+4.5}_{-4.0}$	$-7.9^{+4.6}_{-4.6}$	$-7.8^{+4.3}_{-4.4}$	$-8.6^{+4.4}_{-4.0}$	$-8.5^{+4.1}_{-4.1}$
$L(\mathrm{N_2O})$	-6.495	$-9.8^{+3.6}_{-3.0}$	$-8.3^{+3.5}_{-3.9}$	$-9.3^{+3.2}_{-3.4}$	$-9.5^{+3.2}_{-3.5}$	$-8.5^{+3.7}_{-3.9}$	$-9.0^{+3.6}_{-3.6}$	$-9.4^{+3.2}_{-3.5}$	$-9.7^{+3.2}_{-3.2}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Table D.3: Retrieval results for the 4–18.5 μm wavelength range for the nominal case.

		R=20				*R=35*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.3^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$
a_3	23.12	$39.3^{+28.4}_{-22.0}$	$15.1^{+10.1}_{-7.1}$	$14.9^{+13.0}_{-7.4}$	$11.7^{+10.5}_{-5.7}$	$18.9^{+17.0}_{-9.3}$	$12.1^{+10.1}_{-5.9}$	$12.6^{+9.7}_{-5.8}$	$13.5^{+11.1}_{-6.5}$
a_2	99.70	$192.8^{+137.4}_{-90.8}$	$60.8^{+34.4}_{-28.7}$	$59.1^{+46.5}_{-29.8}$	$45.3^{+37.7}_{-22.1}$	$85.3^{+58.3}_{-41.0}$	$53.4^{+36.6}_{-26.3}$	$52.4^{+28.2}_{-24.2}$	$57.6^{+35.7}_{-24.2}$
a_1	146.63	$176.0^{+176.3}_{-90.0}$	$98.3^{+38.2}_{-31.5}$	$99.4^{+60.6}_{-35.8}$	$80.5^{+56.6}_{-31.1}$	$127.6^{+89.4}_{-60.3}$	$103.3^{+53.6}_{-36.5}$	$84.0^{+47.5}_{-32.6}$	$93.0^{+55.1}_{-33.2}$
a_0	285.22	$272.1^{+93.4}_{-107.3}$	$267.3^{+32.8}_{-30.9}$	$280.4^{+40.9}_{-31.4}$	$275.1^{+41.1}_{-33.0}$	$265.1^{+59.5}_{-53.7}$	$291.7^{+44.0}_{-33.7}$	$270.0^{+33.3}_{-28.4}$	$273.6^{+36.6}_{-26.8}$
$L(P_0\,[\mathrm{bar}])$	0.006	$0.1^{+0.5}_{-0.5}$	$0.2^{+0.2}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.1^{+0.4}_{-0.4}$	$0.2^{+0.4}_{-0.4}$	$-0.1^{+0.4}_{-0.4}$	$0.2^{+0.4}_{-0.3}$	$0.1^{+0.3}_{-0.3}$
$R\,[R_\oplus]$	1.0	$0.86^{+0.1}_{-0.09}$	$0.99^{+0.1}_{-0.1}$	$1.01^{+0.07}_{-0.08}$	$1.03^{+0.06}_{-0.06}$	$0.9^{+0.1}_{-0.1}$	$1.01^{+0.08}_{-0.08}$	$1.02^{+0.06}_{-0.06}$	$1.03^{+0.05}_{-0.05}$
$L(M\,[M_\oplus])$	0.0	$0.1^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-7.6^{+4.2}_{-4.2}$	$-7.4^{+4.4}_{-4.4}$	$-7.6^{+4.7}_{-4.5}$	$-7.1^{+4.2}_{-4.7}$	$-7.0^{+4.0}_{-4.3}$	$-7.2^{+4.5}_{-4.7}$	$-7.2^{+4.5}_{-4.9}$	$-7.4^{+4.6}_{-4.7}$
$L(O_2)$	-0.679	$-7.4^{+4.5}_{-4.5}$	$-7.6^{+4.5}_{-4.5}$	$-7.7^{+4.5}_{-4.3}$	$-7.2^{+4.5}_{-4.5}$	$-8.0^{+4.6}_{-4.2}$	$-8.0^{+4.4}_{-4.4}$	$-8.0^{+4.8}_{-4.4}$	$-7.5^{+4.5}_{-4.8}$
$L(H_2O)$	-3.000	$-9.5^{+4.1}_{-3.2}$	$-3.2^{+0.9}_{-0.8}$	$-2.7^{+0.8}_{-0.8}$	$-2.7^{+0.9}_{-0.9}$	$-3.8^{+1.4}_{-1.6}$	$-2.5^{+0.9}_{-0.9}$	$-2.9^{+0.8}_{-0.8}$	$-2.8^{+0.7}_{-0.7}$
$L(CO_2)$	-3.387	$-9.2^{+3.7}_{-3.5}$	$-3.8^{+0.8}_{-0.8}$	$-3.4^{+0.8}_{-0.7}$	$-3.4^{+0.8}_{-0.7}$	$-4.3^{+1.1}_{-1.3}$	$-3.2^{+0.9}_{-0.8}$	$-3.5^{+0.7}_{-0.7}$	$-3.4^{+0.7}_{-0.6}$
$L(CH_4)$	-5.770	$-10.2^{+2.9}_{-2.8}$	$-9.5^{+3.3}_{-3.4}$	$-9.2^{+3.1}_{-3.5}$	$-9.2^{+3.1}_{-3.5}$	$-9.2^{+3.1}_{-3.3}$	$-9.5^{+3.4}_{-3.4}$	$-9.0^{+2.9}_{-3.5}$	$-8.0^{+2.1}_{-4.3}$
$L(O_3)$	-6.523	$-10.5^{+2.6}_{-2.5}$	$-7.0^{+0.8}_{-3.5}$	$-6.4^{+0.6}_{-0.6}$	$-6.5^{+0.6}_{-0.6}$	$-8.5^{+2.0}_{-3.7}$	$-6.3^{+0.7}_{-0.6}$	$-6.6^{+0.5}_{-0.6}$	$-6.6^{+0.5}_{-0.5}$
$L(CO)$	-6.903	$-10.5^{+3.1}_{-2.7}$	$-7.8^{+4.4}_{-4.4}$	$-8.0^{+4.5}_{-4.3}$	$-8.2^{+4.7}_{-4.2}$	$-8.1^{+4.0}_{-3.9}$	$-7.5^{+4.3}_{-4.5}$	$-7.9^{+4.7}_{-4.5}$	$-7.9^{+4.6}_{-4.4}$
$L(N_2O)$	-6.495	$-10.5^{+2.7}_{-2.6}$	$-8.8^{+3.8}_{-3.6}$	$-8.5^{+3.7}_{-3.8}$	$-8.9^{+3.5}_{-3.8}$	$-9.6^{+3.5}_{-3.1}$	$-8.9^{+4.0}_{-3.9}$	$-9.2^{+3.8}_{-3.6}$	$-9.3^{+3.6}_{-3.5}$

