Open access • Journal Article • DOI:10.1088/0004-637X/749/1/93

Information Content of Exoplanetary Transit Spectra: An Initial Look — Source link

Michael R. Line, Xi Zhang, G. Vaisht, Pin Chen ...+2 more authors

Published on: 10 Nov 2011 - arXiv: Earth and Planetary Astrophysics

Topics: Optimal estimation

The Astrophysical Journal, 749:93 (10pp), 2012 April 10 doi:10.1088/0004-637X/749/1/93

INFORMATION CONTENT OF EXOPLANETARY TRANSIT SPECTRA: AN INITIAL LOOK

Michael R. Line[1], Xi Zhang[1], Gautam Vasisht[2], Vijay Natraj[2], Pin Chen[2], and Yuk L. Yung[1]
[1] Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA; mrl@gps.caltech.edu
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
Received 2011 November 9; accepted 2012 February 10; published 2012 March 26

ABSTRACT

It has been shown that spectroscopy of transiting extrasolar planets can potentially provide a wealth of information about their atmospheres. Herein, we set up the *inverse* problem in spectroscopic retrieval. We use nonlinear optimal estimation to retrieve the atmospheric state (pioneered for Earth sounding by Rodgers). The formulation quantifies the degrees of freedom and information content of the spectrum with respect to geophysical parameters; herein, we focus specifically on temperature and composition. First, we apply the technique to synthetic near-infrared spectra and explore the influence of spectral signal-to-noise ratio and resolution (the two important parameters when designing a future instrument) on the information content of the data. As expected, we find that the number of retrievable parameters increases with increasing signal-to-noise ratio and resolution, although the gains quickly level off for large values. Second, we apply the methods to the previously studied dayside near-infrared emission spectrum of HD 189733b and compare the results of our retrieval with those obtained by others.

Key words: methods: data analysis – planets and satellites: atmospheres – planets and satellites: individual (HD189733b) – radiative transfer

Online-only material: color figures

1. INTRODUCTION

Currently there are about 130 confirmed transiting exoplanets (www.exoplanet.org). Of these planets, several dozen have spectra that have been observed, either through broadband photometry from instruments like the *Spitzer* Infrared Array Camera (IRAC; Deming et al. 2005; Charbonneau et al. 2005, 2008; Knutson et al. 2007, 2008; Harrington et al. 2006, 2007; Stevenson et al. 2010) or higher resolution spectroscopy from the *Hubble Space Telescope* (*HST*) Near Infrared Camera and Multi-Object Spectrometer (NICMOS; Swain et al. 2009a, 2009b), *Spitzer* Infrared Spectrometer (IRS; Grillmair et al. 2008), and recently, from ground-based instruments (Redfield et al. 2007; Snellen et al. 2008, 2010; Swain et al. 2010; Mandell et al. 2011; Waldmann et al. 2012). Although the spectra are of low resolution ($R = \lambda/\Delta\lambda \sim 5$–$50$) and low signal to noise ($S/N \leqslant 10$), they nevertheless provide useful information about the temperature and composition of the exoplanetary atmospheres (Tinetti et al. 2007, 2010a; Madhusudhan & Seager 2009, etc.). A typical approach to retrieving this information is to match the data set with forward models by manually tuning the model abundances and temperatures, until a possible best fit is obtained (Tinetti et al. 2007; Swain et al. 2009a, 2009b). This approach does not provide an optimal solution to the atmospheric state; furthermore, it can be cumbersome and is susceptible to multiple degeneracies (Tinetti et al. 2007; Madhusudhan & Seager 2009)

Others have used multi-dimensional grid models to constrain atmospheric parameters (Madhusudhan & Seager 2009), a method that is well tuned to systematically searching the parameter space given sparse data (as with *Spitzer*/IRAC color photometry). In this approach, an ensemble of forward models are generated using up to 10 gridded free parameters (6 to govern the shape of the temperature profile and 4 scaling factors for uniform mixing ratios of H_2O, CH_4, CO, and CO_2); model families that best describe the data are selected based on a χ^2 statistic criterion. Because of the degeneracies between the different gases, and between gases and temperature, thousands of solutions can exist within a given χ^2 region, thus only giving loose constraints on the atmospheric composition and temperature. Furthermore, the formalism provides no easy way to explore the change in information content associated with a change in the data phase space (e.g., R or S/N).

