The Astronomical Journal, 154:91 (24pp), 2017 September https://doi.org/10.3847/1538-3881/aa7ed8
© 2017. The American Astronomical Society. All rights reserved.

HELIOS-RETRIEVAL: An Open-source, Nested Sampling Atmospheric Retrieval Code; Application to the HR 8799 Exoplanets and Inferred Constraints for Planet Formation

Baptiste Lavie[1,2,3], João M. Mendonça[1], Christoph Mordasini[2], Matej Malik[1], Mickaël Bonnefoy[4,5], Brice-Olivier Demory[1], Maria Oreshenko[2], Simon L. Grimm[1], David Ehrenreich[3], and Kevin Heng[2]
[1] University of Bern, Space Research and Planetary Sciences, Sidlerstrasse 5, CH-3012, Bern, Switzerland; baptiste.lavie@space.unibe.ch
[2] University of Bern, Center for Space and Habitability, Sidlerstrasse 5, CH-3012, Bern, Switzerland; kevin.heng@csh.unibe.ch
[3] Observatoire de l'Université de Genève, 51 chemin des Maillettes, 1290, Sauverny, Switzerland
[4] Université Grenoble Alpes, IPAG, F-38000, Grenoble, France
[5] CNRS, IPAG, F-38000, Grenoble, France
Received 2016 October 9; revised 2017 June 26; accepted 2017 July 4; published 2017 August 11

Abstract

We present an open-source retrieval code named HELIOS-RETRIEVAL, designed to obtain chemical abundances and temperature–pressure profiles by inverting the measured spectra of exoplanetary atmospheres. In our forward model, we use an exact solution of the radiative transfer equation, in the pure absorption limit, which allows us to analytically integrate over all of the outgoing rays. Two chemistry models are considered: unconstrained chemistry and equilibrium chemistry (enforced via analytical formulae). The nested sampling algorithm allows us to formally implement Occam's Razor based on a comparison of the Bayesian evidence between models. We perform a retrieval analysis on the measured spectra of the four HR 8799 directly imaged exoplanets. Chemical equilibrium is disfavored for HR 8799b and c. We find supersolar C/H and O/H values for the outer HR 8799b and c exoplanets, while the inner HR 8799d and e exoplanets have a range of C/H and O/H values. The C/O values range from being superstellar for HR 8799b to being consistent with stellar for HR 8799c and being substellar for HR 8799d and e. If these retrieved properties are representative of the bulk compositions of the exoplanets, then they are inconsistent with formation via gravitational instability (without late-time accretion) and consistent with a core accretion scenario in which late-time accretion of ices occurred differently for the inner and outer exoplanets. For HR 8799e, we find that spectroscopy in the *K* band is crucial for constraining C/O and C/H. HELIOS-RETRIEVAL is publicly available as part of the Exoclimes Simulation Platform (http://www.exoclime.org).

Key words: methods: statistical – planets and satellites: atmospheres – planets and satellites: composition – planets and satellites: formation – planets and satellites: gaseous planets – planets and satellites: individual (HR 8799)

1. Introduction

1.1. Motivation

Traditionally, the masses and radii of brown dwarfs and substellar objects have been inferred by applying evolutionary tracks to measurements of their luminosities and ages (e.g., Burrows et al. 1997; Chabrier et al. 2000; Baraffe et al. 2002). On rare occasions, brown dwarfs and low-mass stars may transit their binary companions and allow for their other properties to be studied (see Burrows et al. 2011 and references therein). A particularly important study was conducted by Konopacky et al. (2010), who were able to obtain dynamical masses for 15 brown dwarfs residing in binaries. By comparing the dynamical and photometric masses, Konopacky et al. (2010) showed that both the Burrows et al. (1997) and Chabrier et al. (2000) models underpredicted the masses of M and L dwarfs and overpredicted the mass of the lone T dwarf in their sample by ∼10% (tens of percent). By studying a sample of 46 L dwarfs, Hiranaka et al. (2016) suggested that a dust haze of sub-micron-sized particles exist in their upper atmospheres, which are neglected by the standard evolutionary tracks.

Taken together, these results suggest that the traditional approach of using self-consistent evolutionary tracks may be incomplete and motivates alternative and complementary ways of interpreting the spectra of brown dwarfs and substellar objects. We expect this train of thought to apply to the recently discovered directly imaged exoplanets as well, since the interpretation of their photometry and spectroscopy is typically performed using the evolutionary tracks computed for brown dwarfs (e.g., Bonnefoy et al. 2016).

1.2. Theoretical Improvements

Self-consistent forward modeling starts with a set of assumptions and computes forward to predict the temperature–pressure profile and synthetic spectrum of an object. Atmospheric retrieval is a complementary approach borrowed from the Earth's remote sensing community, where one applies an inversion method to obtain the temperature–pressure profile and chemical abundances from finding the best-fit solution to the measured spectrum (e.g., Madhusudhan & Seager 2009; Benneke & Seager 2013; Lee et al. 2013; Line et al. 2013, 2016; Barstow et al. 2015; Waldmann et al. 2015). It sacrifices self-consistency and sophistication for simplicity, which allows for a more thorough exploration of parameter space. Atmospheric retrieval is particularly well-suited for addressing questions regarding planet formation, since it allows for the posterior distributions of the carbon-to-oxygen ratio (C/O) and the elemental abundances of carbon (C/H) and oxygen (O/H) to be computed.

The HR 8799 system hosts four exoplanets (Marois et al. 2008, 2010), whose formation mechanisms remain an enigma (Kratter et al. 2010). Spectra with resolutions of about 30 to 4000 have been obtained by, e.g., Barman et al. (2011, 2015), Konopacky et al. (2013), Oppenheimer et al. (2013), Ingraham et al. (2014), and Zurlo et al. (2016). Since these spectra have resolutions that are considerably higher than those obtained for hot Jupiters using WFC3 on the *Hubble Space Telescope* (e.g.,

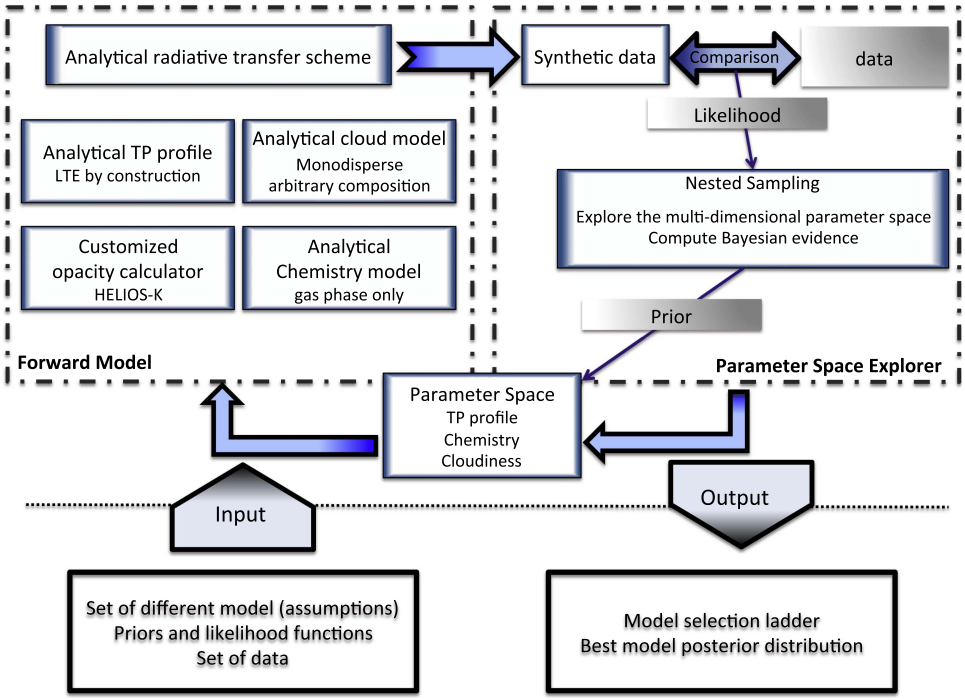

Figure 1. Flow chart for HELIOS-R and a description of its main components. Note that enforcing equilibrium chemistry is optional and our approach for distinguishing between equilibrium and non-equilibrium chemistry is completely data-driven.

Deming et al. 2013; Mandell et al. 2013; Kreidberg et al. 2014; Stevenson et al. 2014), they present an opportunity for performing remote sensing of exoplanetary atmospheres that is similar to what planetary scientists had to work with a few decades ago, before the advent of probes. A key difference is that the radii and masses of these directly imaged exoplanets are unknown,[6] unlike for transiting exoplanets. A recent review of directly imaged exoplanets, which includes the HR 8799 system, may be found in Bowler (2016).

The first atmospheric retrieval analysis of directly imaged exoplanets was performed by Lee et al. (2013), who studied only the HR 8799b exoplanet. In the current study, we collect all of the published spectra of the HR 8799b, c, d, and e exoplanets and subject them to the same retrieval method with the intention of using the retrieved chemistry to constrain planet formation scenarios.

Besides the novelty of our analysis, the current study is also a method paper for our new atmospheric retrieval code named HELIOS-R, which we constructed from scratch to study exoplanetary atmospheres. HELIOS-R is part of the HELIOS radiation package of the Exoclimes Simulation Platform[7] and has the following features (Figure 1).

1. We have implemented a nested sampling algorithm to explore the multi-dimensional parameter space (Skilling 2006; Feroz et al. 2009; Benneke & Seager 2013; Waldmann et al. 2015; Line et al. 2016). Unlike other approaches (e.g., Markov Chain Monte Carlo, nonlinear optimal estimation), nested sampling allows for the Bayesian evidence to be directly calculated, which in turn allows for models with different parametrizations (and number of parameters) to be compared on an equal footing. Models with extra complexity are penalized, which allows for Occam's Razor[8] to be formally enforced. For example, our retrieval analysis allows us to formally determine if chemical equilibrium is favored or disfavored in an atmosphere in a completely data-driven manner. As another example, it allows us to determine the number and types of molecules to be included in the retrieval.
2. Our temperature–pressure profile is taken from Heng et al. (2014), who generalized the work of Guillot (2010) and Heng et al. (2012) to include non-isotropic scattering and non-constant opacities. When stellar irradiation and scattering are omitted, the temperature–pressure profile reduces to the classical solution of Milne for self-luminous objects (Mihalas 1970). By construction, it conserves energy in an analytical and exact sense.
3. Our atmospheric cross-sections are computed using our customized opacity calculator named HELIOS-K, which was previously published by Grimm & Heng (2015).
4. To combine the cross-sections of different molecules, one needs to have a chemistry model that calculates their relative abundances. We use the analytical solutions of Heng et al. (2016), Heng & Lyons (2016), and Heng & Tsai (2016), which have been shown to be accurate at the ~1% level (or better) when benchmarked against numerical solutions using Gibbs free energy minimization. These analytical solutions allow for fast computation if one wishes to enforce chemical equilibrium.
5. Our radiative transfer scheme, which translates cross-sections and temperatures into fluxes (and hence allows us to compute the synthetic spectrum), uses the exact

[6] Meaning that they are typically not directly measured, but rather inferred using evolutionary models, which means the radii and masses are model-dependent.

[7] http://www.exoclime.org

[8] Whether Occam's Razor always yields the correct answer is another matter. In the current study, we are guided by Occam's Razor in the limit of sparse data.

Table 1
Shorthand Notation for the Suite of Models Tested in This Study

Notation	Meaning
U	Unconstrained chemistry
E	Equilibrium chemistry
B	Cloud-free ("blue sky")
C	Cloudy
1	H_2O is included in retrieval
2	CO_2 is included in retrieval
5	CO is included in retrieval
6	CH_4 is included in retrieval
r	Planet radius R is included in retrieval
g	Planet surface gravity g is included in retrieval
d	Distance of the system d is included in retrieval

Note. "1," "2," "5," and "6" refer to the HITRAN/HITEMP labels for these molecules. When no number is specified, it means that all four molecules are included in the retrieval. For example, UBrg16 is a cloud-free model with unconstrained chemistry, where the mixing ratios of water and methane, as well as the planetary radius and surface gravity, are included as fitting parameters. By contrast, the UB model includes all four molecules in the retrieval, but fixes the planetary radius and surface gravity to user-specified values.

Table 2
Priors Used in This Study

Symbol	Prior Used	Value
R	Gaussian	$R = 1.2 \pm 0.1\ R_J$
g	Gaussian	$\log g = 4.1 \pm 0.3$ (cgs)
X_i	Log-uniform	10^{-20} to 10^{-1}
κ_0	Log-uniform	$\log \kappa_0 = 10^{-15}$–10 (mks)
$T_{\rm int}$	Uniform	10–1500 K
Q_0	Uniform	1–100
r_c	Log-uniform	10^{-7}–10^{-3} m
$f_{\rm cloud}$	Log-uniform	10^{-30}–10^{-4}
d	Gaussian	39.4 ± 1.0 pc

Note. cgs: centimeters, grams, and seconds. mks: meters, kilograms, and seconds.

analytical solution in the limit of isothermal model layers and pure absorption (Heng et al. 2014). It allows us to analytically integrate over all of the incoming and outgoing angles associated with every ray.

6. Our cloud model is based on the basic principles of Mie theory (e.g., Pierrehumbert 2010). It assumes a monodisperse set of particles, which may be interpreted as the dominant size in a size distribution of particles (e.g., Burrows et al. 2011). It includes a dimensionless parameter that is a proxy for the cloud composition. When the particles are small compared to the wavelength, it reproduces Rayleigh scattering. By contrast, models that implement a constant cloud-top pressure implicitly assume the cloud particles to be large (compared to the wavelength observed) and preclude Rayleigh scattering by construction.

While each component of HELIOS-R may not be novel by itself, the assembly of all of these components into a single code and retrieval tool is a novel endeavor. Furthermore, we have designed HELIOS-R to run on Graphics Processing Units (GPUs), which affords speed-ups of at least a factor of several compared to the CPU version. With a UCrg model (see Table 1) retrieval performed on the HR 8799b data set, the GPU version is five times faster than the CPU version on a macbook Pro laptop equipped with an NVIDIA GeForce GT 750M GPU card and an Intel Core i7 2.5 GHz CPU. For this analysis, we used our GPU cluster of NVIDIA K20 cards; it takes 10^{-2} s to evaluate one likelihood of this UCrg model.

In Section 2, we provide a detailed description of each component or ingredient of HELIOS-R. In Section 3, we subject HELIOS-R to several tests before applying it to the measured spectra of the HR 8799b, c, d, and e directly imaged exoplanets. In Section 4, we present our retrieval results of the HR 8799 system. In Section 5, we compare our study to previous work and describe opportunities for future work. Table 1 shows the suite of models tested in the current study. Table 2 states the priors used for our fitting parameters. Table 3 summarizes our retrieval results. Appendix A states our fast analytical formulae for evaluating the exponential integral of the first order. Appendix B includes, for completeness, the full posterior distributions of the best models for the atmospheres of HR 8799b, c, d, and e.

2. Methodology

The executive summary is that each model of the retrieval contains up to 11 parameters: the radius, the surface gravity, 2 for the temperature–pressure profile, 2 or 4 for the chemistry (depending on whether one adopts equilibrium or unconstrained chemistry), and 3 for the cloud model. The mean molecular weight is not a parameter and is constructed from the mixing ratios. Each HR 8799 exoplanet typically has between 40 and 120 data points for its measured spectrum: 68 for b, 105 for c, 115 for d, and 48 for e.

To construct an atmospheric retrieval model, we need a forward model. By "forward model," we refer to the temperature–pressure profile, atmospheric opacities, chemistry model, radiative transfer scheme, and cloud model. We also need a method to scan the vast multi-dimensional parameter space of our forward model to locate the highest likelihood region, i.e., the best solution that fits the data (e.g., for a review, see Press et al. 2007).

2.1. Nested Sampling

We use a nested sampling algorithm (Skilling 2006) to scan the diverse, multi-dimensional parameter space describing our one-dimensional model atmospheres. Benneke & Seager (2013) previously gave a detailed overview of the nested sampling method. Waldmann et al. (2015) and Line et al. (2016) also used nested sampling. Here, we provide a concise description of our implementation.

Consider a model with a set of parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, ..., \theta_{N_\theta}\}$, where N_θ is the number of parameters. Consider a set of models labeled by the index i: $\mathcal{M}_i$. The probability density function (PDF) on the parameters for a given model is $\mathcal{P}(\boldsymbol{\theta}|\mathcal{M}_i)$, which is also known as the "prior."

