Astronomy & Astrophysics

Retrieving scattering clouds and disequilibrium chemistry in the atmosphere of HR 8799e

P. Mollière[1,2], T. Stolker[3], S. Lacour[4,5,6], G. P. P. L. Otten[7], J. Shangguan[5], B. Charnay[4], T. Molyarova[17],
M. Nowak[8,9], Th. Henning[1], G.-D. Marleau[20,23,1], D. A. Semenov[1,18], E. van Dishoeck[2,5], F. Eisenhauer[5],
P. Garcia[10,13], R. Garcia Lopez[22,1], J. H. Girard[16], A. Z. Greenbaum[14], S. Hinkley[21], P. Kervella[4], L. Kreidberg[1],
A.-L. Maire[11,1], E. Nasedkin[1], L. Pueyo[16], I. A. G. Snellen[2], A. Vigan[7], J. Wang (王劲飞)[12,⋆],
P. T. de Zeeuw[2,5], and A. Zurlo[15,19,7]

(Affiliations can be found after the references)

ABSTRACT

Context. Clouds are ubiquitous in exoplanet atmospheres and they represent a challenge for the model interpretation of their spectra. When generating a large number of model spectra, complex cloud models often prove too costly numerically, whereas more efficient models may be overly simplified.
Aims. We aim to constrain the atmospheric properties of the directly imaged planet HR 8799e with a free retrieval approach.
Methods. We used our radiative transfer code petitRADTRANS for generating the spectra, which we coupled to the PyMultiNest tool. We added the effect of multiple scattering which is important for treating clouds. Two cloud model parameterizations are tested: the first incorporates the mixing and settling of condensates, the second simply parameterizes the functional form of the opacity.
Results. In mock retrievals, using an inadequate cloud model may result in atmospheres that are more isothermal and less cloudy than the input. Applying our framework on observations of HR 8799e made with the GPI, SPHERE, and GRAVITY, we find a cloudy atmosphere governed by disequilibrium chemistry, confirming previous analyses. We retrieve that C/O = $0.60^{+0.07}_{-0.08}$. Other models have not yet produced a well constrained C/O value for this planet. The retrieved C/O values of both cloud models are consistent, while leading to different atmospheric structures: either cloudy or more isothermal and less cloudy. Fitting the observations with the self-consistent Exo-REM model leads to comparable results, without constraining C/O.
Conclusions. With data from the most sensitive instruments, retrieval analyses of directly imaged planets are possible. The inferred C/O ratio of HR 8799e is independent of the cloud model and thus appears to be a robust. This C/O is consistent with stellar, which could indicate that the HR 8799e formed outside the CO_2 or CO iceline. As it is the innermost planet of the system, this constraint could apply to all HR 8799 planets.

Key words. methods: numerical – planets and satellites: atmospheres – radiative transfer – instrumentation: spectrographs

1. Introduction

The description of clouds in exoplanets and brown dwarfs is among the major uncertainties when modeling the structures and spectra of self-luminous atmospheres (e.g., Marley et al. 2013). Fully modeling the microphysics of clouds is difficult due to modeling uncertainties and long computational timescales, as the chemistry, nucleation process, condensation, particle coalescence, settling, and mixing need to be accurately described. Moreover, even if all these processes are taken into account, it is not straightforward to determine which values to prescribe for the remaining free parameters. Some of the "free parameters" of such elaborate cloud models are likely to not be free at all, but are determined by the full solution of the (multi-dimensional) atmospheric structure, which is a function of the cloud properties itself due to the radiative feedback of the clouds. At the same time, such complicated cloud models are useful and necessary because they allow us to understand the interplay of physical processes during cloud formation and will ultimately have to explain the cloud properties of all exoplanets and brown dwarfs, whether irradiated or self-luminous. Examples for such more complete

cloud models are described in Woitke & Helling (2004); Helling et al. (2008a); Woitke et al. (2020), Gao et al. (2018), and Powell et al. (2018).

When comparing synthetic cloudy spectra to observations, often one-dimensional self-consistent models are used, where the cloud opacity is radiatively coupled to the atmospheric temperature structure. Here iterating the structure is necessary. This requires a faster cloud modeling approach, which parametrizes parts of the cloud formation process. Examples are models based on timescale comparisons as in Allard et al. (2001, 2003), implementing the approach of Rossow (1978), or Ackerman & Marley (2001), which uses the ratio of the cloud particle settling and mixing velocities (f_{sed}) as a free parameter and solves for the particle size by assuming a log-normal particle size distribution and a vertical diffusion coefficient K_{zz}. The model of Charnay et al. (2018) is again different and combines the two previous approaches: as in Ackerman & Marley (2001), the vertical distribution of the cloud mass is determined by assuming a steady state between mixing, settling, and cloud condensation in every layer, while the average particle size is found using the timescale approach of Rossow (1978). Thus no f_{sed} needs to be specified in Charnay et al. (2018) and this model can reproduce the L-T transition, including the effect of low gravity, which moves

---

the transition to lower effective temperatures. Another interesting approach is the recent model by Ormel & Min (2019), which determines the cloud mass fraction and average particle size by solving the steady-state differential equations including cloud settling, mixing, nucleation, condensation, and coagulation as a function of K_{zz} and the nucleation rate. A more general treatment of this setup is presented in Ohno & Okuzumi (2018); Ohno et al. (2020) which tracks the time evolution of the cloud until a steady state is reached, and accounts for the porosity of particles, including their fractal growth and compaction. A summary of other models can be found in Helling et al. (2008b); Helling & Casewell (2014).

The two modeling philosophies described above (full microphysical or simplified for increased speed) are invaluable for understanding both cloud (micro)physics and the self-consistent radiative feedback of clouds. However, they are challenging when one aims at fitting cloudy spectra. The former, more complete models may take prohibitively long when calculating a large number of atmospheric structures and spectra. The latter, more parametrized models, allow for the calculation of larger self-consistent model grids. However, an important question is whether the simplification steps during model construction were all justified and if all explicit (and implicit) free parameters of the model have been varied sufficiently. Moreover, an update of the model requires the calculation of a new grid, the models of which are demanding to produce and converge especially with clouds (see, e.g., Morley et al. 2014; Mollière et al. 2017), whereas considering additional parameters increases the grid size by orders of magnitude.

In the work presented here, we use a different approach. Namely, we attempt to retrieve the characteristics of cloudy, self-luminous atmospheres by means of free retrievals. This is done by parameterizing the temperature profile, as well as the cloud properties, while using chemical equilibrium abundances with a simple quench pressure treatment to account for atmospheric mixing. While clouds in various parameterizations have been included in retrievals of transmission spectra of exoplanets (see MacDonald & Madhusudhan 2017; Fisher & Heng 2018; Tsiaras et al. 2018; Pinhas et al. 2019; Barstow 2020, for some recent examples), cloudy retrievals are still a comparatively novel approach for self-luminous targets. The use of a free retrieval approach for fitting the spectra of brown dwarfs and directly imaged planets is motivated by Line et al. (2015, 2017); Zalesky et al. (2019), who retrieved the atmospheric properties of clear T- and Y-dwarfs, Burningham et al. (2017), who studied cloudy L-dwarfs, and Lee et al. (2013) as well as Lavie et al. (2017), who attempted to retrieve the properties of the cloudy HR 8799 planets for the first time. These pioneering works show the power of free retrievals for constraining condensation physics and clouds in cloud-free and cloudy brown dwarfs, and how retrieved planetary abundances may be connected to planet formation.

The radiative transfer tool used in the retrievals here is *petitRADTRANS* (Mollière et al. 2019), which we update to include the effect of scattering, which can no longer be ignored for cloudy atmospheres. By parametrizing the clouds, rather than making assumptions on how to simplify the cloud modeling process, we let the data constrain basic cloud characteristics such as cloud mass, location, and particle size, provided that the signal-to-noise of the data is high enough. Our model has the advantage that any changes in the (cloud) parametrization approach can be quickly implemented and tested, without the need for recalculating cloudy model grids.

The capabilities of our retrieval model are demonstrated by analyzing new and archival spectra of the cloudy planet

HR 8799e, taken with the GRAVITY (*K* band, see Gravity Collaboration 2019), SPHERE (*YJH* bands, see Zurlo et al. 2016), and GPI (*H* band, see Greenbaum et al. 2018) instruments. While not included in the fit, we also compare our models to archival mid-infrared (MIR) photometry[1]. The HR 8799 system is especially interesting because it hosts four directly imaged planets that orbit their star within a massive debris disk (Marois et al. 2008, 2010; Currie et al. 2011; Su et al. 2009) at distances from 15 to 70 au (e.g., Wang et al. 2018). This allows for the comparative characterization of the planets' spectral properties. In particular the planets' atmospheric abundances may shed light on how they formed from the circumstellar disk. Consequently, the HR 8799 planets have been extensively studied in the literature, and have been classified to bear the hallmarks of thick clouds and disequilibrium chemistry, placing them in the low-gravity, cool end of L spectral sequence (see Sect. 5.2 for a more detailed discussion of the literature). In addition to confirming these findings for HR 8799e, we also derive the planet's metallicity and, for the first time, carbon-to-oxygen number ratio (C/O), to study possible formation pathways.

Our retrieval model is detailed in Sect. 2, including our description of the scattering implementation, the temperature and chemistry parameterization, as well as our cloud model parameterizations. Section 3 describes tests to verify our model setup with mock retrievals. Section 4 contains our retrieval study of the planet HR8799e. Section 5 discusses the implications of our results for the formation of HR 8799e, and compares to the extensive body of literature on the HR 8799 planets. We summarize our findings in Sect. 6 and provide an outlook for the further development and application of our new retrieval model.

2. Forward retrieval model

2.1. Adding scattering to petitRADTRANS

The objects we seek to model are expected to be inherently cloudy. Hence, scattering is an important process that needs to be considered during the radiative transfer calculations. To calculate quantities such as the photon destruction probability, it is necessary to compare the scattering and absorption opacities. For this, the total atmospheric opacity, combined from the individual absorber opacitites, needs to be calculated. Thus the correlated-k treatment of *petitRADTRANS* needs to be adapted to combine the k-tables (opacity tables) of individual atmospheric absorbers[2]. This is in contrast to the case treating purely emission, which can be handled by using the products of the transmissions of individual species (see, e.g., Irwin et al. 2008; Mollière et al. 2019). Because our goal is to run retrievals, the k-table combination has to be as fast and as accurate as possible.

Here we present a newly developed method to quickly combine k-tables of different absorber species, which works by sampling the opacity distribution functions of individual absorbers. Computational time is saved by sampling the indices of the k-table entries, instead of interpolating the k-tables to sampled values of the cumulative probability. The k-table mixing process is described and tested in Appendix A.

---

[1] Our rationale for excluding the photometry from the fit is explained in Sect. 4.1.

[2] Correlated-k means that the radiative transfer is carried out using the cumulative probability of the opacity distribution function as the spectral coordinate, and assuming that different probability values map to the same wavelength in all atmospheric layers (e.g., Lacis & Oinas 1991; Fu & Liou 1992; Marley & Robinson 2015).

For solving the radiative transfer equation we then use the same treatment as described in our self-consistent *petitCODE* (Mollière et al. 2015, 2017), namely, by using the Feautrier method (Feautrier 1964), converging the scattering source function with local Accelerated Lambda Iteration (ALI; Olson et al. 1986) and Ng acceleration (Ng 1974). The scattering process is assumed to be isotropic, with a $(1 - g_a)$ correction factor applied to the scattering cross-sections, where g_a is the scattering anisotropy. The scattering implementation is further described in Appendix A.6 of Mollière et al. (2017). We show a verification of the cloudy spectra of *petitRADTRANS*, including scattering, in Appendix B. For this verification, we used *petitCODE* to calculate a self-consistent HR 8799e model in radiative-convective and chemical equilibrium, which included clouds of $MgSiO_3$ and Fe. The spectra of *petitRADTRANS* and *petitCODE* agree excellently. It takes *petitRADTRANS* a few seconds to calculate a cloudy emission spectrum in the *YJHK* bands, which is fast enough for retrievals on computational clusters.

2.2. Temperature model

Our goal is to parameterize the vertical temperature profile of the atmosphere in a way that imposes as few prior constraints on the solution as possible. It would hence appear to be ideal to retrieve the temperatures in every layer of the discretized atmosphere independently, which is an approach commonly followed in the planetary science community (e.g., Rodgers 2000; Irwin et al. 2008), and which has also been applied to cloud-free brown dwarfs (Line et al. 2014a) and exoplanets (e.g., Lee et al. 2012).

However, if the data are sparse or of low signal-to-noise, a level-by-level retrieval of the temperature can lead to overfitting and thus unphysical oscillations in the inferred temperature profile. One way of reducing such oscillations is by smoothing the resulting P-T profile (as done in Irwin et al. 2008; Lee et al. 2012). Another way to circumvent this problem is to retrieve temperatures at a limited number of altitudes in the atmosphere, which are then connected via (spline) interpolation to yield a temperature at all layers, thereby reducing the number of free parameters (Line et al. 2015; Kitzmann et al. 2020). In addition, Line et al. (2015) included a penalty term on the spatial sum of second derivatives of the temperature profile which further discouraged oscillatory solutions. Because the weight of said penalty can bias the results, its optimal value is also fitted in the inference process. This method has been used to retrieve the temperature profiles of cloud-free T- and Y-dwarfs in Line et al. (2015, 2017); Zalesky et al. (2019).

The most biased class of temperature models are those that use some kind of physical reasoning to parametrize the shape of the temperature profiles. This includes analytical solutions for self-luminous or irradiated atmospheres, assuming a gray or double-gray[3] opacity. The analytical solution (or a modification of it) by Guillot (2010); Parmentier & Guillot (2014) is commonly used, for example in Line et al. (2012, 2013a,b, 2014b), Benneke & Seager (2012), Waldmann et al. (2015), Rocchetto et al. (2016), Kreidberg et al. (2018), Brogi & Line (2019), Mollière et al. (2019). There also exists the temperature parametrization suggested by Madhusudhan & Seager (2009), which allows to parametrize temperature structures with or without inversions, and with or without a deep isothermal layer, as commonly expected for hot Jupiter planets. This parametrization has been used in, for example, Madhusudhan et al.

(2011a, 2014a), Madhusudhan & Seager (2011), MacDonald & Madhusudhan (2017, 2019), Burningham et al. (2017), Gandhi & Madhusudhan (2018), Pinhas et al. (2018).

When investigating different models suitable for retrieving the atmospheres of cloudy self-luminous exoplanets, we settled on a model that uses both freely variable and physically motivated parameterizations, based on the atmospheric altitude. This temperature model, which allows us to retrieve the synthetic structures of cloudy atmospheres, is split into three parts, going from high, to middle, to low altitude. The spatial coordinate of the temperature model is an optical depth τ[4], which we relate to the pressure P by

$$\tau = \delta P^\alpha, \tag{1}$$

where δ and α are free parameters, and P is the atmospheric pressure in units of dyn cm^{-2}. This mapping is required because P is the vertical coordinate of *petitRADTRANS*. We then setup the atmospheric temperature profile, starting with the middle altitudes, that is, the "photosphere".

2.2.1. "Photosphere" (middle altitudes)

This region stretches from $\tau = 0.1$ to the radiative-convective boundary. Here we set the temperature according to the Eddington approximation

$$T(\tau)^4 = \frac{3}{4} T_0^4 \left(\frac{2}{3} + \tau \right), \tag{2}$$

where T_0 is a free parameter. The optical depth τ is obtained from Eq. (1) above. In the original Eddington solution, from which we take the functional form of the temperature profile, this corresponds to the internal temperature. Likewise, we note that the "photospheric" region does not necessarily have to correspond to the true photosphere of the planet. Because the Eddington solution will always lead to an isothermal upper atmosphere, which is not expected to occur in reality, the high-altitude region of the atmosphere is treated separately, described immediately below.

2.2.2. High altitude

This region extends from the top of the atmosphere ($P = 10^{-6}$ bar) to $\tau = 0.1$. Here we split the atmosphere into four equidistant locations in $\log(P)$ space and treat the temperature at the three upper locations as free parameters. The temperature at the lowest altitude, which is at $\tau = 0.1$, is taken from the Eddington approximation of the "photosphere". The temperatures in this atmospheric region are then found from a cubic spline interpolation.

2.2.3. Troposphere (low altitudes)

This region starts from the radiative-convective boundary and extends to the bottom of the atmosphere. The radiative convective boundary is found by comparing the atmospheric temperature gradient of the Eddington approximation with the moist adiabatic temperature gradient of the atmosphere. This is done by interpolating the moist adiabatic gradient in the T-P-[Fe/H]-C/O space of the chemistry table (see Sect. 2.3). As soon as the atmosphere is found to be Schwarzschild-unstable, the atmosphere is forced onto the moist adiabat.

---

[3] That is, taking two (or more) separate gray opacities within given wavelength bands, for example in the optical and infrared.

[4] This optical depth is merely used for parameterization, so not associated to any particular wavelength or mean opacity.

2.2.4. Priors

We restrict α (see Eq. (1)) to vary between 1 and 2, following Robinson & Catling (2012). Moreover, to prevent the formation of temperature inversions, which are not expected in self-luminous objects, we required that the three free temperature points in the high altitude region of the atmosphere are colder than the highest point of the "photosphere" (middle altitude region), and that they decrease in temperature monotonically with increasing altitude. This prior is enforced by setting the upper boundary of the allowed temperature range of such a free temperature point equal to the temperature of the underlying temperature point, identical to the treatment in Kitzmann et al. (2020).

In Gravity Collaboration (2020) we had found that we had to impose further priors on the temperature parameterization based on the structure of the atmospheric opacity of a given forward modeling realization. For example, the power law index of the optical depth, α, was not allowed to deviate too far from the power law index measured from the opacity structure in the forward model, within the spectral range of the retrieved data. We found that this was necessary because of the high dimensionality of our retrieval model, and our inability to make the MCMC sampler find the global maximum of the log-probability in a finite amount of time otherwise. These opacity priors restricted the parameter space for the MCMC sufficiently. In this publication we use nested sampling (Skilling 2004), and using a sufficiently large number of live points made such opacity priors unnecessary.

2.3. Chemistry model

In the retrievals presented in this work, the chemical abundances within the atmosphere are determined by means of interpolation in a chemical equilibrium table, with a simple quench layer approximation used to account for atmospheric mixing. The abundance tables are calculated with *easyCHEM*, our *CEA* (Gordon & McBride 1994; McBride & Gordon 1996) clone described in Mollière et al. (2017). The equilibrium condensation of the following species is included in the abundance calculations: Al_2O_3, Fe, FeO, Fe_2O_3, Fe_2SiO_4, H_2O, H_3PO_4, KCl, $MgSiO_3$, Mg_2SiO_4, Na_2S, SiC, TiO, TiO_2, VO. Because rainout is expected to remove Si from the upper layers of the atmosphere (see, e.g. Lodders 2010), we do not include feldspars, thereby inhibiting the sequestration of Na and K at high temperatures. This is consistent with abundance constraints of the alkalis from systematic retrieval analyses of T- and Y-dwarfs (Line et al. 2017; Zalesky et al. 2019).

The chemical abundances (mass fractions) are tabulated as a function of pressure, P, temperature, T, carbon-to-oxygen number ratio, C/O, and metallicity, [Fe/H]. The pressure ranges from 10^{-8} to 1000 bar, in 100 points spaced equidistantly in $\log(P)$ space. The temperature ranges from 60 to 4000 K, in 100 equidistant points. The C/O values go from 0.1 to 1.6, in 20 equidistant points and the metallicity is tabulated for [Fe/H] values going from -2 to 1.84, in 31 equidistant points. Four-dimensional linear interpolation is used to interpolate the log-abundances of all absorbers. If a T-P-[Fe/H]-C/O coordinate falls outside of the grid, the abundances interpolated to the closest boundary point are used. The C/O ratio is varied by varying the oxygen abundance. For minimizing the Gibbs free energy we use the thermodynamic data of the *CEA* code, or the references detailed in Mollière et al. (2017). The data for FeH were obtained from M. Line (priv. comm.).

We approximate the effect of disequilibrium chemistry by setting the quench pressure P_{quench} as a free parameter. For atmospheric pressures $P < P_{quench}$ we take the abundances of CO, H_2O, and CH_4 to be constant, and equal to the abundances at $P = P_{quench}$. This follows the result from, for example, Zahnle & Marley (2014), namely that the abundances of CO, H_2O, and CH_4 can be taken to be constant above the quenching point. This treatment has been further verified by comparing to the results of the reaction network by Venot et al. (2012, 2015) in Baudino et al. (2017).

The chemical abundance table also contains the value of the adiabatic temperature gradient, ∇_{ad}, which was calculated with *easyCHEM*, using Equations 2.50, 2.59, and 2.75 of Gordon & McBride (1994). The derivatives used for the calculation of the specific heat of the mixture are so-called equilibrium derivatives (Gordon & McBride 1994), meaning that the ∇_{ad} value used in this work accounts for any change in the abundances (and thus heat release) during the adiabatic temperature change. Hence our adiabats are moist adiabats. The interpolation in the ∇_{ad} table is used when constructing the temperature profile in the troposphere of our temperature model.