Here, we present the *inverse* approach (see also Lee et al. 2012) that determines the atmospheric "state" (i.e., its temperature structure and abundances) by minimizing a cost function that simultaneously takes into account new measurements and prior knowledge of atmospheric properties (such as a state retrieved from previous observations). Additionally we determine, within the context of our model, the quality of the spectra and the number of useful retrievable atmospheric properties. This work represents the first attempt at determining the amount of useful information that can be retrieved from typical exoplanet spectra. Furthermore, this paper represents the first attempt at using information theoretic limits for retrievals assuming certain instrument capabilities (such as R and S/N). Ultimately, the theory is general and enables prediction of the advances that can be made with improvements in instrumentation and via more prudent choosing of spectral ranges.

In Section 2, we outline the basics of the classic retrieval theory of Rodgers (2000). We first test the technique on an artificial data set and explore how the number of retrievable parameters depends on R and S/N and discuss how these can be optimized to maximize the usefulness of a measurement in Section 3. We then apply these techniques to the well-studied HD189733b dayside emission spectra in Section 4. This is followed by a discussion and conclusions in Section 5.

2. METHOD

2.1. Retrieval Theory

The retrieval problem is well known in the field of Earth atmospheric studies (Rodgers 1976; Chahine 1968; Twomey et al. 1977) and in studies of planetary atmospheres (see, e.g.,

Nixon et al. 2007). The fundamental problem is to determine the state vector, $\mathbf{x}$ of dimension n, often a vector of temperatures and mixing ratios at different altitudes (but could be other desirable variables), given some set of observations, $\mathbf{y}$ of dimension m, usually a vector of flux values at each wavelength. In the absence of any noise, they can be related through $\mathbf{y} = \mathbf{F}(\mathbf{x})$, where $\mathbf{F}(\mathbf{x})$ is a model that simulates the measurement at each wavelength given a representative atmosphere. In an idealized scenario, if the relationship between $\mathbf{x}$ and $\mathbf{y}$ is linear, we can linearize $\mathbf{F}(\mathbf{x})$ and write

$$\mathbf{y} = \mathbf{F}(\mathbf{x}_a) + \mathbf{K}(\mathbf{x} - \mathbf{x}_a), \tag{1}$$

where $\mathbf{K}$ is the $m \times n$ Jacobian matrix whose elements are given by the Frechet derivative

$$K_{ij} = \frac{\partial F_i(\mathbf{x})}{\partial x_j}, \tag{2}$$

with F_i being the measurement in the ith channel, and x_j the value of the jth parameter. The vector $\mathbf{x}_a$ is the prior (a priori) state, which represents our best initial guess of the true state before the observations are made. The Jacobian describes the sensitivity of the measurement at each wavelength in a spectrum to a perturbation of a given parameter in the forward model. If the lengths of $\mathbf{x}$ and $\mathbf{y}$ are the same then Equation (1) may be readily inverted to

$$\mathbf{x} = \mathbf{x}_a + \mathbf{K}^{-1}(\mathbf{y} - \mathbf{F}(\mathbf{x}_a)). \tag{3}$$

Real data are often noisy and usually have a large number of measurements that overconstrain the atmospheric state. For this we must use a more sophisticated scheme to invert the data to determine the atmospheric properties. This can be readily achieved by using a Bayesian framework. In the remainder of this section, we present the basic formalism and useful equations and algorithms that we can use to retrieve atmospheric properties from spectra as well as their information content, following the derivations in Rodgers (2000). For further details, see either Rodgers (2000) or Jacob (2007).

Bayes theorem can be written as

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x}), \tag{4}$$

where $P(\mathbf{x})$ is the prior probability distribution, which is knowledge of the atmospheric state before making a measurement, $P(\mathbf{y}|\mathbf{x})$ is the likelihood function, that is, the probability that the data exists within the context of a particular model, and $P(\mathbf{x}|\mathbf{y})$ is the posterior probability distribution density function which can be interpreted as the probability that some state $\mathbf{x}$, in our case atmospheric state, exists given the observations, $\mathbf{y}$. If we assume Gaussian probability distributions for the observational error and for the a priori information, we can write

$$P(\mathbf{y}|\mathbf{x}) \propto e^{-\frac{1}{2}(\mathbf{y}-\mathbf{K}\mathbf{x})^T\mathbf{S}_e^{-1}(\mathbf{y}-\mathbf{K}\mathbf{x})} \tag{5}$$

$$P(\mathbf{x}) \propto e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x}-\mathbf{x}_a)}, \tag{6}$$

where $\mathbf{S}_e$ is the $m \times m$ diagonal error covariance matrix (assuming no correlation between measurements) and $\mathbf{S}_a$ is the $n \times n$ a priori covariance matrix. The a priori covariance matrix represents our prior knowledge of the natural variability of the system and like $\mathbf{S}_e$, it is assumed to be diagonal. It essentially defines our "trust" region, or how far from the prior state we think the actual state can exist. In general, the prior constraint should be loose enough to allow flexibility in the retrieval but not so loose that the retrieval fails when a measurement contributes no information.