Discussions of any Bayesian method necessarily start with Bayes's rule, which states that the PDF of a model given the data (denoted by $\boldsymbol{D}$) is (e.g., Skilling 2006)

$$\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{D}, \mathcal{M}_i) = \frac{\mathcal{P}(\boldsymbol{\theta}|\mathcal{M}_i)\ \mathcal{L}\ (\boldsymbol{D}|\boldsymbol{\theta}, \mathcal{M}_i)}{\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_i)}. \quad (1)$$

Table 3
Summary of Retrieved Results

Property	Exoplanet	Value
X_{H_2O}	HR 8799b	$-2.89^{+0.09}_{-0.09}$
X_{H_2O}	HR 8799c	$-2.60^{+0.12}_{-0.05}$
X_{H_2O}	HR 8799d	[−2.29]
X_{H_2O}	HR 8799e	[−1.84]
X_{CO_2}	HR 8799b	$-6.70^{+1.52}_{-6.33}$
X_{CO_2}	HR 8799c	$-4.63^{+0.13}_{-0.11}$
X_{CO_2}	HR 8799d	[−18.84]
X_{CO_2}	HR 8799e	[−19.13]
X_{CO}	HR 8799b	$-1.86^{+0.10}_{-0.09}$
X_{CO}	HR 8799c	$-2.48^{+0.14}_{-0.20}$
X_{CO}	HR 8799d	[−16.32]
X_{CO}	HR 8799e	[−17.36]
X_{CH_4}	HR 8799b	$-5.03^{+0.14}_{-0.16}$
X_{CH_4}	HR 8799c	$-5.03^{+0.17}_{-0.18}$
X_{CH_4}	HR 8799d	[−28.11]
X_{CH_4}	HR 8799e	[−27.59]
μ	HR 8799b	$2.18^{+0.00}_{-0.00}$
μ	HR 8799c	$2.19^{+0.00}_{-0.00}$
μ	HR 8799d	[2.28]
μ	HR 8799e	[2.42]
C/O	HR 8799b	$0.92^{+0.01}_{-0.01}$
C/O	HR 8799c	$0.55^{+0.10}_{-0.12}$
C/O	HR 8799d	$0.00^{+0.00}_{-0.00}$
C/O	HR 8799e	$0.00^{+0.00}_{-0.00}$
C/H	HR 8799b	$-2.11^{+0.10}_{-0.09}$
C/H	HR 8799c	$-2.73^{+0.14}_{-0.20}$
C/H	HR 8799d	$-16.62^{+4.08}_{-2.04}$
C/H	HR 8799e	$-11.93^{+4.62}_{-4.64}$
O/H	HR 8799b	$-2.07^{+0.09}_{-0.09}$
O/H	HR 8799c	$-2.47^{+0.09}_{-0.11}$
O/H	HR 8799d	$-3.20^{+0.19}_{-0.15}$
O/H	HR 8799e	$-2.75^{+0.57}_{-0.57}$
Q_0	HR 8799b	$1.21^{+0.49}_{-0.73}$
Q_0	HR 8799c	$0.79^{+0.25}_{-0.26}$
Q_0	HR 8799d	$1.39^{+0.27}_{-0.34}$
Q_0	HR 8799e	$0.95^{+0.53}_{-0.52}$
r_c	HR 8799b	$-4.37^{+0.49}_{-0.46}$
r_c	HR 8799c	$-4.44^{+0.25}_{-0.20}$
r_c	HR 8799d	$-6.68^{+0.18}_{-0.18}$
r_c	HR 8799e	$-4.69^{+0.79}_{-0.76}$
X_c	HR 8799b	$-21.22^{+0.80}_{-0.85}$
X_c	HR 8799c	$-20.55^{+0.41}_{-0.49}$
X_c	HR 8799d	$-15.96^{+1.20}_{-0.95}$
X_c	HR 8799e	$-20.56^{+1.55}_{-1.37}$
d	HR 8799b	$40.30^{+0.66}_{-0.79}$
d	HR 8799c	$39.73^{+0.22}_{-0.20}$
d	HR 8799d	$40.81^{+0.42}_{-0.50}$
d	HR 8799e	$39.40^{+0.70}_{-0.78}$

Note. We have listed the 1σ uncertainties, which were computed by locating the 15.87th and 84.13th percentile points on the horizontal axis. In the limit of a symmetric Gaussian function, these would yield the full-width at half-maximum of the Gaussian. For planets d and e, the molecule abundances and the mean molecular weight are given at 1 bar. Values are in $\log_{10}$ (except for C/O) and dimensionless (except for r_c, which is in meters).

The quantity $\mathcal{L}(\boldsymbol{D}|\boldsymbol{\theta}, \mathcal{M}_i)$ is the "likelihood." We assume $\mathcal{L}(\boldsymbol{D}|\boldsymbol{\theta}, \mathcal{M}_i)$ to be the same Gaussian function as Equation (5) of Benneke & Seager (2013).

We will term $\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{D}, \mathcal{M}_i)$ the "posterior." Since it normalizes to unity, the Bayesian evidence is given by the multi-dimensional integral,

$$\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_i) = \int \mathcal{P}(\boldsymbol{\theta}|\mathcal{M}_i)\,\mathcal{L}(\boldsymbol{D}|\boldsymbol{\theta}, \mathcal{M}_i)\,d\boldsymbol{\theta}. \quad (2)$$

Fitting a model to a measured spectrum is an exercise in which a better fit is obtained when more free parameters (e.g., more molecules) are introduced. Model selection is essentially the enforcement of Occam's Razor, meaning that we select the model that has a level of sophistication or complexity that is commensurate with the quality of data available. It prevents the over-fitting of data by a model that is too complex. For example, Hansen et al. (2014) find that, for some of the exoplanets, the photometric data of *Spitzer* alone may be fitted with a Planck function and a more complex model is unnecessary. As the data quality improves, so does the complexity of the best model.

The essence of the nested sampling is to reduce the computation of the Bayesian evidence to a one-dimensional integral (Skilling 2006),

$$\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_i) = \int_0^1 \mathcal{L}'(\mathcal{X})\,d\mathcal{X}, \quad (3)$$

where the likelihood now only depends on a single variable and is denoted by $\mathcal{L}'$. This variable $\mathcal{X}$ is termed the "prior mass" and is bounded between 0 and 1. A visualization of the prior mass and its relationship to the Bayesian evidence is given in Figure 3 of Skilling (2006) and Figure 1 of Benneke & Seager (2013). Numerically, we use the trapezoid rule to compute the Bayesian evidence as a finite sum,

$$\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_i) = \sum_j \frac{\mathcal{X}_{j+1} - \mathcal{X}_j}{2}(\mathcal{L}'_{j+1} + \mathcal{L}'_j). \quad (4)$$

We begin by randomly drawing $N_{\rm live}$ points from the parameter space(θ) subjected to the constraint of the chosen prior. We use either Gaussian (radius, logarithm of gravity, distance), log-uniform (mixing ratios, mean opacity, cloud particle radius, cloud mixing ratio), or uniform (temperature, cloud composition parameter) priors. For a set of points drawn, we compute their likelihood values. At each step of the algorithm, we discard the worst point and replace it with a newly drawn point until the convergence criteria is met (see Skilling 2006). This newly drawn point needs to have a higher likelihood than the worst point that was just discarded. Specifically, we use the open-source software named `PyMultiNest`[9] (Buchner et al. 2014), which is a `Python` wrapper for the open-source `MultiNest`[10] program written in `Fortran 90` (Feroz & Hobson 2008; Feroz et al. 2009, 2013). For each model, we run the nested sampling algorithm using 40,000 living points parallelized into 20 runs of 2000 "living points" each. For comparison, Waldmann et al. (2015) uses $N_{\rm live} = 4000$ living points. Benneke & Seager (2013) use between $N_{\rm live} = 50$ and 10,000 living points. Line et al. (2016) do not specify the number of living points used. Equation (4) is used to compute the Bayesian evidence. As a byproduct of this

[9] https://github.com/JohannesBuchner/PyMultiNest/
[10] https://ccpforge.cse.rl.ac.uk/gf/project/multinest/

procedure, one also obtains posterior-distribution samples of the model parameters.

For the purpose of comparing two models, which we denote by $\mathcal{M}_i$ and $\mathcal{M}_{i+1}$, it is useful to define a quantity known as the Bayes factor, which is the ratio of the Bayesian evidences (Trotta 2008),

$$\mathcal{B} = \frac{\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_i)}{\mathcal{Z}(\boldsymbol{D}|\mathcal{M}_{i+1})}. \quad (5)$$

The Bayes factor is equal to the posterior odds when both models are considered equally likely. As shown in Table 2 of Trotta (2008), which is reproduced in Table 2 of Benneke & Seager (2013), there is a relationship between the Bayes factor, the p-value of the frequentists and the significance in terms of the number of standard deviations. We use the Jeffreys scale (Kass & Raftery 1995) to evaluate model significances. Weak, moderate, and strong evidence for favoring the ith model over the $(i+1)$th model correspond to $\ln \mathcal{B} = 1$, 2.5, and 5, respectively.

2.2. Temperature–Pressure Profile

For the temperature–pressure profile, we assume a one-dimensional, plane–parallel model atmosphere. Its layers are evenly spaced in the logarithm of pressure between 1 μbar and 1 kbar. We implement equation (126) of Heng et al. (2014), who previously generalized the work of Guillot (2010; pure absorption limit and constant opacities) and Heng et al. (2012; isotropic scattering, constant shortwave/optical opacity) to include non-isotropic scattering and a non-constant shortwave/optical opacity. Since the HR 8799 exoplanets are non-irradiated, we essentially use a reduced version of Equation (126) of Heng et al. (2014),

$$T^4 = \frac{T_{\rm int}^4}{4}\left[\frac{8}{3} + \frac{3\tilde{m}}{\beta_{\rm L}^2}\left(\kappa_0 + \frac{\kappa_{\rm CIA}\tilde{m}}{2\tilde{m}_0}\right)\right], \quad (6)$$

where $T_{\rm int}$ is the internal/interior temperature, $\beta_{\rm L}$ is the longwave/infrared scattering parameter, κ_0 is the constant component of the longwave/infrared opacity, and $\kappa_{\rm CIA}$ is the opacity associated with collision-induced absorption (CIA). The column mass is denoted by $\tilde{m}$, while $\tilde{m}_0$ is the column mass referenced to the bottom of the model atmosphere. We set $P_0 = \tilde{m}_0 g = 1$ kbar, where g is the surface gravity.

Equation (6) is essentially a generalization of the classical Milne's solution (Mihalas 1970) to include scattering and CIA. In the limit of pure absorption ($\beta_{\rm L} = 1$) and in the absence of CIA, we obtain $T = T_{\rm int}$ when $\kappa_0\tilde{m} = 4/9$, which is somewhat different from the classical Milne value of 2/3. It is worth emphasizing that Equation (6) is, by construction, a temperature–pressure profile in radiative equilibrium, which implies that both local and global energy conservation are guaranteed in an exact, analytical sense (Heng et al. 2014; Heng & Lyons 2016). By contrast, the versatile fitting function used by Madhusudhan & Seager (2009) does not, by construction, obey energy conservation and this has to be enforced as a separate numerical condition. However, by using a mean opacity, Equation (6) sacrifices accuracy for simplicity, which makes the temperature–pressure profile more isothermal at high altitudes than if a more realistic radiative transfer calculation was performed.

In principle, κ_0 and $\kappa_{\rm CIA}$ are mean opacities that may be calculated directly from the spectroscopic line lists. However, while deriving these analytical temperature–pressure profiles Guillot (2010) and Heng et al. (2012, 2014) have assumed that the absorption, flux, Planck, and Rosseland mean opacities are equal, which makes it unclear how to exactly compute κ_0 and $\kappa_{\rm CIA}$. Therefore, we opt to use κ_0 and $\kappa_{\rm CIA}$ as fitting parameters instead. In other words, our temperature–pressure profile is not self-consistent with the atmospheric opacities used.

We find that using $\kappa_{\rm CIA}$ and $\beta_{\rm L}$ as fitting parameters has a negligible effect on our results (not shown). In practice, the use of Equation (6) with only $T_{\rm int}$ and κ_0 as fitting parameters (i.e., setting $\beta_{\rm L} = 1$ and $\kappa_{\rm CIA} = 0$) is sufficient for our retrieval calculations.

We use a constant value of the surface gravity, as we are sensing 6 orders of magnitude in pressure at most, which corresponds to 13.8 scale heights. This means that the region of the atmosphere being sensed is only several percent of the radius of the exoplanet. A constant surface gravity is thus not unreasonable.

2.3. Atmospheric Cross-sections

We first distinguish between our use of the terms "cross-section" and "opacity." The former has units of area. The latter is the cross-section per unit mass. We previously designed and wrote an open-source opacity calculator (Grimm & Heng 2015), based on implementing Algorithm 916 (Zaghloul & Ali 2012) to perform fast computations of the Voigt profile by recasting it as a Faddeeva function. Typically, `HELIOS-K` is able to compute an opacity or cross-section function with $\sim 10^5$ spectral lines in $\sim$1 s on an NVIDIA K20 GPU. In principle, it is agnostic about the spectroscopic line list being used and is able to take any line list as an input. The details of how to take the inputs of a line list and use them to compute the integrated line strengths and line shapes have previously been summarized in Grimm & Heng (2015), and we will not repeat them here.

We restrict ourselves to only four molecules: carbon monoxide (CO), carbon dioxide (CO_2), water (H_2O), and methane (CH_4). For CO and CO_2, we use the HITEMP database (Rothman et al. 1996, 2010, 2013). For H_2O and CH_4, we use the ExoMol line list (Barber et al. 2006; Yurchenko & Tennyson 2014). Acetylene, ammonia, ethylene, and hydrogen cyanide have been omitted because they are subdominant at the photospheric temperatures of the HR 8799 exoplanets (Madhusudhan 2012; Heng & Tsai 2016; Moses et al. 2016). In particular, see Figure 10 of Moses et al. (2016).

In the current study, we choose to deal with cross-sections instead of opacities. For our `HELIOS` self-consistent radiative transfer code, we chose to use opacities instead (Malik et al. 2017). There are various strategies to construct the cross-section function of the atmosphere. By "cross-section function," we refer to the function that depends on temperature, pressure, wavenumber, and type of molecule. The cross-section function is a theoretical construction: it may be defined continuously or be sampled at an arbitrary number of discrete points. We consider the way in which the cross-section function is sampled as an issue of implementation, which we will now discuss. Regardless of the approach used to construct and sample the cross-section function, the end goal is the same: to use them to construct transmission functions and ultimately integrate fluxes over a waveband.

The first approach is to use the "k-distribution method," which resamples the highly erratic cross-section function into a

monotonically increasing cumulative distribution function (Lacis & Oinas 1991; Fu & Liou 1992; Grimm & Heng 2015). Since the k-distribution method is only exact for a homogeneous atmosphere with one molecule (Grimm & Heng 2015), one has to apply the "correlated-k approximation" as well, which assumes that the spectral lines are perfectly correlated (see Chapter 4.4.5 of Pierrehumbert 2010).

The second approach is to use "opacity sampling," which is to discretely sample the opacity function, typically at a smaller number of points than there are lines. In our context, it is perhaps more accurate to use the term "cross-section sampling."

The "line-by-line" limit occurs when the integrated fluxes over a waveband is exact (to machine precision). It is essentially the second approach, but where the cross-section function is sampled at more wavenumber points than there are lines. Since there are $\sim 10^9$ (or more) lines for the water molecule alone, this is a formidable computational challenge and is currently infeasible for any retrieval code dealing with hot exoplanetary atmospheres. We note that a cross-section function that includes all of the lines of a given line list does not qualify it as being "line-by-line," if the sampling is not fine enough to resolve each line profile.

In the current study, we adopt the second approach, which is also used by Madhusudhan & Seager (2009), Benneke & Seager (2013), Line et al. (2013), and Waldmann et al. (2015). Our spectral resolution used is 1 cm^{-1}, evenly sampled across the wavenumber. We note that Line et al. (2013, 2015) and Waldmann et al. (2015) also used a spectral resolution of 1 cm^{-1}. Some authors do not specify the spectral resolution of their atmospheric cross-section function (e.g., Madhusudhan & Seager 2009; Benneke & Seager 2013; Barstow et al. 2015; Line et al. 2016). We precompute our cross-sections on a grid across wavenumber, pressure, and temperature: 100–2900 K (in increments of 200 K) and 1 μbar to 1 kbar (with two points per dex in pressure) for CO, CO_2, CH_4, and H_2O. The grid is then interpolated to obtain values of the cross-sections for any temperature and pressure within the stated ranges.

A lingering issue, which stems from an unsolved physics problem, is that the far line wings of Voigt profiles do not accurately represent the wings of real lines. Various groups have adopted different ad hoc approaches to truncating the Voigt profiles (see Grimm & Heng 2015 and references therein). Hedges & Madhusudhan (2016) discuss this issue, but do not provide any solution for it. In the current study, we adopt a 100 cm^{-1} cutoff.

2.4. Chemistry

Once the cross-sections have been computed, they may be used to compute the optical depth of each model layer for all of the molecules,

$$\Delta\tau = \sum_i \frac{X_i \sigma_i}{\bar{m} g} \Delta P, \tag{7}$$

where X_i and σ_i are the mixing ratio and cross-section of the ith molecule, respectively. ΔP is the thickness of the layer in terms of the difference in pressure. The mean molecular mass is given by $\bar{m} = \mu m_\mathrm{u}$, where μ is the mean molecular weight and m_u is the atomic mass unit. The preceding expression assumes hydrostatic equilibrium, isothermal layers, and that the surface gravity is constant throughout our model atmosphere.