2.4. Cloud model 1

As discussed in Sect. 1, our goal is to impose as few prior assumptions on the cloud properties as possible. Hence the ideal setup would be to freely retrieve the altitude-dependent distribution of cloud particle radii, as well as the vertical cloud density for every layer, independently. Such an approach would require many free parameters, which in turn requires enough data points of sufficiently high signal-to-noise to prevent over-fitting. Instead, we start with a more modest approach. As the Ackerman & Marley (2001) model is already implemented in *petitRADTRANS*, we use its three free parameters to control the mean particle size, cloud mass fraction and particle size distribution independently.

Hence, for the cloud to be retrieved, we set the settling parameter f_{sed} as a free parameter, which controls the altitude-dependent cloud mass fraction X^c above the cloud base via

$$X^c(P) = X_0^c \left(\frac{P}{P_{base}} \right)^{f_{sed}} , \tag{3}$$

where the cloud mass fraction at the cloud base X_0^c is an additional free parameter. The pressure at the cloud base P_{base} is found by intersecting the saturation vapor pressure curve of the considered cloud species with the temperature profile of the atmosphere. In our retrievals below, we only consider $MgSiO_3$ and Fe clouds, because silicates and iron likely dominate the atmosphere at the temperatures and surface gravities reported for HR 8799 planets (see, e.g., Fig. 7 in Morley et al. 2012, and Marley et al. 2012; Charnay et al. 2018). The recent findings of Gao et al. (2020), based on micro-physical cloud modeling, corroborate the importance of especially silicate clouds at these temperatures.

We also let the vertical eddy diffusion coefficient K_{zz} vary as a free parameter, which effectively sets the particle size, given an f_{sed} value. In self-consistent calculations, the K_{zz} parameter is usually set by mixing length theory with some lower limit assumption (see, e.g., Ackerman & Marley 2001; Morley et al. 2014; Mollière et al. 2017; Charnay et al. 2018), or fixed to a constant value (e.g., Marley et al. 2012; Samland et al. 2017). Here we let it float as a free parameter (and take it to be vertically constant), so as to determine the average particle size independently

from f_{sed}. We note that a retrieved K_{zz} value could be inconsistent with the derived chemical quench pressure (see Sect. 2.3). This could imply either a true shortcoming of our retrieval model, a shortcoming in a disequilibrium kinetic network, or a deviation from how K_{zz} actually sets the average particle size to how it is implemented in the Ackerman & Marley (2001) model.

Lastly, the particle size distribution is fitted by letting the width of the log-normal size distribution, σ_{g}, be a free parameter. This parameter is usually not varied if the Ackerman & Marley (2001) cloud model is used. We note that it has been shown that a log-normal particle size distribution can be a poor choice when compared to the often bi-modal particle size distributions found from microphysics (Gao et al. 2018; Powell et al. 2018). However, in Gao et al. (2018), the K_{zz} and σ_{g} values were both fixed when fitting f_{sed} to the cloud structure of the microphysics result. Here we let f_{sed}, K_{zz}, and σ_{g} vary independently, so the retrieval should be flexible enough to determine the values of these three parameters that describe the cloud mass fraction, effective particle size, and dispersion of sizes around that value, independently. In principle, this model can also describe monodisperse particle distributions, if a retrieval were to favor cases with σ_{g} close to 1. In the limit $\sigma_{\text{g}} \rightarrow 1$ a log-normal particle size distribution approaches a delta function. In general, the retrieved cloud parameter values are expected to describe those visible atmospheric layers which are most affected by the clouds.

2.5. Cloud model 2

The choice of f_{sed}, K_{zz} and σ_{g} in Cloud Model 1 may just be a glorified way of setting the cloud spectral slope and single scattering albedo. It is also questionable, for example, whether a retrieved K_{zz} does actually correspond to the true vertical diffusion coefficient of the atmosphere. Rather, it could effectively be a nuisance parameter of the retrieval, varied to mimic the true properties of the cloud opacity, as alluded to above.

To test for a less physically motivated treatment, we constructed Cloud model 2. This approach is motivated by the cloud parameterization of Burningham et al. (2017). Our model retrieves the spectral slope ξ (which we take to be vertically constant) of the cloud opacity directly. Specifically, we set

$$\kappa_{\text{tot}} = \kappa(P) \left(\frac{\lambda}{1\,\mu\text{m}} \right)^{\xi}, \tag{4}$$

where κ_{tot} is the total (scattering + absorption) cloud opacity, $\kappa(P)$ is its value at $1\,\mu\text{m}$, at pressure P, and λ the wavelength. The $\kappa(P)$ we describe as

$$\kappa(P) = \kappa_0 \left(\frac{P}{P_{\text{base}}} \right)^{f_{\text{sed}}} \text{ for } P < P_{\text{base}}, \tag{5}$$

and set it to zero for pressures larger than the cloud base pressure P_{base}. The f_{sed} again describes the power law decrease of the cloud with altitude. Additionally, we set the single-scattering albedo, ω, as a free parameter, which we also take to be vertically constant. For the absorption opacity it then holds that

$$\kappa_{\text{abs}} = (1 - \omega)\kappa_{\text{tot}}. \tag{6}$$

In summary, the five free parameters of Cloud model 2 are κ_0, ξ, f_{sed}, P_{base}, and ω. Because there is a degeneracy between κ_0 and P_{base} (see Eq. (5)) if the atmosphere below the cloud deck cannot be probed by the observations, we put a prior on P_{base}

Table 1. Parameters for generating the synthetic observations of the cloudy exoplanet spectrum retrieved for verification purposes in Sect. 3.

Parameter	Value	Parameter	Value
T_1	330.6 K	$\log(X_0^{\text{Fe}}/X_{\text{eq}}^{\text{Fe}})$	-0.86
T_2	484.7 K	$\log(X_0^{\text{MgSiO}_3}/X_{\text{eq}}^{\text{MgSiO}_3})$	-0.65
T_3	687.6 K	f_{sed}	3
$\log(\delta)$	-7.51	$\log(K_{zz}/\text{cm}^2\text{s}^{-1})$	8.5
α	1.39	σ_{g}	2
T_0	1063.6 K	R_{P}	$1\,R_{\text{J}}$
C/O	0.55	$\log(g/\text{cm s}^{-2})$	3.75
[Fe/H]	0	$\log(P_{\text{quench}}/\text{bar})$	-10

Notes. X_{eq} is the mass fraction predicted for the cloud species when assuming equilibrium condensation at the cloud base location.

such that

$$\log(L_{\text{base}}) = \log\left[\exp\left(-\frac{\log_{10}^2(P_{\text{base}}/P_{\text{Fe}})}{2 \cdot (0.5\,\text{dex})^2} \right) \right.$$
$$\left. + \exp\left(-\frac{\log_{10}^2(P_{\text{base}}/P_{\text{MgSiO}_3})}{2 \cdot (0.5\,\text{dex})^2} \right) \right], \tag{7}$$

where P_{Fe} and P_{MgSiO_3} are the cloud base positions of Fe and MgSiO$_3$, respectively, obtained from intersecting their saturation vapor pressure curves with the atmospheric temperature profile. Clearly, this will also tend to favor clouds that lie close to the expected cloud base positions for these condensate species. We note that this prior choice was made to not introduce parameters which are degenerate by construction. As mentioned, silicate and iron clouds may be the dominant cloud opacity carriers for the temperatures and surface gravities generally inferred for the HR 8799 planets. Removing this prior would allow to better probe situations where different cloud species dominate in the planetary atmosphere. This will be tested in our future work.

To test how well such a cloud description may be suited to describe more "physically consistent" clouds we carried out the following test. The cloud properties of the atmospheric model we used to create the synthetic observation for the verification retrieval (see Sect. 3) were generated with Cloud model 1. Fitting these cloud opacities with Eq. (5), and the spectral slope with Eq. (4), we found that the cloud parameters could indeed be well represented with Cloud model 2. In addition, the single-scattering albedo, taking the spectral average over the 0.9–2.5 μm range, was consistent with a large ($\omega \sim 0.85$), vertically constant value.

3. Verification

3.1. Retrieval with Cloud model 1

Here we present our retrieval tests when using Cloud model 1 (see Sect. 2.4), that is, MgSiO$_3$ and Fe clouds parameterized using the Ackerman & Marley (2001) model, varying all of its three free parameters. We generated a synthetic observation as follows: for the input parameters of the temperature profile, we fitted our temperature model to the self-consistent atmospheric structure used for verifying our scattering implementation (see Sect. 2.1 and Appendix B). The values of all input parameters of the model are shown in Table 1.
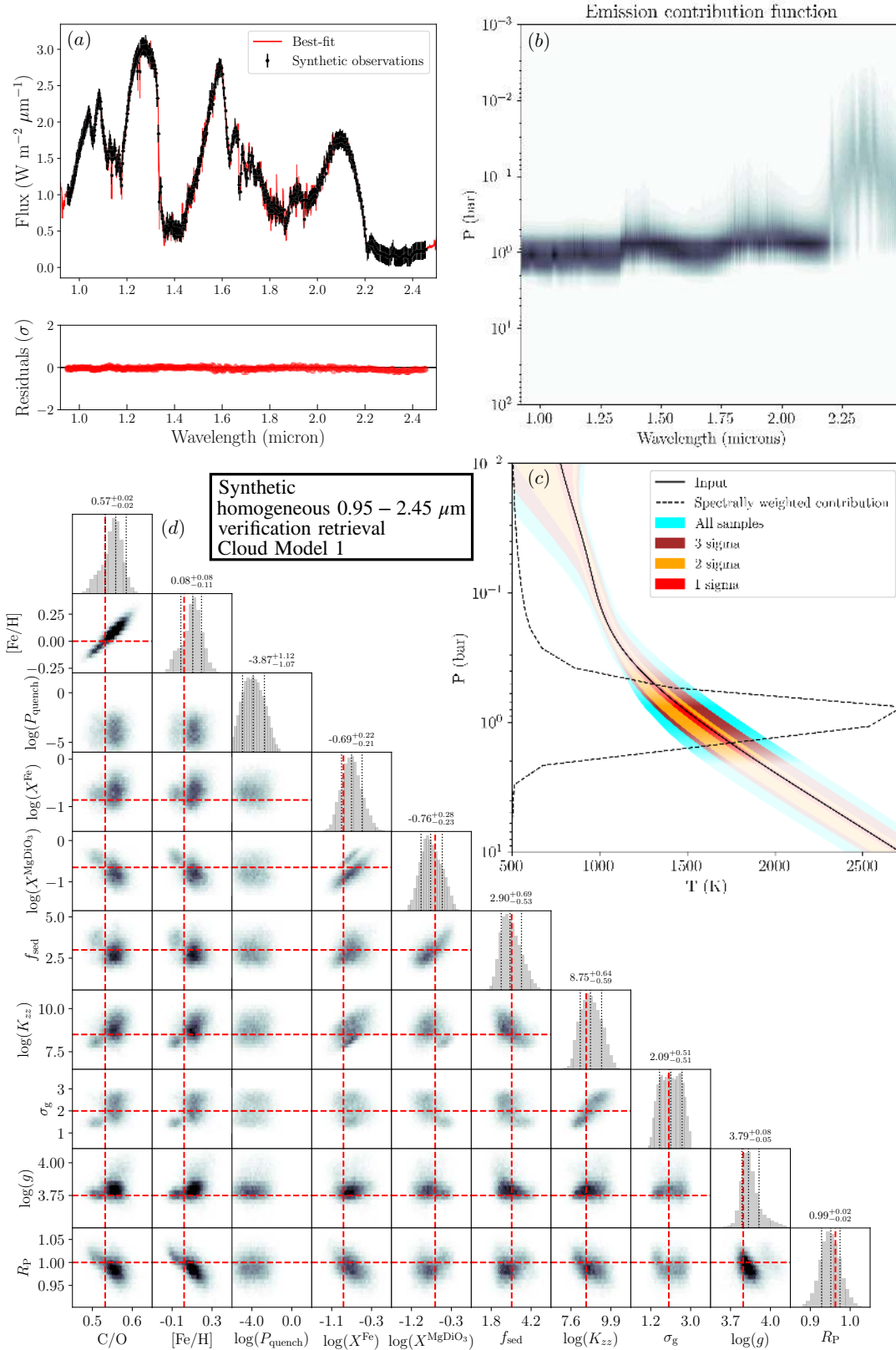
Fig. 1. Results of the verification retrieval using Cloud model 1. *Panel a*: synthetic observation, best-fit spectrum and residuals. *Panel b*: emission contribution function. Due to the clouds, pressures larger than 1–2 bar cannot be probed. *Panel c*: retrieved pressure-temperature confidence envelopes. The black dashed line shows the flux average of the emission contribution function that is shown in *panel b*. The opaqueness of the temperature uncertainty envelopes has been scaled by this contribution function, with a minimum value of 10%. *Panel d*: 2d posterior plot of the (non-nuisance) retrieved atmospheric parameters. The red dashed lines denote the input values. The values of the cloud mass fractions at the cloud base have been divided by the mass fractions predicted when assuming equilibrium condensation at the cloud base location.

To derive the posterior abundances of our fit we used the PyMultiNest[5] package (Buchner et al. 2014), which is a Python wrapper of the MultiNest method (Feroz & Hobson 2008; Feroz et al. 2009, 2019) for nested sampling (Skilling 2004). Nested sampling has the benefit of being able to approximate model evidences (i.e., the probability of the model, given the data), which allows for the pair-wise vetting of different models. Moreover, it is sampling the parameter space more thoroughly. This minimizes the problem of sampling the posterior distribution around a local, but not the global, maximum of the log-probability. It does not fully alleviate this problem, however (see discussion below). To ensure a high sample acceptance fraction of our high-dimensional model, we ran MultiNest in the constant efficiency mode, with a prescribed sampling efficiency of 5%. When using MultiNest in Importance Nested Sampling mode, evidences can still be calculated, even when prescribing a sampling efficiency[6]. We used 4000 live points in our retrievals, which we found necessary to cover the parameter space sufficiently.

To initially test our retrieval framework under idealized conditions, a synthetic observation was created, assuming a continuous wavelength spacing of $\lambda/\Delta\lambda = 400$ between 0.95 and 2.45 μm. We focused on this spectral region because it overlaps with the *YJHK*-bands of the SPHERE, GPI and GRAVITY instruments. The flux error was chosen to be constant across this wavelength range, with a mean signal-to-noise ratio (S/N) value of 10 per wavelength step. For comparison, the S/N per wavelength step of HR 8799e is about 4 ($\lambda/\Delta\lambda \approx 70$), 7 ($\lambda/\Delta\lambda \approx 200$), and 11 ($\lambda/\Delta\lambda \approx 1000$) for the SPHERE YJH (Zurlo et al. 2016), GPI H (Greenbaum et al. 2018) and GRAVITY K (Gravity Collaboration 2019) band data, respectively. In order to not be affected by a given noise instantiation during the verification retrieval, we took the observational errors into account for calculating the log-likelihood, but did not perturb the mock observations using these error bars.

The results of this verification retrieval are shown in Fig. 1. panel a shows the synthetic observation, best-fit model, and the residuals between the two, scaled by the error bars. The residuals are flat and consistent with zero.

The emission contribution function of the best-fit model is shown in panel b of Fig. 1. Regions between 0.004 and 2 bar are accessible, with the Fe and $MgSiO_3$ cloud blocking the flux from the deeper regions. Methane absorption blocks most of the flux longward of 2.2 μm, probing cool regions as high as 0.004 bar. Shortward of 2.2 μm most of the flux originates from a narrow pressure region from 0.2 to 2 bar.

In panel c the retrieved pressure temperature structure of the atmosphere is shown, with the percentiles setting the boundaries of the uncertainty envelopes corresponding to the 1, 2 and 3σ ranges of a Gaussian distribution. In order to illustrate which altitudes of the atmosphere can actually be probed by the observation, the opaqueness of the temperature uncertainty envelopes has been scaled by the atmospheric contribution function, with a minimum value of 10%. For this the contribution function was flux-averaged. As can be seen, the uncertainty envelopes follow the input P-T profile. The input profile lies within the 1σ envelope.

Finally, panel d shows the corner plot of the remaining parameters. The planetary radius, surface gravity, metallicity and C/O ratio can all be well retrieved. Only an upper limit is found for the quench pressure, which is as expected, as no quenching

was considered. The cloud parameters are well constrained. We note that the retrieval appears to have found a bi-modal parameter distribution, but the one-dimensional posteriors constrain the input parameters well. A clear positive correlation can be seen between the metallicity and C/O ratio. This is attributed to the fact that we vary the oxygen abundance when changing C/O, such that a higher C/O corresponds to less oxygen, and hence water.

In summary, the retrieval verification test presented here can be regarded as successful. We were able to retrieve the temperature, composition, and cloud properties of the atmosphere. Nonetheless, the following challenges can be identified: even though an excellent fit to the spectrum has been achieved (see panel a of Fig. 1) the median values of the retrieved parameters were not exactly at the input values (although within the 1σ envelope). This is unexpected because the synthetic observation was not perturbed by the assumed error bars, in order to not be sensitive to stochastic noise of a given noise instantiation. Another point is the bi-modality of the posterior.

In our initial tests we found that using 400 live points in PyMultiNest resulted in biased retrieval results, such that median parameter values could be more than 1σ away from the input values, and residuals in the spectral fit that were larger. Increasing the number of live points to 4000 ameliorated this issue: the input parameters were retrieved at higher accuracy (within 1σ), and the residuals to the best-fit spectrum became smaller. We deduce from this that for input models of high dimensionality a sufficient number of live points has to be used, so as to increase chances that the positions of the live points sampled during the early stages of the nested sampling run will fall into the vicinity of the global maximum of the likelihood. Because the nested sampling method will zero-in on the highest likelihood regions during the retrieval, the danger exists that the true maximum location in parameter space will be missed. This problem is especially pressing for observations of high S/N, such as used in our example presented here, because the high-likelihood volume of the parameter space will shrink. This means that higher quality observations require a larger number of live points, and thus more computational time. This is especially important for the large spectral coverage, high S/N data to be taken with JWST.

Further exploring the bi-modality of the posterior shown in Fig. 1, we ran a second retrieval with smaller prior ranges. They were restricted by the high-likelihood regions of the posterior from the initial fit, enclosing its bi-modal posterior distribution. The resulting posterior is shown in the corner plot in Fig. C.1. It is unimodal and consistent with the input parameters. In general, the bimodality and the offset between the median and input parameters of the posteriors indicate that multiple parameter combinations can lead to excellent spectral fits, while within 1σ of the true parameters. This may indicate that a retrieval model with fewer free parameters could be favored, leading to unique solutions, which we will explore in future studies.

3.2. Retrieving Cloud Model 1 with Cloud model 2

In this section we describe what happens when retrieving a mock observation made with Cloud Model 1 using Cloud Model 2. The retrieval was thus set up identically to the one described in the section immediately above, but used Cloud Model 2, while the mock observation was identical to the test above, that is made with Cloud Model 1. Figure D.1 shows the corresponding results (see Appendix D).

We find that the spectral fit is again very good, although there are a few regions of systematic residuals of 1σ. Hence, Cloud model 2 appears to satisfactorily describe the properties of the

---

[5] https://johannesbuchner.github.io/PyMultiNest/
[6] Also see the corresponding discussion in the MultiNest manual at https://github.com/farhanferoz/MultiNest/blob/master/README.md

synthetic observations generated with Cloud model 1. However, we find significant differences in the retrieved atmospheric properties. C/O is constrained to $0.58^{+0.01}_{-0.01}$ (input was 0.55), [Fe/H] is retrieved to be $0.14^{+0.08}_{-0.08}$ (input was 0), and $\log(g) = 4.01^{+0.10}_{-0.10}$ (input was 3.75). Thus, we find values close to, but offset from the true input values.

Moreover, instead of probing down to 2 bar at most, the atmosphere can now be probed down to 10 bar, and is more isothermal than the input temperature profile: here the retrieval mimics the effect of a thick cloud. Instead of the cloud hiding the deep hot regions from view, these regions are erroneously constrained to be less hot by the retrieval. At the same time, the cloud is thus too optically thin and deep, with the retrieved cloud position at 8 bar (which is well constrained). The emission contribution consequently shows that the emission stems from a more extended region than in the retrieval described in the section immediately above. Hence Cloud Model 2 was not able to describe the clouds made with Cloud Model 1 accurately enough, such that the retrieval modified the P-T profile instead.

We conclude that a good fit to the spectrum alone is a dangerous measure when assessing whether or not a fit result is reasonable. All retrieved parameter values need to be carefully vetted, the retrieved P-T profile should also be compared to that of a self-consistent atmospheric code, when running the latter using the best-fit or median parameters of the free retrieval. We note that with free retrievals alone it may also be challenging to determine whether the atmospheric temperature profile is truly shallower than expected, while being less cloudy at the same time. This has been suggested by Tremblin et al. (2015, 2016, 2017), challenged in Leconte (2018), and defended in Tremblin et al. (2019). As they suggested, we find here that a shallow P-T profile can indeed result in an excellent fit to the observations, even though we know that in this case the input model was actually more cloudy than retrieved. At the same time, it is somewhat reassuring that in our example shown here the absolute deviation between the input and retrieved parameters such as C/O, [Fe/H], $\log(g)$, etc. is not large. The values, however, are biased, and we stress that a more detailed study needs to be performed on how strongly cloud model assumptions can affect the retrieved best-fit parameters.