Using Bayes theorem from Equation (4) we can write the posterior probability distribution as a product of Equations (5) and (6):

$$P(\mathbf{x}|\mathbf{y}) \propto e^{-\frac{1}{2}J(\mathbf{x})}, \tag{7}$$

where $J(\mathbf{x})$ is the cost function and is given by

$$J(\mathbf{x}) = (\mathbf{y} - \mathbf{K}\mathbf{x})^T\mathbf{S}_e^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}) + (\mathbf{x} - \mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x} - \mathbf{x}_a). \tag{8}$$

The first term in the cost function represents the contribution from the data. The second term represents the contribution from the prior knowledge. If the data are of good quality (high-S/N, and high R), then the data term will dominate. Since the product of two Gaussians is a Gaussian, Equation (8) can be equivalently written as

$$J(\mathbf{x}) = (\mathbf{x} - \hat{\mathbf{x}})^T\hat{\mathbf{S}}^{-1}(\mathbf{x} - \hat{\mathbf{x}}), \tag{9}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{S}}$ are the mean and covariance, respectively, of the posterior probability distribution. A diagonal element of $\hat{\mathbf{S}}$ is the variance in the jth component of the state vector, $\hat{S}_{jj} = \hat{\sigma}_j^2$, where $\hat{\sigma_j}$ is the retrieval uncertainty in the jth parameter.

The goal of any retrieval is to obtain the most likely set of atmospheric parameters given the data. This is achieved when Equation (7) is maximized which occurs at the mean of the posterior probability function. Equating Equations (8) and (9) we can solve for $\hat{\mathbf{x}}$ and $\hat{\mathbf{S}}$ to get

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{G}(\mathbf{y} - \mathbf{K}\mathbf{x}), \tag{10}$$

where $\mathbf{G}$ is the gain matrix that describes the sensitivity of the retrieval to the observations (if $\mathbf{G} = 0$, no sensitivity, then the measurements do not contribute toward the retrieved state), given by

$$\mathbf{G} = \frac{\partial\hat{\mathbf{x}}}{\partial\mathbf{y}} = \hat{\mathbf{S}}\mathbf{K}^T\mathbf{S}_e^{-1}, \tag{11}$$

with

$$\hat{\mathbf{S}} = \left(\mathbf{K}^T\mathbf{S}_e^{-1}\mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}. \tag{12}$$

As the elements of $\mathbf{S}_a$ approach ∞ or the elements of $\mathbf{S}_e$ approach 0, then $\mathbf{G}$ approaches $\mathbf{K}^{-1}$ which is identically the sensitivity of the state vector to the observations, and thus the retrieval is fully characterized by the data.

If the forward model is linear, then Equation (10) can be solved to obtain the desired state vector. Often, the forward model is nonlinear, generally the case in radiative transfer; it is then best to use a numerical iteration scheme to determine the state vector. In the nonlinear case, the $\mathbf{K}\mathbf{x}$ terms in the cost function in Equation (8) are replaced with $\mathbf{F}(\mathbf{x})$. The Levenberg–Marquardt iteration scheme is used to find the minimum of the nonlinear cost function. The prescribed scheme is given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \left[(1+\gamma)\mathbf{S}_a^{-1} + \mathbf{K}_k^T\mathbf{S}_e^{-1}\mathbf{K}_k\right]^{-1} \left\{\mathbf{K}_k^T\mathbf{S}_e^{-1}[\mathbf{y} - \mathbf{F}(\mathbf{x}_k)] - \mathbf{S}_a^{-1}[\mathbf{x}_k - \mathbf{x}_a]\right\}, \tag{13}$$

where $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$ are the state vectors for the kth and $k+1$st iterations, and $\mathbf{K_k}$ is the Jacobian matrix calculated at the kth iteration. γ is a factor that controls the rate of convergence and

is adjusted at each iteration (Press et al. 1995). Equation (13) is iterated until convergence, when

$$(\mathbf{x}_k - \mathbf{x}_{k+1})^T \hat{\mathbf{S}}^{-1} (\mathbf{x}_k - \mathbf{x}_{k+1}) \ll n. \quad (14)$$

Upon convergence, we obtain the retrieved state, $\hat{\mathbf{x}}$ and its precision $\hat{\mathbf{S}}$.