		R=50				*R=100*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$17.4^{+12.5}_{-8.8}$	$16.5^{+14.5}_{-8.9}$	$14.4^{+9.1}_{-6.6}$	$14.4^{+12.1}_{-6.5}$	$14.5^{+12.9}_{-7.4}$	$13.5^{+9.5}_{-6.0}$	$14.9^{+10.7}_{-6.1}$	$19.8^{+13.1}_{-8.1}$
a_2	99.70	$77.0^{+54.3}_{-39.5}$	$62.6^{+56.0}_{-29.8}$	$63.4^{+39.7}_{-26.0}$	$60.2^{+32.0}_{-21.5}$	$61.2^{+44.3}_{-29.6}$	$59.6^{+37.7}_{-23.1}$	$67.2^{+32.6}_{-24.4}$	$86.9^{+35.4}_{-27.1}$
a_1	146.63	$127.5^{+93.3}_{-56.3}$	$100.0^{+78.3}_{-38.0}$	$109.7^{+85.1}_{-44.5}$	$90.2^{+46.2}_{-29.3}$	$105.6^{+68.9}_{-40.4}$	$109.0^{+61.0}_{-51.3}$	$107.0^{+60.5}_{-43.6}$	$135.3^{+47.2}_{-43.5}$
a_0	285.22	$286.7^{+57.0}_{-47.0}$	$280.9^{+39.3}_{-30.5}$	$286.4^{+58.7}_{-35.9}$	$264.2^{+30.3}_{-20.0}$	$287.3^{+46.9}_{-37.2}$	$286.0^{+41.3}_{-41.4}$	$276.4^{+38.4}_{-30.8}$	$284.5^{+31.4}_{-26.5}$
$L(P_0\,[\mathrm{bar}])$	0.006	$0.0^{+0.4}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.4}_{-0.4}$	$0.2^{+0.2}_{-0.3}$	$-0.0^{+0.4}_{-0.3}$	$-0.0^{+0.5}_{-0.3}$	$0.1^{+0.4}_{-0.3}$	$0.0^{+0.2}_{-0.2}$
$R\,[R_\oplus]$	1.0	$1.0^{+0.1}_{-0.1}$	$1.02^{+0.07}_{-0.07}$	$1.03^{+0.05}_{-0.05}$	$1.02^{+0.04}_{-0.04}$	$1.00^{+0.08}_{-0.08}$	$1.03^{+0.05}_{-0.05}$	$1.01^{+0.04}_{-0.04}$	$1.00^{+0.03}_{-0.03}$
$L(M\,[M_\oplus])$	0.0	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-7.1^{+4.3}_{-4.5}$	$-7.5^{+4.6}_{-4.7}$	$-7.4^{+4.5}_{-4.6}$	$-7.5^{+4.6}_{-4.7}$	$-7.5^{+4.7}_{-4.7}$	$-7.4^{+4.5}_{-4.6}$	$-7.5^{+4.5}_{-4.7}$	$-7.6^{+4.8}_{-4.5}$
$L(O_2)$	-0.679	$-7.6^{+4.4}_{-4.3}$	$-7.8^{+4.6}_{-4.2}$	$-7.5^{+4.5}_{-4.6}$	$-7.4^{+4.7}_{-4.9}$	$-7.5^{+4.6}_{-4.5}$	$-7.2^{+4.4}_{-4.7}$	$-7.5^{+4.7}_{-4.5}$	$-7.9^{+4.9}_{-4.6}$
$L(H_2O)$	-3.000	$-3.1^{+1.0}_{-1.1}$	$-2.8^{+0.7}_{-0.8}$	$-2.6^{+0.8}_{-0.8}$	$-3.1^{+0.6}_{-0.6}$	$-2.8^{+0.9}_{-0.9}$	$-2.7^{+0.7}_{-0.9}$	$-3.0^{+0.6}_{-0.7}$	$-2.9^{+0.5}_{-0.5}$
$L(CO_2)$	-3.387	$-3.7^{+0.9}_{-0.9}$	$-3.3^{+0.7}_{-0.7}$	$-3.1^{+0.8}_{-0.7}$	$-3.6^{+0.6}_{-0.5}$	$-3.3^{+0.8}_{-0.8}$	$-3.2^{+0.6}_{-0.8}$	$-3.4^{+0.5}_{-0.6}$	$-3.3^{+0.4}_{-0.5}$
$L(CH_4)$	-5.770	$-8.8^{+3.0}_{-3.5}$	$-9.3^{+3.1}_{-3.5}$	$-7.9^{+2.3}_{-4.2}$	$-6.9^{+1.1}_{-4.3}$	$-9.2^{+3.2}_{-3.4}$	$-7.9^{+2.3}_{-4.2}$	$-6.3^{+0.9}_{-2.0}$	$-5.8^{+0.6}_{-0.6}$
$L(O_3)$	-6.523	$-7.2^{+1.1}_{-4.3}$	$-6.5^{+0.5}_{-0.5}$	$-6.4^{+0.6}_{-0.6}$	$-6.7^{+0.4}_{-0.4}$	$-6.5^{+0.7}_{-0.7}$	$-6.5^{+0.5}_{-0.6}$	$-6.6^{+0.4}_{-0.5}$	$-6.5^{+0.4}_{-0.4}$
$L(CO)$	-6.903	$-7.6^{+4.4}_{-4.2}$	$-7.5^{+4.5}_{-4.6}$	$-8.0^{+4.8}_{-4.4}$	$-7.9^{+4.6}_{-4.3}$	$-7.7^{+4.5}_{-4.3}$	$-7.8^{+4.4}_{-4.5}$	$-8.0^{+4.4}_{-4.3}$	$-8.3^{+4.3}_{-4.2}$
$L(N_2O)$	-6.495	$-8.3^{+3.8}_{-3.7}$	$-8.9^{+3.9}_{-3.6}$	$-8.6^{+3.5}_{-3.8}$	$-9.0^{+3.2}_{-3.7}$	$-8.5^{+4.0}_{-4.0}$	$-8.8^{+3.6}_{-3.6}$	$-9.2^{+3.3}_{-3.6}$	$-9.4^{+3.3}_{-3.4}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Table D.4: Retrieval results for the 4–18.5 µm wavelength range for the optimized case.