4. Retrieving HR 8799e

In this section we describe how we used GRAVITY and archival SPHERE and GPI data to retrieve the atmospheric properties of HR 8799e, which is located at 15 au from its host star (e.g., Wang et al. 2018). We also compare our results to archival photometry for this planet, while not including these data in the retrieval itself.

4.1. Data

4.1.1. GRAVITY K band spectroscopy

We use three separate GRAVITY (Gravity Collaboration 2017) observations of HR 8799e. First, the observation presented in Gravity Collaboration (2019). In addition, we here report on two new observations, taken November 9th, 2019 and November 11th. 2019, as part of the ExoGRAVITY Large Program. The log of the observations is presented in Table 2. The observations on the 9th were made from a short observing block, with a total integration time of 180 s. The phase referencing was done on the star using the fringe tracker (Lacour et al. 2019). The zero point of the metrology was obtained by directly observing the star on

Table 2. Log of the GRAVITY observations.

Date	Target	Exp	NDIT	DIT	Seeing
2018 Aug. 28	HR 8799 e	7	100	10 s	0.5–0.8″
·	HR 8799 A	2	50	1 s	
2019 Nov. 9	HR 8799 e	3	60	8 s	0.8–1.0″
·	HR 8799 A	3	64	1 s	
2019 Nov. 11	HR 8799 e	3	100	8 s	0.8–1.1″
·	HR 8799 A	1	64	1 s	
·	HD 25535 AB	8	64	1 s	

the spectrometer. The observations on the 11th were done using the roof mirror as a field splitter: 100% of the planetary flux could be used, but such an observation needs a binary to calibrate the zero point. This zero point was obtained on the binary system HD 25335.

The data were reduced analogously to the data reduction presented for β Pic b in Gravity Collaboration (2020). The flux of stellar origin is removed, and the spectra were obtained from the ratio between the coherent flux on the planet from the coherent flux on the star. This ratio is then multiplied by a theoretical BT-NextGen spectra of the star (Allard et al. 2012a). The spectra is therefore calibrated from the telluric absorption. For the retrieval, the full spectral covariance is considered when deriving log-likelihoods.

4.1.2. SPHERE and GPI archival data

We use the *YJH*-band spectroscopy of SPHERE reported in Zurlo et al. (2016). In addition, we use the GPI *H*-band spectroscopy reported in Greenbaum et al. (2018). We do not take the spectral covariance into account for GPI or SPHERE. For both SPHERE and GPI we fit for a scaling factor with respect to the GRAVITY observation during the retrieval, as was also done in Gravity Collaboration (2020). This also appears necessary given the noticeable shift between the SPHERE and GPI observations in their overlapping region at ~1.6 micron. Similar shifts between GPI and SPHERE observations have been reported in Samland et al. (2017). However, this difference could also be caused by variability, because spectral template brown dwarfs that reproduce the spectral properties of HR 8799e well have been found to exhibit considerable variability, possibly up to a 20–30% peak-to-peak amplitude (see the discussion of Mace et al. 2013; Biller et al. 2015 in Bonnefoy et al. 2016).

4.1.3. Archival photometry

Although we do not include it in the fit, we compare our results with archival photometry in the mid-infrared. We consider the 3.3 μm LBT photometry reported in Skemer et al. (2012), the *L'* band and [4.05]-Brα photometry reported in Currie et al. (2014), as well as the *M'* band upper limit of Galicher et al. (2011). The photometry was converted from magnitudes to flux using th *species*[7] toolkit, which has been described in Stolker et al. (2020). We decided against including the photometric fluxes in the fit because their relatively low signal-to-noise would add little constraining power to the retrieval, when compared to the spectra, but would double the run-time of our retrievals due to the increased spectral range.

---

[7] https://species.readthedocs.io

Table 3. Priors of the HR 8799e retrieval.

Parameter	Prior	Parameter	Prior
T_1	$\mathcal{U}(0, T_2)$	$\log(\tilde{X}_{Fe})$ [c]	$\mathcal{U}(-2.3, 1)$
T_2	$\mathcal{U}(0, T_3)$	$\log(\tilde{X}_{MgSiO_3})$	$\mathcal{U}(-2.3, 1)$
T_3	$\mathcal{U}(0, T_{connect})$ [a]	f_{sed}	$\mathcal{U}(0, 10)$
$\log(\delta)$	$P_{phot} \in [10^{-3}, 100]$ [b]	$\log(K_{zz})$	$\mathcal{U}(5, 13)$
α	$\mathcal{U}(1, 2)$	σ_g	$\mathcal{U}(1.05, 3)$
T_0	$\mathcal{U}(300, 2300)$	R_P	$\mathcal{U}(0.9, 2)$
C/O	$\mathcal{U}(0.1, 1.6)$	$\log(g)$	$\mathcal{U}(2, 5.5)$
[Fe/H]	$\mathcal{U}(-1.5, 1.5)$	$\log(P_{quench})$	$\mathcal{U}(-6, 3)$
f_{SPHERE}	$\mathcal{U}(0.8, 1.2)$	f_{GPI}	$\mathcal{U}(0.8, 1.2)$

Notes. $\mathcal{U}$ stands for a uniform distribution, with the two parameters being the range boundaries. The units for the parameters are the same as the ones used for Table 1. f_{SPHERE} and f_{GPI} are the scaling factors retrieved for the SPHERE and GPI data, respectively. [a] and [b]: please see Sect. 4.2 for a definition of P_{phot} and $T_{connect}$. [c] It holds that $\tilde{X}_i = X_0^i / X_{eq}^i$, where the latter quantity has been defined in Table 1.

4.2. Retrieval model setup

We set up our nominal retrieval model with 18 free parameters, using Cloud Model 1, which we describe in the following. Additionally, we make a comparison with retrievals using Cloud Model 2; see Sect. 4.4. The free parameters and prior ranges are listed in Table 3. As in Gravity Collaboration (2020), we fitted multiplicative scaling factors f_{SPHERE} and f_{GPI} to account for systematic biases in the flux normalization of these datasets. The $T_{connect}$ quantity referenced in the prior range of T_3 is the uppermost temperature of the "photospheric" layer, and was calculated by setting $\tau = 0.1$ in Eq. (2). Like the priors for T_2 and T_1, this ensures a temperature profile that is monotonically decreasing with altitude, also see Sect. 2.2.4. δ was sampled from the prior by assuming a log-uniform prior on P_{phot}, where we defined P_{phot} as the pressure where $\tau = 1$ in Eq. (1). This allowed to solve for δ for a given P_{phot} value. The following absorber species were included: CO, CO_2 and H_2O (from Rothman et al. 2010), CH_4 (Yurchenko & Tennyson 2014), NH_3 (Yurchenko et al. 2011), H_2S (Rothman et al. 2013), Na and K (Piskunov et al. 1995, with Allard wings, see Mollière et al. 2019 for more details), PH_3 (Sousa-Silva et al. 2015), VO and TiO (Plez line lists, see Mollière et al. 2019 for more details), FeH (Wende et al. 2010) as line absorbers, H_2, He as Rayleigh scatterers, the collision induced absorption of H_2–H_2, H_2–He, and the scattering and absorption cross sections of crystalline, irregularly shaped Fe and $MgSiO_3$(c) cloud particles. See Mollière et al. (2019) for the full reference list and description of the opacity sources. The FeH opacity has been multiplied by a factor $1/2$ due to the partition function correction described in Charnay et al. (2018). We convolved the synthetic spectra using a Gaussian kernel, in order to approximate the line spread function of the SPHERE, GPI, and GRAVITY instruments. The instrumental resolving power was assumed to be 30, 45 and 500 for SPHERE, GPI and GRAVITY, respectively. The resolution element $\Delta\lambda$ of the spectrograph was assumed to be the FWHM of the line spread function. This means that the standard deviation σ_{LSF} of the Gaussian kernel used for convolution is defined by $\Delta\lambda = 2\sqrt{2\ln 2}\,\sigma_{LSF}$. The retrieval was run with PyMultiNest, using 4000 live points.

4.2.1. Pressure grid

We use an adaptive spacing for the atmosphere's pressure grid. For cloud-free calculations the atmosphere would be separated into 60 points, spaced equidistantly in log-pressure between 10^{-6} and 1000 bar. For cloudy calculations the retrieval code considers all cloud base pressures P_{base} and increases the spatial resolution for $P \in [0.5\,P_{base}, 1.12\,P_{base}]$ (corresponding to a pressure range of -0.3 and 0.05 dex) by a factor of 12. This better resolves the placement of the cloud base within the atmosphere, in case a P_{base} does not fall into the immediate proximity of a grid point of the coarse pressure grid. In addition, the abrupt increase in atmospheric opacity at the cloud deck position, and its decline $\propto P^{f_{sed}}$ for lower pressures, is resolved better. For two spatially separated cloud decks this leads to 104 grid points. We found that this treatment is as accurate as running the whole calculation at a 12 times-increased resolution, which would result in 720 grid points. For this we compared to a baseline calculation made at a 24 times higher resolution, using 1440 grid points.

4.2.2. Testing the retrieval model

Because all test retrievals mentioned in Sect. 3 were carried out on data sets of homogeneous wavelength coverage in *YJHK* bands, we verified our retrieval setup by running a mock retrieval that had the same wavelength spacing and error properties as the actual HR 8799e data sets of SPHERE, GPI, and GRAVITY. As input for the synthetic observation we used a posterior sample of the actual HR 8799e retrieval, the result is shown in Fig. E.1. We find that we can retrieve all parameters well, except for [Fe/H] and K_{zz}, which are biased, as are the scaling parameters. Running a second fit that neglected the scaling parameters lead to a well retrieved [Fe/H], but slightly too small radius. This behavior could be due to the random noise instantiation used in the retrieval, and the fact that especially the scaling may introduce biases in the retrieved atmospheric parameters in cases where differences are introduced between model and observation. This has also been described in Kitzmann et al. (2020). In their case the differences arose from using two different models for generating the mock observation and the retrieval, here the differences arise from the random noise properties. We note that the scaling value we retrieve for the actual HR 8799e data below is consistent with unity.

To test the impact of the random noise instantiation further we also ran test retrievals for the same synthetic observation, but neglecting the random perturbation of the data due to the noise. Similar to our noise-free test retrieval presented in Sect. 3.1, we found that the noise-free test led to a bi-modal posterior, with the modes bracketing the input values. An analogous approach (zooming in on the prior volume populated by the bi-modal posterior) lead to a uni-modal posterior, retrieving the input parameters. Thus we reconfirm our observation that noiseless test retrievals can lead to multi-modal posteriors if the ratio of the prior volume and the number of live points is large. This indicates that multiple parameter combinations can lead to excellent spectral fits, while within 1σ of the true parameters. As stated in the manuscript before this may indicate that a retrieval model with fewer free parameters could be favored, leading to more unique solutions. We note that in these test retrievals described here, the atmospheric C/O ratio was a robustly retrieved parameter in all retrieval setups, and that our two retrievals for the real HR 8799e data presented below, using either Cloud Model 1 or 2, lead to consistent C/O, [Fe/H] and $\log(g)$ constraints.
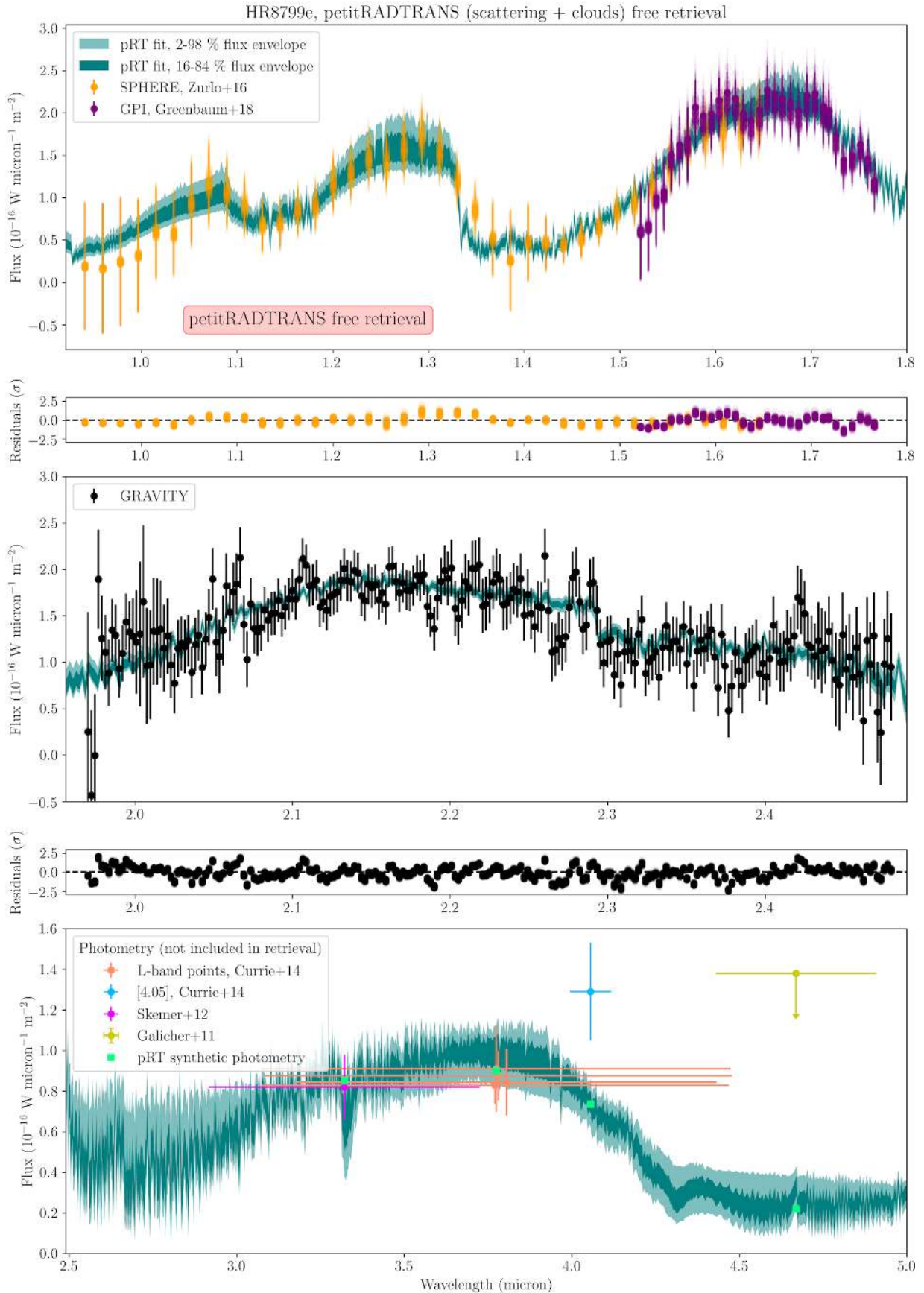
Fig. 2. Spectral fit of HR 8799e. *Upper panel*: *YJH*-band observations of SPHERE and GPI, *middle panel*: GRAVITY *K*-band observations. *Lowest panel*: photometry of the planet, which was not included during the retrieval. The 16–84 and 2–98% flux envelopes of the sampled *petitRADTRANS* retrieval models are shown in all panels. Because also the SPHERE and GPI scaling factors were sampled 100 times for making this plot, there are multiple points visible at every wavelength.
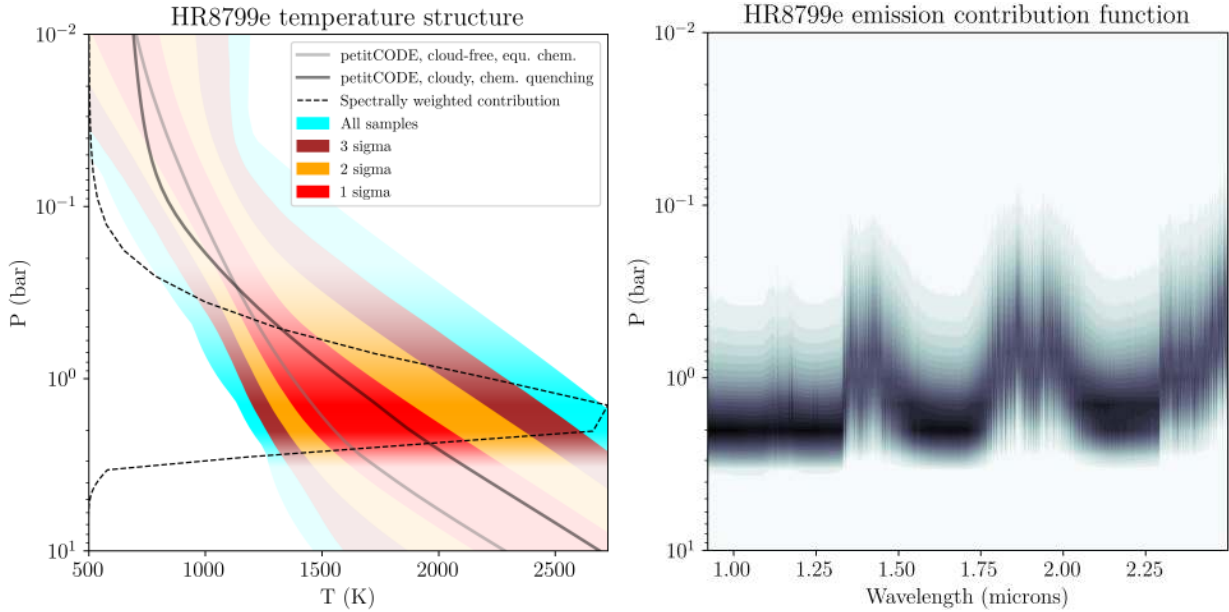
Fig. 3. *Left panel*: temperature distribution of the atmosphere of HR 8799e, retrieved with the *petitRADTRANS* free retrieval setup. See the caption of Fig. 1 for an explanation of how to read this plot. In addition the self-consistent P-T curves derived from *petitCODE*, assuming chemical equilibrium and no clouds, or chemical quenching with clouds, are shown as gray and black solid lines, respectively. *Right panel*: emission contribution function of the best-fit model of the HR 8799e retrieval.

4.3. Free retrieval results of HR 8799e

In this section, we describe the retrieval results. The results are also discussed in view of the possible planet formation history, and compared to existing literature studies of HR 8799e, in Sects. 5.1 and 5.2, respectively.

The spectral fit for HR 8799e is presented in Fig. 2. In general, the retrieval model is able to explain the observations well. The residuals scatter around zero, with some systematic differences visible at 1.325, 1.525, 1.725, 2.07, 2.11, 2.18, 2.275 and 2.425 micron. These differences could be due to the model missing absorbers, or not being flexible enough to fit intricacies in the atmospheric temperature, abundance, or cloud structure. Another likely possibility is that there are remaining systematics in the observations: the difference between the SPHERE and GPI observations in their overlap region (1.525 micron), as well as the overall wiggly appearance of the GRAVITY observation may indicate this. We note that also the photometric flux measurements in the MIR are fit well, except for the narrow [4.05] band measurement by Currie et al. (2014)[8]. This is especially interesting as these points were not included in the retrieval. Especially the 3.3 μm LBT - *L'* and *L'*-[4.05] colors have been noted to be difficult to explain with self-consistent models, see discussion in Sect. 5.2.

From the spectral appearance of the *H* and *K* band observations it is already clear that CH_4 is not an important absorber in the atmosphere: the flux decrease expected from CH_4 absorption at 1.6 and 2.2 micron is absent. The fact that only an upper limit is found by observations in the *M*-band, together with the comparatively high flux in the 3.3 μm LBT and *L'* bands, also speaks for a CH_4-poor atmosphere.

The retrieved pressure-temperature uncertainty envelopes of HR 8799e are shown in the left panel of Fig. 3. As in the analogous plot shown in Fig. 1, the opacity of the uncertainty

envelopes is scaled by the flux-averaged emission contribution function of the best-fit model, to show where in the atmosphere the observations are probing. This contribution function is also shown in the right panel of Fig. 3. For comparison, we also show the P-T curves derived with our self-consistent *petitCODE* using the best-fit parameters of the retrieval. This is further discussed in Sect. 4.5.1.