2.2. Information Content and Degrees of Freedom

The information content (Shannon & Weaver 1962) and the total number of degrees of freedom are useful quantities that can help diagnose the quality and ability of a spectral data set to contribute to our knowledge of the atmospheric state. The number of degrees of freedom represents how many independent parameters can be retrieved from the spectrum, and the information content is a metric of how much the precision in the retrieved parameters has improved as a result of the observation. In the simplest sense, if there are m independent measurements with no error (e.g., fluxes at m different wavelengths), then there will be at most be m independent pieces of information (degrees of freedom) that can be obtained from the observations. If m is fewer than the number of model parameters, n, the exact values of $n-m$ parameters cannot be obtained from the observations. We do not discuss those cases in this article, we choose only cases for which $m > n$. For a given forward model, with n parameters, the maximum number of obtainable degrees of freedom will be the smaller of n and m. In an ideal case, the total number of degrees of freedom will be close to n, meaning that the observations can be fully characterized by those n parameters.

In reality, measurements are susceptible to error, and the total number of degrees of freedom in the observed signal (denoted by d_s), and thus the number of parameters accessible to our retrieval, may be fewer than the number of independent measurements, n. Some degrees of freedom, d_n, can be lost in the noise. The sum of d_s and d_n must add up to the total number of parameters we are seeking, n.

Before calculating the degrees of freedom it is useful to first introduce the averaging kernel, $\mathbf{A}$. The averaging kernel tells us which of the parameters in the state vector have the greatest impact on the retrieval, that is, the sensitivity of the retrieval to a given parameter, given by

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{GK}. \quad (15)$$

$\mathbf{A}$ is an $n \times n$ matrix whose elements are given by

$$A_{ij} = \frac{\partial \hat{x}_i}{\partial x_j}. \quad (16)$$

If a diagonal element of $\mathbf{A}$ is unity, or close to it, then that means for a given change in the true atmospheric state, there is identically the same change in the retrieved state. This suggests that the parameter, x_j, is fully characterized by the data. If that diagonal element is less than unity, meaning that the data itself are not of a high enough quality to constrain that parameter, then some fraction of the a priori information must have been used in determining the value of that parameter. If each parameter is fully characterized by the data, that is, if all of the diagonal elements of $\mathbf{A}$ are unity, then we would expect to be able to retrieve all n parameters. If the diagonal elements are less than unity, then the sum of the diagonals would be less than n. In essence, the diagonal elements of the averaging kernel can be thought of as the degrees of freedom per parameter. If the value of a particular diagonal element is 1, then that parameter is well characterized by the data. If it is much less than 1, then the data contribute little to our knowledge of that parameter. The total degrees of freedom from the signal can be determined by calculating the trace of $\mathbf{A}$. The difference between n and the trace of $\mathbf{A}$ is the number of degrees of freedom lost to the noise.

The total degrees of freedom, again, tell us how many independent parameters can be determined from the observations. The information content, H, tells us quantitatively how well the observations increased our confidence in our estimate of the atmospheric state relative to the a priori knowledge. In a more precise language, the information content of a measurement is the reduction in the entropy of the probability that an atmospheric state exists given some set of observations, or

$$H = \text{entropy}(P(\mathbf{x})) - \text{entropy}(P(\mathbf{x}|\mathbf{y})). \quad (17)$$

The entropy of a Gaussian distribution of width σ, which the prior and a posterior distributions are assumed to be, can be shown to be proportional to $\ln(\sigma)$. Using this fact and Equations (17), (6), and (9),

$$H = \frac{1}{2} \ln(|\hat{\mathbf{S}}^{-1} \mathbf{S}_a|). \quad (18)$$

From this we can see that if the data are good (small error bars), then the elements of $\hat{\mathbf{S}}$ will be small, resulting in a large H. Thus, H is a quantitative measure of the reduction in our uncertainty in the retrieved atmospheric state as a result of the observations. The larger the value of H, the more useful the observations are in constraining the atmospheric state.