		R=20				*R*=35			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.3}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$
a_3	23.12	$22.6^{+18.3}_{-11.5}$	$13.1^{+9.8}_{-6.3}$	$13.0^{+7.8}_{-6.1}$	$11.7^{+8.5}_{-5.6}$	$16.4^{+15.1}_{-8.0}$	$10.8^{+9.2}_{-5.4}$	$12.2^{+10.8}_{-5.8}$	$12.1^{+11.2}_{-5.4}$
a_2	99.70	$101.3^{+51.4}_{-42.6}$	$61.5^{+42.6}_{-31.9}$	$51.7^{+34.6}_{-26.4}$	$49.0^{+32.9}_{-25.0}$	$71.5^{+49.6}_{-34.7}$	$39.5^{+29.5}_{-20.3}$	$55.6^{+43.9}_{-27.4}$	$47.8^{+39.1}_{-20.3}$
a_1	146.63	$135.8^{+57.9}_{-56.7}$	$132.1^{+62.6}_{-47.6}$	$94.7^{+71.5}_{-39.4}$	$94.7^{+61.3}_{-32.5}$	$113.8^{+53.7}_{-44.9}$	$65.7^{+36.2}_{-23.4}$	$112.4^{+81.2}_{-45.4}$	$89.8^{+47.5}_{-40.3}$
a_0	285.22	$211.1^{+60.6}_{-79.7}$	$304.6^{+45.6}_{-36.4}$	$283.3^{+55.6}_{-37.6}$	$290.9^{+48.1}_{-28.4}$	$246.6^{+56.3}_{-67.9}$	$259.4^{+30.6}_{-28.2}$	$310.0^{+55.4}_{-44.4}$	$275.8^{+29.8}_{-28.1}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.5^{+0.3}_{-0.3}$	$-0.1^{+0.3}_{-0.3}$	$0.1^{+0.5}_{-0.4}$	$-0.0^{+0.3}_{-0.4}$	$0.4^{+0.3}_{-0.4}$	$0.4^{+0.3}_{-0.4}$	$-0.2^{+0.5}_{-0.4}$	$0.1^{+0.4}_{-0.3}$
$R\ [R_\oplus]$	1.0	$0.9^{+0.1}_{-0.1}$	$0.98^{+0.09}_{-0.09}$	$0.99^{+0.07}_{-0.07}$	$1.01^{+0.06}_{-0.05}$	$0.9^{+0.1}_{-0.1}$	$1.00^{+0.07}_{-0.08}$	$1.01^{+0.05}_{-0.05}$	$1.01^{+0.04}_{-0.04}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-6.6^{+4.0}_{-4.4}$	$-7.2^{+4.1}_{-4.6}$	$-7.0^{+4.1}_{-4.6}$	$-7.1^{+4.3}_{-4.7}$	$-8.0^{+4.3}_{-4.1}$	$-7.8^{+4.6}_{-4.4}$	$-7.8^{+4.7}_{-4.4}$	$-7.6^{+4.6}_{-4.6}$
$L(\mathrm{O_2})$	-0.679	$-7.3^{+4.4}_{-4.4}$	$-7.8^{+4.4}_{-4.4}$	$-7.9^{+4.5}_{-4.2}$	$-7.3^{+4.3}_{-4.5}$	$-7.7^{+4.6}_{-4.5}$	$-7.3^{+4.2}_{-4.5}$	$-7.5^{+4.5}_{-4.5}$	$-7.3^{+4.6}_{-4.8}$
$L(\mathrm{H_2O})$	-3.000	$-3.9^{+0.9}_{-1.4}$	$-2.1^{+0.9}_{-0.9}$	$-2.5^{+1.1}_{-0.9}$	$-2.2^{+0.8}_{-0.8}$	$-3.3^{+1.1}_{-0.8}$	$-3.1^{+0.9}_{-0.8}$	$-2.0^{+0.9}_{-0.9}$	$-2.7^{+0.7}_{-0.8}$
$L(\mathrm{CO_2})$	-3.387	$-6.3^{+1.9}_{-4.3}$	$-3.5^{+1.0}_{-1.1}$	$-3.5^{+1.0}_{-0.8}$	$-3.1^{+0.7}_{-0.8}$	$-5.0^{+1.4}_{-3.1}$	$-3.9^{+0.9}_{-0.8}$	$-2.8^{+0.8}_{-0.9}$	$-3.4^{+0.7}_{-0.7}$
$L(\mathrm{CH_4})$	-5.770	$-9.6^{+2.8}_{-3.1}$	$-9.5^{+3.3}_{-3.2}$	$-9.4^{+3.0}_{-3.4}$	$-8.8^{+2.9}_{-3.7}$	$-10.1^{+2.9}_{-2.9}$	$-9.9^{+3.0}_{-3.1}$	$-7.8^{+2.5}_{-4.4}$	$-6.7^{+1.1}_{-3.9}$
$L(\mathrm{O_3})$	-6.523	$-9.7^{+2.4}_{-3.1}$	$-6.2^{+0.7}_{-0.7}$	$-6.5^{+0.7}_{-0.6}$	$-6.2^{+0.6}_{-0.6}$	$-7.2^{+0.9}_{-2.1}$	$-6.8^{+0.6}_{-0.6}$	$-6.0^{+0.7}_{-0.7}$	$-6.5^{+0.5}_{-0.5}$
$L(\mathrm{CO})$	-6.903	$-8.9^{+3.9}_{-3.6}$	$-7.8^{+4.6}_{-4.5}$	$-7.2^{+4.3}_{-4.4}$	$-7.7^{+4.3}_{-4.5}$	$-8.2^{+4.5}_{-4.1}$	$-8.1^{+4.7}_{-4.3}$	$-7.8^{+4.6}_{-4.5}$	$-8.0^{+4.3}_{-4.3}$
$L(\mathrm{N_2O})$	-6.495	$-9.7^{+3.5}_{-3.1}$	$-8.2^{+3.8}_{-4.0}$	$-8.4^{+3.7}_{-3.8}$	$-8.5^{+3.6}_{-3.8}$	$-8.7^{+3.6}_{-3.5}$	$-8.9^{+3.4}_{-3.6}$	$-9.3^{+3.7}_{-3.5}$	$-9.2^{+3.2}_{-3.4}$