Generally, the mixing ratios of molecular hydrogen and helium only show up via CIA as a contribution to the continuum of a spectrum, which implies that they cannot be as definitively determined as that of the molecules. Our CIA opacities are obtained from Richard et al. (2012). In the range of $X_{\mathrm{H_2}} \approx 0.8$–$0.9$, the effect on the continuum of the synthetic spectrum is very similar (not shown). The effects of H_2-He CIA are even more subtle. As such, we adjust $X_{\mathrm{H_2}}$ to render the sum of the mixing ratio's unity,

$$1.1 X_{\mathrm{H_2}} + \sum_i X_i = 1, \tag{8}$$

where we have assumed that $X_{\mathrm{He}} = 0.1 X_{\mathrm{H_2}}$ to reflect cosmic abundance. By denoting the mass of the ith molecule by m_i, the corresponding mean molecular weight is calculated using

$$\mu = 2X_{\mathrm{H_2}} + 4X_{\mathrm{He}} + \sum_i \frac{m_i X_i}{m_\mathrm{u}}. \tag{9}$$

For example, if we have $X_{\mathrm{H_2}} = 0.85$, $X_{\mathrm{He}} = 0.085$, and $X_{\mathrm{CO}} = 0.065$, then we have $\mu = 3.86$. In models with equilibrium chemistry, the mean molecular weight changes slightly for each layer, because the molecular abundances vary from layer to layer even for the same metallicity.

In the current study, we consider two chemistry models. "Unconstrained chemistry" refers to using each X_i as a fitting parameter in the retrieval. "Equilibrium chemistry" means that the X_i may be determined using only the elemental abundances of carbon (f_C) and oxygen (f_O), if C–H–O gaseous chemistry is considered. In this case, the four-parameter system of unconstrained chemistry reduces to two parameters. To compute the four X_i values given f_C and f_O, we use the validated analytical formulae of Heng et al. (2016), Heng & Lyons (2016), and Heng & Tsai (2016). Specifically, we implement Equations (12), (20), and (21) of Heng & Lyons (2016) for gaseous C–H–O chemistry. The benchmarking of these formulae against calculations of Gibbs free energy minimization was previously performed by Heng & Tsai (2016), who showed that they are accurate at the $\sim$1% level or better. Further validation of these formulae comes from matching the trends found by Madhusudhan (2012) and Moses et al. (2013).

For unconstrained chemistry, the carbon-to-oxygen ratio is computed using

$$\mathrm{C/O} = \frac{X_{\mathrm{CO}} + X_{\mathrm{CO_2}} + X_{\mathrm{CH_4}}}{X_{\mathrm{CO}} + 2X_{\mathrm{CO_2}} + X_{\mathrm{H_2O}}}. \tag{10}$$

The elemental abundances are inferred using

$$\mathrm{C/H} = \frac{X_{\mathrm{CO}} + X_{\mathrm{CO_2}} + X_{\mathrm{CH_4}}}{2X_{\mathrm{H_2}} + 4X_{\mathrm{CH_4}} + 2X_{\mathrm{H_2O}}},$$
$$\mathrm{O/H} = \frac{X_{\mathrm{CO}} + 2X_{\mathrm{CO_2}} + X_{\mathrm{H_2O}}}{2X_{\mathrm{H_2}} + 4X_{\mathrm{CH_4}} + 2X_{\mathrm{H_2O}}}. \tag{11}$$

Each mixing ratio is assumed to be constant over the entire model atmosphere. The alternative, which is to have a different value of the mixing ratio for each of the 100 model layers we assume, would result in 400 free parameters. This is unwarranted given the sparseness of the data, i.e., we have less than 400 data points.

For equilibrium chemistry, the carbon-to-oxygen ratio is simply

$$\mathrm{C/O} = \frac{f_\mathrm{C}}{f_\mathrm{O}}, \tag{12}$$

and $f_\mathrm{C} \equiv \mathrm{C/H}$ and $f_\mathrm{O} \equiv \mathrm{O/H}$ are directly the fitting parameters of the retrieval. Since the mixing ratios of all of the molecules can be exactly specified for each layer, which has its own temperature and pressure, the assumption of constant mixing ratios across height/pressure is unnecessary for the models with equilibrium chemistry. The 400 values of the mixing ratios are specified by just two free parameters: f_C and f_O.

Unlike in previous studies, we do not manually decide whether to pick unconstrained or equilibrium chemistry. Rather, we compute both of these models and select between them based on the Bayesian evidence computed.

We note that, as part of the ESP, we have previously developed a chemical kinetics solver named `VULCAN` (Tsai et al. 2017).

2.5. Radiative Transfer Scheme

With the cross-sections and temperature–pressure profiles in hand, one may compute the optical depth and hence the transmission function for each layer of the model atmosphere. To propagate fluxes through the atmosphere and thus obtain the synthetic spectrum, we need a radiative transfer scheme. Beer's law[11] is the simplest example of such a scheme, where incident radiation through a passive medium is exponentially attenuated. A more sophisticated radiative transfer scheme needs to account for both the fluxes incident upon a layer and the thermal emission associated with the layer itself, since each layer has a finite temperature. To this end, we use Equation (B4) of Heng et al. (2014),

$$F_{\uparrow_{j+1}} = F_{\uparrow_j}\mathcal{T} + \pi B(1 - \mathcal{T}), \tag{13}$$

where the fluxes are computed at the j and $(j + 1)$th interfaces. The Planck function (B) is evaluated within each layer. The transmission function is given by Equation (B5) of Heng et al. (2014)

$$\mathcal{T} = (1 - \Delta\tau)\exp(-\Delta\tau) + (\Delta\tau)^2\mathcal{E}_1, \tag{14}$$

with $\mathcal{E}_1(\Delta\tau)$ being the exponential integral of the first order. The optical thickness of each layer is given by $\Delta\tau$. Appendix A describes an analytical fitting formula for $\mathcal{E}_1$ that is highly accurate and allows for the computation to be significantly sped up.

We use Equation (13) to propagate the boundary condition at the bottom of the atmosphere (i.e., the internal/interior heat flux), which is the Planck function with a temperature given by the temperature–pressure profile at the bottom boundary. The outgoing flux at the top of the atmosphere is then the synthetic spectrum.

We emphasize that Equation (13) is an exact solution of the radiative transfer solution in the limit of isothermal layers and pure absorption. It is an improvement over using approximate solutions (e.g., two-stream solutions) and allows us to implement a radiative transfer scheme without taking any approximations besides assuming pure absorption. Equation (13) is equivalent to the approach of Line et al. (2016), who used four-point Gaussian quadrature to account for angle-dependent flux propagation. In our exact solution, the integration over angle has been performed analytically and is encapsulated in the exponential integral of the first order. We gain computational efficiency both by bypassing the need for performing Gaussian quadrature and also by evaluating $\mathcal{E}_1$ using an analytical approach (Appendix A). The overall accuracy is relegated to the number of discrete layers used.

The radius of the exoplanet (R) only appears as a scaling factor between the observed flux (F_obs) and the flux escaping from the top of the atmosphere (F_TOA),

$$F_\mathrm{obs} = \left(\frac{R}{d}\right)^2 F_\mathrm{TOA}, \tag{15}$$

where d is the distance between the observer and the object. The HR 8799 system is located at 39.4 ± 1.0 pc (van Leeuwen 2007), but the measured fluxes are usually reported as if it were located at $d = 10$ pc (i.e., absolute fluxes).

2.6. Cloud Model

The need for a cloud model is motivated by previous suggestions that the atmospheres of the HR 8799 exoplanets are cloudy (Barman et al. 2011; Madhusudhan et al. 2011; Marley et al. 2012), and also by the finding that each cloud configuration essentially corresponds to a different mass–radius relationship (Burrows et al. 2011; Lee et al. 2013). Our cloud model is based on the notion that, while cloud formation is challenging to model from first principles (e.g., Helling & Woitke 2006), once clouds do form it is somewhat easier to describe their effects on the synthetic spectrum, since this is derived from our knowledge of classical optics and Mie theory (Pierrehumbert 2010).

Following Lee et al. (2013), we consider the presence of clouds to add an extra contribution to the optical depth,

$$\Delta\tau_\mathrm{c} = Q_\mathrm{ext}\pi r_\mathrm{c}^2 n_\mathrm{cloud}\,\Delta z = Q_\mathrm{ext}\pi r_\mathrm{c}^2 f_\mathrm{cloud}\frac{\Delta P}{\bar{m}g}, \tag{16}$$

where Q_ext is the extinction efficiency, r_c is the radius of the (spherical) particles, n_cloud is the number density of clouds, and $\Delta z = \Delta P/n\bar{m}g$ is the spatial thickness of the layer. The cloud mixing ratio is $f_\mathrm{cloud} = n_\mathrm{cloud}/n$ and it is this quantity that we set a prior on (see Table 2 for its range of values). We assume the cloud to be uniformly distributed throughout the atmosphere.

In a departure from the approach of Lee et al. (2013), we do not use a specific composition of the cloud (e.g., enstatite). Specifically, we adopt their approximate fitting formula (listed in the appendix of Lee et al. 2013 but not used in their analysis),

$$Q_\mathrm{ext} = \frac{5}{Q_0 x^{-4} + x^{0.2}}, \tag{17}$$

where $x \equiv 2\pi r_\mathrm{c}/\lambda$ and λ is the wavelength. When the particles are small ($x \ll 1$), we recover Rayleigh scattering: $Q_\mathrm{ext} \propto \lambda^{-4}$. Large particles ($x \gg 1$) produce a roughly constant Q_ext. By contrast, Benneke & Seager (2013) assume their clouds to be described by only one number, which is the cloud-top pressure. Their model carries the implicit assumption that the cloud

[11] Also known as the Beer–Lambert–Bouguer law.

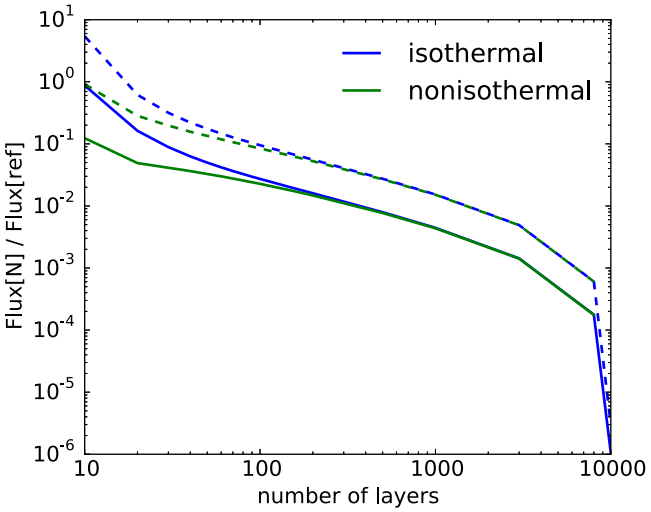

Figure 2. Mean (solid curves) and maximum (dashed curves) errors in the synthetic spectrum as a function of the number of model layers used, computed by performing retrievals on the measured spectrum of HR 8799b. The reference used is the retrieval with 10,000 model layers (see the text). When about 100 layers are used, the models with isothermal and non-isothermal layers yield the same answers.

particles are large compared to the range of wavelengths examined.

The dimensionless quantity Q_0 serves as a proxy for the cloud composition. Refractory species (e.g., silicates) have $Q_0 \approx 10$, while volatile species (e.g., ammonia, methane, and water) have $Q_0 \approx 40$–80. By using Q_0 as a fitting parameter in the retrieval, we can constrain the composition of the clouds. The other fitting parameters in our cloud model are r_c and f_{cloud}.

Since we do not self-consistently treat the cloud physics and gaseous chemistry, the caveat is that our retrieved C/O values are representative of only the gaseous component of the atmosphere. It is conceivable that the true C/O values, which must account for the material sequestered in the cloud particles, are different.

2.7. Data Selection: Spectra of HR 8799b, c, d, and e

The spectra and photometric data points of the HR 8799b, c, d, and e exoplanets have been taken from Bonnefoy et al. (2016) and Zurlo et al. (2016). The new SPHERE data were presented in Zurlo et al. (2016), while Bonnefoy et al. (2016) unified all of the previous data of the four exoplanets. Specifically, we use the data from Figure 4 of Bonnefoy et al. (2016).

To compute the flux in a photometric waveband, we simply integrate the synthetic spectrum over the range of wavelengths of the filter and assume a Heaviside function with a value of unity throughout. Unlike Lee et al. (2013), we do not apply filter functions with non-unity values to our synthetic spectrum, because this correction has already been done en route to reporting the observed fluxes in Bonnefoy et al. (2016) and Zurlo et al. (2016). It is unclear what has been done in previous studies. Madhusudhan et al. (2011) display filter functions in their Figure 1, but do not describe whether these filter functions were applied to their synthetic spectra. Line et al. (2013) state that, "For the broadband points we simply integrate the flux from the high-resolution model spectrum with the appropriate filter function for that point," but do not provide quantitative descriptions of their filter functions. It cannot be ruled out that these filter functions have values of unity throughout. For the spectroscopic data points, we do not convolve the synthetic spectrum with the instrument's response function, because the impact is minor for low-resolution spectra.

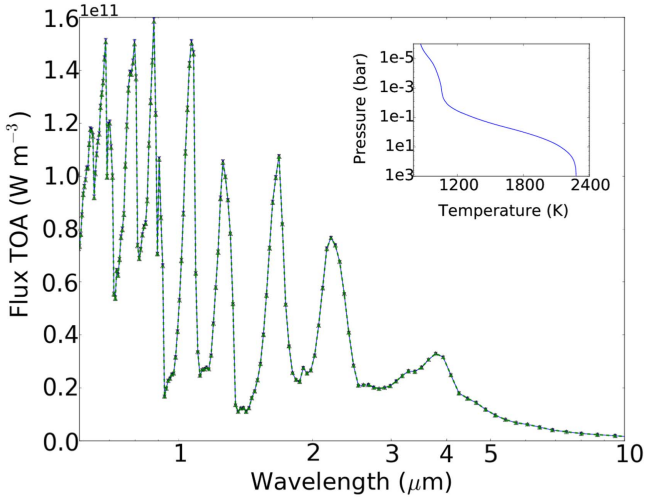

Figure 3. Validation of our HELIOS-R forward model (green dashed curve and upward-facing triangles) against that used in the HELIOS self-consistent radiative transfer code (blue solid curve and downward-facing triangles). The insert shows the temperature–pressure profile used as an input.

3. Tests

Before analyzing the measured spectra of the HR 8799b, c, d, and e directly imaged exoplanets, we subject HELIOS-R to a battery of tests.

3.1. Number of Atmospheric Layers

The number of layers used in a one-dimensional model atmosphere is a critical but often overlooked or unexplored detail. We wish to quantify the mean and maximum errors associated with assuming a specific number of model layers. We use the measured spectrum of HR 8799b as an illustration. We consider an ensemble of 10^3 cloud-free models with unconstrained chemistry. For each model, we randomly select our parameter values: two parameters for the temperature–pressure profile, four parameters for the mixing ratios, and one for the surface gravity. The range of parameter values used is listed in Table 2. No model selection is performed for this test. We consider forward models with both isothermal and non-isothermal layers. For the latter, we use Equation (B6) of Heng et al. (2014).

For each of the 10^3 models, the spectrum computed with 10,000 non-isothermal layers is used as a reference. We then compute coarser models with between 10 and 8000 isothermal or non-isothermal layers and calculate the fractional error on the synthetic spectrum compared to the reference model. In Figure 2, we show both the mean and maximum errors associated with the synthetic spectrum. With 100 layers, we see that models with isothermal and non-isothermal layers have the same mean and maximum errors of about 2.5% and 8%, respectively. For the rest of the paper, we will use 100 isothermal layers. For comparison, Madhusudhan & Seager (2009), Lee et al. (2013), and Line et al. (2013) used 100, 43, and 90 layers, respectively.