In summary, both d_s and H are quantitative measures of the quality and usefulness of the observations in determining the atmospheric state, within the context of a given forward model. From their definitions we would expect that a spectrum with a higher S/N, or a higher R, would result in higher values. We will show this in Section 3.

2.3. Forward Model

A relatively simple forward model, $\mathbf{F}(\mathbf{x})$, which nonetheless captures the basic physics and the measurement process, is at the core of our retrieval. We assume a simplified understanding of the physical and chemical state of the exoplanet atmosphere, i.e., a parameterized temperature structure, the major volatile constituents, the important radiative processes, and the instrument line profiles, etc. Our forward model, as most such models, is an approximation because the data are of limited quality, the underlying physics is relatively ill-understood, and simplifying approximations are necessary. Examples of physics missing in our $\mathbf{F}(\mathbf{x})$ include absent species, inaccurate line lists, clouds, aerosols, three-dimensional effects, etc., or possibly insufficient parameterization of the atmosphere. Therefore, our retrievals must be taken in context of our chosen forward model. Herein, we only consider the dayside spectra of hot-Jupiters with near solar metallicity, though the methods are easily be extended to other kinds of observations (transmission spectra) and exoplanets (hot-Neptunes, mini-Neptunes, super-Earths, etc.) with relatively minor modifications to the forward model. For future instruments, with broader spectral coverage and higher spectral resolution, the forward models can increase in sophistication.

Lacking sufficient data (these are low signal-to-noise, low-resolution spectra), we simplify our atmosphere to eight parameters that characterize the temperature structure and gas concentrations. For the sake of simplicity, we use an analytic temperature profile formulated by Guillot (2010), and since then modified by V. Parmentier & T. Guillot (in preparation), to include three channels. The profile, derived using a three-channel approximation, is given by

$$T^4(\tau) = \frac{3T_{\rm int}^4}{4}\left(\frac{2}{3}+\tau\right) + \frac{3T_{\rm irr}^4}{4}(1-\alpha)\xi_{\gamma_1}(\tau) + \frac{3T_{\rm irr}^4}{4}\alpha\xi_{\gamma_2}(\tau), \tag{19}$$

where

$$\xi_{\gamma_i} = \frac{2}{3} + \frac{2}{3\gamma_i}\left[1+\left(\frac{\gamma_i\tau}{2}-1\right)e^{-\gamma_i\tau}\right] + \frac{2\gamma_i}{3}\left(1-\frac{\tau^2}{2}\right)\mathrm{E}_2(\gamma_i\tau), \tag{20}$$

with $\gamma_1 = \kappa_{v_1}/\kappa_{\rm IR}$ and $\gamma_2 = \kappa_{v_2}/\kappa_{\rm IR}$, where κ_{v_1}, κ_{v_2}, and $\kappa_{\rm IR}$ are the visible and infrared (thermal) opacities, respectively. The parameter α (range 1–0) partitions the flux between the two visible streams, and $E_2(\gamma\tau)$ is the second-order exponential integral function. The internal heat flux (from the net cooling history) is represented by the temperature $T_{\rm int}$, while the solar flux at the top of the atmosphere is represented by $T_{\rm irr}$; these two temperatures are fixed. Assuming zero albedo and unit emissivity, $T_{\rm irr}$ is

$$T_{\rm irr} = \left(\frac{R_*}{2a}\right)^{1/2} T_*, \tag{21}$$

where R_* and T_* are the stellar radius and temperature, a is the star planet separation, and τ is the infrared (thermal) optical depth:

$$\tau = \frac{\kappa_{\rm IR} P}{g}, \tag{22}$$

with P the pressure and g the surface gravity (at 1 bar). In total, there are four free parameters governing the temperature structure, $\kappa_{\rm IR}$, κ_{v_1}, κ_{v_2}, and α. We choose this parameterization with two visible streams as opposed to the traditional one visible stream (Hansen 2008; Guillot 2010) because the extra stream allows more freedom for a temperature inversion, though in some cases (as we shall see below) the second visible stream does not matter.