		R=50				*R*=100			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$13.1^{+10.4}_{-6.5}$	$11.5^{+7.8}_{-5.5}$	$13.9^{+10.5}_{-5.8}$	$15.7^{+13.0}_{-6.7}$	$13.2^{+9.8}_{-6.3}$	$17.0^{+13.7}_{-7.5}$	$16.3^{+13.5}_{-7.2}$	$19.7^{+14.4}_{-8.7}$
a_2	99.70	$59.2^{+44.2}_{-31.6}$	$36.8^{+22.5}_{-17.5}$	$62.0^{+42.8}_{-25.8}$	$74.7^{+44.0}_{-28.0}$	$53.6^{+36.4}_{-26.3}$	$81.4^{+57.7}_{-35.6}$	$73.6^{+41.0}_{-27.1}$	$84.6^{+40.5}_{-27.7}$
a_1	146.63	$125.9^{+69.8}_{-51.9}$	$60.8^{+29.0}_{-20.1}$	$113.3^{+92.3}_{-41.9}$	$130.1^{+68.4}_{-48.3}$	$97.0^{+55.4}_{-33.6}$	$156.6^{+101.0}_{-70.0}$	$118.9^{+62.2}_{-43.0}$	$132.7^{+51.0}_{-47.1}$
a_0	285.22	$298.7^{+56.7}_{-49.8}$	$257.1^{+24.7}_{-20.7}$	$299.9^{+63.9}_{-37.3}$	$298.7^{+40.7}_{-37.0}$	$280.5^{+41.7}_{-31.4}$	$325.9^{+71.9}_{-53.1}$	$282.2^{+43.2}_{-28.6}$	$283.1^{+31.3}_{-27.3}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$-0.1^{+0.4}_{-0.4}$	$0.4^{+0.3}_{-0.3}$	$-0.1^{+0.4}_{-0.4}$	$-0.1^{+0.3}_{-0.2}$	$0.1^{+0.3}_{-0.3}$	$-0.3^{+0.4}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.2}$
$R\ [R_\oplus]$	1.0	$1.0^{+0.1}_{-0.1}$	$1.01^{+0.07}_{-0.06}$	$1.01^{+0.05}_{-0.05}$	$1.00^{+0.04}_{-0.03}$	$1.00^{+0.08}_{-0.08}$	$1.01^{+0.05}_{-0.05}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.02}_{-0.02}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-7.6^{+4.3}_{-4.2}$	$-7.6^{+4.5}_{-4.3}$	$-7.3^{+4.3}_{-4.3}$	$-7.5^{+4.6}_{-4.6}$	$-7.9^{+4.7}_{-4.4}$	$-7.6^{+4.6}_{-4.8}$	$-7.5^{+4.6}_{-4.4}$	$-7.4^{+4.7}_{-4.6}$
$L(\mathrm{O_2})$	-0.679	$-7.7^{+4.5}_{-4.3}$	$-7.4^{+4.4}_{-4.6}$	$-7.4^{+4.5}_{-4.6}$	$-7.2^{+4.4}_{-4.7}$	$-7.1^{+4.3}_{-4.7}$	$-8.1^{+4.9}_{-4.3}$	$-7.4^{+4.5}_{-4.7}$	$-7.8^{+4.9}_{-4.7}$
$L(\mathrm{H_2O})$	-3.000	$-2.4^{+1.0}_{-1.1}$	$-3.1^{+0.8}_{-0.7}$	$-2.3^{+0.8}_{-0.9}$	$-2.7^{+0.6}_{-0.7}$	$-2.6^{+0.8}_{-0.8}$	$-2.2^{+0.8}_{-0.8}$	$-2.9^{+0.6}_{-0.6}$	$-3.0^{+0.5}_{-0.5}$
$L(\mathrm{CO_2})$	-3.387	$-3.6^{+1.1}_{-1.3}$	$-3.8^{+0.7}_{-0.6}$	$-3.0^{+0.8}_{-0.7}$	$-3.1^{+0.5}_{-0.6}$	$-3.5^{+0.8}_{-0.8}$	$-2.8^{+0.7}_{-0.7}$	$-3.3^{+0.6}_{-0.6}$	$-3.4^{+0.5}_{-0.5}$
$L(\mathrm{CH_4})$	-5.770	$-9.9^{+3.2}_{-3.1}$	$-9.1^{+2.6}_{-3.7}$	$-6.2^{+1.2}_{-3.9}$	$-5.7^{+0.6}_{-0.7}$	$-8.4^{+2.5}_{-3.8}$	$-5.5^{+0.9}_{-1.1}$	$-5.8^{+0.6}_{-0.6}$	$-5.8^{+0.5}_{-0.5}$
$L(\mathrm{O_3})$	-6.523	$-6.4^{+0.8}_{-1.0}$	$-6.8^{+0.5}_{-0.5}$	$-6.2^{+0.6}_{-0.6}$	$-6.3^{+0.4}_{-0.5}$	$-6.5^{+0.6}_{-0.6}$	$-6.1^{+0.6}_{-0.5}$	$-6.5^{+0.5}_{-0.4}$	$-6.5^{+0.4}_{-0.4}$
$L(\mathrm{CO})$	-6.903	$-8.2^{+4.5}_{-3.9}$	$-8.0^{+4.6}_{-4.2}$	$-8.0^{+4.5}_{-4.4}$	$-8.3^{+4.6}_{-4.1}$	$-7.4^{+4.6}_{-4.4}$	$-7.8^{+4.3}_{-4.5}$	$-7.9^{+4.3}_{-4.4}$	$-8.7^{+4.2}_{-4.0}$
$L(\mathrm{N_2O})$	-6.495	$-8.5^{+3.9}_{-3.9}$	$-9.2^{+3.4}_{-3.3}$	$-9.0^{+3.5}_{-3.5}$	$-9.3^{+3.2}_{-3.5}$	$-8.7^{+3.7}_{-3.7}$	$-9.0^{+3.7}_{-3.7}$	$-9.5^{+3.2}_{-3.4}$	$-9.9^{+3.1}_{-3.2}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Table D.5: Retrieval results for the 6–17 μm wavelength range for the nominal case.