The corner plot of the one- and two-dimensional projections of the 18-dimensional posterior distribution of the retrieval is shown in Fig. 4. We summarize a few of the most striking results here, while the implications of the retrieved parameter values will be discussed in Sects. 5.1 and 5.2. In general, we note that none of the retrieved parameters ran into its prior boundaries and that all parameters (except for the Fe mass fraction at the cloud base) are well constrained. In addition, we derive that the atmospheric C/O ratio is C/O = $0.60^{+0.07}_{-0.08}$ and the retrieved metallicity is [Fe/H] = $0.48^{+0.25}_{-0.29}$. Together with the planet's mass, which we derive to be $4.81^{+8.78}_{-3.33}$ M_J, this has important implications for how the planet could have formed, see Sect. 5.1. We note that the surface gravity and radius retrieved for HR 8799e, $\log(g) = 4.00^{+0.46}_{-0.52}$ and $R_P = 1.12^{+0.09}_{-0.09}$ R_J, are constrained with a symmetric, uni-modal peak. Hence also the logarithm of the planet mass is constrained with a symmetric, uni-modal peak, but the distribution of the mass itself is skewed towards lower masses, with large mass uncertainties due to the large uncertainty on $\log(g)$. The atmosphere is clearly affected by disequilibrium chemistry, with a large quench pressure of $\log(P_{quench}/1\,\text{bar}) = 1.35^{+0.50}_{-0.56}$. We find that disabling quenching at the best-fit parameters leads to strong CH_4 absorption features in the spectrum, which are inconsistent with the data. Moreover, the scaling value retrieved for SPHERE is consistent with one, $1.01^{+0.08}_{-0.07}$. The scaling retrieved for GPI is significantly smaller than 1, namely $0.90^{+0.06}_{-0.05}$. Only when applying this scaling on GPI do the SPHERE and GPI datasets agree in their overlapping region, see Fig. 2. From sampling the posterior distribution 300 times, and calculating the spectra between

---

[8] We used *species* to convert the *petitRADTRANS* flux to photometric fluxes.
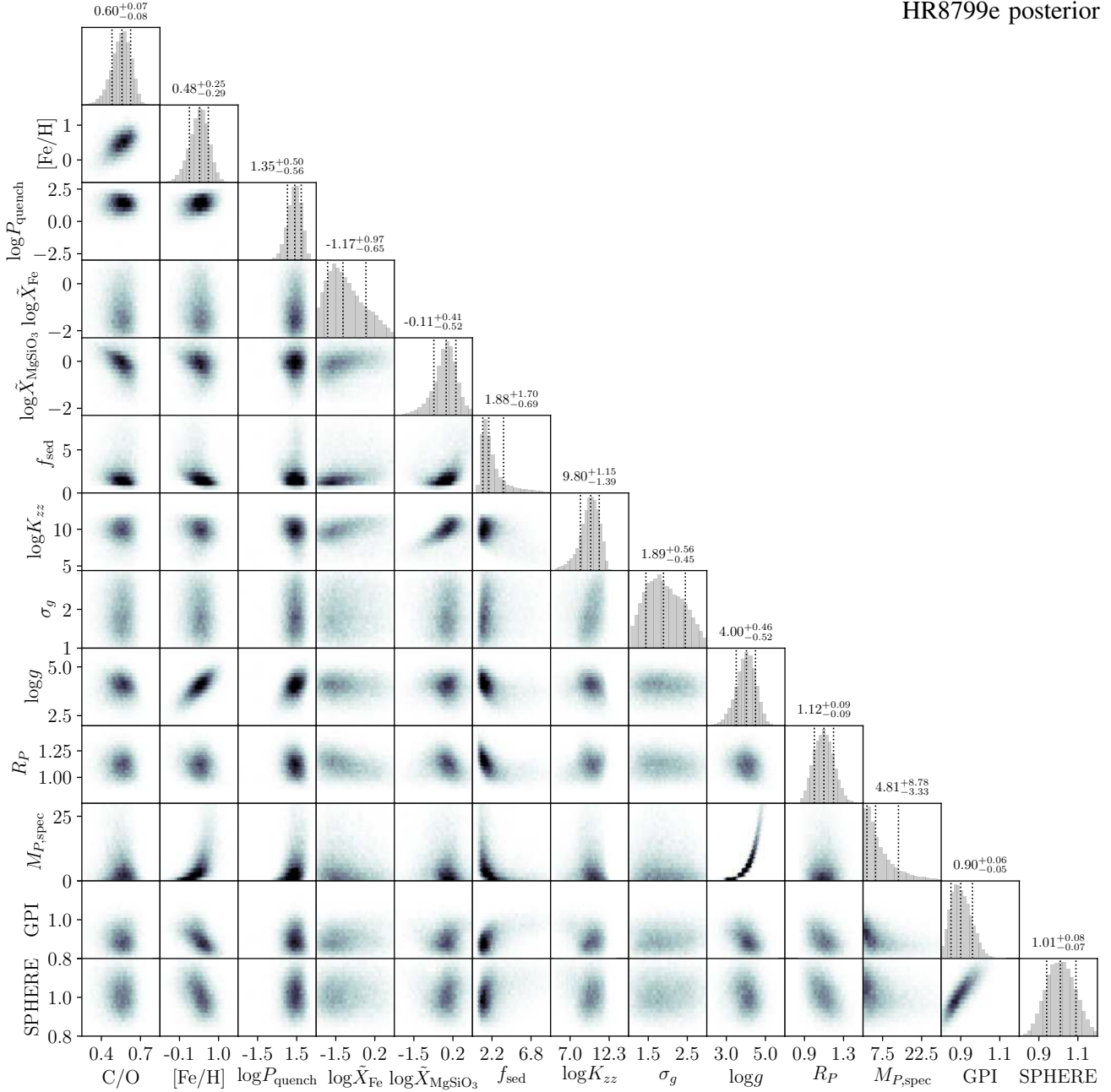
HR8799e posterior

Fig. 4. Corner plot of the free retrieval of HR 8799e with *petitRADTRANS*. This plot shows the one and two-dimensional projections of the 18-dimensional posterior distribution for all but the six nuisance parameters of the temperature structure (see the left panel of Fig. 3 for the P-T confidence envelopes). $M_{P,spec}$ is the mass of HR 8799e in units of Jupiter masses, derived from the retrieval posterior distributions of $\log(g)$ and R_P.

0.5 and 28 micron we derive an effective temperature of $T_{eff} = 1154^{+49}_{-48}$ K. Using this derived temperature, our retrieved surface gravity, and Eq. (4) of Zahnle & Marley (2014), we find that the upper limit for the atmospheric mixing is $\log(K_{zz,max}) = 10.2$, which assumes that all flux is transported by convection. Our derived value, which is used solely for determining the particle size for a given f_{sed}, is $\log(K_{zz}) = 9.80^{+1.15}_{-1.39}$, so below the theoretical upper limit (while still large).

4.4. Retrieval with the non-nominal Cloud Model 2

To test the robustness of the constrained atmospheric properties, we also retrieved HR 8799e with Cloud Model 2. Like before, we

find that Cloud Model 2 leads to retrieved temperature gradients which are too shallow when compared to physically consistent solutions. The retrieved atmospheric solution is bi-modal, with one solution corresponding to a P-T structure with a shallow temperature gradient and intermediate cloudiness, and a second solution corresponding to an even more isothermal, cloud-free atmosphere. The spectral fit, full posterior distribution, and P-T uncertainty envelopes are shown in Appendix F, where we also describe the prior setup of this retrieval.

Focusing on the bulk atmospheric properties ($\log(g)$, C/O, [Fe/H]), we find that the cloudy mode of the retrieval with Cloud Model 2 is fully consistent with the results from our nominal retrieval with Cloud Model 1. The one-dimensional posterior
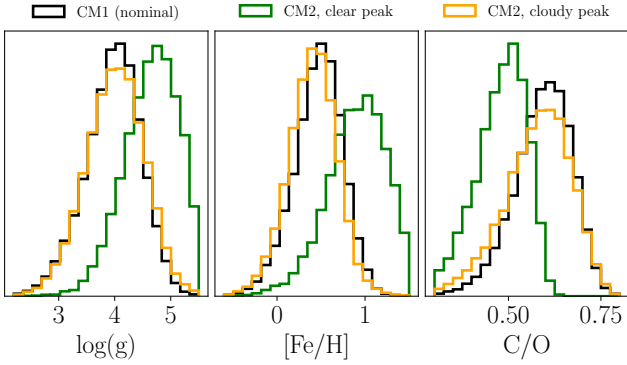
Fig. 5. Marginalized one-dimensional posterior distributions of HR 8799e's gravity, metallicity, and C/O, shown for our nominal retrieval with Cloud Model 1 (black solid line), and the retrieval with Cloud Model 2, which lead to a bi-modal solution of a cloudy (orange solid line) and ~isothermal, clear (green solid line) atmospheric state.

distributions look almost identical. Thus, even though the atmospheric temperature and cloud structure retrieved with different cloud models can differ, quantities such as C/O, [Fe/H] and the atmospheric gravity may be very robust. The clear, isothermal solution of Cloud Model 2 leads to different values for these retrieved parameters, but their 1σ uncertainty regions overlap. Figure 5 shows the marginalized one-dimensional posterior distributions of HR 8799e's gravity, metallicity, and C/O, derived with Cloud Model 1 and Cloud Model 2.

4.5. Comparison of the results with self-consistent atmospheric models

Free retrievals allow to deviate from the rigidity of (potentially imperfect) physical assumptions in self-consistent codes. This can be both boon and bane. On one hand it allows accounting for effects that influence the atmospheric structure which are not adequately captured by the physical assumptions made in the self-consistent codes. On the other hand, the free retrieval may converge on parameter results which lead to a seemingly good fit to the data, but are unphysical. Because of this, it is crucial to verify retrieval results by comparing to constraints that can be obtained from self-consistent codes, as was also done in, for example, Line et al. (2017); Gandhi & Madhusudhan (2018). We describe such tests below: we compared our derived atmospheric structure with self-consistent results and compared our free retrieval to a grid interpolation retrieval with self-consistent atmospheric spectra.

4.5.1. Self-consistent P-T structures of *petitCODE*

petitCODE (Mollière et al. 2015, 2017) is a self-consistent code for calculating atmospheric structures and spectra. It assumes radiative-convective and chemical equilibrium to calculate the atmospheric structure. The radiative transfer includes scattering, and the code can include gas line, continuum, and cloud opacities.

We carried out two tests: for the overall planet parameters (log (g), T_{eff}, [Fe/H], C/O), we used the median of the retrieved values. We then calculated a cloud-free atmospheric structure, assuming chemical equilibrium. As a second test we included clouds, prescribed the median values of the retrieved cloud parameters (f_{sed}, K_{zz}, σ_g, $\tilde{X}_{\text{Fe}}$, $\tilde{X}_{\text{MgSiO3}}$), and enforced that the H_2O, CH_4 and CO abundances be held constant below the retrieved best-fit quench pressure. The resulting structures

are shown as gray and black solid lines in the left panel of Fig. 3 for the cloud-free chemical equilibrium and the cloudy non-equilibrium structure, respectively.

Overall, it can be seen that the self-consistent structures follow the uncertainty envelopes of the retrieved pressure temperature structure well, in terms of absolute temperature and slope. Interestingly, we notice that the cloud-free structure in chemical equilibrium falls within the 1σ envelope at all pressures, while the self-consistent cloudy structure which included quenching moves out into the 2σ envelope between 0.05 and 0.3 bar. This is above the region of maximum emission as measured by the contribution function of the best-fit model, however. We conclude that we do not see any clear deviation of our retrieved temperature envelopes when compared to physical expectations, both the overall shape and absolute temperatures appear to be close to what is predicted in a self-consistent model.

4.5.2. Spectral fit with Exo-REM

In addition to the free retrieval with *petitRADTRANS* described above, we carried out a grid-interpolation retrieval to obtain the atmospheric parameters of HR8799e from its spectrum. The grid of self-consistent model spectra was obtained with *Exo-REM* (Baudino et al. 2015, 2017; Charnay et al. 2018), in the version by Charnay et al. (2018), which includes scattering and disequilibrium chemistry. The chemical disequilibrium and cloud scale height is determined through taking into account the vertical atmospheric mixing, which is set through the atmospheric eddy diffusion parameter K_{zz}. In *Exo-REM*, K_{zz} is determined from the atmospheric structure in the convective region consistently, using mixing length theory. Above the convective region, K_{zz} is determined from a convective overshooting description. *Exo-REM* includes the cloud opacities of spherical, amorphous Fe and Mg_2SiO_4 grains, as well as the gas opacities of Na, K, H_2O, CH_4, CO, CO_2, NH_3, PH_3, TiO, VO, and FeH, in addition to H_2–H_2 and H_2–He collision-induced absorption (CIA). The *Exo-REM* grid used here ranged in T_{eff} from 1000–2000 K (ΔT_{eff} = 50 K), and in C/O from 0.3–0.75 (ΔC/O = 0.05). The [Fe/H] grid points were −0.5, 0, 0.5, while the log (g) points were at 3.5, 4, 4.5.

Notable points of difference in the opacity treatment between *Exo-REM* and *petitRADTRANS* are the use of different alkali wing profiles (Burrows & Volobuyev 2003 for *Exo-REM*, Allard et al. 2003, 2012b for *petitRADTRANS*) and that *petitRADTRANS* assumed irragularly shaped, crystalline Fe and $MgSiO_3$ cloud particles instead of spherical, amorphous Fe and Mg_2SiO_4 ones. Moreover, *Exo-REM* varies C/O by changing C, whereas *petitRADTRANS* varies C/O by changing O, which can make a difference (Lodders 2010).

Applying a LSF convolution and rebinning identically to the *petitRADTRANS* retrieval, we used *species* to carry out a grid-based retrieval with *PyMultiNest*, where the model spectra were obtained by linearly interpolating within the *Exo-REM* grid. Hundred spectra sampled from the Exo-REM posterior are shown in Fig. 6, analogous to the spectra shown for the free retrieval with *petitRADTRANS* in Fig. 2.

In general the quality of the spectral fit is very similar to the free retrieval with *petitRADTRANS*, with the difference that the residuals in the near-IR are a somewhat larger. Like *petitRADTRANS*, *Exo-REM* can fit the photometry in the mid-IR well, even though it was not included in the fit. Analogous to the *petitRADTRANS* fit, also *Exo-REM* does not reproduce the [4.05] band photometric point well, but is somewhat more consistent than *petitRADTRANS*.
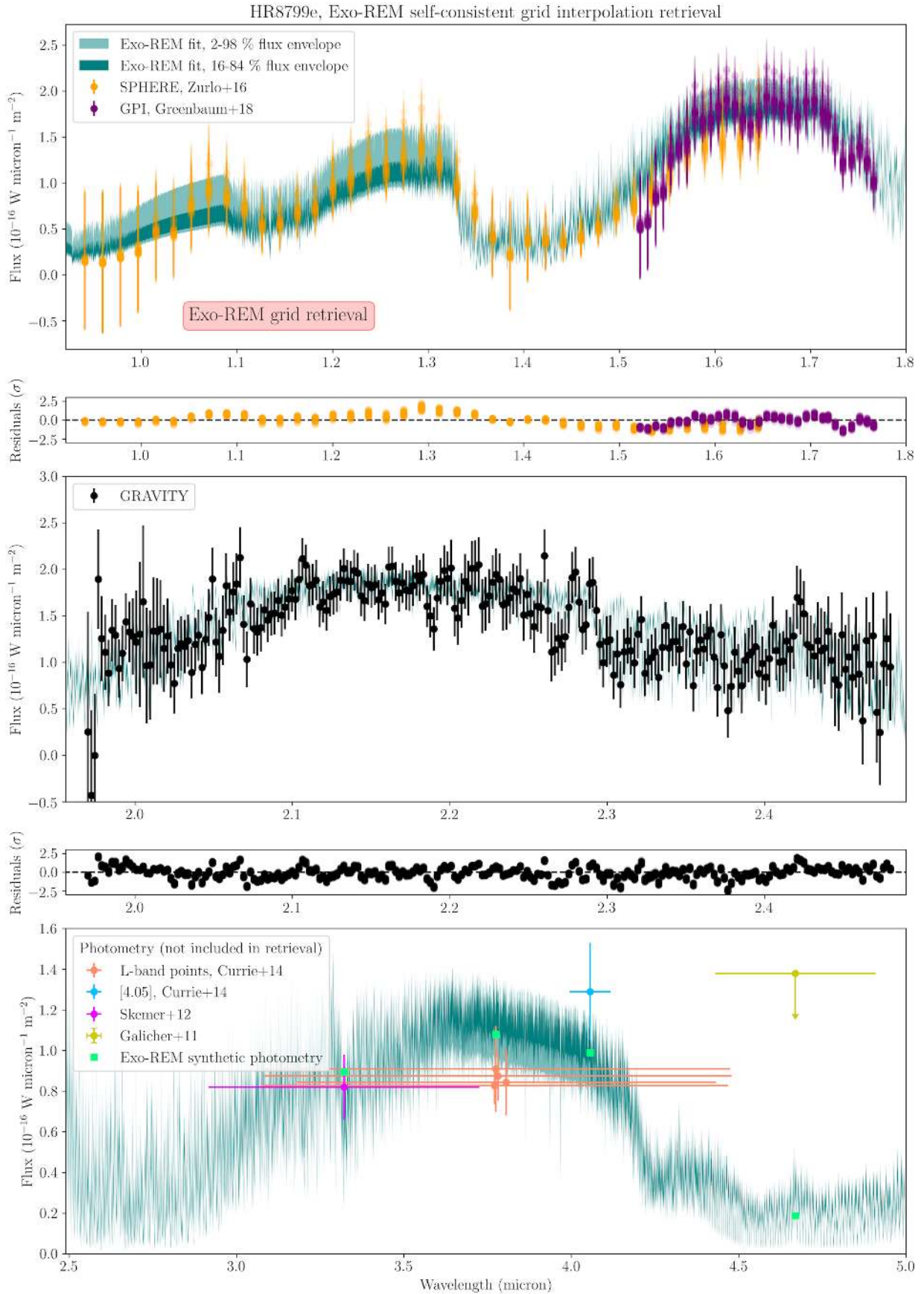
Fig. 6. Same as Fig. 2, but for a grid-based interpolation retrieval with the self-consistent code *Exo-REM*.
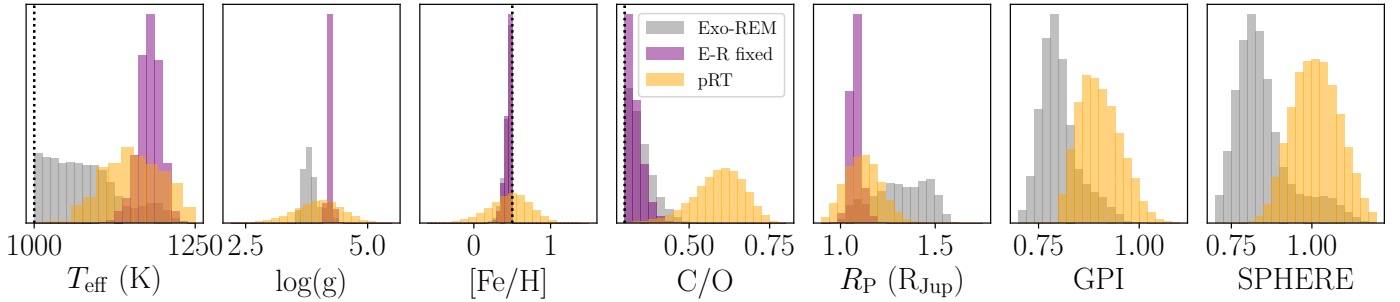
Fig. 7. One-dimensional marginalization of the HR 8799e retrieval posteriors for the self-consistent *Exo-REM* grid when retrieving the SPHERE/GPI flux scaling (gray), or fixing it to the median values of *petitRADTRANS* (purple). The results of the freely parameterized *petitRADTRANS* retrieval are shown in orange. The vertical dashed lines denote the upper [Fe/H], and lower C/O and T_{eff} grid boundaries of the *Exo-REM* models.

We show the one-dimensional marginalizations of the *Exo-REM* posterior distribution in Fig. 7, together with the marginalized posteriors of the corresponding parameters of the free *petitRADTRANS* retrieval. Because the effective temperature T_{eff} is not a free parameter of the *petitRADTRANS* retrieval, the T_{eff} distribution was obtained from sampling the posterior 300 times, and calculating the effective temperature from the spectrum ranging from 0.5 to 28 μm. We find two modes in the *Exo-REM* retrieval. The first mode has lower temperatures, smaller GPI and SPHERE scaling factors, and larger planet radii when compared to the *petitRADTRANS* retrieval. For the second mode the scaling factors, T_{eff} and radius are consistent with the *petitRADTRANS* results. The surface gravity agrees with the broad *petitRADTRANS* posterior, for both modes. The lower temperature mode of the *Exo-REM* fits runs into the lower grid boundary for the effective temperature. Both modes run into to lower grid boundary for C/O, and in the upper grid boundary for [Fe/H]. Due to the inherent difficulty of converging self-consistent cloudy models the *Exo-REM* grid is incomplete, that is, models in the grid are missing especially at lower temperatures. This could affect the fit at low effective temperatures. Therefore, as a second test, we fixed the GPI and SPHERE scaling factors to the best-fit values of the *petitRADTRANS* fit, which is also shown in Fig. 7. This leads to effective temperatures and radii consistent with the *petitRADTRANS* fit, while still running into the C/O and [Fe/H] boundaries.