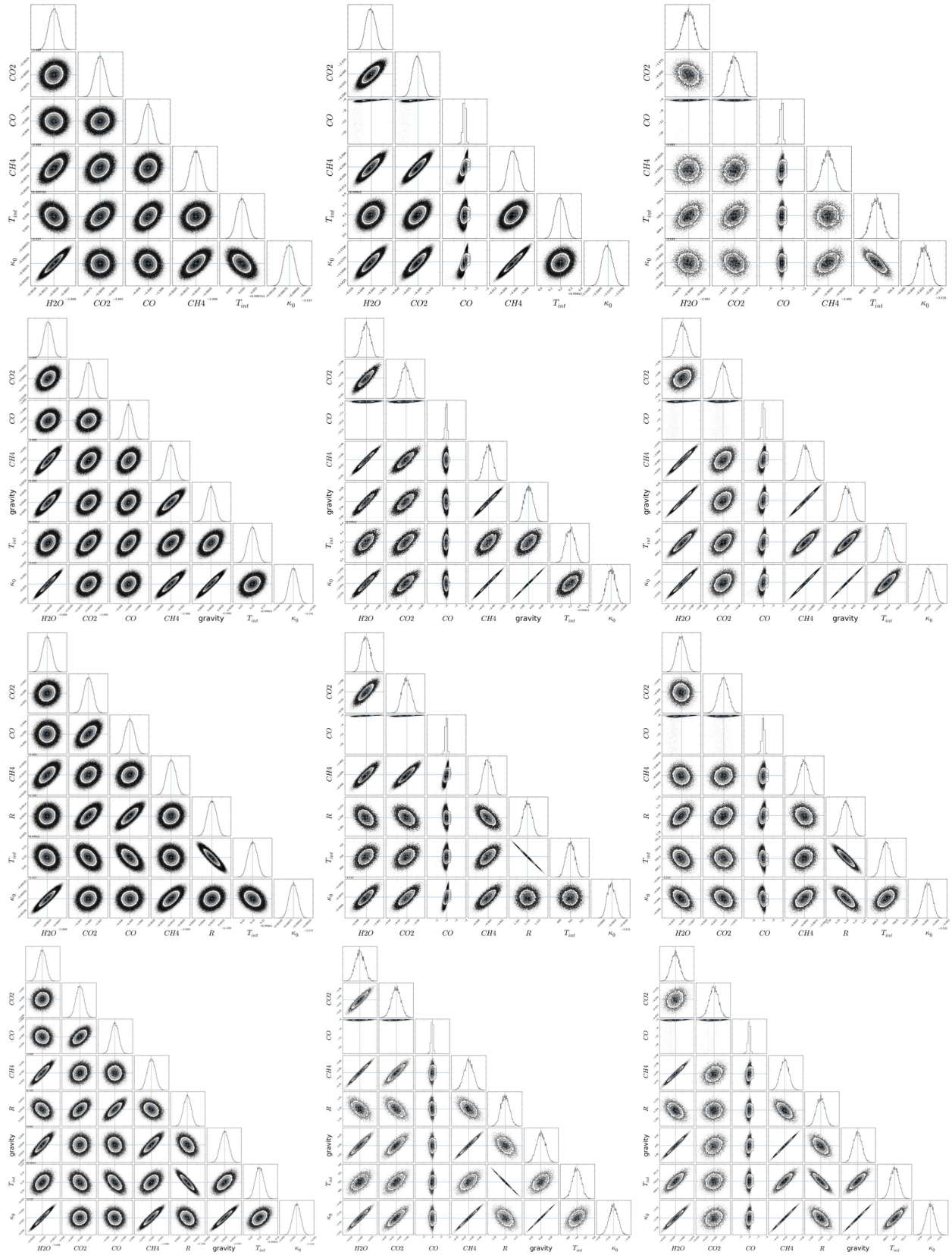

Figure 4. Performing retrieval on a mock data set, where the "ground truth" is known (see the text for input parameters). We focus on performing retrieval with the UB model (see Table 1). The columns represent the different wavelength coverages: 0.7–5 μm with 0.01 μm resolution (first column), HR 8799b-like data (second column) and HR 8799e-like data (third column). The rows represent different assumptions for R and g: fixed R and g (first row), fixed R but g is a fitting parameter (second row), fixed g but R is a fitting parameter (third row), both R and g are fitting parameters (fourth row). The labels "CO," "CO2," "CH4," and "H2O" refer to the mixing ratios of carbon monoxide, carbon dioxide, methane, and water, respectively. $T_{\rm int}$ is in units of K. The labels "κ_0" and "gravity" are for $\log \kappa_0$ and $\log g$ in mks and cgs units, respectively.

3.2. Validating the Forward Model

We previously developed a self-consistent radiative transfer code named HELIOS, which solves the radiative transfer equation in tandem with the first law of thermodynamics to obtain one-dimensional model atmospheres in radiative equilibrium (Malik et al. 2017). HELIOS was validated against the radiative transfer model of Miller-Ricci & Fortney (2010). In the limit of pure absorption, we also demonstrated that the two-stream and exact solutions produce excellent agreement if the diffusivity factor is set to two (Heng et al. 2014; Malik et al. 2017).

The forward model of HELIOS uses the same equation as HELIOS-R, but was implemented independently by the first author of each study. Here, we compare the forward models of HELIOS-R and HELIOS to verify that our implementation is bug-free. In Figure 3, we constructed a cross-section function consisting purely of water and used the k-distribution method to compute the fluxes. Malik et al. (2017) used an opacity function instead, but theirs also consisted purely of water, and they used the k-distribution method as well. The k-distribution tables were constructed using a resolution of $10^{-5}\,\mathrm{cm}^{-1}$ evenly distributed across the wavenumber (not shown). Other assumptions include a hydrogen-dominated atmosphere ($\mu = 2$), a water mixing ratio of 10^{-3}, and a surface gravity of $\log g = 3.3$ in cgs units ($\approx$19.5 m s^{-2}). We then assumed an input temperature–pressure profile, as shown in the insert of Figure 3, in tandem with the k-distribution tables to compute the synthetic spectrum using both HELIOS-R and HELIOS. The excellent agreement validates our implementation of the forward model.

3.3. Retrieval on a Mock Data Set

A useful test is to create a mock data set in which we know what the "ground truth" is concerning the synthetic spectrum, temperature–pressure profile, molecular abundances, surface gravity, etc. We assume a cloud-free model with unconstrained chemistry, which has the following input parameters.

$$\begin{aligned} X_{\mathrm{CO}} = X_{\mathrm{CO_2}} = X_{\mathrm{CH_4}} = X_{\mathrm{H_2O}} = 10^{-4}, \\ R = 1.2\,R_{\mathrm{J}},\ \log g = 4.0(\mathrm{cgs}), \\ T_{\mathrm{int}} = 700\ \mathrm{K},\ \ \kappa_0 = 2.9 \times 10^{-4}\ \mathrm{m^2\,kg^{-1}}, \end{aligned} \tag{18}$$

where R_{J} is the radius of Jupiter. Using this setup, we create three mock data sets: a full mock spectrum from 0.7 to 5 μm with 0.01 μm resolution, HR 8799b-like and HR 8799e-like data coverage. We assume this mock object to be located $d = 10$ pc away.

Such a test serves three purposes. First, if R and g are fixed to their input values (and excluded from being fitting parameters in the retrieval), then it is a test of the ability of our nested sampling algorithm to correctly recover the molecular abundances and temperature–pressure profile. Second, if we now include g and R as fitting parameters, it allows us to study the degeneracies associated with our ignorance of the surface gravity and/or radius. Third, by adapting and degrading the mock spectrum to the data resolution and spectral coverage of HR 8799b and HR 8799e, we may study the effects of incomplete or sparse data on the retrieved molecular abundances. The key difference between the currently available data for HR 8799b and HR 8799e is that the latter does not have K-band spectroscopy.

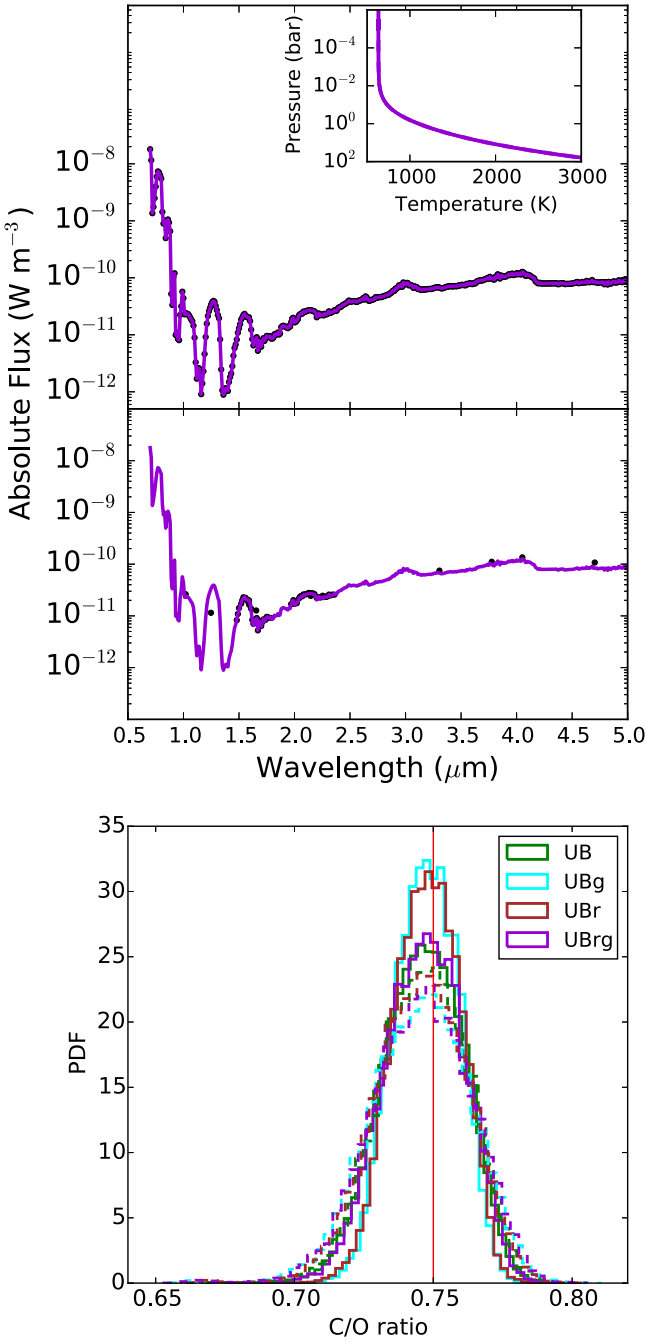

Figure 5. Further results from the retrievals on the mock data set. The top panel shows the mock data set at full resolution (0.01 μm), and also with HR 8799b-like data coverage (circles). The mock data set and best-fit spectrum overlap exactly (to within the resolution of the plot). The insert shows the retrieved temperature–pressure profile. The bottom panel shows the retrieved posterior distributions of C/O, assuming different models (see Table 1). The solid and dotted curves are for HR 8799b-like and HR 8799e-like data coverage. The broader posterior distributions of C/O associated with HR 8799e-like data coverage are primarily due to the lack of K-band data.

Figure 4 shows the outcomes of these tests. When R and g are fixed to their input values, HELIOS-R correctly recovers the input values of the mixing ratios and T–P profile parameters from the full mock spectrum (first row, first column). Surprisingly, our ability to recover these input values appears to be insensitive to whether the mock spectrum is degraded or not (first row, second and third columns), if R and g are known.

Figure 6. Montage of posterior distributions from performing retrieval on the measured spectrum of HR 8799b and allowing the radius, surface gravity, and mean molecular weights to be uniform or log-uniform priors. The retrieved value of the radius ($R \approx 0.5\ R_J$) is unphysical (see the text for a discussion). The labels "CO," "CO2," "CH4," and "H2O" refer to the mixing ratios of carbon monoxide, carbon dioxide, methane, and water, respectively. R and $T_{\rm int}$ are in units of R_J and K, respectively. The labels "κ_0" and "gravity" are for $\log \kappa_0$ and $\log g$ in mks and cgs units, respectively.

When the radius is implemented as a uniform prior, its value is correctly recovered, though the posterior distributions of the other fitting parameters become a little broader (third row of Figure 4). With HR 8799e-like data coverage, we see clear signs of degeneracies being introduced into the posterior distributions. It suggests that the *K*-band spectrum contains important information on the molecular abundances, an issue we will explore further in Section 4.4.

Allowing the surface gravity to be a fitting parameter has more serious consequences, as it introduces degeneracies into all of the other fitting parameters (second row of Figure 4). Even full data coverage does not lift these degeneracies (second row, first column). It suggests that an informative prior needs to be set on the surface gravity.

Surprisingly, the retrieved posterior distribution of C/O appears to be robust to the different model assumptions (Figure 5). It suggests that the C/O is a robust outcome of the retrieval.

Overall, these exercises teach us that wavelength coverage and spectral resolution are generally not as important as

knowledge of the surface gravity, though the *K*-band spectrum appears to encode crucial information on the molecular abundances. In Section 4.1, we will argue for setting Gaussian priors on $\log g$ as well as R when analyzing real data from the HR 8799 exoplanets.

4. Results

4.1. Setting Priors on Radius and Surface Gravity

The strongest demonstration of why our assumptions for the prior distributions of input parameters are important comes from examining a model where the radius and surface gravity are implemented as uniform priors in the retrieval. Specifically, we perform a retrieval on the measured spectrum of HR 8799b using model UBrg in Figure 6, where R and g are specified as uniform priors. We see that the retrieved solution is $R \approx 0.5\ R_J$, which is physically unreasonable. The surface gravity takes on unphysical values of $\log g \approx 5.5$–6. As we have learned from the mock-retrieval exercises in Section 3.3, these difficulties stem from specifying the radius and surface gravity as unconstrained fitting parameters.

We now discuss why the values for R are physically unreasonable. There are indirect arguments for why retrieved solutions with radii well below a Jupiter radius should be rejected. First, brown dwarfs and low-mass stars with masses between 20 and 100 Jupiter masses have transit radii that are at least $0.8\ R_J$ (see Burrows et al. 2011 and references therein), including CoRoT-3b, which is a low-mass brown dwarf with a dynamical mass of $M = 21.66 \pm 1.0\ M_J$ and a transit radius of $R = 1.01 \pm 0.07\ R_J$ (Deleuil et al. 2008). Second, a review of the data for all of the transiting Jupiter-like exoplanets also reveals that objects with radii below $0.8\ R_J$ do not exist (Figure 7). When a cut is made to only include objects with zero-albedo equilibrium temperatures below 1000 K (to exclude objects that are "inflated" by some unknown mechanism related to stellar heating, e.g., Demory & Seager 2011), we find that the radii are bound between 0.8 and $1.2\ R_J$. The single outlier with $R = 1.65^{+0.59}_{-0.56}\ R_J$ is Kep-447b, which has an extremely grazing transit (Lillo-Box et al. 2015) that may render its radius measurement unreliable. Third, objects with a mass of Jupiter (or higher) are partially degenerate and it is theoretically challenging to get their radius to be less than that of Jupiter's (Burrows & Liebert 1993).

While it may be tempting to fix our model radius at between 0.8 and $1.2\ R_J$, we should be reminded of the fact that these radii are measured for >1 Gyr old objects, whereas the HR 8799 exoplanets are estimated to be ~10–100 Myr old. Guided by evolutionary models (Mordasini et al. 2012; Spiegel & Burrows 2012), we set $R = 1.2 \pm 0.1\ R_J$ as a Gaussian prior of our retrievals. The uncertainty of $0.1\ R_J$ is the full-width at half-maximum of the Gaussian. We note that Moses et al. (2016) assume a fixed value of $R = 1.2\ R_J$ for their self-consistent model of HR 8799b.

The bottom panel of Figure 7 is also revealing, as it shows the measured surface gravities of transiting Jupiter-sized exoplanets to be hovering around $\log g \approx 4$ for objects with masses of $M > 2\ M_J$, where M_J is the mass of Jupiter. Since we expect the HR 8799 exoplanets to have radii that are slightly larger than Jupiter's, we expect their surface gravities to also be $\log g \approx 4$. Surface gravities of $\log g \approx 4.5$–5.0 are only appropriate when one crosses over into the brown dwarf

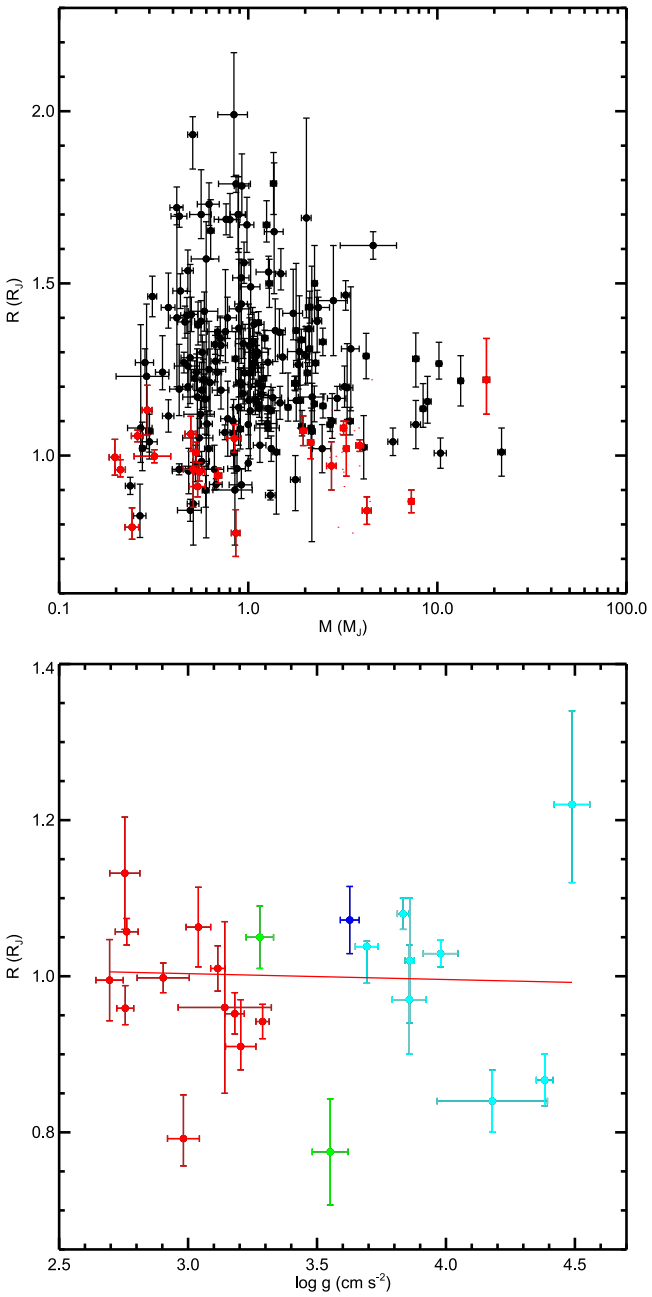

Figure 7. Top panel: transit radius vs. surface gravity of a sample of transiting Jupiter-sized exoplanets around main-sequence stars (black data points). The red data points are the sub-sample of transiting exoplanets with zero-albedo equilibrium temperatures below 1000 K. The single outlier is Kep-447b (see the text). Bottom panel: the same sub-sample, but color-coded by mass. The red, green, blue, and cyan points are for $<0.8\ M_J$, $0.8 < M < 1.2\ M_J$, $1.2 < M < 2\ M_J$, and $>2\ M_J$, respectively. Data taken from http://www.exoplanets.org (Han et al. 2014).

regime ($\gtrsim 13\ M_J$), e.g., CoRoT-3b has $\log g = 4.72 \pm 0.07$. The photometric masses of HR 8799b, c, d, and e are less than half that of CoRoT-3b (Marois et al. 2008, 2010). Based on the evolutionary calculations of Marleau & Cumming (2014), who estimated $M \approx 4$–$13\ M_J$ for the HR 8799 exoplanets, we set a Gaussian prior of $\log g = 4.1 \pm 0.3$ on the surface gravity (taking into account $R = 1.2 \pm 0.01\ R_J$). This range of surface gravities is somewhat higher than the $\log g = 3.5 \pm 0.5$ values considered by Barman et al. (2015).