		R=20				*R=35*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.3}_{-0.3}$	$1.2^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.2^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$43.0^{+29.6}_{-23.8}$	$30.0^{+29.9}_{-17.0}$	$17.9^{+17.6}_{-9.1}$	$15.4^{+11.3}_{-7.3}$	$35.5^{+26.6}_{-18.3}$	$15.5^{+11.6}_{-7.5}$	$16.8^{+16.3}_{-8.7}$	$16.8^{+14.4}_{-7.8}$
a_2	99.70	$167.4^{+96.9}_{-77.4}$	$120.6^{+96.3}_{-62.6}$	$71.6^{+50.3}_{-33.6}$	$59.6^{+31.5}_{-25.6}$	$130.7^{+97.7}_{-62.0}$	$63.4^{+34.7}_{-27.9}$	$61.4^{+57.6}_{-28.2}$	$63.7^{+41.9}_{-26.0}$
a_1	146.63	$148.2^{+122.3}_{-79.8}$	$118.5^{+71.1}_{-47.4}$	$101.1^{+59.8}_{-37.0}$	$80.6^{+41.1}_{-28.1}$	$103.7^{+117.6}_{-54.7}$	$96.6^{+40.7}_{-32.6}$	$84.1^{+81.8}_{-34.6}$	$86.8^{+55.1}_{-38.3}$
a_0	285.22	$217.6^{+92.3}_{-88.8}$	$235.7^{+47.4}_{-70.0}$	$269.6^{+37.1}_{-34.3}$	$258.0^{+26.7}_{-26.5}$	$213.8^{+53.4}_{-61.3}$	$269.6^{+35.4}_{-32.5}$	$261.2^{+45.5}_{-27.0}$	$258.2^{+29.4}_{-22.8}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.4^{+0.3}_{-0.4}$	$0.3^{+0.2}_{-0.3}$	$0.2^{+0.3}_{-0.3}$	$0.3^{+0.2}_{-0.3}$	$0.5^{+0.2}_{-0.3}$	$0.2^{+0.3}_{-0.3}$	$0.2^{+0.3}_{-0.4}$	$0.3^{+0.3}_{-0.3}$
$R\ [R_\oplus]$	1.0	$0.9^{+0.1}_{-0.1}$	$1.01^{+0.1}_{-0.09}$	$1.00^{+0.09}_{-0.08}$	$1.01^{+0.07}_{-0.07}$	$0.9^{+0.1}_{-0.1}$	$0.99^{+0.09}_{-0.08}$	$1.01^{+0.07}_{-0.07}$	$1.01^{+0.06}_{-0.06}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-7.5^{+4.2}_{-4.3}$	$-7.4^{+4.3}_{-4.4}$	$-7.4^{+4.4}_{-4.5}$	$-7.5^{+4.8}_{-4.6}$	$-8.1^{+4.4}_{-4.1}$	$-7.6^{+4.7}_{-4.3}$	$-7.1^{+4.4}_{-4.8}$	$-7.6^{+4.7}_{-4.8}$
$L(O_2)$	-0.679	$-7.7^{+4.5}_{-4.1}$	$-7.5^{+4.4}_{-4.1}$	$-7.2^{+4.4}_{-4.5}$	$-7.8^{+4.8}_{-4.7}$	$-6.8^{+4.2}_{-4.9}$	$-7.3^{+4.4}_{-4.5}$	$-7.7^{+4.7}_{-4.6}$	$-7.6^{+4.7}_{-4.6}$
$L(H_2O)$	-3.000	$-9.5^{+3.3}_{-3.1}$	$-5.5^{+2.2}_{-5.1}$	$-3.2^{+0.9}_{-1.1}$	$-3.2^{+0.8}_{-0.8}$	$-8.6^{+3.4}_{-3.6}$	$-3.2^{+1.0}_{-1.2}$	$-3.2^{+0.9}_{-0.8}$	$-3.2^{+0.7}_{-0.7}$
$L(CO_2)$	-3.387	$-5.1^{+1.0}_{-2.2}$	$-4.0^{+0.6}_{-0.5}$	$-3.6^{+0.7}_{-0.6}$	$-3.8^{+0.6}_{-0.6}$	$-4.5^{+0.6}_{-0.7}$	$-3.6^{+0.8}_{-0.7}$	$-3.7^{+0.7}_{-0.7}$	$-3.8^{+0.6}_{-0.6}$
$L(CH_4)$	-5.770	$-9.6^{+2.8}_{-3.1}$	$-9.1^{+3.1}_{-3.5}$	$-8.4^{+2.8}_{-3.6}$	$-9.3^{+3.0}_{-3.5}$	$-9.6^{+3.1}_{-3.2}$	$-8.9^{+3.2}_{-3.9}$	$-9.2^{+3.0}_{-3.5}$	$-9.2^{+2.9}_{-3.5}$
$L(O_3)$	-6.523	$-10.4^{+2.7}_{-2.7}$	$-7.8^{+1.1}_{-3.5}$	$-6.7^{+0.6}_{-0.6}$	$-6.8^{+0.5}_{-0.5}$	$-9.6^{+2.3}_{-3.1}$	$-6.7^{+0.7}_{-0.7}$	$-6.8^{+0.6}_{-0.5}$	$-6.8^{+0.5}_{-0.4}$
$L(CO)$	-6.903	$-7.8^{+4.4}_{-4.1}$	$-7.1^{+4.5}_{-4.4}$	$-7.8^{+4.4}_{-4.3}$	$-7.4^{+4.7}_{-4.8}$	$-7.7^{+4.4}_{-4.2}$	$-7.4^{+4.4}_{-4.4}$	$-7.6^{+4.7}_{-4.8}$	$-7.4^{+4.7}_{-4.7}$
$L(N_2O)$	-6.495	$-9.7^{+3.0}_{-3.0}$	$-9.4^{+3.3}_{-3.3}$	$-8.9^{+3.9}_{-3.7}$	$-8.5^{+3.6}_{-4.0}$	$-10.0^{+2.9}_{-2.9}$	$-8.7^{+3.8}_{-3.7}$	$-9.0^{+3.7}_{-3.8}$	$-8.9^{+3.6}_{-3.7}$