For both the metallicity and C/O ratio the comparison between the *petitRADTRANS* and *Exo-REM* results is difficult: the retrieved best-fit metallicity value from the *Exo-REM* grid is consistent with the *petitRADTRANS* peak metallicity. However, the *Exo-REM* retrieval runs into the upper boundary of the grid. For the C/O ratio the *petitRADTRANS* fit peaks at around 0.6, while it is driven into the lower grid boundary (0.3) for the *Exo-REM* fit. The behavior of the *Exo-REM* posterior for these two quantities therefore makes a more detailed comparison of the *petitRADTRANS* and *Exo-REM* retrievals difficult. It appears as if the *Exo-REM* grid may not be able to fully reproduce the near-IR photometry, and trends into the boundaries of the grid in search of the true probability maximum that is not contained within its grid boundaries.

We can speculate at the near-IR region being the cause for the difficulties we face in the *Exo-REM* retrieval: here the residuals are larger than in the *petitRADTRANS* retrieval. Because the *YJH* bands are the most strongly affected by clouds, this could hint at a difference in the description of clouds, but also the alkalis could play a role, especially in the SPHERE *Y* band. Indeed, when only retrieving the atmospheric properties using

the GRAVITY *K*-band with *Exo-REM*, we find that the C/O is constrained at $0.43^{+0.07}_{-0.07}$, while [Fe/H] still trends into the upper grid boundary (0.5). A larger grid extend with additional [Fe/H] and C/O values, may therefore alleviate the problems we face in our analysis here. However, we expect that a grid that also varies the cloud parameters of *Exo-REM* may help, as we find that decreasing the C/O (or increasing the metallicity) leads to stronger water absorption features across the *YJH* bands. This speaks for C/O being used to counteract too strong cloud absorption.

We conclude that the *petitRADTRANS* and *Exo-REM* retrievals are consistent with each other in the parameters that can be easily compared. However, parameters trending into the grid boundary in the *Exo-REM* retrieval make the comparison difficult for [Fe/H] and C/O.

5. Discussion

5.1. Implication of the retrieved C/O and [Fe/H] for the formation location of HR 8799e

Measuring a planet's C/O has been suggested as a powerful tool to trace where in the protoplanetary disk a planet may have formed (e.g., Öberg et al. 2011; Madhusudhan et al. 2014b; Mordasini et al. 2016; Cridland et al. 2016). The general idea is to compare a planet's C/O to the C/O predicted for the disk's solid and gas phases, and using the planet's bulk enrichment to determine whether the planet's metal enrichment is dominated by solid or gas accretion. From this it may be possible to infer where in the disk a planet formed. If this formation location were to be conclusively shown to be further away from the star than the planet's current location, this would be a proof for orbital migration. In what follows we attempt such an analysis using the C/O ratio inferred for HR 8799e. We neglect the, potentially very important, effect of compositional gradients in the planet, that may lead to atmospheric abundances being different from the bulk of the planet (e.g., Leconte & Chabrier 2012; Vazan et al. 2018).

The C/O value that has been reported for the host star HR 8799 is C/O$_{\text{star}}$ = 0.56 (Sadakane 2006). The author finds that C and O have roughly solar abundances. The star is a λ-Boötis-type star, meaning that its iron peak elements are subsolar ([Fe/H] measured for iron specifically is -0.55 ± 0.1, Sadakane 2006). Because the analysis presented in this sub-section hinges on C/O and the total metal content of the planet, we neglect this additional information for now. We note, however, that if the composition of the photosphere of HR 8799 is representative

for the disk from which its planets formed, assuming solar abundance ratios for all elements except O[9] during the atmospheric modeling is problematic.

To constrain HR 8799e's formation history, we make the assumption that the disk from which the HR 8799 planets formed had a composition as specified in Table 1 of Öberg & Wordsworth (2019). The authors of the paper present this as a model for the young solar nebula. Because the C/O ratio of HR 8799 (C/O = 0.56) is essentially solar (C/O = 0.55, see Asplund et al. 2009) and we are most interested in the molecular volatiles, which represent the largest mass reservoir for solid planetary building blocks, we deem this assumption acceptable, even though HR 8799 is a λ-Boötis-type star.

Using the planetary mass of $4.81^{+8.78}_{-3.33}$ M_J, the metallicity of $0.48^{+0.25}_{-0.29}$ and the C/O of $0.60^{+0.07}_{-0.08}$, inferred from our spectral retrieval with *petitRADTRANS*, we find that HR 8799e is likely heavily enriched in ices, accreted in forms of pebbles and planetesimals, and most likely formed outside of the CO iceline. This conclusion was obtained from fitting the O/H and C/H content of HR 8799e, derived from our spectral retrieval, with an abundance model of a planet that forms in a disk as defined in Öberg & Wordsworth (2019). For this we treated the planets mass, accreted solid mass, and the accretion locations of the solids and gas as free parameters. For the planet mass a prior based on our spectrally retrieved mass was assumed. The planetary O/H and C/H was then fitted with *PyMultiNest*. Specifically, due to the increased planetary metallicity of $0.48^{+0.25}_{-0.29}$, we find that the planet has accreted between 65 and 360 $M_\oplus$ of ices (1σ range) that are mixed into its envelope and atmosphere. The large uncertainty stems from the large mass and metallicity uncertainties from our spectral retrieval. The atmospheric metal content of HR 8799e appears to be dominated by solids due to the large inferred atmospheric enrichment (e.g., Espinoza et al. 2017). If the metal content were dominated by gas accretion the atmospheric metallicity is expected to be close to, or smaller than, stellar. As we find that the planet has a C/O ratio consistent with its host star (we find C/O = $0.60^{+0.07}_{-0.08}$ for HR 8799e) this could mean that the planet has formed outside of the CO iceline. In particular, we derive that a formation location outside the CO iceline is more than twice as likely compared to a formation inside the CO iceline (more details on this analysis will be published in an upcoming study). This is explained by the fact that only outside the CO iceline the solid material in the disk will have the stellar value, see Öberg et al. (2011). In contrast, C/O values of 0.3, which is the lower boundary of the atmospheric grid approached in our *Exo-REM* retrieval, are possible if the planet formed within the CO iceline, where H_2O and CO_2 dominate the ice composition. If HR 8799e did form outside the CO iceline, this could mean that all HR 8799 planets formed outside the CO iceline, because HR 8799e is the innermost planet of the HR 8799 system. This will have to be tested by deriving C/Os and metallicities for all HR 8799 planets. The C/O analysis of the HR 8799 planets by Lavie et al. (2017) is consistent with this assessment: for the HR 8799b and c, where the authors succeeded in deriving C/O values, they found C/O ≥ C/O$_{star}$. Likewise, Konopacky et al. (2013); Barman et al. (2015) derive a C/O of $0.65^{+0.10}_{-0.05}$ for HR 8799c and a value between 0.58 and 0.7 for HR 8799b, which is consistent with the value we derive for HR 8799e.

If HR 8799e truly formed outside the CO iceline, this also allows to put constraints on the planet's possible migration. In Öberg & Wordsworth (2019), the CO iceline is situated at ~20 au

for the young solar nebula. In the disk around HR 8799, the same temperature due to the irradiation of the star would have been reached at ~45 au (Marois et al. 2010), neglecting the evolution of the stellar luminosity at young ages. Because HR 8799e resides at ~15 au (Wang et al. 2018), this could imply that the planet migrated significantly. This would be consistent with the finding by Wang et al. (2018) that the HR 8799 planets needed to migrate in the gas disk after formation to get locked into a stable resonant orbit.

The model for the disk composition in Öberg et al. (2011); Öberg & Wordsworth (2019) is strongly simplified. The disk's properties such as temperature, surface density, and abundance profiles are assumed to be static. The iceline positions are determined from a simple thermodynamic stability analysis of the ice species. Processes such as the viscous and chemical evolution of the disk are neglected. However, chemical evolution and ionization of the disk material can be of crucial importance, as well as the initial composition of the disk at the start of the evolution, as shown by Eistrup et al. (2016, 2018). Importantly, these studies describe how gas-grain chemistry may deplete CO from the gas phase within the CO iceline, condensing it in the form of CO_2, at the expense of also H_2O. Thus our conclusion regarding migration, based on the C/O ratio of HR 8799e, may be based on oversimplified disk chemistry assumptions.

To assess the effect of properly treating the disk chemistry we used the ANDES physical-chemical code to compute a 2D steady-state disk physical structure and time-dependent chemistry for the HR 8799 disk (Akimkin et al. 2013; Molyarova et al. 2017). The detailed setup of the model is described in Appendix G. For this setup of the disk chemical model we found that the CO iceline lies at around 100 au, which would indicate that HR 8799e migrated even further after formation. However, due to the above-mentioned gas-grain chemistry, CO is converted into CO_2 ice effectively starting from around 20 au. This makes the solid C/O in the disk approach stellar values already at 20 au, such that HR 8799e may have formed as close as 20 au from the star. Hence realistic disk chemistry modeling could indicate that the planet migrated much less than when compared to simplistic disk abundance models.

Finally, we note that the nitrogen content may be a better way of constraining a planets formation location in the disk (Öberg & Wordsworth 2019; Bosman et al. 2019), where a large N-content corresponds to a formation in the outer parts of the disk. However, this would require to study planets cooler than HR 8799e, for which most of its accreted nitrogen is in the form of N_2, and therefore invisible due to the low N_2 opacity.

5.2. Comparison of retrieval results with literature studies

Since their discovery (Marois et al. 2008, 2010; Currie et al. 2011) the HR 8799 planets have been extensively studied. We provide a summary of the studies that exist on the HR 8799 planets below and how they relate to our results for HR 8799e. We start with the qualitative properties of the planet, before comparing our retrieved values with the ones reported by others.

5.2.1. Clouds

In general, studies find that all HR 8799 planets have comparable surface gravities and temperatures. In addition, all studies find that the HR 8799 planets are dominated by thick clouds, as indicated by their red near-infrared (NIR) colors. For example, using NIR and MIR photometry and a grid of self-consistent atmospheric models, Madhusudhan et al. (2011b)

---

[9] We vary C/O by varying O.

find that HR 8799bcd are dominated by clouds much thicker than expected for field brown dwarfs. This is similar to an assessment already made by Bowler et al. (2010), studying HR 8799b. Marley et al. (2012) came to the conclusion that thick clouds are required when studying HR 8799bcd, but emphasized that this is not peculiar, and rather a consequence of cloud formation being gravity-dependent, a result that has also been borne out by the model calculations of Charnay et al. (2018).

Considering the apparent lack of comparison brown dwarfs that resemble the HR 8799 planets, it is important to remember that the earlier results in the literature are highly heterogeneous in terms of the available data (and models) that were used to arrive at a given conclusion. High quality spectra are important to truly unlock the planetary characteristics. For example, additional SPHERE spectroscopy obtained by Zurlo et al. (2016) allowed Bonnefoy et al. (2016) to show that especially HR 8799de can be well fit with low-gravity cloudy brown dwarfs of the late L spectral type. For HR 8799bc the picture is less clear, and Bonnefoy et al. (2016) find that good comparison objects can only be identified when reddening T-dwarf spectra with iron or silicate extinction. Given the abundance of literature reporting on the cloudiness of the HR 8799 planets, our finding that HR 8799e is cloudy does not come out of the blue (sky).

5.2.2. Disequilibrium chemistry

Disequilibrium chemistry has also been reported in the HR 8799 planets by a variety of studies. For planets such as HR 8799bcde, disequilibrium chemistry means that CO is more and CH_4 is less abundant then predicted from chemical equilibrium, due to atmospheric mixing overruling the chemical reactions in the upper atmosphere. Using OSIRIS H and K band spectra at low resolution, Barman et al. (2011) report on HR 8799b exhibiting weaker CH_4 absorption than expected, and that disequilibrium chemistry may be at play in order to decrease the CH_4 abundance. Similar findings were also reported for the medium-resolution OSIRIS data for planets b and c, where CH_4 was not detectable in c (Konopacky et al. 2013; Barman et al. 2015). Madhusudhan et al. (2011b) found that they had to neglect the 3.3 μm band containing too strong CH_4 absorption when fitting their chemical equilibrium models to the HR 8799bc data obtained by Currie et al. (2011) (there was only an upper limit available for d in the 3.3 μm band). Marley et al. (2012) reported disequilibrium chemistry to be important for HR 8799bcd. Skemer et al. (2014) report on disequilibrium chemistry for HR 8799cd, based on NIR and MIR (narrow and broad band) photometry. Based on photometry, Currie et al. (2014) report the need for disequilibrium in planets b and c, but less strongly for d and e. Lavie et al. (2017) presented the first free retrieval analysis of the HR 8799 planets and found that bc require disequilibrium chemistry, but not planets de. We note, however, that no K-band spectroscopy was available for planet e, and that the K-band seemed heavily affected by systematics for planet d. Because the K band is important for detecting the presence of CO, and hence for detecting disequilibrium chemistry, their finding for d and e to be in chemical equilibrium has to be taken with caution. Using SPHERE and the same GPI K band data as in Lavie et al. (2017), Bonnefoy et al. (2016) report that HR 8799de can be fit well with chemical equilibrium models when using a non-scattering-cloud Exo-REM (Baudino et al. 2015) grid. With the new Exo-REM models by Charnay et al. (2018), which include chemical disequilibrium and scattering clouds, it is found that HR 8799e requires disequilibrium to explain the GRAVITY K band data reported in Gravity Collaboration (2019). In summary,

while the consensus in the literature regarding HR 8799e is not entirely clear, we corroborate the finding of Gravity Collaboration (2019), which is based on a high-S/N spectrum in the K-band, that HR 8799e is affected by disequilibrium chemistry.

5.2.3. Reported modeling challenges

When comparing models to observations of the HR 8799 planets a few problems have been identified. First, the models can converge toward high effective temperatures and thus small radii, in order to conserve the total flux of the planet within the model. Some radii that have been reported are smaller than expected from theoretical standpoints. Due to electron degeneracy pressure, the radii of gas giant planets and brown dwarfs are always above 0.75 R_J and even this lowest limit is only reached after 10 Gyr of contraction, for brown dwarfs of 70 M_J (Chabrier et al. 2009). For ages up to 5 Gyr, radii are above 1 R_J for masses up to 30 M_J (Mordasini et al. 2012). For objects with ages below 100 Myr, the minimum radius is therefore expected to be above 1 R_J (Marley et al. 2012). A radius that is too small when compared to these theoretical constraints hints at shortcomings in the atmospheric model being used. Examples for such small inferred radii are found in Barman et al. (2011) (who report $R = 0.75$ R_J for HR 8799b), Greenbaum et al. (2018) (who report that some, but not all, of the model grids they tried have radii below 1 R_J for HR 8799cde) and Bonnefoy et al. (2016) (who reported the same for some but not for all of the models they applied to HR 8799bcde). One approach is to reject such unphysical radii right from the start, by tying the spectral model fitting to evolutionary models, which guarantees that physically consistent radii are used for the spectral analysis (Marley et al. 2012). Because this can worsen the spectral fit and thus lead to additional biases concerning inferred properties of the atmosphere, it is questionable how much is gained from such an approach (Barman et al. 2015). In this regard, limiting the radius in our HR 8799e retrieval to a minimum value of 0.9 R_J can be seen as a compromise between these two approaches, and our best-fit radius of $1.12^{0.09}_{-0.09}$ R_J is above the 1 R_J limit described above.

In a similar vein, it is useful to check whether the retrieved values of the planet's effective temperature, surface gravity and radius (hence also its luminosity and mass) are consistent with evolutionary models. Considering the evolutionary plots presented in Marley et al. (2012; their Figs. 8 and 11) and following their analysis, places our median $\log(g)$ and $T_{\rm eff}$ values between their 10 and 30 Myr isochrones, although our $\log(g)$ uncertainties also allow for ages in excess of 100 Myr. These values are consistent with the ages that can be inferred for the HR 8799 system (between 30 and 60 Myr, see Marley et al. 2012, for a more complete discussion). Similarly, our median $\log(g)$ and $T_{\rm eff}$ values lie between the evolutionary tracks of 5 and 10 M_J models, and again our large uncertainties on $\log(g)$ allow for masses well below 5 M_J and in excess of 10 M_J. The mass we derive from HR 8799e's spectrum, $M_P = 4.81^{+8.78}_{-3.33}$ M_J is certainly consistent with this assessment. We note here that Marley et al. (2012) made the assumption of a hot start evolution in their work, and that the HR 8799 system is young enough for hot and cold start differences to play a role. It was found by Spiegel & Burrows (2012); Marley et al. (2012); Marleau & Cumming (2014), however, that at least a classical cold start assumption (Marley et al. 2007) is ruled out for these planets. In any case, recent theoretical modeling of the physics of the accretion shock (Marleau et al. 2017, 2019) and of the structure of accreting planets (Berardo et al. 2017; Berardo & Cumming 2017; Cumming et al. 2018) suggests

Table 4. Comparison of reported properties of HR 8799e, derived from spectral/photometric analyses.

Parameter (unit)	This study *petitRADTRANS*	This study *Exo-REM2*	This study *Exo-REM2* fix SPHERE/ GPI scaling	B16 *Exo-REM1* [a]	L17 *Helios-r*	G18 *PHOENIX* [a]	C14 *COOLTLUSTY*
$\log(g)$ (cgs)	$4.00^{+0.46}_{-0.52}$	$3.82^{+0.26}_{-0.13}$	$4.23^{+0.04}_{-0.03}$	3.7	$3.9^{+0.05}_{-0.05}$ [b]	3.5	4
$T_{\rm eff}$ (K)	1154^{+49}_{-48}	1071^{+61}_{-50}	1180^{+16}_{-17}	1200	$-$ [c]	1100	1000
$R_{\rm P}$ ($R_{\rm J}$)	$1.12^{+0.09}_{-0.09}$	$1.32^{+0.15}_{-0.17}$	$1.08^{+0.03}_{-0.02}$	1.0	$1.2^{+0.05}_{-0.1}$ [b]	1.3	$R_{\rm evo}$ [d]
C/O	$0.60^{+0.07}_{-0.08}$	<0.3 [e]	<0.3 [e]	$-$	$\rightarrow 0$ [f]	$-$	$-$
[Fe/H]	$0.48^{+0.25}_{-0.29}$	>0.5 [g]	>0.5 [g]	0.5	0.4 [h]	$-$	$-$

Notes. The second *Exo-REM2* retrieval from our study fixed the SPHERE/GPI scaling factors to the best-fit values of *petitRADTRANS*. [a]These publications compared their observations to more than one model, we report the best-fit values of the models that provide the best fit to the data, either as stated by the authors or by visual inspection. [b]As read of by eye from their corner plot. [c]Not specified. [d]Consistent with radii derived from evolutionary models. [e]The best-fit model was trending into the boundaries of the Exo-REM grid, and only a boundary value can be reported for C/O. [f]No K-band spectrum was available for their analysis, such that the CO abundance could not be constrained. Their retrieved C/O ratio is pushed against the 0 boundary. [g]Same as (e), but for the [Fe/H] parameter. [h]As derived from their stated O/H ratio, in comparison to the solar O/H ratio.

References. B16 (Bonnefoy et al. 2016), L17 (Lavie et al. 2017), G18 (Greenbaum et al. 2018), C14 (Currie et al. 2014). Code references: *Exo-REM2* (version of Charnay et al. 2018), *Exo-REM1* (version of Baudino et al. 2015), *Helios-r* (Lavie et al. 2017), *PHOENIX* (models reported in Barman et al. 2011), *COOLTLUSTY* (Sudarsky et al. 2003; Hubeny et al. 2003; Burrows et al. 2006).

that hot starts are more likely. The spectroscopic mass derived in our study is also consistent with models studying the orbital evolution of the HR 8799 system, where it was found that the mass of HR 8799e has to be below 7.6 $M_{\rm J}$ to ensure orbital stability (Wang et al. 2018).

In addition, models that assume a homogeneous cloud cover have been reported to have trouble at reproducing all the MIR photometry simultaneously, especially the 3.3 μm, L' and [4.05] bands. The use of patchy cloud models, or models mixing clouds of different vertical extent, has been shown to be one promising way of solving this problem (see, e.g., Currie et al. 2011, 2014; Skemer et al. 2012, 2014). With this in mind it is interesting to see that our high likelihood retrieval models reproduce both the 3.3 μm and L' band photometry, without including these data points in the retrieval. This would speak against a heterogeneous cloud cover being necessary to explain the data. A similar result was found by Bonnefoy et al. (2016) when fitting the HR 8799 planets with the Exo-REM code (Baudino et al. 2015) in the non-scattering, chemical equilibrium version. It is important to note, however, that neither Bonnefoy et al. (2016) nor we can reproduce the [4.05] band photometry of Currie et al. (2014), which is consistently higher than our best-fit spectra, and all the best-fit models presented in Bonnefoy et al. (2016), for all of the HR 8799 planets. Thus, a heterogeneous cloud cover cannot be ruled out. Given the prevalence of variability of, especially, L-T dwarf transition objects (see, e.g., Apai et al. 2013; Crossfield 2014), such a heterogeneous cloud cover could be expected for the HR 8799 planets.