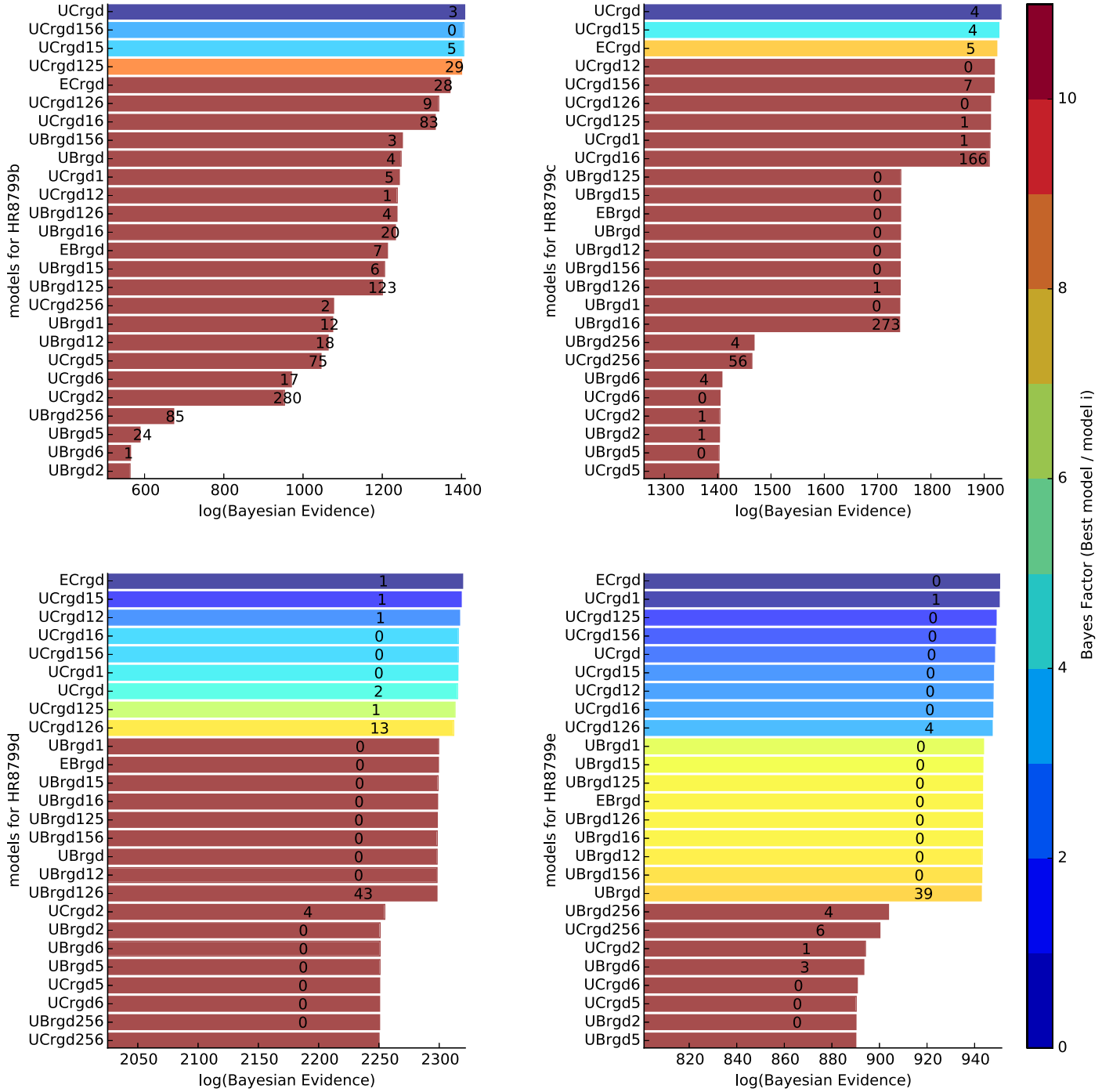

Figure 8. Bayes factors from a suite of models for each HR 8799 exoplanet. See Table 1 for an explanation of the shorthand notation used to mark each model. All of the models assume Gaussian priors on R and $\log g$ (see Table 2). For HR 8799b and c, the Bayesian evidence clearly favors cloudy models with non-equilibrium (unconstrained) chemistry. For HR 8799e, the lack of K-band spectroscopy implies that none of the models are strongly favored. The number associated with each histogram is the logarithm of the Bayes factor between the model in question and its neighbor below. The color bar shows the logarithm of the Bayes factor between the model in question and the best model, which is the model placed at the top of each panel.

In summary, we find that what we assume for the prior distributions of the input quantities is critical to the outcome of the retrieval. Uniform or log-uniform priors may not always be the best choice because they may lead to unphysical or even nonsensical outcomes. Gaussian priors are better choices in these instances, but only when they are guided by physics. We find our retrievals to be physically meaningful only when Gaussian priors are set on the radius and surface gravity, which is a departure from the HR 8799b analysis of, e.g., Lee et al. (2013).

4.2. Model Selection Using Bayesian Evidence

Traditionally, model selection is performed manually by the modeler or theorist. One starts with a set of assumptions, computes forward and arrives at a prediction for the thermal structure and synthetic spectrum. These assumptions include chemical equilibrium or disequilibrium, a value for the strength of atmospheric mixing, the number of atoms and molecules included in the model, the metallicity, and C/O, etc. Other assumptions are more closely related to technique, e.g., the

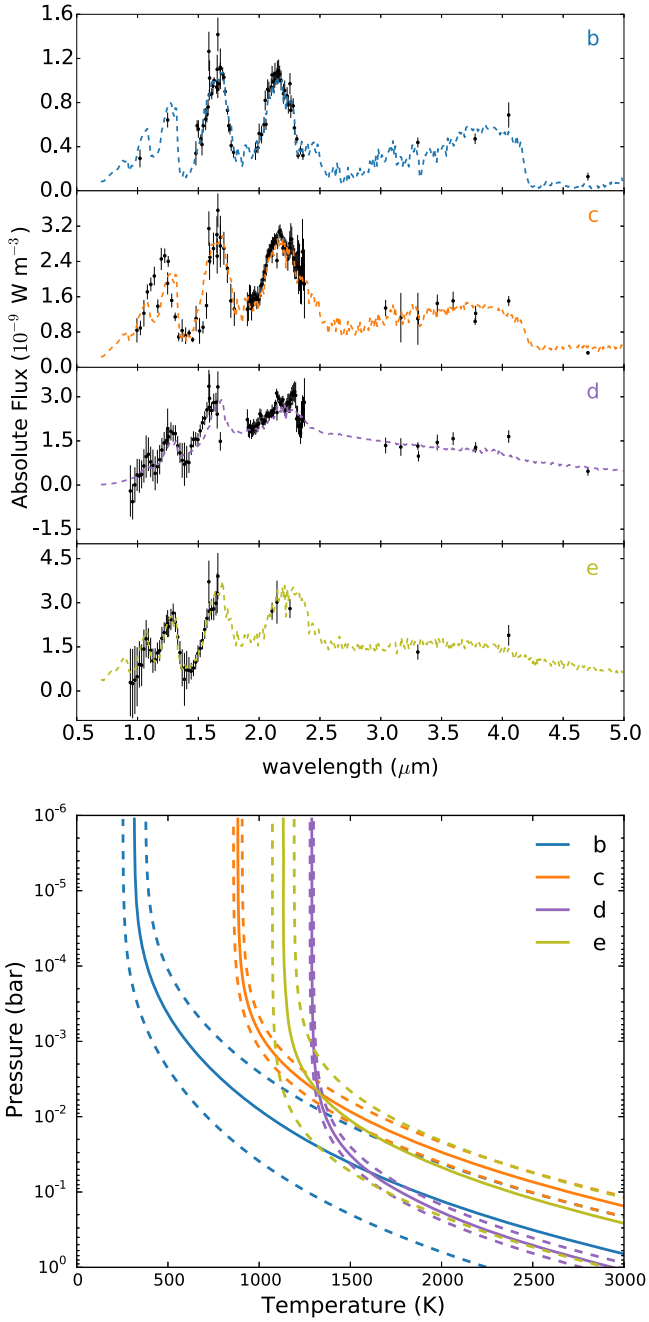

Figure 9. Best-fit spectra and temperature–pressure profiles for HR 8799b, c, d, and e.

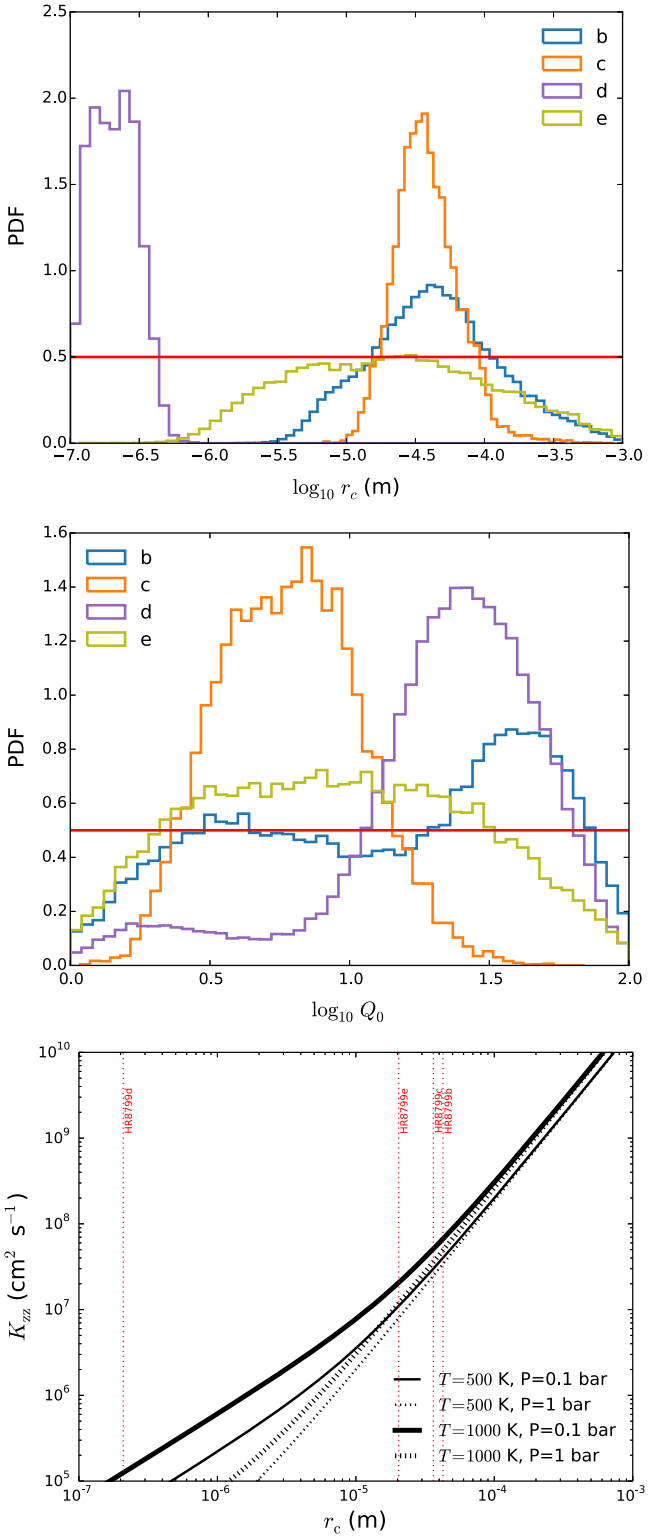

Figure 10. Top and middle panels show the retrieved cloud particle radius and composition parameter for the HR 8799b, c, d, and e directly imaged exoplanets, where the flat line is the prior. The bottom panel shows the eddy diffusion coefficient and demonstrates that its inferred value corresponding to the peak value of the posterior distribution for r_c are not unreasonable and broadly consistent with the assumed K_{zz} values in previous studies. PDF stands for "probability density function."

approximate or limiting form of the radiative transfer equation being solved.

Like all of the other previous studies involving both forward modeling and retrieval, we inevitably make a set of both physical and technical assumptions. However, we use our nested sampling approach to go a step further: we compute the Bayesian evidence for models with and without equilibrium chemistry. We then compare them in order to formally quantify whether equilibrium chemistry is a warranted assumption. Instead of assuming a fixed set of cloud parameters for each retrieval, as was done by Lee et al. (2013), we allow our cloud model to be part of the retrieval and also compare its Bayesian evidence to a retrieval that assumes a cloud-free atmosphere. In these ways, we allow model selection based on the Bayesian evidence to inform us of whether the atmosphere is cloudy or cloud-free and in chemical equilibrium or disequilibrium.

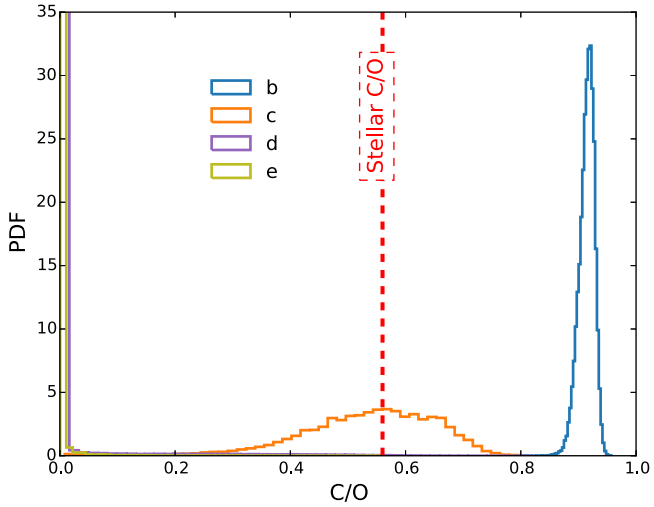

Figure 11. Retrieved C/O values for the HR 8799b, c, d, and e directly imaged exoplanets. The stellar C/O value is about 0.56. PDF stands for "probability density function."

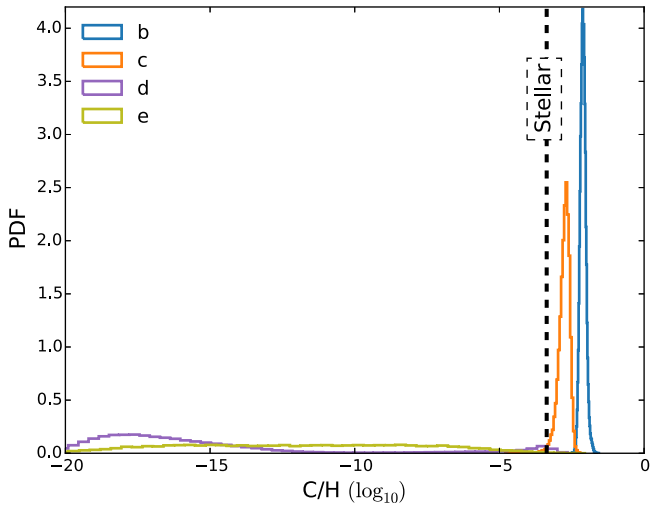

Figure 12. Retrieved C/H values for the HR 8799b, c, d, and e directly imaged exoplanets. The stellar C/H value is about 4.3×10^{-4}. PDF stands for "probability density function."

Figure 8 shows a montage of all of the models tested for all four HR 8799 exoplanets. Table 1 explains what the labels of the models correspond to. For HR 8799b, c, and d, we see that the Bayesian evidence favors model atmospheres that are not in chemical equilibrium and are cloudy. For HR 8799e, the relative lack of data, compared to the other HR 8799 exoplanets, means that we are unable to strongly select between the different models.

Figure 9 shows the best-fit spectra. Our retrieval procedure generally manages to find good fits to the data, except for the band-head near 1 μm for HR 8799c. We speculate that this mismatch could be due to the influence of an additional molecule we have not included in our analysis, but we deem it beyond the scope of the present paper to identify it. In the Appendix, Figure 16 elucidates the effects of using ExoMol methane and water versus HITRAN methane and HITEMP water.