		R=50				*R=100*			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.2^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.3}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$33.5^{+22.8}_{-17.3}$	$15.0^{+14.3}_{-7.6}$	$16.3^{+13.6}_{-7.8}$	$15.5^{+12.2}_{-6.6}$	$17.8^{+18.9}_{-9.1}$	$16.2^{+13.6}_{-7.7}$	$19.0^{+16.5}_{-8.5}$	$19.4^{+15.5}_{-8.7}$
a_2	99.70	$127.0^{+75.9}_{-57.5}$	$62.9^{+39.9}_{-28.5}$	$63.1^{+48.4}_{-28.5}$	$66.3^{+30.7}_{-24.3}$	$70.1^{+43.5}_{-31.0}$	$67.0^{+35.9}_{-26.8}$	$78.3^{+52.3}_{-26.8}$	$81.7^{+48.2}_{-30.0}$
a_1	146.63	$118.8^{+109.9}_{-58.6}$	$94.5^{+47.3}_{-31.0}$	$92.7^{+82.6}_{-48.3}$	$92.8^{+50.9}_{-33.7}$	$102.3^{+45.1}_{-35.8}$	$98.4^{+56.9}_{-39.0}$	$126.9^{+60.6}_{-54.5}$	$119.9^{+72.1}_{-50.9}$
a_0	285.22	$232.6^{+49.7}_{-50.4}$	$271.5^{+34.5}_{-30.5}$	$266.7^{+48.3}_{-32.9}$	$261.4^{+32.4}_{-22.0}$	$269.0^{+37.2}_{-32.6}$	$271.0^{+36.1}_{-29.7}$	$281.0^{+35.3}_{-32.5}$	$275.9^{+43.7}_{-30.8}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.3^{+0.2}_{-0.3}$	$0.1^{+0.3}_{-0.3}$	$0.2^{+0.5}_{-0.4}$	$0.2^{+0.3}_{-0.3}$	$0.1^{+0.3}_{-0.3}$	$0.1^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.2}$	$0.1^{+0.3}_{-0.3}$
$R\ [R_\oplus]$	1.0	$1.0^{+0.1}_{-0.1}$	$1.01^{+0.08}_{-0.08}$	$1.01^{+0.06}_{-0.06}$	$1.02^{+0.05}_{-0.05}$	$1.0^{+0.1}_{-0.1}$	$1.01^{+0.07}_{-0.06}$	$1.02^{+0.05}_{-0.05}$	$1.01^{+0.04}_{-0.04}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(N_2)$	-0.107	$-8.0^{+4.5}_{-4.1}$	$-7.4^{+4.6}_{-4.8}$	$-7.6^{+4.7}_{-4.5}$	$-7.5^{+4.6}_{-4.6}$	$-7.4^{+4.3}_{-4.4}$	$-7.3^{+4.5}_{-4.7}$	$-7.7^{+4.9}_{-4.7}$	$-7.4^{+4.7}_{-4.7}$
$L(O_2)$	-0.679	$-7.3^{+4.4}_{-4.4}$	$-7.4^{+4.4}_{-4.5}$	$-7.6^{+4.6}_{-4.6}$	$-7.7^{+4.9}_{-4.7}$	$-7.4^{+4.5}_{-4.6}$	$-7.7^{+4.7}_{-4.5}$	$-7.3^{+4.5}_{-4.6}$	$-7.5^{+4.8}_{-4.6}$
$L(H_2O)$	-3.000	$-7.5^{+3.2}_{-4.5}$	$-3.0^{+0.9}_{-0.9}$	$-3.1^{+0.8}_{-0.9}$	$-3.2^{+0.7}_{-0.6}$	$-3.5^{+1.1}_{-1.6}$	$-3.0^{+0.7}_{-0.7}$	$-2.9^{+0.6}_{-0.6}$	$-3.0^{+0.6}_{-0.6}$
$L(CO_2)$	-3.387	$-4.1^{+0.6}_{-0.6}$	$-3.5^{+0.7}_{-0.7}$	$-3.6^{+0.7}_{-0.8}$	$-3.6^{+0.6}_{-0.6}$	$-3.6^{+0.7}_{-0.6}$	$-3.5^{+0.7}_{-0.6}$	$-3.4^{+0.5}_{-0.6}$	$-3.5^{+0.6}_{-0.6}$
$L(CH_4)$	-5.770	$-8.8^{+2.9}_{-3.5}$	$-9.1^{+3.1}_{-3.6}$	$-9.3^{+3.1}_{-3.5}$	$-8.2^{+2.2}_{-3.9}$	$-9.0^{+3.1}_{-3.5}$	$-8.4^{+2.5}_{-4.1}$	$-7.1^{+1.4}_{-4.5}$	$-6.3^{+0.9}_{-2.6}$
$L(O_3)$	-6.523	$-8.4^{+1.5}_{-3.7}$	$-6.6^{+0.6}_{-0.5}$	$-6.7^{+0.6}_{-0.6}$	$-6.7^{+0.5}_{-0.5}$	$-6.8^{+0.7}_{-1.3}$	$-6.6^{+0.5}_{-0.5}$	$-6.6^{+0.5}_{-0.4}$	$-6.6^{+0.4}_{-0.5}$
$L(CO)$	-6.903	$-8.0^{+4.7}_{-4.1}$	$-7.4^{+4.5}_{-4.7}$	$-7.4^{+4.5}_{-4.7}$	$-7.9^{+4.6}_{-4.4}$	$-7.2^{+4.4}_{-4.5}$	$-7.7^{+4.8}_{-4.6}$	$-7.2^{+4.4}_{-4.6}$	$-7.2^{+4.5}_{-4.8}$
$L(N_2O)$	-6.495	$-9.5^{+3.0}_{-3.1}$	$-8.9^{+3.9}_{-3.8}$	$-8.8^{+3.8}_{-3.6}$	$-9.3^{+3.6}_{-3.5}$	$-8.5^{+3.5}_{-3.8}$	$-8.5^{+3.6}_{-3.9}$	$-9.2^{+3.8}_{-3.6}$	$-9.3^{+3.6}_{-3.6}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Table D.6: Retrieval results for the 6–17 μm wavelength range for the optimized case.