5.2.4. Quantitative comparison to the literature

Finally, we compare our retrieved parameter values for HR8799e with those reported in the literature. All values are listed in Table 4. Our derived gravity value, $\log(g) = 4.00^{+0.46}_{-0.52}$, falls within the values reported in the literature, which range from 3.5 to 4. The same holds for the effective temperature, for which we retrieve 1154^{+49}_{-48} K. This value is bracketed by the reported values, ranging between 1000 and 1200 K. Similarly, our retrieved

radius value ($1.12^{+0.09}_{-0.09}$ $R_{\rm J}$) falls between the reported values of 1 to 1.3 $R_{\rm J}$. Including ours, four out of six data–model comparison studies also varied the metallicity. Reported values are between 0.4 and 0.5, and our value ([Fe/H] $= 0.48^{+0.25}_{-0.29}$) is consistent with their assessments. C/O deserves a more detailed discussion. Only one other study considered the atmospheric C/O. The retrieved values range from 0 to $0.60^{+0.07}_{-0.08}$, the latter being the value we derive with *petitRADTRANS* in this study. As discussed in Sect. 4.5.2, the Exo-REM fit runs into the grid boundaries for both C/O and [Fe/H], making it difficult for us to assess whether this trend to low C/O ratios is actually merited by the data or whether this parameter is used to compensate for the too stringent cloud description. The value reported by Lavie et al. (2017; C/O = 0) suffers from the fact that K-band spectroscopy was not available at the time of their study, such that the CO abundance could not be determined. They only report an upper limit for the atmospheric C/H value, and their best-fit spectrum does not show any CO absorption in the K-band, which we detect in the GRAVITY data. Given these caveats we conclude that a comparison to the C/O ratios derived in other studies is at this point inconclusive.

6. Summary

We present a new version of our retrieval radiative transfer code *petitRADTRANS* (Mollière et al. 2019) to which we added the effect of multiple scattering. This enables us to run free retrievals on cloudy self-luminous objects such as directly imaged planets and brown dwarfs. This updated version of *petitRADTRANS* will be available on the *petitRADTRANS* website soon[10] and is already available now, upon request.

Running verification retrievals on synthetic observations, we found that we can retrieve the input parameters. The high dimensionality of the input model can lead to small offsets within the observational uncertainties, however. Increasing the number of live points in our retrievals with the nested sampling method

---

[10] https://petitradtrans.readthedocs.io

improves this, but we expect this to be a persisting problem for models with a large number of free parameters, especially for high S/N observations which lead to narrow posterior distributions and thus let a smaller fraction of the prior volume be of interest.

We tested two different cloud models. The first is the physically motivated Ackerman & Marley (2001) cloud model. Our second cloud model simply retrieves the wavelength properties of the cloud opacities. When running retrievals with Cloud Model 2 on mock observations made with Cloud model 1, we find that we can get an excellent fit but with biased atmospheric parameters. In particular the planet's photosphere is found to be more isothermal and less cloudy than the input, which mimics the shallow temperature gradients predicted by Tremblin et al. (2015, 2016, 2017, 2019). Thus, retrievals alone will likely not be enough to investigate whether shallow temperature gradients indeed occur in such thought-to-be-very-cloudy atmospheres. Retrieval analyses could be aided by longer wavelength data in the mid-IR, however, which could reveal the spectral features of cloud particles at ~10 micron (Cushing et al. 2006).

We ran our retrieval setup on archival GPI, SPHERE, and partially new GRAVITY data for the directly imaged planet HR 8799e, using Cloud Model 1. Applying such data-driven, free retrievals for directly imaged planets becomes possible with our high S/N observations, especially GRAVITY's K band spectra at a spectral resolution of $R = 500$. In addition, observations in the K band can probe H_2O, CH_4 and CO features and are therefore crucial for constraining atmospheric disequilibrium chemistry and the atmospheric C/O.

We are able to fit the observations well, and confirm a cloudy atmosphere dominated by disequilibrium chemistry. We also compare our retrieved atmospheric spectra to archival photometric observations in the L', $3.3\,\mu m$, [4.05] and M band. Our spectra are consistent with all points, except for the [4.05] band point, although the photometry has not been included in the retrieval. The $L'-3.3\,\mu m$ and $L'-$[4.05] band colors have been suggested to require heterogeneous cloud coverage to explain the data, and we cannot confirm this for the $L'-3.3\,\mu m$ color.

The posterior parameter values we retrieve for the atmospheric $\log(g)$, T_{eff} and [Fe/H] are consistent with previous studies, and hot start evolutionary calculations. For the first time, we successfully constrain the C/O of HR 8799e and find that it is $0.60^{+0.07}_{-0.08}$, which is consistent with stellar. Running additional retrievals on HR 8799e with Cloud Model 2, we find that the retrieved planetary C/O, [Fe/H] and $\log(g)$ are identical to the values found with the nominal Cloud Model 1, therefore independent of our cloud model choice. This is noteworthy as the Cloud Model 2 retrievals lead to less cloudy, more isothermal atmospheres. This indicates that C/O may be a quite robust outcome of the retrievals. We also fit the HR 8799e spectrum with the self-consistent code *Exo-REM*, which uses a state-of-the-art one-dimensional cloud model, scattering, and disequilibrium chemistry. Our free retrieval results compare well to the *Exo-REM* fit, except for the C/O ratio. With *Exo-REM* the C/O ratio is driven into the grid boundary (C/O < 0.3), as is the metallicity derived from *Exo-REM* ([Fe/H] > 0.5). This makes the comparison between the free retrieval with *petitRADTRANS* and *Exo-REM* difficult for C/O. A larger *Exo-REM* grid, which also varies the free parameters of its cloud prescription, may resolve this issue.

Using our retrieved C/O and metallicity, and a highly simplified disk model, we find that HR 8799e could have formed outside the CO iceline. This would imply that the planet migrated significantly. Because HR 8799e is the innermost planet of the HR 8799 system, this could indicate that all HR 8799 planets formed outside of the CO iceline. Similar formation distances, relative to the icelines, have been theorized for Jupiter in the Solar System (Öberg & Wordsworth 2019; Bosman et al. 2019). Using sophisticated gas-grain chemical modeling for the protoplanetary disk we find that the planet could also have formed more closely to the star, but outside the CO_2 iceline. This would require less migration.

7. Outlook

Here we introduce the first version of our retrieval framework for cloudy scattering atmospheres. With this we introduce a versatile tool for interpreting the spectra of directly imaged planets and brown dwarfs. At the same time, it is clear that there are many avenues for improving and testing our method.

One could explore different pressure-temperature parameterizations, cloud model setups, or abundance models. The latter could mean, for example, retrieving absorber abundances independently, as is often done in retrieval studies. Alternatively, the assumption of chemistry could be kept, while retrieving not just C/O (which we varied by changing O) and metallicity, but by retrieving C/H, O/H, and other atomic abundance ratios instead (as was done in, e.g., Lavie et al. 2017; Spake et al. 2019). Another important addition will be to include a parameterization for heterogeneous cloud coverage.

It is also crucial to test what happens when retrieving atmospheric parameters using a model setup that is different from the one used to generate a synthetic observation. In this way one can begin to quantify the uncertainties and biases of retrieved parameters given the model choices, and how robust certain parameters are against using a wrong model. An example is the robustness of C/O that we found in our results here, when using different cloud models. In principle, the most likely among a set of models can be found using the Bayes factor, computed with the model evidences derived from nested sampling. However, a given model may be worse than another in terms of model assumptions, which could be highly unphysical, while still being favored by a Bayes factor analysis. Thus, such comparison retrieval studies offer additional insight regarding the real parameter uncertainties, including the modeling choices.

On the observational side, additional data in the mid-IR, ideally spectroscopy, is necessary to explore the properties of clouds further. This is because the L', $3.3\,\mu m$ and [4.05] band may encode information about a heterogeneous cloud cover. Excitingly, the outer planets (HR 8799bcd) will be studied in the mid-IR, using NIRSpec IFU spectroscopy with the JWST[11]. HR 8799bcde will also be studied with photometry in the mid-IR with JWST[12]. ESO's ERIS[13] instrument, to be mounted at the VLT, is a promising option, as well as KPIC, to be mounted on Keck II (Mawet et al. 2018). ESO's imminent CRIRES+ instrument may offer the possibility to study the HR 8799 planets at high spectral resolution, where the S/N of the planetary flux measurement could be boosted using the cross-correlation method (e.g., Hoeijmakers et al. 2018), which also allows for carrying out retrievals (Brogi & Line 2019). Further in the future, the METIS spectrograph (Brandl et al. 2014) of ESO's upcoming ELT telescope as well as the PSI instrument (Skemer et al. 2018) on the TMT will be excellent instruments for mid-IR observations. For studying the potential absorption feature of silicate

---

[11] http://www.stsci.edu/jwst/phase2-public/1188.pdf
[12] http://www.stsci.edu/jwst/phase2-public/1194.pdf
[13] http://www.eso.org/sci/facilities/develop/instruments/eris.html

clouds at 10 micron, JWST will again be an excellent instrument. Here one approach could be to obtain mid-IR spectra of HR 8799 planet analogs such as PSO J318 (Liu et al. 2013)[14].

Lastly, also constraining basic parameters of the HR 8799 planets better could be of great help. If astrometry were to give a mass estimate for the HR 8799 planets, a prior on the planet mass would lead to a better log (*g*) and therefore [Fe/H] inference during the retrievals (these two parameters being correlated). This, in turn, would also help to understand the planet's formation better, because the planetary metallicity can be regarded as a measure for the relative importance of the solid body accretion of a planet.

References

Ackerman, A. S., & Marley, M. S. 2001, ApJ, 556, 872
Akimkin, V., Zhukovska, S., Wiebe, D., et al. 2013, ApJ, 766, 8
Allard, F., Hauschildt, P. H., Alexander, D. R., Tamanai, A., & Schweitzer, A. 2001, ApJ, 556, 357
Allard, N. F., Allard, F., Hauschildt, P. H., Kielkopf, J. F., & Machin, L. 2003, A&A, 411, L473
Allard, F., Homeier, D., & Freytag, B. 2012a, Phil. Trans. R. Soc. London, Ser. A, 370, 2765
Allard, N. F., Kielkopf, J. F., Spiegelman, F., Tinetti, G., & Beaulieu, J. P. 2012b, A&A, 543, A159
Amundsen, D. S., Tremblin, P., Manners, J., Baraffe, I., & Mayne, N. J. 2017, A&A, 598, A97
Apai, D., Radigan, J., Buenzli, E., et al. 2013, ApJ, 768, 121
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Barman, T. S., Macintosh, B., Konopacky, Q. M., & Marois, C. 2011, ApJ, 733, 65
Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, ApJ, 804, 61
Barstow, J. K. 2020, MNRAS, 497, 4183
Baudino, J.-L., Bézard, B., Boccaletti, A., et al. 2015, A&A, 582, A83
Baudino, J.-L., Mollière, P., Venot, O., et al. 2017, ApJ, 850, 150
Benneke, B., & Seager, S. 2012, ApJ, 753, 100
Berardo, D., & Cumming, A. 2017, ApJ, 846, L17
Berardo, D., Cumming, A., & Marleau, G.-D. 2017, ApJ, 834, 149
Biller, B. A., Vos, J., Bonavita, M., et al. 2015, ApJ, 813, L23
Bonnefoy, M., Zurlo, A., Baudino, J. L., et al. 2016, A&A, 587, A58
Bosman, A. D., Cridland, A. J., & Miguel, Y. 2019, A&A, 632, L11
Bowler, B. P., Liu, M. C., Dupuy, T. J., & Cushing, M. C. 2010, ApJ, 723, 850
Brandl, B. R., Feldt, M., Glasse, A., et al. 2014, SPIE, 9147, 747
Brogi, M., & Line, M. R. 2019, AJ, 157, 114
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Burningham, B., Marley, M. S., Line, M. R., et al. 2017, MNRAS, 470, 1177
Burrows, A., & Volobuyev, M. 2003, ApJ, 583, 985
Burrows, A., Sudarsky, D., & Hubeny, I. 2006, ApJ, 640, 1063
Chabrier, G., Baraffe, I., Leconte, J., Gallardo, J., & Barman, T. 2009, AIP Conf. Ser., 1094, 102
Charnay, B., Bézard, B., Baudino, J. L., et al. 2018, ApJ, 854, 172
Cridland, A. J., Pudritz, R. E., & Alessi, M. 2016, MNRAS, 461, 3274
Crossfield, I. J. M. 2014, A&A, 566, A130
Cumming, A., Helled, R., & Venturini, J. 2018, MNRAS, 477, 4817
Cuppen, H. M., Walsh, C., Lamberts, T., et al. 2017, Space Sci. Rev., 212, 1
Currie, T., Burrows, A., Itoh, Y., et al. 2011, ApJ, 729, 128
Currie, T., Burrows, A., Girard, J. H., et al. 2014, ApJ, 795, 133
Cushing, M. C., Roellig, T. L., Marley, M. S., et al. 2006, ApJ, 648, 614
Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2016, A&A, 595, A83
Eistrup, C., Walsh, C., & van Dishoeck, E. F. 2018, A&A, 613, A14
Espinoza, N., Fortney, J. J., Miguel, Y., Thorngren, D., & Murray-Clay, R. 2017, ApJ, 838, L9
Feautrier, P. 1964, C. R. Acad. Sci. Paris, 258, 3189
Feroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, Open J. Astrophys., 2, 10
Fisher, C., & Heng, K. 2018, MNRAS, 481, 4698
Fu, Q., & Liou, K. N. 1992, J. Atmos. Sci., 49, 2139
Galicher, R., Marois, C., Macintosh, B., Barman, T., & Konopacky, Q. 2011, ApJ, 739, L41
Gandhi, S., & Madhusudhan, N. 2018, MNRAS, 474, 271
Gao, P., Marley, M. S., & Ackerman, A. S. 2018, ApJ, 855, 86
Gao, P., Thorngren, D. P., Lee, G. K. H., et al. 2020, Nat. Astron., in press
Gordon, S., & McBride, B. J. 1994, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications. Part 1: Analysis (Washington, DC, USA: National Aeronautics and Space Administration)
Gravity Collaboration (Abuter, R., et al.) 2017, A&A, 602, A94
Gravity Collaboration (Lacour, S., et al.) 2019, A&A, 623, L11
Gravity Collaboration (Nowak, M., et al.) 2020, A&A, 633, A110
Greenbaum, A. Z., Pueyo, L., Ruffio, J.-B., et al. 2018, AJ, 155, 226
Guillot, T. 2010, A&A, 520, A27
Helling, C., & Casewell, S. 2014, A&ARv, 22, 80
Helling, C., Woitke, P., & Thi, W. F. 2008a, A&A, 485, 547
Helling, C., Ackerman, A., Allard, F., et al. 2008b, MNRAS, 391, 1854
Hoeijmakers, H. J., Schwarz, H., Snellen, I. A. G., et al. 2018, A&A, 617, A144
Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011
Irwin, P. G. J., Teanby, N. A., de Kok, R., et al. 2008, J. Quant. Spectr. Rad. Transf., 109, 1136
Kitzmann, D., Heng, K., Oreshenko, M., et al. 2020, ApJ, 890, 174
Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, Science, 339, 1398
Kreidberg, L., Line, M. R., Thorngren, D., Morley, C. V., & Stevenson, K. B. 2018, ApJ, 858, L6
Lacis, A. A., & Oinas, V. 1991, J. Geophys. Res., 96, 9027
Lacour, S., Dembet, R., Abuter, R., et al. 2019, A&A, 624, A99
Lavie, B., Mendonça, J. M., Mordasini, C., et al. 2017, AJ, 154, 91
Leconte, J. 2018, ApJ, 853, L30
Leconte, J., & Chabrier, G. 2012, A&A, 540, A20
Lee, H. H., Roueff, E., Pineau des Forets, G., et al. 1998, A&A, 334, 1047
Lee, J.-M., Fletcher, L. N., & Irwin, P. G. J. 2012, MNRAS, 420, 170
Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, ApJ, 778, 97
Line, M. R., Zhang, X., Vasisht, G., et al. 2012, ApJ, 749, 93
Line, M. R., Wolf, A. S., Zhang, X., et al. 2013a, ApJ, 775, 137
Line, M. R., Knutson, H., Deming, D., Wilkins, A., & Desert, J.-M. 2013b, ApJ, 778, 183
Line, M. R., Fortney, J. J., Marley, M. S., & Sorahana, S. 2014a, ApJ, 793, 33
Line, M. R., Knutson, H., Wolf, A. S., & Yung, Y. L. 2014b, ApJ, 783, 70
Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S. 2015, ApJ, 807, 183
Line, M. R., Marley, M. S., Liu, M. C., et al. 2017, ApJ, 848, 83
Liu, M. C., Magnier, E. A., Deacon, N. R., et al. 2013, ApJ, 777, L20
Lodders, K. 2010, Formation and Evolution of Exoplanets (Berlin: Wiley), 157
MacDonald, R. J., & Madhusudhan, N. 2017, MNRAS, 469, 1979
MacDonald, R. J., & Madhusudhan, N. 2019, MNRAS, 486, 1292
Mace, G. N., Kirkpatrick, J. D., Cushing, M. C., et al. 2013, ApJS, 205, 6
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Madhusudhan, N., & Seager, S. 2011, ApJ, 729, 41
Madhusudhan, N., Harrington, J., Stevenson, K. B., et al. 2011a, Nature, 469, 64
Madhusudhan, N., Burrows, A., & Currie, T. 2011b, ApJ, 737, 34

---

[14] https://www.stsci.edu/jwst/phase2-public/1275.pdf

Madhusudhan, N., Crouzet, N., McCullough, P. R., Deming, D., & Hedges, C. 2014a, ApJ, 791, L9

Madhusudhan, N., Amin, M. A., & Kennedy, G. M. 2014b, ApJ, 794, L12

Marleau, G. D., & Cumming, A. 2014, MNRAS, 437, 1378

Marleau, G.-D., Klahr, H., Kuiper, R., & Mordasini, C. 2017, ApJ, 836, 221

Marleau, G.-D., Mordasini, C., & Kuiper, R. 2019, ApJ, 881, 144

Marley, M. S., & Robinson, T. D. 2015, ARA&A, 53, 279

Marley, M. S., Fortney, J. J., Hubickyj, O., Bodenheimer, P., & Lissauer, J. J. 2007, ApJ, 655, 541

Marley, M. S., Saumon, D., Cushing, M., et al. 2012, ApJ, 754, 135

Marley, M. S., Ackerman, A. S., Cuzzi, J. N., & Kitzmann, D. 2013, Comparative Climatology of Terrestrial Planets, eds. S. J. Mackwell, A. A. Simon-Miller, J. W. Harder, & M. A. Bullock (Tucson: University of Arizona Press), 367

Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348

Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, Nature, 468, 1080

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425

Mawet, D., Bond, C. Z., Delorme, J. R., et al. 2018, Proc. SPIE, 10703, 1070306

McBride, B. J., & Gordon, S. 1996, Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications II. User's Manual and Program Description (Washington, DC, USA: National Aeronautics and Space Administration)

Min, M., Hovenier, J. W., & de Koter, A. 2005, A&A, 432, 909

Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, ApJ, 813, 47

Mollière, P., van Boekel, R., Bouwman, J., et al. 2017, A&A, 600, A10

Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, A&A, 627, A67

Molyarova, T., Akimkin, V., Semenov, D., et al. 2017, ApJ, 849, 130

Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012, A&A, 547, A111

Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, ApJ, 832, 41

Morley, C. V., Fortney, J. J., Marley, M. S., et al. 2012, ApJ, 756, 172

Morley, C. V., Marley, M. S., Fortney, J. J., et al. 2014, ApJ, 787, 78

Ng, K.-C. 1974, J. Chem. Phys., 61, 2680

Öberg, K. I., & Wordsworth, R. 2019, AJ, 158, 194

Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJ, 743, L16

Ohno, K., & Okuzumi, S. 2018, ApJ, 859, 34

Ohno, K., Okuzumi, S., & Tazaki, R. 2020, ApJ, 891, 131

Olson, G. L., Auer, L. H., & Buchler, J. R. 1986, J. Quant. Spectr. Rad. Transf., 35, 431

Ormel, C. W., & Min, M. 2019, A&A, 622, A121

Parmentier, V., & Guillot, T. 2014, A&A, 562, A133

Pinhas, A., Rackham, B. V., Madhusudhan, N., & Apai, D. 2018, MNRAS, 480, 5314

Pinhas, A., Madhusudhan, N., Gandhi, S., & MacDonald, R. 2019, MNRAS, 482, 1485

Piskunov, N. E., Kupka, F., Ryabchikova, T. A., Weiss, W. W., & Jeffery, C. S. 1995, A&AS, 112, 525

Powell, D., Zhang, X., Gao, P., & Parmentier, V. 2018, ApJ, 860, 18

Robinson, T. D., & Catling, D. C. 2012, ApJ, 757, 104

Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P.-O., & Tinetti, G. 2016, ApJ, 833, 120

Rodgers, C. D. 2000, Inverse Methods for Atmospheric Sounding: Theory and Practice (Singapore: World Scientific)

Rossow, W. B. 1978, Icarus, 36, 1

Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, J. Quant. Spectr. Rad. Transf., 111, 2139

Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, J. Quant. Spectr. Rad. Transf., 130, 4