For the rest of this paper, we will discuss the retrieved properties of the HR 8799b, c, and d exoplanets based on the best-fit models only. For HR 8799e, we will discuss results from the model with equilibrium chemistry and that includes all four molecules in the retrieval. In Figures 18–21 of

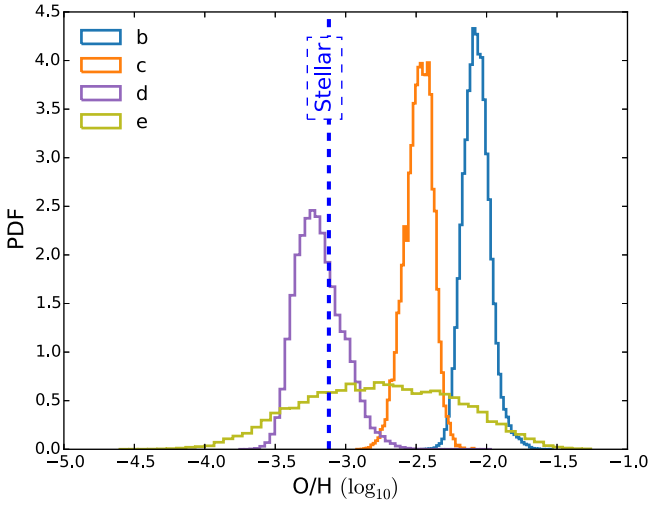

Figure 13. Retrieved O/H values for the HR 8799b, c, d, and e directly imaged exoplanets. The stellar O/H value is about 7.6×10^{-4}. PDF stands for "probability density function."

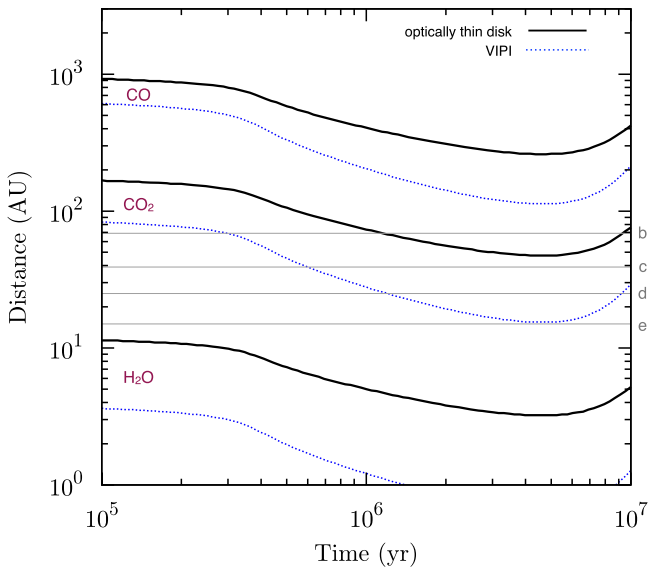

Figure 14. Positions of the CO, CO_2, and H_2O snowlines or icelines as functions of the stellar age of the HR 8799 system. The solid curves are calculated assuming an optically thin disk. The dotted curves are calculations from a vertically isothermal, passively irradiated (VIPI) disk.

Appendix B, we provide the full posterior distributions of the best models for all exoplanets for completeness.

4.3. Retrieving the Cloud Properties and Inferring K_{zz}

Figure 10 shows the retrieved posterior distributions of the cloud particle radius (r_c) and composition parameter (Q_0). Unsurprisingly, the retrieved values of Q_0 span a broad enough range (three to four orders of magnitude) that they are uninformative with regards to distinguishing between different compositions, consistent with the expectation that the absorption and scattering properties of the cloud are mainly determined by the particle size and less by the composition (Heng & Demory 2013).

The inferred values of r_c span a broad range and lie between about 1 and 100 μm. The presence of these cloud particles implies that they are being held aloft by atmospheric motion. Since these exoplanets are not being heavily irradiated (unlike

for hot Jupiters), we can safely assume that the underlying mechanism driving this motion is convection (Burrows et al. 1997; Chabrier et al. 2000; Baraffe et al. 2002) and estimate approximate values for the associated "eddy diffusion coefficient," which we denote with K_{zz}. We use Equations (15) and (17) of Spiegel et al. (2009), as well as Equations (6) and (8) of Heng & Demory (2013), to calculate the terminal speed associated with a particle of radius r_c, which we denote by $v_{\rm terminal}$. The eddy diffusion coefficient is roughly

$$K_{zz} \sim 0.1 v_{\rm terminal} H, \quad (19)$$

where the pressure scale height is $H = k_B T/\tilde{m} g$ and k_B is the Boltzmann constant. We follow the prescription of Smith (1998) and use $0.1H$ as the characteristic length scale, which is more conservative than what was assumed in Lee et al. (2013). We note that the preceding expression for K_{zz} has no dependence on g, as it appears in the numerator of $v_{\rm terminal}$ and the denominator of H. We assume the intrinsic density of the particles to be 3 g cm^{-3}.

In Figure 10, we see that K_{zz} spans a broad range of values from $\sim 10^5$ cm^2 s^{-1} to $\sim 10^{10}$ cm^2 s^{-1} as r_c increases from 1 μm to 1 mm. The deviation in the curves between $P = 0.1$ and 1 bar arises from the Cunningham–Millikan–Davies "slip factor correction" kicking in when the mean-free path for collisions between the hydrogen molecules becomes comparable to the cloud particle radius. If we place the retrieved values of r_c corresponding to the peak of each posterior distribution on the plot, we infer $K_{zz} \sim 10^5$–10^8 cm^2 s^{-1}, in agreement with Barman et al. (2015). Madhusudhan et al. (2011) assume $K_{zz} = 10^2$–10^6 cm^2 s^{-1}, while Barman et al. (2011) and Marley et al. (2012) assume $K_{zz} = 10^4$ cm^2 s^{-1}.

4.4. Retrieving C/O, C/H, and O/H for the HR 8799b, c, d, and e Exoplanets and Implications for Planet Formation

4.4.1. The Star of HR 8799

We refer to the "metallicity" as the set of elemental abundances with atomic mass numbers that are larger than that of hydrogen and helium. In our current study, these would be $f_C \equiv$ C/H and $f_O \equiv$ O/H. For comparison, their values in the solar photosphere are $f_C \approx 3 \times 10^{-4}$ and $f_O \approx 6 \times 10^{-4}$, such that C/O ≈ 0.5 (Lodders 2003). For the star of the HR 8799 system, Sadakane (2006) has found that

$$\mathrm{C/H}_\star \approx 4.3 \times 10^{-4}, \ \mathrm{O/H}_\star \approx 7.6 \times 10^{-4}, \ \mathrm{C/O}_\star \approx 0.56. \quad (20)$$

4.4.2. Retrieved C/O, C/H, and O/H Values

Given the interest in the possibility of carbon-rich exoplanets (Gaidos 2000; Kuchner & Seager 2005), our retrieval analysis yields the posterior distributions of C/O, C/H, and O/H for the atmospheres of HR 8799b, c, d, and e in Figures 11–13, respectively, which we then compare to the values for the star listed in Equation (20). A caveat is that the retrieved values are only for the gaseous phase and the true C/O ratio may be hidden in a condensed phase such as graphite (Moses et al. 2013). The retrieved posterior distributions of C/O and C/H for HR 8799e are not as definitive as for the other three exoplanets, because its K-band spectrum has not been measured.

4.4.3. Locations of Snowlines/Icelines

Konopacky et al. (2013) have previously estimated that the H_2O, CO_2, and CO snowlines or icelines are located at about 10, 90, and 600 au, respectively. We wish to point out that the iceline locations depend on the formation history of the HR 8799 exoplanets.

In Figure 14, we show calculations of the locations of the CO, CO_2, and H_2O icelines as functions of the age of the HR 8799 system. We consider two scenarios: an optically thin disk and a vertically isothermal, passively irradiated disk. For the optically thin disk, the temperatures are simply the zero-albedo equilibrium temperatures at a given distance from the star informed by the Pisa stellar evolution models (Tognelli et al. 2007). By "passively irradiated," we mean that viscous heating associated with turbulence is neglected (Chiang & Goldreich 1997). Both models consider the evolution of stellar heating as the star ages. We expect more sophisticated calculations that involve temperature gradients, photoevaporation, and viscous heating to produce iceline curves that are intermediate between these two scenarios. The calculations are shown for $t = 10^5$–10^7 years because this encompasses the gas-clearing phase of the protoplanetary disk. Curiously, the CO_2 iceline sits between different pairs of HR 8799 exoplanets as its location evolves during the gas-clearing phase ($t \sim 10^6$ years), implying that a variation in the C/O, C/H, and O/H values of these exoplanets may be a natural outcome of the planet formation process.

4.4.4. Implications for Planet Formation

Our findings have implications for planet formation, if we assume the retrieved C/O, C/H, and O/H values to be representative of the bulk composition of each exoplanet. Öberg et al. (2011) previously elucidated the chemical signatures associated with the planet formation mechanism and history of an exoplanet. If an exoplanet forms by gravitational instability, the zeroth-order expectation is that its C/O, C/H, and O/H values mirror that of the star, unless late-time accretion occurred. This is clearly at odds with our inferred values of C/O, C/H, and O/H for the HR 8799b, c, d, and e exoplanets.

In the context of the core accretion formation mechanism, all four exoplanets should have C/O values that are enhanced above stellar, but below unity, if they formed in situ and in between the water and carbon dioxide snowlines/icelines (Öberg et al. 2011). Our retrieved values of C/O for HR 8799b and c are consistent with this scenario, whereas HR 8799d and e have sub-solar C/O values. Öberg et al. (2011) have suggested that substellar C/O values are still consistent with core accretion if the late-time accretion of planetesimals has occurred to pollute the atmospheres. The link between late-time planetesimal accretion and atmospheric composition has been emphasized by Mordasini et al. (2016). The HR 8799b and c exoplanets have superstellar C/H and O/H values, which suggests that they accreted both carbon- and oxygen-rich ices. The HR 8799d and e exoplanets, which reside closer to the star, have substellar C/H values but stellar to superstellar O/H values, which suggests the accretion of only oxygen-rich ices.

Overall, our retrieved values of C/O, C/H, and O/H appear to be consistent with the core accretion formation mechanism and inconsistent with gravitational instability without late-time accretion, as has been suggested by, e.g., Kratter et al. (2010).

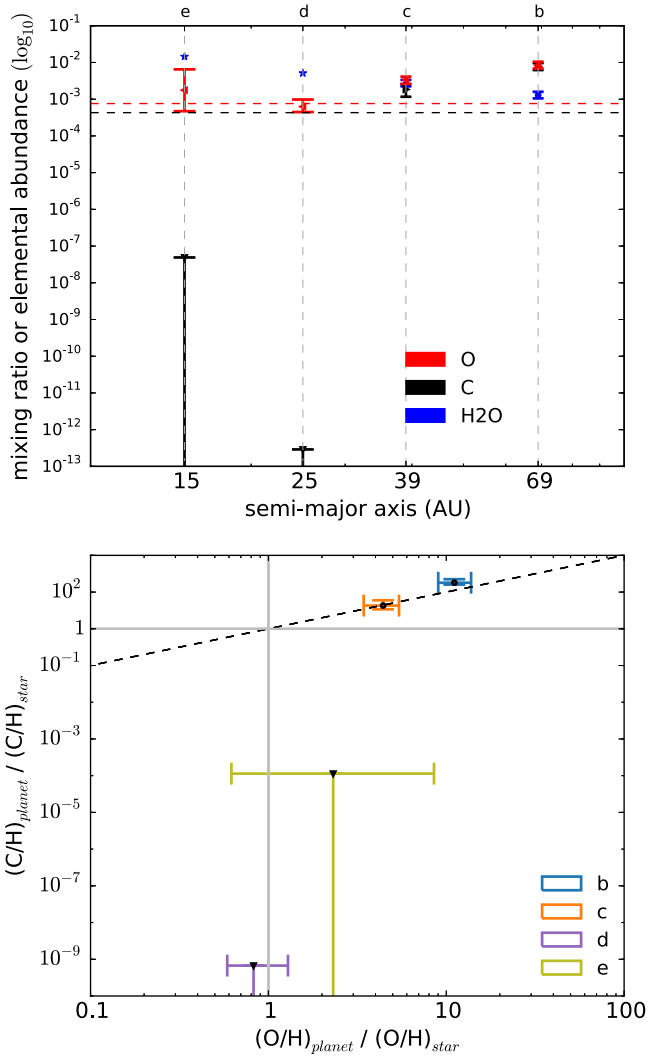

Figure 15. Summary of our main results. The top panel shows the retrieved water mixing ratios and elemental abundances of carbon and oxygen for all four HR 8799 exoplanets. For HR 8799d and e, we show the water abundance in chemical equilibrium at 1 bar (represented by the blue stars). For C/H and O/H, we also show the corresponding values of the HR 8799 star (horizontal dashed lines). The bottom panel shows the exoplanetary elemental abundances normalized to their stellar values with the dashed line denoting parity.

4.4.5. Why Spectroscopy in the K Band is Crucial

A lesson we have learned from our analysis is that spectroscopy in the *K* band is crucial for obtaining meaningful constraints on C/H and C/O, as it affects the ability of the retrieval approach to constrain the abundances of CO and/or CH_4. The lack of *K*-band spectroscopy for HR 8799e hampers our ability to make stronger statements on its C/H and C/O values. These findings have implications for the design of future instruments on the European Extremely Large Telescope (ELT). Furthermore, multiple wavebands should be monitored simultaneously in order to detect variability (Apai et al. 2016).

5. Discussion

5.1. Summary and Comparison to Previous Work

We have presented the complete methodology for a nested sampling atmospheric retrieval code named HELIOS-R, which allows us to insert arbitrary prior distributions of parameters and also compute the full posterior distributions of the retrieved quantities. In its current implementation, we used analytical formulae for the forward model, temperature–pressure profile, and equilibrium chemistry, as well as a customized opacity calculator (HELIOS-K). By computing the Bayesian evidence, we can compare models that assume equilibrium versus unconstrained chemistry and determine which scenario is favored by the data.

We apply HELIOS-R to the measured spectra of the HR 8799b, c, d, and e directly imaged exoplanets. We find that the outer HR 8799b and c exoplanets are enriched in carbon and have superstellar and stellar C/O values, respectively. The inner HR 8799d and e exoplanets are diminished in carbon and C/O. All four exoplanets are possibly enriched in oxygen relative to the star, which is a clear signature of late-time accretion of water-rich planetesimals. Figure 15 provides a summary of our findings. We note that our retrieved water abundances are about two to three orders of magnitude higher than what was found by Madhusudhan et al. (2014) for three hot Jupiters, though it should be noted that these authors do not include a cloud model in their retrievals. The inclusion of a cloud model should worsen the discrepancy between these outcomes. Our retrieved molecular abundances and C/O for HR 8799b are in broad agreement with Lee et al. (2013), despite differences in our retrieval techniques. Table 3 summarizes the properties of the four exoplanets inferred from the retrieval.

Our conclusions differ somewhat from previous studies, which reach a diversity of conclusions. Barman et al. (2011) used self-consistent models to interpret the H- and *K*-band spectra of HR 8799b. They infer $R = 0.75^{+0.17}_{-0.12}\,R_J$ and $M = 0.72^{+2.6}_{-0.6}\,M_J$. We deem this radius value to be unphysical for the reasons described in Section 4.1. Marley et al. (2012) also used self-consistent models and found that if the theoretical interpretation is made of the photometry alone, then the inferred radius for the HR 8799b exoplanet is 1.11 R_J but with a surface gravity of $\log g = 4.75$, considerably higher than the $\log g = 3.5 \pm 0.5$ value of Barman et al. (2011). Madhusudhan et al. (2011) used self-consistent models[12] with various cloud configurations to conclude that the HR 8799b, c, and d exoplanets have masses of 2–12, 6–13, and 3–11 M_J, respectively, and surface gravities $\log g \approx 4$. In these studies, solar abundance is assumed. The diversity of reported results from these studies already hint at the difficulty of using photometry and spectroscopy to infer the radius and mass of a directly imaged exoplanet from the traditional use of forward modeling.

Barman et al. (2015) performed a manual fitting of the *H*- and *K*-band spectra of HR 8799b and HR 8799c. They first held the CO and CH_4 abundances fixed to their solar values, then fitted for the abundance of H_2O. The bandheads involving CO and CH_4 are masked or excluded from the fit. Next, the H_2O abundance is held at its best-fit value (and CH_4 is again held fixed at its solar value) and the abundance of CO is inferred. The final step involves fitting for CH_4. Such an approach is plausible as a first step, but does not explore the model degeneracies. It is likely that the reported value of $C/O = 0.61 \pm 0.05$ for HR 8799b has uncertainties that are underestimated. Barman et al. (2015) themselves remark that, "The various sources of uncertainty in the models (are) not accounted for in the formal mole fraction error-bars." Building on the work of Barman et al. (2015), Moses et al. (2016) assumed fixed values for the equilibrium temperature, surface gravity, radius, C/O, metallicity, and K_{zz}, as well as a fixed temperature–pressure profile. They explored thermo- and photochemical models of HR

[12] Strictly speaking, these are parametric models because the cloud physics is not treated self-consistently with the gaseous chemistry and is instead parametrized.