		R=20				*R*=35			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.3^{+0.3}_{-0.3}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$32.1^{+27.0}_{-16.6}$	$14.7^{+11.3}_{-7.7}$	$11.4^{+8.5}_{-5.5}$	$12.6^{+9.3}_{-5.9}$	$18.4^{+12.0}_{-9.5}$	$13.3^{+10.0}_{-6.7}$	$12.2^{+10.3}_{-5.9}$	$14.3^{+10.4}_{-6.9}$
a_2	99.70	$113.7^{+76.8}_{-48.6}$	$64.6^{+43.4}_{-35.0}$	$46.9^{+33.7}_{-23.7}$	$54.6^{+38.5}_{-26.9}$	$80.2^{+41.0}_{-36.5}$	$55.2^{+32.8}_{-28.9}$	$49.5^{+32.5}_{-22.9}$	$54.5^{+37.8}_{-23.4}$
a_1	146.63	$114.4^{+79.8}_{-57.1}$	$114.1^{+71.8}_{-44.2}$	$87.6^{+55.1}_{-31.7}$	$110.4^{+63.1}_{-37.0}$	$125.6^{+64.4}_{-51.4}$	$97.7^{+55.9}_{-36.3}$	$91.4^{+48.6}_{-36.8}$	$88.7^{+65.9}_{-31.1}$
a_0	285.22	$194.6^{+68.0}_{-72.8}$	$277.5^{+53.6}_{-44.3}$	$276.4^{+46.8}_{-30.8}$	$310.2^{+46.0}_{-35.3}$	$247.1^{+47.8}_{-59.5}$	$281.9^{+44.8}_{-34.2}$	$283.4^{+38.2}_{-31.7}$	$277.0^{+42.8}_{-26.1}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.5^{+0.3}_{-0.4}$	$0.1^{+0.3}_{-0.4}$	$0.2^{+0.3}_{-0.4}$	$-0.2^{+0.4}_{-0.3}$	$0.3^{+0.3}_{-0.3}$	$0.1^{+0.3}_{-0.4}$	$0.1^{+0.4}_{-0.3}$	$0.1^{+0.3}_{-0.4}$
$R\ [R_\oplus]$	1.0	$1.0^{+0.1}_{-0.1}$	$0.96^{+0.1}_{-0.09}$	$0.99^{+0.07}_{-0.08}$	$0.99^{+0.06}_{-0.06}$	$0.9^{+0.1}_{-0.1}$	$0.98^{+0.08}_{-0.08}$	$1.00^{+0.06}_{-0.06}$	$1.00^{+0.04}_{-0.05}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.1^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-7.7^{+4.2}_{-4.2}$	$-7.4^{+4.5}_{-4.4}$	$-7.4^{+4.4}_{-4.6}$	$-7.5^{+4.4}_{-4.7}$	$-7.4^{+4.5}_{-4.4}$	$-7.2^{+4.4}_{-4.6}$	$-7.8^{+4.7}_{-4.7}$	$-7.4^{+4.7}_{-4.6}$
$L(\mathrm{O_2})$	-0.679	$-7.0^{+4.8}_{-4.6}$	$-7.5^{+4.4}_{-4.4}$	$-7.4^{+4.4}_{-4.5}$	$-7.9^{+4.5}_{-4.3}$	$-6.7^{+4.3}_{-4.8}$	$-7.4^{+4.6}_{-4.5}$	$-7.8^{+4.7}_{-4.5}$	$-7.6^{+4.5}_{-4.5}$
$L(\mathrm{H_2O})$	-3.000	$-5.1^{+1.5}_{-5.5}$	$-2.7^{+1.0}_{-0.9}$	$-2.6^{+1.0}_{-0.8}$	$-1.9^{+0.8}_{-0.8}$	$-3.3^{+0.9}_{-1.0}$	$-2.4^{+0.8}_{-0.9}$	$-2.5^{+0.8}_{-0.8}$	$-2.6^{+0.8}_{-0.7}$
$L(\mathrm{CO_2})$	-3.387	$-5.6^{+1.5}_{-3.0}$	$-4.0^{+1.0}_{-1.1}$	$-3.7^{+0.9}_{-0.8}$	$-3.0^{+0.7}_{-0.8}$	$-4.8^{+1.2}_{-2.0}$	$-3.5^{+0.8}_{-0.8}$	$-3.4^{+0.8}_{-0.7}$	$-3.4^{+0.7}_{-0.7}$
$L(\mathrm{CH_4})$	-5.770	$-9.9^{+2.7}_{-2.9}$	$-9.9^{+3.1}_{-2.9}$	$-9.5^{+2.9}_{-3.4}$	$-9.0^{+3.1}_{-3.6}$	$-9.6^{+2.8}_{-3.0}$	$-9.5^{+3.0}_{-3.4}$	$-9.1^{+2.8}_{-3.5}$	$-7.2^{+1.4}_{-4.2}$
$L(\mathrm{O_3})$	-6.523	$-10.6^{+2.7}_{-2.6}$	$-6.6^{+0.7}_{-0.7}$	$-6.5^{+0.6}_{-0.6}$	$-6.0^{+0.6}_{-0.6}$	$-7.2^{+0.8}_{-3.1}$	$-6.4^{+0.6}_{-0.6}$	$-6.4^{+0.6}_{-0.6}$	$-6.5^{+0.5}_{-0.5}$
$L(\mathrm{CO})$	-6.903	$-7.5^{+4.3}_{-4.3}$	$-7.2^{+4.2}_{-4.5}$	$-7.2^{+4.3}_{-4.7}$	$-7.5^{+4.3}_{-4.5}$	$-6.7^{+4.0}_{-4.6}$	$-7.5^{+4.6}_{-4.6}$	$-7.5^{+4.4}_{-4.7}$	$-7.8^{+4.7}_{-4.5}$
$L(\mathrm{N_2O})$	-6.495	$-10.1^{+3.1}_{-2.8}$	$-8.7^{+4.0}_{-3.8}$	$-8.9^{+3.8}_{-3.7}$	$-8.7^{+3.7}_{-3.7}$	$-8.7^{+3.6}_{-3.6}$	$-8.9^{+3.8}_{-3.8}$	$-9.1^{+3.5}_{-3.6}$	$-9.4^{+3.4}_{-3.4}$