Sadakane, K. 2006, PASJ, 58, 1023

Samland, M., Mollière, P., Bonnefoy, M., et al. 2017, A&A, 603, A57

Semenov, D., & Wiebe, D. 2011, ApJS, 196, 25

Skemer, A. J., Hinz, P. M., Esposito, S., et al. 2012, ApJ, 753, 14

Skemer, A. J., Marley, M. S., Hinz, P. M., et al. 2014, ApJ, 792, 17

Skemer, A. J., Stelter, D., Mawet, D., et al. 2018, Proc. SPIE Conf. Ser., 10702, 10702A5

Skilling, J. 2004, AIP Conf. Ser., 735, 395

Sousa-Silva, C., Al-Refaie, A. F., Tennyson, J., & Yurchenko, S. N. 2015, MNRAS, 446, 2337

Spake, J. J., Sing, D. K., Wakeford, H. R., et al. 2019, MNRAS, submitted [arXiv:1911.08859]

Spiegel, D. S., & Burrows, A. 2012, ApJ, 745, 174

Stolker, T., Quanz, S. P., Todorov, K. O., et al. 2020, A&A, 635, A182

Su, K. Y. L., Rieke, G. H., Stapelfeldt, K. R., et al. 2009, ApJ, 705, 314

Sudarsky, D., Burrows, A., & Hubeny, I. 2003, ApJ, 588, 1121

Toon, O. B., & Ackerman, T. P. 1981, Appl. Opt., 20, 3657

Tremblin, P., Amundsen, D. S., Mourier, P., et al. 2015, ApJ, 804, L17

Tremblin, P., Amundsen, D. S., Chabrier, G., et al. 2016, ApJ, 817, L19

Tremblin, P., Chabrier, G., Baraffe, I., et al. 2017, ApJ, 850, 46

Tremblin, P., Padioleau, T., Phillips, M. W., et al. 2019, ApJ, 876, 144

Tsiaras, A., Waldmann, I. P., Zingales, T., et al. 2018, AJ, 155, 156

Vazan, A., Helled, R., & Guillot, T. 2018, A&A, 610, L14

Venot, O., Hébrard, E., Agúndez, M., et al. 2012, A&A, 546, A43

Venot, O., Hébrard, E., Agúndez, M., Decin, L., & Bounaceur, R. 2015, A&A, 577, A33

Waldmann, I. P., Rocchetto, M., Tinetti, G., et al. 2015, ApJ, 813, 13

Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, AJ, 156, 192

Wende, S., Reiners, A., Seifahrt, A., & Bernath, P. F. 2010, A&A, 523, A58

Woitke, P., & Helling, C. 2004, A&A, 414, 335

Woitke, P., Helling, C., & Gunn, O. 2020, A&A, 634, A23

Yorke, H. W., & Bodenheimer, P. 2008, ASP Conf. Ser., 387, 189

Yurchenko, S. N., & Tennyson, J. 2014, MNRAS, 440, 1649

Yurchenko, S. N., Barber, R. J., & Tennyson, J. 2011, MNRAS, 413, 1828

Zahnle, K. J., & Marley, M. S. 2014, ApJ, 797, 41

Zalesky, J. A., Line, M. R., Schneider, A. C., & Patience, J. 2019, ApJ, 877, 24

Zurlo, A., Vigan, A., Galicher, R., et al. 2016, A&A, 587, A57

[1] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
e-mail: molliere@mpia.de

[2] Leiden Observatory, Leiden University, Postbus 9513, 2300 RA Leiden, The Netherlands

[3] Institute for Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland

[4] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cit, 5 place Jules Janssen, 92195 Meudon, France

[5] Max Planck Institute for extraterrestrial Physics, Giessenbachstraße 1, 85748 Garching, Germany

[6] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany

[7] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

[8] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

[9] Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

[10] CENTRA, Centro de Astrofísica e Gravitação, Instituto Superior Técnico, Avenida Rovisco Pais 1, 1049 Lisboa, Portugal

[11] STAR Institute, Université de Liège, Allée du Six Août 19c, 4000 Liège, Belgium

[12] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

[13] Universidade do Porto, Faculdade de Engenharia, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

[14] Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

[15] Núcleo de Astronomía, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejercito 441, Santiago, Chile

[16] Space Telescope Science Institute, Baltimore, MD 21218, USA

[17] Institute of Astronomy, Russian Academy of Sciences, 48 Pyatnitskaya St., Moscow, 119017, Russia

[18] Department of Chemistry, Ludwig-Maximilians-Universität, Butenandtstraße 5-13, 81377 Munich, Germany

[19] Escuela de Ingeniería Industrial, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejercito 441, Santiago, Chile

[20] Institut für Astronomie und Astrophysik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany

[21] University of Exeter, Physics Building, Stocker Road, Exeter, EX4 4QL, UK

[22] School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

[23] Center for Space and Habitability, Universität Bern, Gesellschaftsstrasse 6, 3012 Bern, Switzerland

Appendix A: K-table mixing

This section describes our process for obtaining the total cumulative opacity distribution function, hereafter called k-table, of the atmosphere. The resulting k-table contains the contribution of all absorbers. This combined k-table is required when including scattering during the computation of emission spectra. The combination is achieved by sampling the k-tables of the individual absorber species.

A.1. Gaussian quadrature grid definition

For every species, *petitRADTRANS* stores k-tables as a function of pressure and temperature. The spectral bins for which the k-tables are stored have the width $\Delta\lambda$. The width varies as a function of wavelength, because it is chosen such that $\lambda/\Delta\lambda = 1000$. Within these bins, the opacity of each species is stored as a function of the cumulative probability g, where $g = 0$ denote the lowest opacity values, and $g = 1$ the highest values. These are the k-tables. An introduction to correlated-k, and why such a k-table treatment is useful to speed up calculations when compared to line-by-line calculations, is given in Irwin et al. (2008), their Sect. 2. Especially their Fig. 1 illustrates the definition of g as the spectral coordinate, when compared to the wavelength. At low pressures, where the effect of pressure broadening is small, the k-tables of a line absorber will have a low-opacity tail extending over most of the g values. In addition, the line cores will give rise to a steep increase in opacity (by orders of magnitude), when approaching g values of unity. In order to sufficiently resolve the cumulative opacity distribution for low pressures, where the rise to highest opacity values will occur over a very narrow range at high g values, we split our g grid in two parts. The first is a Gaussian quadrature grid with coordinate g^{low} extending from $g^{\mathrm{low}} = 0$ to $g^{\mathrm{low}} = 0.9$, while the second is a Gaussian grid extending from $g^{\mathrm{high}} = 0.9$ to $g^{\mathrm{high}} = 1$. Both grids have eight points, and their weights w have been rescaled such that

$$\sum_{i=1}^{8} w_i^{\mathrm{low}} = 0.9 \quad \text{and} \quad \sum_{i=1}^{8} w_i^{\mathrm{high}} = 0.1. \tag{A.1}$$

This weight rescaling guarantees that, for any λ- and g-dependent function f,

$$
\begin{aligned}
\langle f \rangle &= \frac{1}{\Delta\lambda} \int_{\lambda-\Delta\lambda/2}^{\lambda+\Delta\lambda/2} f(\lambda)\mathrm{d}\lambda \\
&= \int_0^1 f(g)\mathrm{d}g \\
&\approx \sum_{i=1}^{8} \left[w_i^{\mathrm{low}} f(g_i^{\mathrm{low}}) + w_i^{\mathrm{high}} f(g_i^{\mathrm{high}}) \right].
\end{aligned}
\tag{A.2}
$$

A.2. Sampling

For obtaining the total k-table, we use the standard assumption for the on-the-fly combination of the opacities of different absorbers (see, e.g., Lacis & Oinas 1991; Irwin et al. 2008; Mollière et al. 2015; Amundsen et al. 2017), namely that their opacity distribution functions are independent (also called "random overlap"). Making this assumption, an opacity value of the total k-table can be sampled by drawing opacity samples from the k-tables of each individual species, scaling them according to the respective abundances of the absorbers, and adding them.
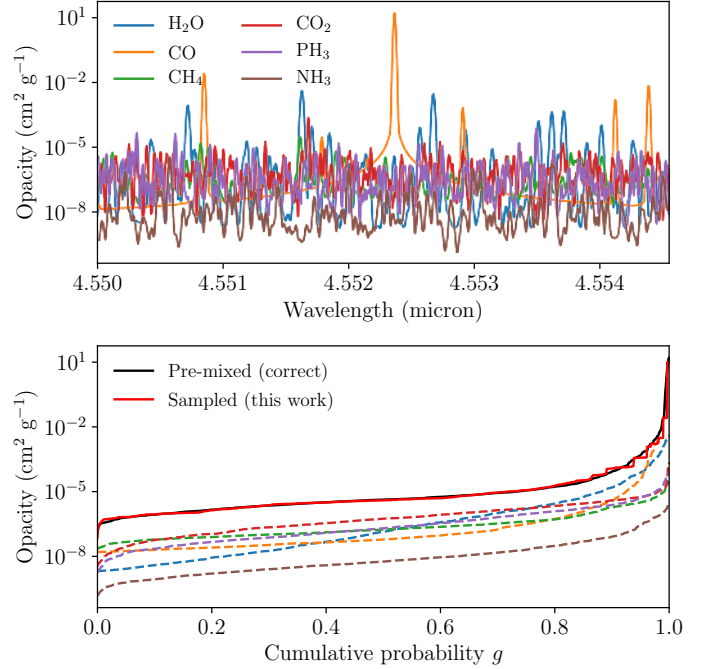
Fig. A.1. *Upper panel*: wavelength-dependent opacities of H$_2$O, CO, CH$_4$, CO$_2$, PH$_3$, and NH$_3$, in the spectral range between 4.55 and 4.5545 micron. *Lower panel*: K-table curves of the individual species (dashed lines, same color coding as in the upper panel), the total k-table obtained from premixing the opacities of all species in wavelength space (black solid line), and total k-table obtained with the method used in this work (red solid line).

By sampling the total k-table in this fashion many times, a good approximation of the total opacity distribution function can be constructed: the sampled values are simply sorted in magnitude, with the lowest value corresponding to $g = 0$, and the highest value corresponding to $g = 1$. This approximated k-table can then be interpolated back to the g^{low} and g^{high} values, to be ready for use in *petitRADTRANS*.

The sampling of a species' k-table could be done by drawing a uniformly distributed number between 0 and 1, and obtaining the opacity value by interpolating in the k-table of the species, to the so-drawn g value. However, this step needs to be done a number of times, and a numerically less costly option is to randomly selected an index numbering the opacities of the species' 16-point k-table, and treating the tabulated opacity at this index as the sampled value. Doing this for all species, and scaling the opacities by the abundances, yields again a sample of the total k-table, when the samples are added.

One subtlety is that not every index is equally likely, because we store the opacities on two eight-point Gaussian quadrature grids for every species. Hence, when drawing samples κ^{tot} of the total opacity distribution function, the corresponding opacity value and its not-yet-normalized weight w^{tot} are

$$\kappa^{\mathrm{tot}} = \sum_{i=1}^{N_{\mathrm{spec}}} X_i \kappa_{j(i)}^i, \tag{A.3}$$

$$w^{\mathrm{tot}} = \prod_{i=1}^{N_{\mathrm{spec}}} w_{j(i)}, \tag{A.4}$$

where X_i is the mass fraction of species i, and $j(i)$ denotes the k-table index sampled for species i.
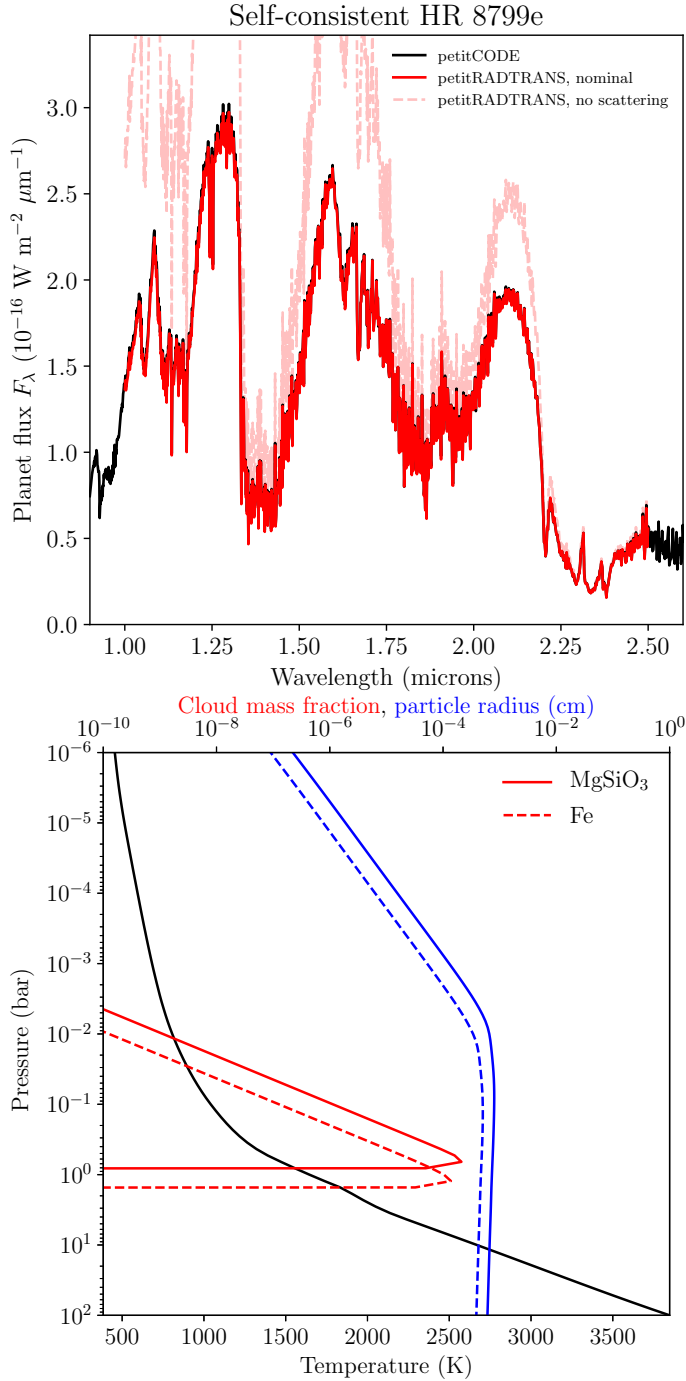
Fig. A.2. *Upper panel*: emission spectrum of a synthetic, self-consistent model of HR 8799e. The *petitCODE* calculation is shown as a black solid line, whereas the *petitRADTRANS* calculations, with and without scattering, are shown as red solid, and light red dashed lines, respectively. *Lower panel*: self-consistent atmospheric structure used for generating the spectra, showing the temperature (black solid line), cloud mass fractions (red lines), and cloud particle radii (blue lines) for $MgSiO_3$ (solid lines) and Fe (dashed lines).

The total sampled k-table can then be obtained by sorting the sampled (κ^{tot}, w^{tot}) pairs by their κ^{tot} values, and normalizing all sampled w^{tot} weights such that their sum equals unity. The corresponding g coordinate of the sorted, sampled points is then equal to the cumulative sum of the rescaled w^{tot} weights. The

thus-constructed k-table can then be used in *petitRADTRANS*, after interpolating back to the g^{low} and g^{high} values.

In order to achieve a sufficient sampling of the opacity values in the g^{low} grid, we set up a sampling that draws the low indices three times as often[15]. To conserve their actual weight, we multiply the weights w^{low} by a factor $1/3$ during this process.

Lastly, we speed up the k-table computation by neglecting those species i for which

$$X_i \kappa_i(g = 1) < 0.01 \times \max_j \left[X_j \kappa_j(g = 0) \right]. \tag{A.5}$$

That is, a species whose maximum opacity value (at $g = 1$) is more than 100 times smaller than the minimum opacity value (at $g = 0$) that is the largest among all species, is deemed negligible.

A.3. Example of the k-table mixing

In Fig. A.1, we show an example of the k-table mixing outlined above. As in all spectral calculations presented in this work, each species is sampled 1000 times. Here we used the HR 8799e model calculated with *petitCODE* (see Appendix B) to show the k-table combination at a pressure of 10^{-5} bar, that is, at pressures so low that the opacities of each species will vary by multiple orders of magnitude in a given spectral bin. We reduced the abundances of PH_3 and CH_4 by factors of 0.001 and 0.01, respectively, to show an example of a k-table combination where no species is clearly dominating the opacity.

The upper panel of Fig. A.1 shows the wavelength-dependent opacities of H_2O, CO, CH_4, CO_2, PH_3, and NH_3, in the spectral range between 4.55 and 4.5545 micron. This spectral width corresponds to the width of a typical *petitRADTRANS* wavelength bin in the mid-infrared. The lower panel shows the k-table (i.e., $\kappa(g)$) curves of the individual species (dashed lines, same color coding as in the upper panel). It also shows the total k-table, obtained from premixing the opacities in wavelength space (black solid line), as well as the total k-table obtained with the method used in this work (red solid line), both of which agree very well.

Appendix B: Verification of scattering treatment in cloudy emission spectra

Here we compare the radiative transfer results of *petitCODE* and *petitRADTRANS*, when using a self-consistent atmospheric structure of *petitCODE* as the input for a *petitRADTRANS* calculation. This is done to verify the k-table mixing (see Appendix A) and scattering implementation (see Sect. 2.1) of *petitRADTRANS*.

The model we use as an input is broadly motivated by the properties of HR 8799e, with the following input parameters: $T_{eff} = 1200$ K, $\log(g) = 3.75$ (with g in cm s^{-2}), [Fe/H] = 0, C/O = 0.55, $f_{sed} = 3$, $\sigma_g = 2$, $K_{zz} = 10^{8.5}$ cm^{2}s^{-1}, where f_{sed} is the settling parameter as defined in Ackerman & Marley (2001), σ_g is the width of the log-normal particle size distribution, and K_{zz} is the atmospheric eddy diffusion coefficient. The rest of the symbols have their usual meaning. The clouds were assumed to consist of irregularly-shaped particles. For this we used cloud opacities calculated with the code by Min et al. (2005), which also makes uses of software by Toon & Ackerman (1981). The code assumes a distribution of hollow spheres (DHS)

---

[15] This is done by randomly drawing the indices from an array containing (1,1,1,2,2,2,3,3,3,4,4,4,5,5,5,6,6,6,7,7,7,8,8,8,9,10,11,12,13,14,15,16).

for the particles. Moreover, the condensates were assumed to be crystalline.

The upper panel of Fig. A.2 shows the comparison of the resulting spectra of *petitCODE* and *petitRADTRANS*, which agree excellently. In addition, we also show the *petitRADTRANS* spectrum which would result from neglecting scattering. It is evident that scattering is an important process that needs to be accounted for. For completeness, the lower panel of Fig. A.2 shows the self-consistent structures of the atmospheric temperature, cloud mass fraction and particle radius, resulting from the *petitCODE* calculation, which is used as an input for producing the spectrum with *petitRADTRANS*.

Appendix C: Verification retrieval with Cloud Model 1 and zoomed-in prior ranges

Table C.1. Priors of the "zoomed-in" verification retrieval of Cloud Model 1.

Parameter	Prior	Parameter	Prior
T_1	$\mathcal{U}(0, T_2)$	$\log(\tilde{X}_{Fe})$	$\mathcal{U}(-1.3, 0)$
T_2	$\mathcal{U}(0, T_3)$	$\log(\tilde{X}_{MgSiO_3})$	$\mathcal{U}(-1.3, 0)$
T_3	$\mathcal{U}(0, T_{connect})$	f_{sed}	$\mathcal{U}(1.8, 4.2)^{(*)}$
$\log(\delta)$	$P_{phot} \in [10^{-3}, 100]$	$\log(K_{zz})$	$\mathcal{U}(7.6, 9.9)^{(*)}$
α	$\mathcal{U}(1, 2)$	σ_g	$\mathcal{U}(1.2, 3)^{(*)}$
T_0	$\mathcal{U}(500, 2000)$	R_P	$\mathcal{U}(0.85, 1.05)^{(*)}$
C/O	$\mathcal{U}(0.45, 0.6)^{(*)}$	$\log(g)$	$\mathcal{U}(3.65, 4.2)^{(*)}$
[Fe/H]	$\mathcal{U}(-0.1, 0.3)^{(*)}$	$\log(P_{quench})$	$\mathcal{U}(-6, -1)^{(*)}$

Notes. Prior ranges marked with $^{(*)}$ have been changed when compared to the initial run, so as to "zoom in" on the high likelihood region identified by the initial retrieval. The symbols have the same meaning as in Table 3.

Figure C.1 shows the follow-up retrieval of the verification retrieval with Cloud Model 1, discussed in Sect. 3. For this retrieval the uniform prior boundaries were chosen to enclose the regions of highest likelihood inferred in the original Cloud Model 1 retrieval, so as to test the effects of using the same number of live points in a smaller prior volume. The prior ranges are listed in Table C.1.

Appendix D: Retrieving Cloud Model 1 with Cloud Model 2

In Fig. D.1 we show the result of the retrieval when a synthetic spectrum made with Cloud Model 1 is retrieved using Cloud Model 2. See Sect. 3.2 for the discussion.