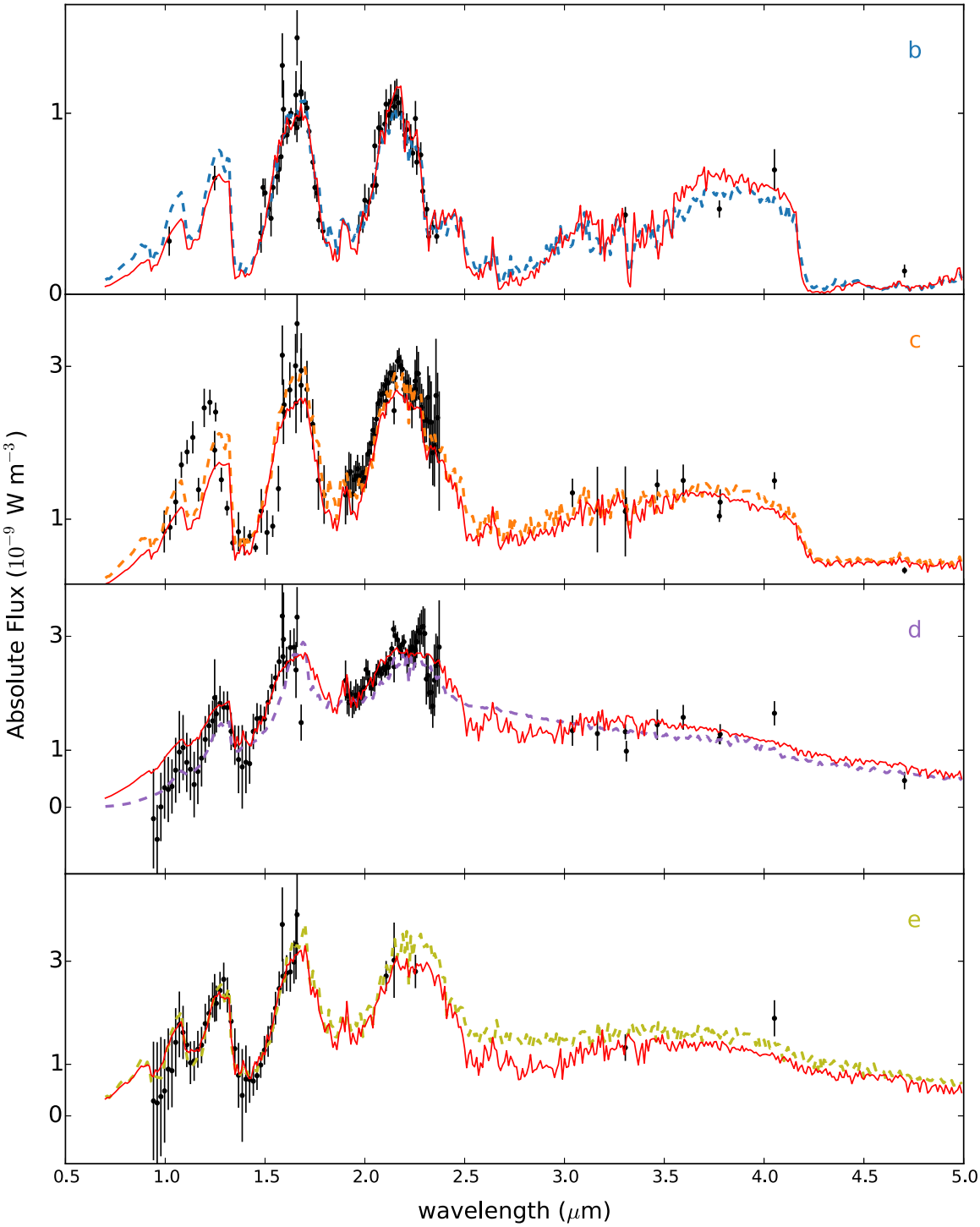

Figure 16. Elucidating the effects of using different spectroscopic line lists. The dashed curves in each panel show the retrievals using ExoMol data for water and methane. The red, continuous curves use the retrieved parameters to produce model spectra but using HITEMP water and HITRAN methane (post-processing).

8799b and produced synthetic spectra that somewhat match the measured spectrum (see their Figure 14).

Lee et al. (2013) analyzed the HR 8799b exoplanet and reported supersolar metallicities for their best fits, consistent with the present study. They considered two cloud models, where the monodisperse cloud particle radius is fixed manually and not formally included as part of the retrieval. The cloud composition is also assumed to be enstatite, whereas we have allowed the cloud composition to be part of the retrieval. The models of Lee et al. (2013) allowed for R and g to be uniform or log-uniform priors, whereas in the current study we have chosen R and $\log g$ to be Gaussian priors. Somewhat surprisingly, despite these differences, they retrieve a C/O value that is similar to what we find (see Figure 11). On the technical side, Lee et al. (2013) used the NEMESIS code, which implements a nonlinear optimal estimation (versus the nested sampling algorithm we implemented). This technique, which is also used by Barstow et al. (2015), assumes that the priors and posteriors are Gaussian and is unable to formally perform model

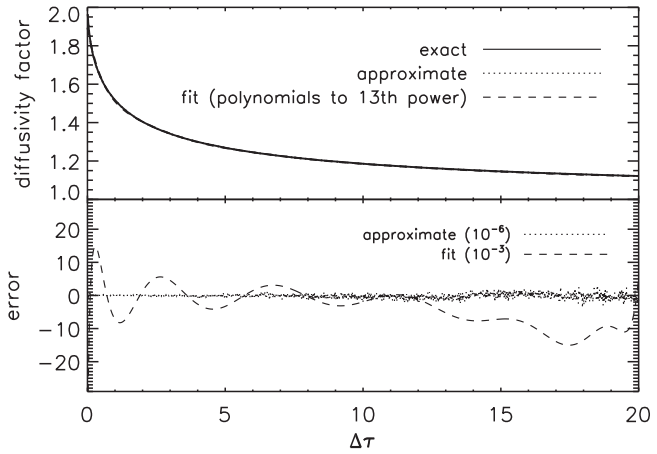

Figure 17. Calculations of the diffusivity factor using a canned routine ("exact") vs. those performed using our fitting formulae for $\mathcal{E}_1$ in Equation (21) ("approximate"). The calculation labeled "fit" is a 13th order polynomial fit to the exact solution, which performs poorly even at the $\sim 10^{-3}$ level.

selection via Bayesian evidence comparison. Lee et al. (2013) also do not consider equilibrium chemistry in their comparison of models. (See Line et al. 2013 for a comparison of these optimization methods.) Overall, `HELIOS-R` implements a number of improvements over `NEMESIS` that are more appropriate for the sparse data regime of exoplanetary atmospheres (compared to the remote sensing data of solar system objects) and is able to more rigorously explore a broader range of parameter space.

5.2. Opportunities for Future Work

There are ample opportunities for future work. Instead of unconstrained chemistry, disequilibrium chemistry may be described by some form of atmospheric mixing (e.g., eddy diffusion). More molecules may be added to the analysis, including acetylene, ethylene, and hydrogen cyanide, which are known to be spectroscopically active in the infrared at temperatures higher than for the photospheres of the HR 8799 exoplanets. Ultimately, it is our hope that the collective body of work on atmospheric retrieval will stimulate and connect to work on disk chemistry (e.g., Cridland et al. 2016). It will also be insightful to train `HELIOS-R` on a large sample of brown dwarf photometry and spectra, as Line et al. (2015) have done for two T dwarfs.

B.L., M.M., J.M., S.G., M.O., and K.H. thank the Center for Space and Habitability (CSH), the Swiss National Science Foundation (SNSF), the PlanetS National Center of Competence in Research (NCCR), and the MERAC Foundation for partial financial support.

Software: Corner (Foreman-Mackey et al. 2016), HELIOS-K (Grimm & Heng 2015), MultiNest (Feroz & Hobson 2008; Feroz et al. 2009, 2013), PyMultiNest (Buchner et al. 2014), VULCAN (Tsai et al. 2017).

Appendix A
Analytical Formula for the Exponential Integral of the First Order

We may avoid the numerical integration of the exponential integral of the first order by using the approximate, but highly accurate, analytical formulae presented in Abramowitz & Stegun (1970),

$$\mathcal{E}_1 = \begin{cases} -\ln \Delta\tau + \sum_{j=0}^{5} \mathcal{A}_j (\Delta\tau)^j, & \Delta\tau \leqslant 1, \\ (\Delta\tau)^{-1} \exp(-\Delta\tau) \dfrac{\sum_{j=0}^{4} \mathcal{B}_j (\Delta\tau)^{4-j}}{\sum_{j=0}^{4} \mathcal{C}_j (\Delta\tau)^{4-j}}, & \text{otherwise.} \end{cases} \quad (21)$$

The fitting coefficients $\mathcal{A}_j$, $\mathcal{B}_j$, and $\mathcal{C}_j$ are given in equations (5.1.53) and (5.1.56) of Abramowitz & Stegun (1970), but we reproduce them here for convenience: $\mathcal{A}_0 = -0.57721566$, $\mathcal{A}_1 = 0.99999193$, $\mathcal{A}_2 = -0.24991055$, $\mathcal{A}_3 = 0.05519968$, $\mathcal{A}_4 = -0.00976004$, and $\mathcal{A}_5 = 0.00107857$; $\mathcal{B}_0 = \mathcal{C}_0 = 1$, $\mathcal{B}_1 = 8.5733287401$, $\mathcal{B}_2 = 18.059016973$, $\mathcal{B}_3 = 8.6347608925$, $\mathcal{B}_4 = 0.2677737343$, $\mathcal{C}_1 = 9.5733223454$, $\mathcal{C}_2 = 25.6329561486$, $\mathcal{C}_3 = 21.0996530827$, and $\mathcal{C}_4 = 3.9584969228$. As originally stated by Abramowitz & Stegun (1970), the formula involving $\mathcal{A}_j$ has a precision better than 2×10^{-7}, while that involving $\mathcal{B}_j$ and $\mathcal{C}_j$ is precise to better than 2×10^{-8}. In Figure 17, we check these claims by evaluating $\mathcal{E}_1$ using a canned routine (`expint` in `IDL`) and computing the diffusivity factor using

$$\mathcal{D} = -\frac{1}{\Delta\tau} \ln[(1 - \Delta\tau)\exp(-\Delta\tau) + (\Delta\tau)^2 \mathcal{E}_1]. \quad (22)$$

We label these calculations as "exact." The calculations labeled "approximate" were performed using the fitting formulae in Equation (21). We see that the error is better than 10^{-6}. By contrast, a 13th order polynomial fit to the exact solution incurs large errors ($\gtrsim 10^{-3}$).

Appendix B
Full Posterior Distributions for Best Models of HR 8799b, c, d, and e

For completeness, in Figures 18–21, we show the full posterior distributions for our best models of HR 8799b, c, d, and e, which elucidate the model degeneracies between each pair of parameters.

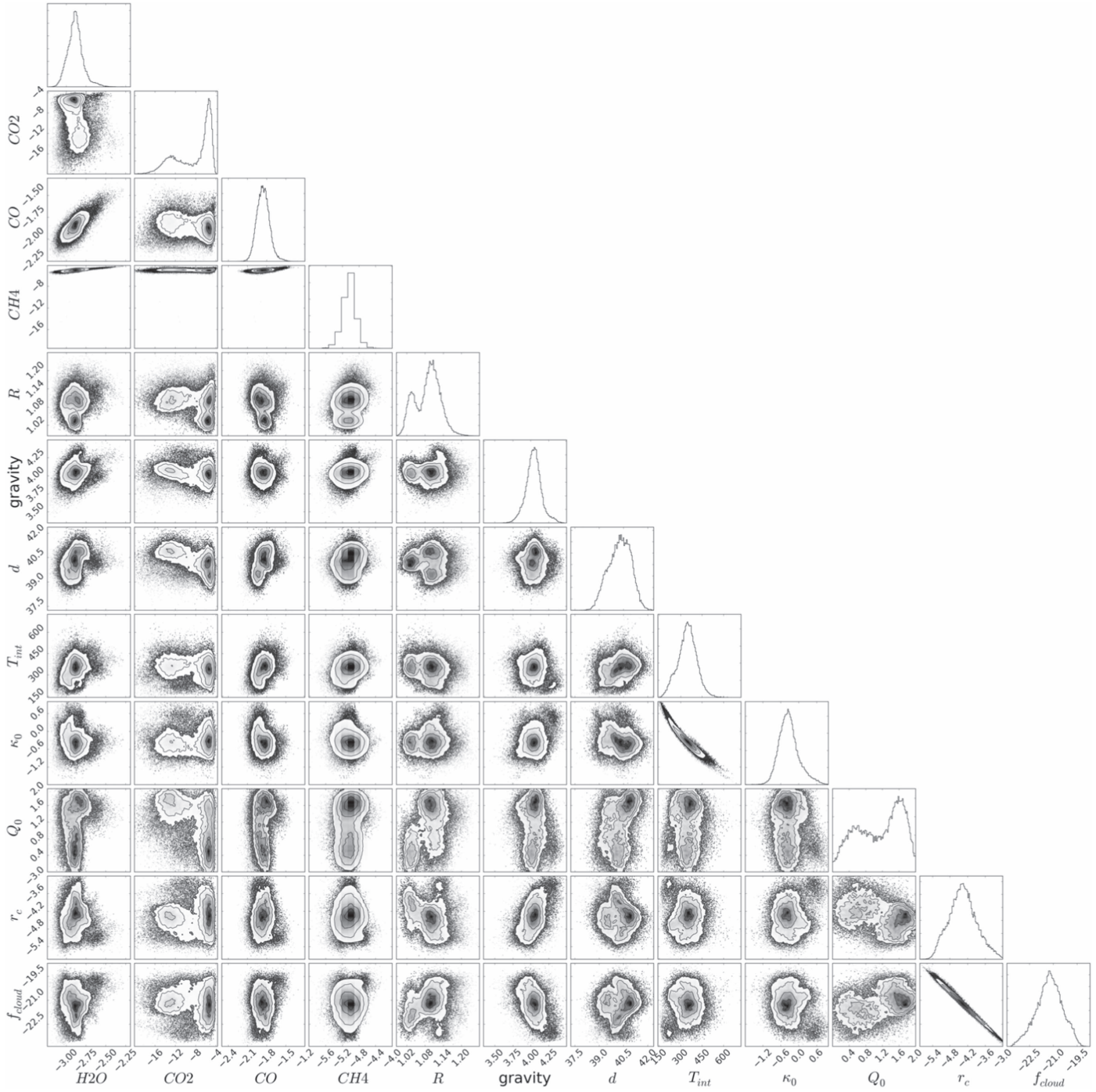

Figure 18. Montage of posterior distributions from the best-fit retrieval model of HR 8799b.

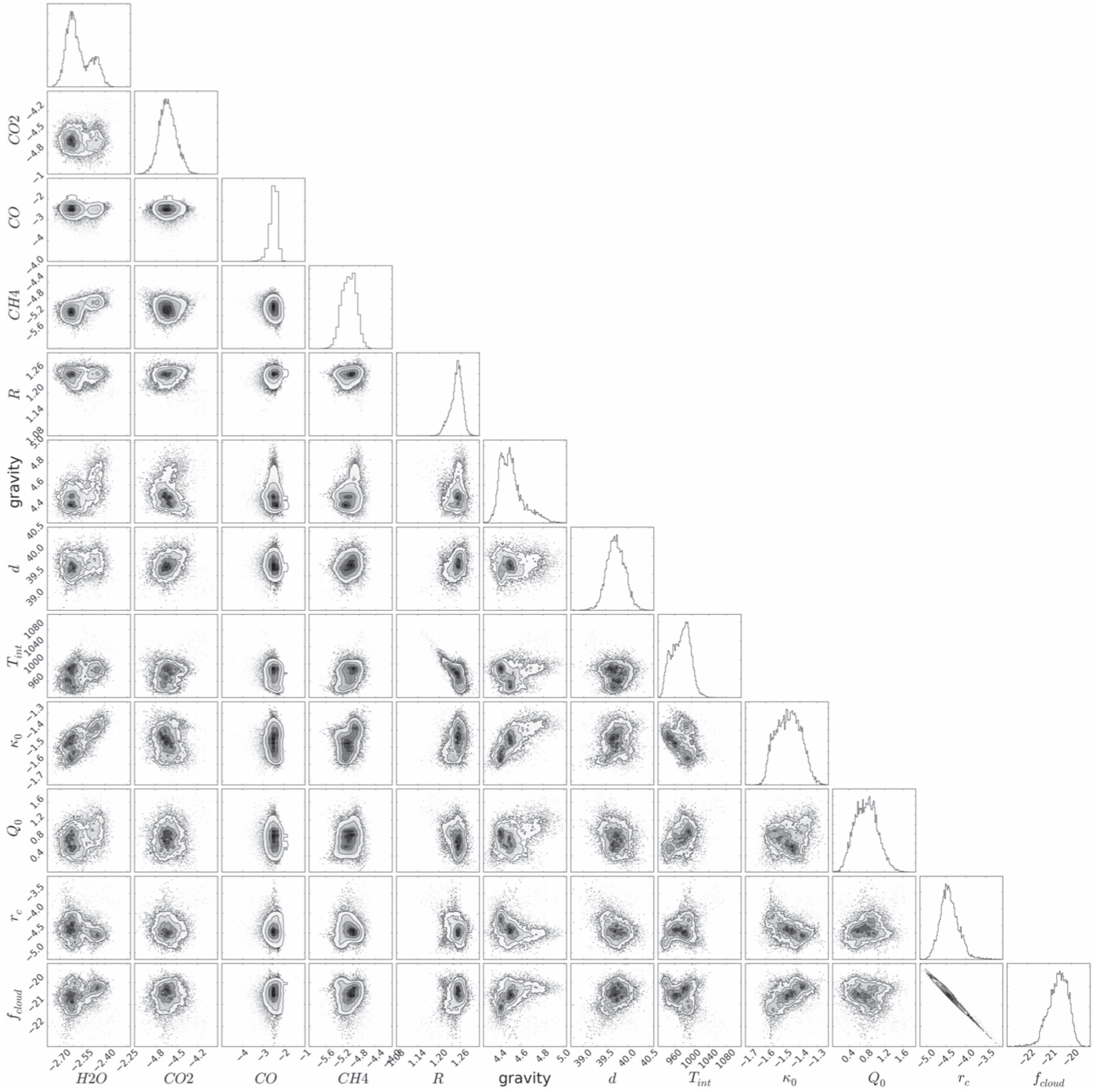

Figure 19. Same as Figure 18, but for HR 8799c.