		R=50				*R*=100			
Parameter	Input	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20	*S/N*=5	*S/N*=10	*S/N*=15	*S/N*=20
$\sqrt[4]{a_4}$	1.14	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.0^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$	$1.1^{+0.2}_{-0.2}$
a_3	23.12	$16.4^{+13.3}_{-8.1}$	$14.0^{+10.3}_{-7.1}$	$13.0^{+12.7}_{-6.5}$	$14.9^{+11.3}_{-6.8}$	$12.5^{+10.9}_{-6.0}$	$15.0^{+11.7}_{-7.5}$	$19.4^{+18.3}_{-8.9}$	$21.8^{+16.8}_{-9.8}$
a_2	99.70	$76.1^{+42.9}_{-37.6}$	$57.8^{+43.0}_{-31.1}$	$49.6^{+40.7}_{-23.1}$	$67.8^{+36.7}_{-30.5}$	$56.3^{+38.3}_{-28.2}$	$68.5^{+53.1}_{-33.7}$	$88.9^{+66.8}_{-37.4}$	$93.1^{+56.3}_{-32.7}$
a_1	146.63	$130.4^{+65.2}_{-49.4}$	$105.9^{+77.0}_{-43.8}$	$84.5^{+56.4}_{-32.7}$	$119.0^{+60.0}_{-59.8}$	$99.9^{+66.8}_{-37.2}$	$126.8^{+106.8}_{-59.3}$	$161.8^{+94.6}_{-83.2}$	$149.2^{+80.3}_{-59.5}$
a_0	285.22	$268.3^{+44.3}_{-45.0}$	$293.2^{+53.4}_{-34.0}$	$274.2^{+35.7}_{-27.1}$	$286.0^{+47.4}_{-37.7}$	$277.1^{+47.8}_{-33.1}$	$308.5^{+68.5}_{-51.4}$	$307.7^{+59.3}_{-51.1}$	$295.9^{+49.0}_{-37.8}$
$L\,(P_0\,[\mathrm{bar}])$	0.006	$0.2^{+0.3}_{-0.3}$	$-0.0^{+0.4}_{-0.4}$	$0.1^{+0.4}_{-0.3}$	$0.0^{+0.4}_{-0.3}$	$0.1^{+0.3}_{-0.4}$	$-0.2^{+0.5}_{-0.4}$	$-0.1^{+0.4}_{-0.3}$	$-0.1^{+0.3}_{-0.3}$
$R\ [R_\oplus]$	1.0	$0.9^{+0.1}_{-0.1}$	$0.99^{+0.07}_{-0.07}$	$1.00^{+0.05}_{-0.05}$	$1.00^{+0.04}_{-0.04}$	$0.99^{+0.09}_{-0.09}$	$1.00^{+0.06}_{-0.05}$	$1.00^{+0.03}_{-0.03}$	$1.00^{+0.03}_{-0.03}$
$L\,(M\ [M_\oplus])$	0.0	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$-0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$	$0.0^{+0.3}_{-0.3}$
$L(\mathrm{N_2})$	-0.107	$-7.6^{+4.5}_{-4.2}$	$-7.2^{+4.4}_{-4.7}$	$-7.6^{+4.7}_{-4.7}$	$-7.6^{+4.6}_{-4.6}$	$-7.8^{+4.5}_{-4.1}$	$-7.4^{+4.6}_{-4.6}$	$-7.7^{+4.9}_{-4.5}$	$-7.5^{+4.7}_{-4.5}$
$L(\mathrm{O_2})$	-0.679	$-7.7^{+4.6}_{-4.3}$	$-7.3^{+4.4}_{-4.5}$	$-7.4^{+4.6}_{-4.8}$	$-7.3^{+4.5}_{-4.6}$	$-7.5^{+4.5}_{-4.5}$	$-7.2^{+4.5}_{-4.7}$	$-7.5^{+4.6}_{-4.5}$	$-7.9^{+4.7}_{-4.4}$
$L(\mathrm{H_2O})$	-3.000	$-2.9^{+0.8}_{-0.7}$	$-2.4^{+0.9}_{-0.8}$	$-2.8^{+0.8}_{-0.8}$	$-2.8^{+0.7}_{-0.8}$	$-2.8^{+0.9}_{-0.9}$	$-2.4^{+0.9}_{-0.9}$	$-2.7^{+0.6}_{-0.7}$	$-2.8^{+0.6}_{-0.6}$
$L(\mathrm{CO_2})$	-3.387	$-4.3^{+1.1}_{-1.9}$	$-3.3^{+0.8}_{-0.8}$	$-3.4^{+0.7}_{-0.7}$	$-3.3^{+0.7}_{-0.7}$	$-3.7^{+0.9}_{-0.8}$	$-3.0^{+0.8}_{-0.8}$	$-3.1^{+0.6}_{-0.6}$	$-3.2^{+0.6}_{-0.6}$
$L(\mathrm{CH_4})$	-5.770	$-10.0^{+3.1}_{-2.9}$	$-8.7^{+2.9}_{-3.9}$	$-6.5^{+1.0}_{-3.7}$	$-5.9^{+0.8}_{-0.9}$	$-8.6^{+2.6}_{-3.7}$	$-5.8^{+1.0}_{-1.6}$	$-5.6^{+0.6}_{-0.7}$	$-5.6^{+0.6}_{-0.6}$
$L(\mathrm{O_3})$	-6.523	$-6.8^{+0.6}_{-0.6}$	$-6.3^{+0.6}_{-0.6}$	$-6.5^{+0.5}_{-0.5}$	$-6.5^{+0.6}_{-0.6}$	$-6.6^{+0.6}_{-0.6}$	$-6.2^{+0.7}_{-0.6}$	$-6.3^{+0.5}_{-0.5}$	$-6.4^{+0.4}_{-0.4}$
$L(\mathrm{CO})$	-6.903	$-7.1^{+4.3}_{-4.5}$	$-7.4^{+4.6}_{-4.4}$	$-7.5^{+4.7}_{-4.7}$	$-7.4^{+4.5}_{-4.6}$	$-7.4^{+4.4}_{-4.6}$	$-7.7^{+4.8}_{-4.5}$	$-7.1^{+4.3}_{-4.7}$	$-7.6^{+4.7}_{-4.6}$
$L(\mathrm{N_2O})$	-6.495	$-8.8^{+3.6}_{-3.7}$	$-8.4^{+3.5}_{-3.8}$	$-9.2^{+3.4}_{-3.4}$	$-9.4^{+3.2}_{-3.5}$	$-8.7^{+3.8}_{-3.7}$	$-8.8^{+3.5}_{-3.7}$	$-9.2^{+3.3}_{-3.6}$	$-9.6^{+3.3}_{-3.3}$

Notes. Here, $L(\cdot)$ stands for $\log_{10}(\cdot)$.

Appendix E: Observation time estimates

Table E.1: Required observation time in days.

(a) Nominal case

			Observation Time [days]			
	R		20	35	50	100
S/N	1 m Mirrors	5	48.6	85.7	122.5	249.1
		10	194.6	342.7	490.1	996.2
		15	437.8	771.0	1102.7	2241.5
		20	778.3	1370.6	1960.4	3984.8
	2 m Mirrors	5	4.7	8.2	11.7	23.9
		10	18.6	32.7	46.7	95.4
		15	41.9	73.5	105.0	214.7
		20	74.5	130.7	186.6	381.7
	3.5 m Mirrors	5	1.0	1.7	2.4	4.9
		10	3.8	6.7	9.5	19.6
		15	8.6	15.1	21.5	44.2
		20	15.3	26.8	38.2	78.5

(b) Optimized case

			Observation Time [days]			
	R		20	35	50	100
S/N	1m Mirrors	5	57.7	102.5	147.1	294.7
		10	230.6	410.0	588.5	1178.7
		15	518.9	922.2	1324.2	2652.2
		20	922.5	1639.5	2354.1	4714.9
	2m Mirrors	5	4.0	7.1	10.2	20.6
		10	16.0	28.5	41.0	82.3
		15	36.2	64.2	92.1	185.2
		20	64.4	114.2	163.8	329.2
	3.5m Mirrors	5	0.6	1.0	1.4	2.9
		10	2.2	3.9	5.7	11.4
		15	5.0	8.9	12.7	25.7
		20	8.9	15.8	22.6	45.8

Notes. Observation times necessary for LIFE to investigate an Earth-twin at 10 pc in orbit around a G2 star for the grid of R and S/N discussed in this publication. The S/N in the Table corresponds to the S/N at 11.2 µm. We consider three different aperture diameters (1m, 2m, and 3.5m) and conservatively assume a total instrument throughput of 5% (cf. Dannert et al. 2022). (a): Nominal case; (b): Optimized case.