Appendix E: Synthetic retrieval assuming HR 8799e-like input data

Figure E.1 shows the results of the synthetic retrieval using HR 8799e-like data, described in Sect. 4.2.2. The black points in the *K*-band correspond to synthetic observations with the same data quality as the GRAVITY observations reported in Gravity Collaboration (2019), while the gray points show the data at the same quality as found for the two new GRAVITY observations presented in this work. All three synthetic GRAVITY data sets were fitted simultaneously.

Appendix F: Retrieval of HR8799e with the (non-nominal) Cloud Model 2

Table F.1. Priors of the non-nominal HR 8799e retrieval with Cloud Model 2.

Parameter	Prior	Parameter	Prior
T_1	$\mathcal{U}(0, T_2)$	$\log(\kappa_0)$	$\mathcal{U}(-8, 7)$
T_2	$\mathcal{U}(0, T_3)$	ξ	$\mathcal{U}(-6, 1)$
T_3	$\mathcal{U}(0, T_{connect})^{(a)}$	$\log(P_{base})$	$\mathcal{U}(-6, 3)$ & Eq. (7)
$\log(\delta)$	$P_{phot} \in [10^{-3}, 100]^{(b)}$	ω	$\mathcal{U}(0, 1)$
α	$\mathcal{U}(1, 2)$	f_{sed}	$\mathcal{U}(0, 10)$
T_0	$\mathcal{U}(300, 2300)$	R_P	$\mathcal{U}(0.9, 2)$
C/O	$\mathcal{U}(0.1, 1.6)$	$\log(g)$	$\mathcal{U}(2, 5.5)$
[Fe/H]	$\mathcal{U}(-1.5, 1.5)$	$\log(P_{quench})$	$\mathcal{U}(-6, 3)$
f_{SPHERE}	$\mathcal{U}(0.8, 1.2)$	f_{GPI}	$\mathcal{U}(0.8, 1.2)$

Notes. $\mathcal{U}$ stands for a uniform distribution, with the two parameters being the range boundaries. The units for the parameters are the same as the ones used for Table 1. f_{SPHERE} and f_{GPI} are the scaling factors retrieved for the SPHERE and GPI data, respectively. $^{(a)}$ and $^{(b)}$: please see Sect. 4.2 for a definition of P_{phot} and $T_{connect}$.

In Fig. F.1 we show the best-fit spectrum, two- and one-dimensional marginalized posterior, and *P-T* uncertainty envelopes resulting from retrieving the HR 8799e observations with Cloud Model 2. This retrieval is discussed in Sect. 4.4. The priors used for this retrieval are given in Table F.1.

Appendix G: Setup of the ANDES disk model

Here we describe the setup of the ANDES disk model (Akimkin et al. 2013; Molyarova et al. 2017), used to study the impact of sophisticated disk chemical modeling on the inferences made about the formation location of planet HR 8799e, see Sect. 5.1. The gas surface density distribution was described by a tapered power law with an exponent $\gamma = 1$ and a characteristic radius $R_c = 100$ au, the total gas mass in the disk was 0.1 $M_\odot$. The disk midplane temperature was calculated from the stellar and accretion luminosities. The thermal structure of the disk upper layers was calculated using ray tracing in the UV-optical wavelengths. The vertical disk density structure was derived from the hydrostatic equilibrium and was iteratively made consistent with the temperature structure. The gas and dust temperatures were assumed to be equal. The chemical model is based on the gas-grain code ALCHEMIC (Semenov & Wiebe 2011) with desorption energies updated according to Cuppen et al. (2017). An MRN-like power law size distribution (that is, following Mathis, Rumpl, & Nordsieck 1977) with the maximum dust size of 25 micron was used. The average dust radius of 0.37 micron and a dust-to-gas ratio of 0.01 were used for chemical simulations. The assumed stellar parameters representative of HR 8799 were defined using an evolutionary track model for 1.47 $M_\odot$ star (Yorke & Bodenheimer 2008). Accretion on the star was described by adding an accretion region with a temperature of 15 000 K, contributing to accretion luminosity and UV radiation field. The assumed accretion rate was 10^{-8} $M_\odot$/yr. Cosmic rays, X-rays and radioactive nuclides were included as additional sources of ionization. The chemical evolution of the disk was run for 1 Myr, starting from a pre-calculated composition of a 1 Myr old molecular cloud with "low metal" initial abundances (Lee et al. 1998).
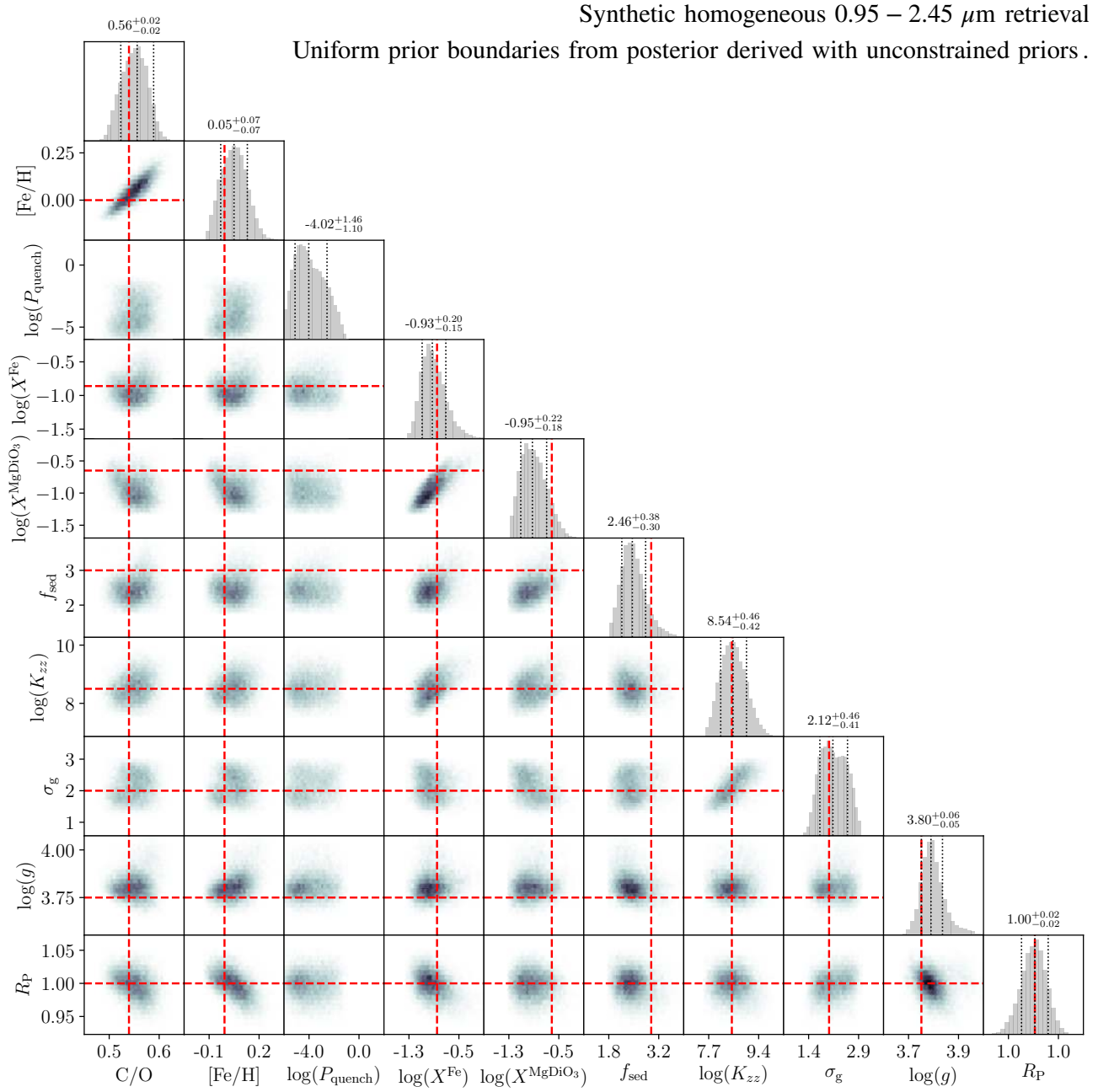
Fig. C.1. Posterior distribution of the follow-up retrieval with Cloud Model 1 (see Sect. 3). For this retrieval the uniform prior boundaries were chosen to enclose the regions of highest likelihood inferred in the original Cloud Model 1 retrieval, so as to test the effects of using the same number of live points in a smaller prior volume. The uniform prior boundaries are described in Appendix C.
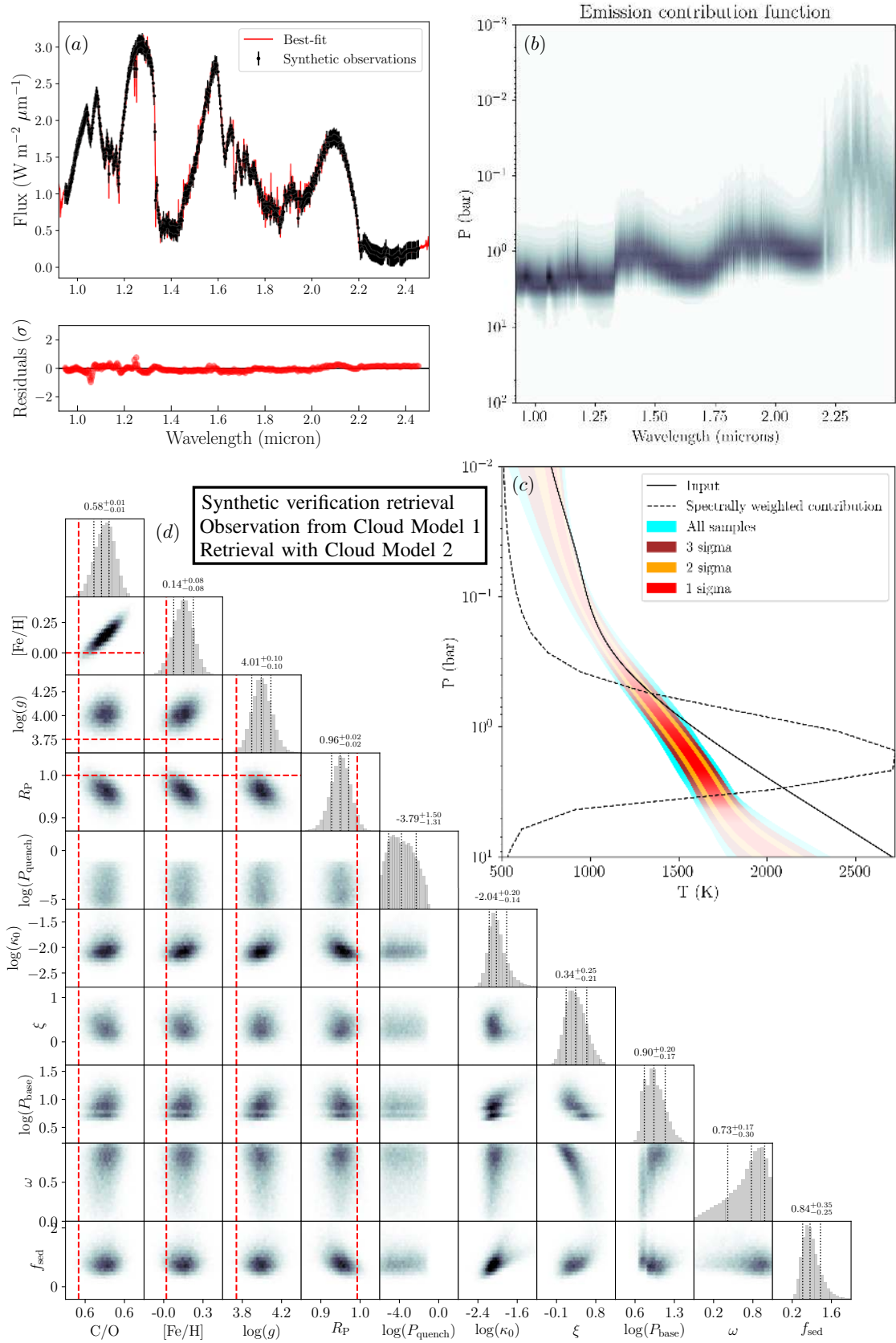
Fig. D.1. Same as Fig. 1, but showing the results when a synthetic spectrum made with Cloud Model 1 is retrieved using Cloud Model 2. This test is described further in Sect. 3.2.
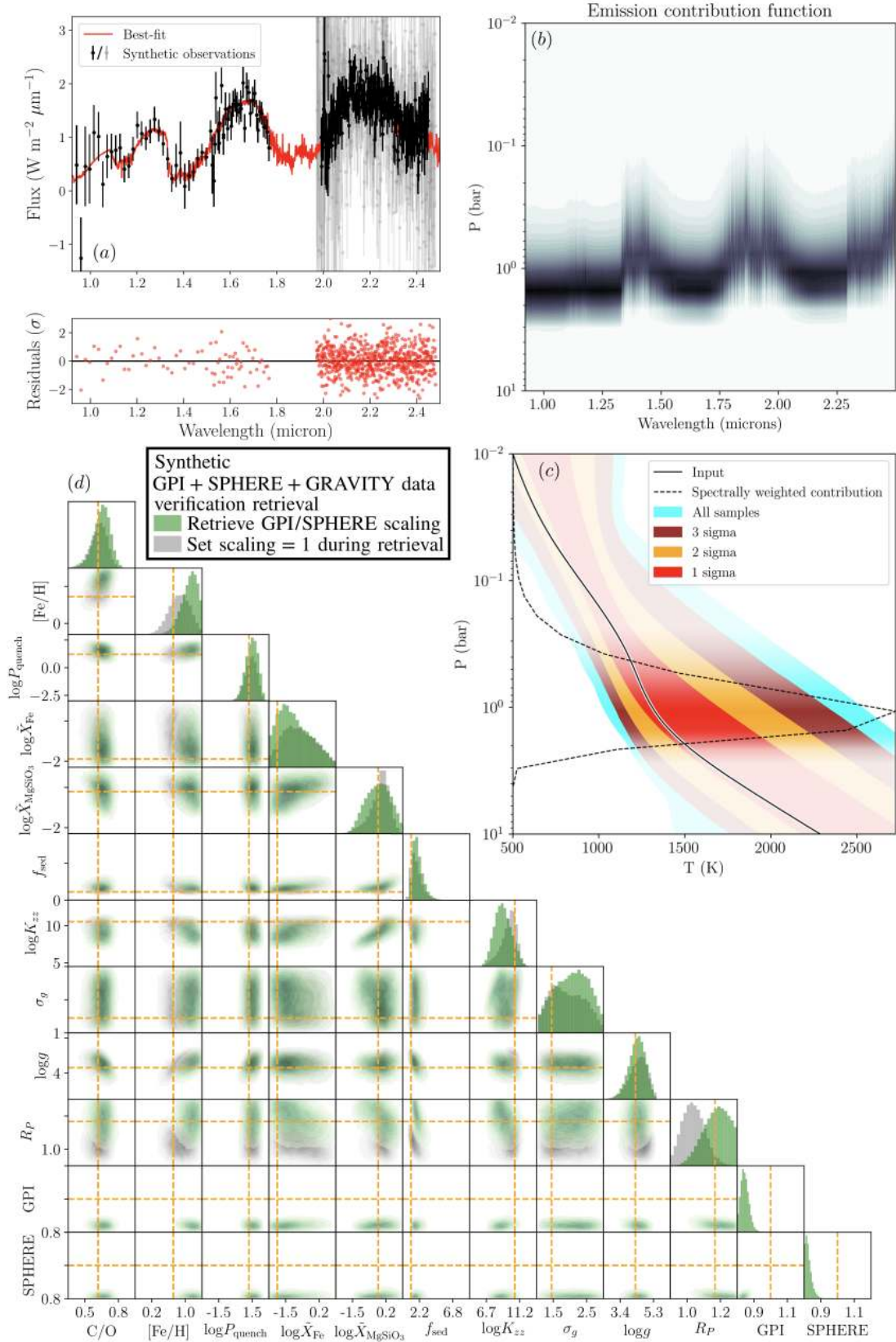
Fig. E.1. Same as Fig. 1, but showing the results when a synthetic spectrum of the same wavelength spacing and noise properties as the actual HR 8799e data is retrieved. Green posteriors indicate the retrieval where the scaling parameters of the GPI and SPHERE spectra were retrieved, while gray posteriors indicate the results that neglected the scaling parameters. This test is described further in Sect. 4.2.2. In *panel a*, the black points in the *K*-band correspond to synthetic observations with the same data quality as the GRAVITY observations reported in Gravity Collaboration (2019), while the gray points show the data at the same quality as found for the two new GRAVITY observations presented in this work. All three synthetic GRAVITY data sets were fitted simultaneously.
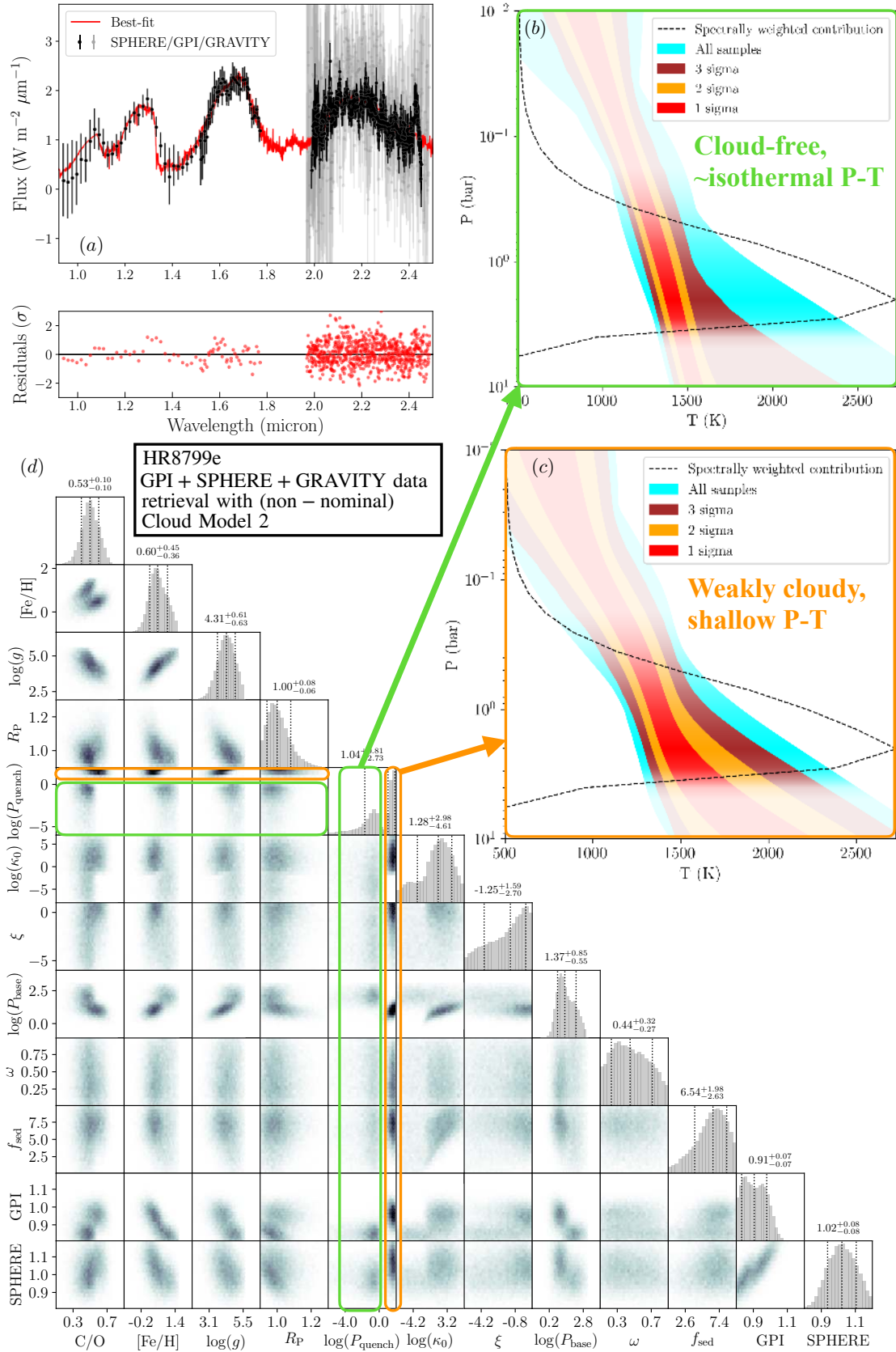
Fig. F.1. Results of the non-nominal retrieval of HR 8799e, using Cloud model 2. *Panel a*: observations, best-fit spectrum and residuals. *Panel b*: retrieved pressure-temperature confidence envelopes of the cloud-free isothermal solution of the retrieval. *Panel c*: retrieved pressure-temperature confidence envelopes of the cloudy solution of the retrieval. The black dashed line shows the flux average of the emission contribution function of the global best-fit (clear and cloudy) from the posterior. *Panel d*: 2d posterior plot of the (non-nuisance) retrieved atmospheric parameters. In *panel a*, the black points in the *K*-band correspond to the GRAVITY observations reported in Gravity Collaboration (2019), while the gray points show the data of the two new GRAVITY observations presented in this work. All three GRAVITY data sets were fitted simultaneously.