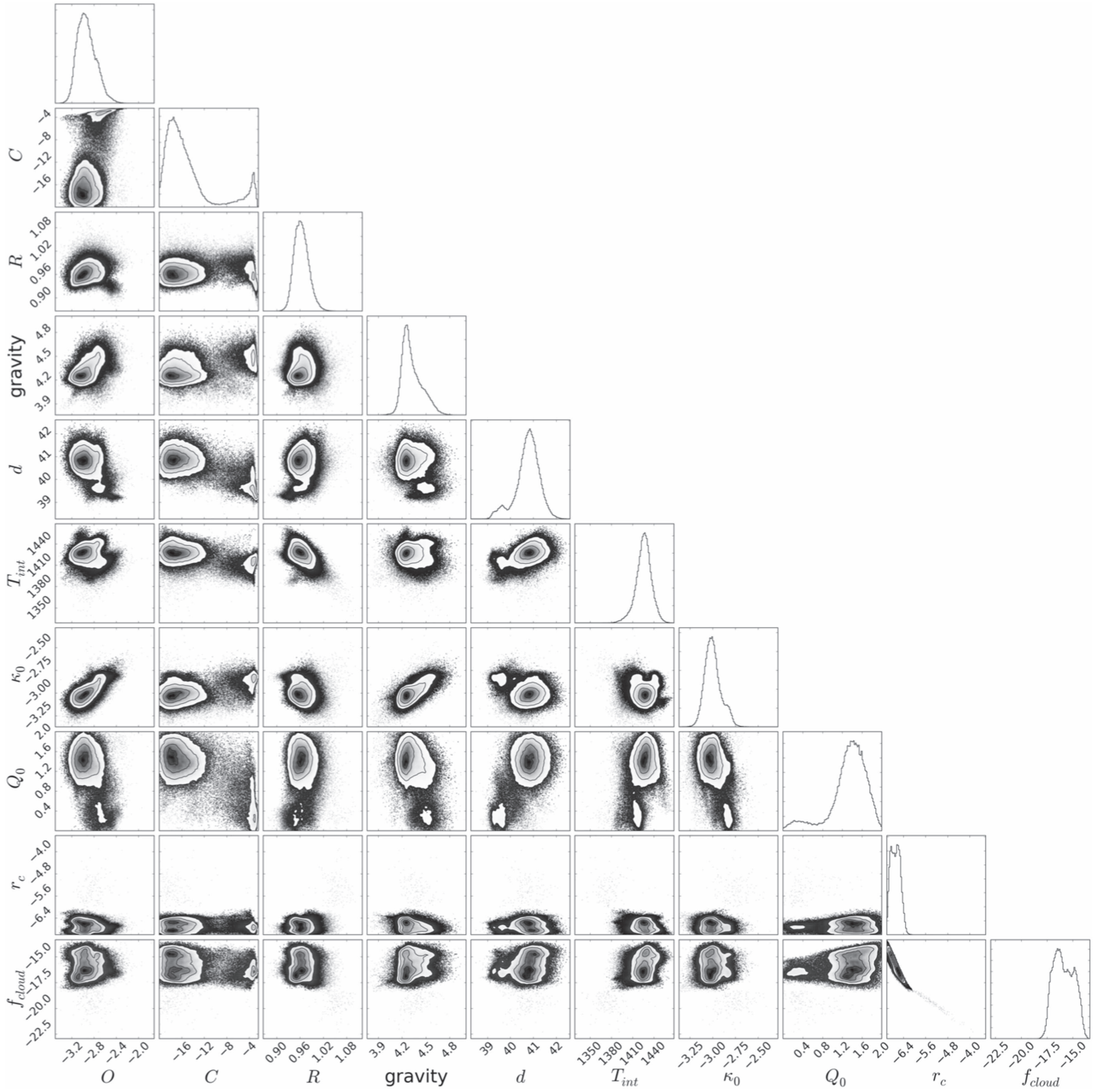

Figure 20. Same as Figure 18, but for HR 8799d.

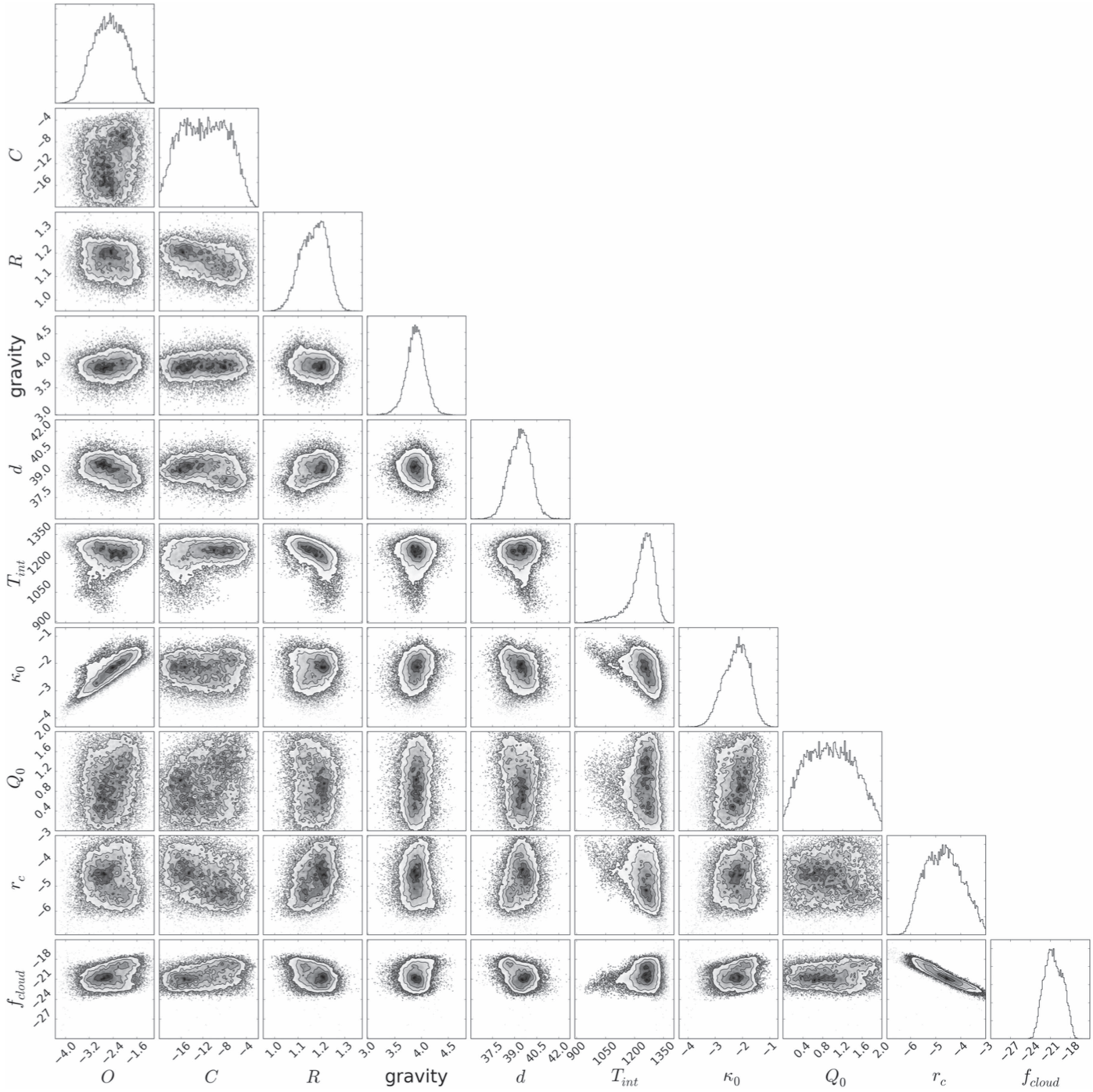

Figure 21. Same as Figure 18, but for HR 8799e.

ORCID iDs

Baptiste Lavie https://orcid.org/0000-0001-8884-9276
João M. Mendonça https://orcid.org/0000-0002-6907-4476
Christoph Mordasini https://orcid.org/0000-0002-1013-2811
Matej Malik https://orcid.org/0000-0002-2110-6694
Brice-Olivier Demory https://orcid.org/0000-0002-9355-5165
Simon L. Grimm https://orcid.org/0000-0002-0632-4407
David Ehrenreich https://orcid.org/0000-0001-9704-5405
Kevin Heng https://orcid.org/0000-0003-1907-5910

References

Abramowitz, M., & Stegun, I. A. 1970, Handbook of Mathematical Functions, 9th Printing (New York: Dover)
Apai, D., Kasper, M., Skemer, A., et al. 2016, ApJ, 820, 40
Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 2002, A&A, 382, 563
Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, MNRAS, 368, 1087
Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, ApJ, 804, 61

Barman, T. S., Macintosh, B., Konopacky, Q. M., & Marois, C. 2011, ApJ, 733, 65
Barstow, J. K., Aigrain, S., Irwin, P. G. J., Kendrew, S., & Fletcher, L. N. 2015, MNRAS, 448, 2546
Benneke, B., & Seager, S. 2013, ApJ, 778, 153
Bonnefoy, M., Zurlo, A., Baudino, J. L., et al. 2016, A&A, 587, A58
Bowler, B. P. 2016, PASP, 128, 102001
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Burrows, A., Heng, K., & Nampaisarn, T. 2011, ApJ, 736, 47
Burrows, A., & Liebert, J. 1993, RvMP, 65, 301
Burrows, A., Marley, M., Hubbard, W. B., et al. 1997, ApJ, 491, 856
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368
Cridland, A. J., Pudritz, R. E., & Alessi, M. 2016, MNRAS, 461, 3274
Deleuil, M., Deeg, H. J., Alonso, R., et al. 2008, A&A, 491, 889
Deming, D., Wilkins, A., McCullough, P., et al. 2013, ApJ, 774, 95
Demory, B.-O., & Seager, S. 2011, ApJS, 197, 12
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. 2013, arXiv:1306.2144
Foreman-Mackey, D., Vousden, W., Price-Whelan, A., et al. 2016 doi:10.5281/zenodo.53155
Fu, Q., & Liou, K. N. 1992, JAtS, 49, 2139
Gaidos, E. J. 2000, Icar, 145, 637
Grimm, S. L., & Heng, K. 2015, ApJ, 808, 182
Guillot, T. 2010, A&A, 520, A27
Han, E., Wang, S. X., Wright, J. T., et al. 2014, PASP, 943, 827
Hansen, C. J., Schwartz, J. C., & Cowan, N. B. 2014, MNRAS, 444, 3632
Hedges, C., & Madhusudhan, N. 2016, MNRAS, 458, 1427
Helling, Ch., & Woitke, P. 2006, A&A, 455, 325
Heng, K., & Demory, B.-O. 2013, ApJ, 777, 100
Heng, K., Hayek, W., Pont, F., & Sing, D. K. 2012, MNRAS, 420, 20
Heng, K., & Lyons, J. R. 2016, ApJ, 817, 149
Heng, K., Lyons, J. R., & Tsai, S.-M. 2016, ApJ, 816, 96
Heng, K., Mendonça, J. M., & Lee, J.-M. 2014, ApJS, 215, 4
Heng, K., & Tsai, S.-M. 2016, ApJ, 829, 104
Hiranaka, K., Cruz, K. L., Douglas, S. T., Marley, M. S., & Baldassare, V. F. 2016, arXiv:1606.09485
Ingraham, P., Marley, M. S., Saumon, D., et al. 2014, ApJL, 794, L15
Kass, R. E., & Raftery, A. E. 1995, JASA, 90, 773
Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, Sci, 339, 1398
Konopacky, Q. M., Ghez, A. M., Barman, T. S., et al. 2010, ApJ, 711, 1087
Kratter, K. M., Murray-Clay, R. A., & Youdin, A. N. 2010, ApJ, 710, 1375
Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, ApJL, 793, L27
Kuchner, M. J., & Seager, S. 2005, arXiv:astro-ph/0504214
Lacis, A. A., & Oinas, V. 1991, JGR, 96, 9027
Lee, J.-M., Heng, K., & Irwin, P. G. J. 2013, ApJ, 778, 97
Lillo-Box, J., Barrado, D., Santos, N. C., et al. 2015, A&A, 577, A105
Line, M. R., Stevenson, K. B., Bean, J., et al. 2016, arXiv:1605.08810
Line, M. R., Teske, J., Burningham, B., Fortney, J. J., & Marley, M. S. 2015, ApJ, 807, 183
Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, ApJ, 775, 137
Lodders, K. 2003, ApJ, 591, 1220
Madhusudhan, N. 2012, ApJ, 758, 36
Madhusudhan, N., Burrows, A., & Currie, T. 2011, ApJ, 737, 34
Madhusudhan, N., Crouzet, N., McCullough, P. R., Deming, D., & Hedges, C. 2014, ApJL, 791, L9
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Malik, M., Grosheintz, L., Mendonça, J. M., et al. 2017, AJ, 153, 56
Mandell, A. M., Haynes, K., Sinukoff, E., et al. 2013, ApJ, 779, 128
Marleau, G.-D., & Cumming, A. 2014, MNRAS, 437, 1378
Marley, M. S., Saumon, D., Cushing, M., et al. 2012, ApJ, 754, 135
Marois, C., Macintosh, B., Barman, T., et al. 2008, Sci, 322, 1348
Marois, C., Zuckerman, B., Konopacky, Q. M., et al. 2010, Natur, 468, 1080
Mihalas, D. 1970, Stellar Atmospheres (San Francisco: Freeman)
Miller-Ricci, E., & Fortney, J. J. 2010, ApJL, 716, L74
Mordasini, C., Alibert, Y., Georgy, C., et al. 2012, A&A, 547, A112
Mordasini, C., van Boekel, R., Mollière, P., Henning, T., & Benneke, B. 2016, ApJ, 832, 41
Moses, J. I., Line, M. R., Visscher, C., et al. 2013, ApJ, 777, 34
Moses, J. I., Marley, M. S., Zahnle, K., et al. 2016, ApJ, 829, 66
Öberg, K. I., Murray-Clay, R., & Bergin, E. A. 2011, ApJL, 743, L16
Oppenheimer, B. R., Baranec, C., Beichman, C., et al. 2013, ApJ, 768, 24
Pierrehumbert, R. T. 2010, Principles of Planetary Climate (New York: Cambridge Univ. Press)
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, Numerical Recipes: The Art of Scientific Computing (3rd ed.; New York: Cambridge Univ. Press)
Richard, C., Gordon, I. E., Rothman, L. S., et al. 2012, JQSRT, 113, 1276
Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, JQSRT, 130, 4
Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, JQSRT, 111, 2139
Rothman, L. S., Rinsland, C. P., Goldman, A., et al. 1996, JQSRT, 60, 665
Sadakane, K. 2006, PASJ, 58, 1023
Skilling, J. 2006, BayAn, 1, 833
Smith, M. D. 1998, Icar, 132, 176
Spiegel, D. S., & Burrows, A. 2012, ApJ, 745, 174
Spiegel, D. S., Silverio, K., & Burrows, A. 2009, ApJ, 699, 1487
Stevenson, K. B., Désert, J.-M., Line, M. R., et al. 2014, Sci, 346, 838
Tognelli, E., Prada Moroni, P. G., & Degl'Innocenti, S. 2007, A&A, 533, A109
Trotta, R. 2008, ConPh, 49, 71
Tsai, S.-M., Lyons, J. R., Grosheintz, L., et al. 2017, ApJS, 228, 20
van Leeuwen, F. 2007, A&A, 474, 653
Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, ApJ, 802, 107
Yurchenko, S. N., & Tennyson, J. 2014, MNRAS, 440, 1649
Zaghloul, M. R., & Ali, A. N. 2012, ACM Transactions on Mathematical Software, 38, 15
Zurlo, A., Vigan, A., Galicher, R., et al. 2016, A&A, 587, A57