The remaining four parameters are the uniform mixing ratios for H_2O, CH_4, CO, CO_2, expected to be the major molecular opacity sources (Tinetti et al. 2007; Swain et al. 2009b). We choose vertically uniform mixing ratios for two reasons. First, the data lack sufficient information content to actually help resolve vertical structure in abundances, and second, chemical kinetics models (Moses et al. 2011; Line et al. 2010, 2011) show that vertical mixing leads to constant vertical mixing ratios for these species within the IR photosphere, so even if we could resolve detailed vertical information, we would most likely find that the abundances remain fairly constant.

Since many of these parameters may vary over many orders of magnitude we find it convenient with the above formalism to solve for the logarithm of the atmospheric state. With that, the state vector of parameters that we would like to retrieve can be given by

$$\mathbf{x} = \begin{bmatrix} \log(\kappa_{v_1}) \\ \log(\kappa_{v_1}) \\ \log(\kappa_{\rm IR}) \\ \alpha \\ \log(f_{\rm H2O}) \\ \log(f_{\rm CH4}) \\ \log(f_{\rm CO}) \\ \log(f_{\rm CO2}) \end{bmatrix},$$

where f_i is the mixing ratio of species i in parts per million (ppm) and the opacities are in cm^2 g^{-1}.

We also include H_2–H_2 and H_2–He collision-induced opacities. The mixing ratios of H_2 and He vary little with the atmospheric levels that produce the bulk of the dayside thermal emission (500–2000 K, 10–10^{-4} bar). We fix $f_{\rm H_2}$ and $f_{\rm He}$ to thermochemical abundances (assuming solar elemental abundances) of 0.86 and 0.14, respectively. These values may change on the tens-of-percent level in enriched atmospheres, however, this variation has negligible effect on the resultant infrared spectra. Also, we do not include NH_3 as an opacity source as it has little influence in the spectral region we consider.

We use the Reference Forward Model (RFM),[3] a line-by-line radiative transfer code, to calculate the disk-integrated dayside emission spectra, modified to handle H_2-H_2 and H_2–He collisionally induced opacities. The collisionally induced opacity tables are taken from Borysow et al. (2001), Borysow (2002), and Jørgensen et al. (2000). The molecular line strengths for H_2O, CO_2, and CO, are from the HITEMP (Rothman et al. 2010) database and CH_4[4] is from the HITRAN 2008 database (Rothman et al. 2009). In order to keep the molecular line lists from becoming too unwieldy we make an intensity cutoff at 298 K of 10^{-40} cm molecule^{-1}, as recommended by Sharp & Burrows (2007).

3. TEST ON SYNTHETIC DATA

First, we test the retrieval method on a synthetic data set for which we know the answer. Using this synthetic spectrum, we explore the effect that signal-to-noise ratio and spectral resolution have on the degrees of freedom and information content.

A hypothetical hot-Jupiter atmosphere is generated using $\kappa_{v_1} = \kappa_{v_2} = 4\times10^{-3}$ cm^2 g^{-1}, $\kappa_{\rm IR} = 1\times10^{-2}$ cm^2 g^{-1}, $\alpha = 0.5$, and fixed vertical mixing ratios of $f_{\rm H2O} = 5\times10^{-4}$, $f_{\rm CH4} = 1\times10^{-6}$, $f_{\rm CO} = 3\times10^{-4}$, and $f_{\rm CO_2} = 1\times10^{-7}$. The planet orbits around a G0V host star (e.g., HD 209458a) with $T_* = 6000$ K, $R_* = 1.14\,R_\odot$ at a separation of $a = 0.064$ AU. The planetary properties are a radius of $1.35R_J$, an internal temperature of $T_{\rm int} = 200$ K, and $g = 21.1$ m s^{-2} (at 1 bar pressure). Using Equation (21) we find $T_{\rm irr} = 1223$ K. The emission spectrum of the exoplanet (see Figure 1) is initially generated with a one wavenumber resolution (resolving power, $R \simeq 5000$ at 2 μm).

For the initial test, the synthetic spectrum (Figure 1) is degraded by convolving it with an instrumental profile matching the defocused *HST*/NIC3 camera with a spectral FWHM of

[3] see http://www.atm.ox.ac.uk/RFM/

[4] Upon completion of our initial investigation it was also brought to light that there exists more appropriate high temperature based line lists for methane such as the STDS (http://icb.u-bourgogne.fr/OMR/SMA/SHTDS/HTDS.html). Using this line list over HITRAN makes absolutely no difference for our synthetic work since the synthetic data were produced using the HITRAN methane. We have also compared our HD189733b retrieval results for both methane line lists and found no difference.

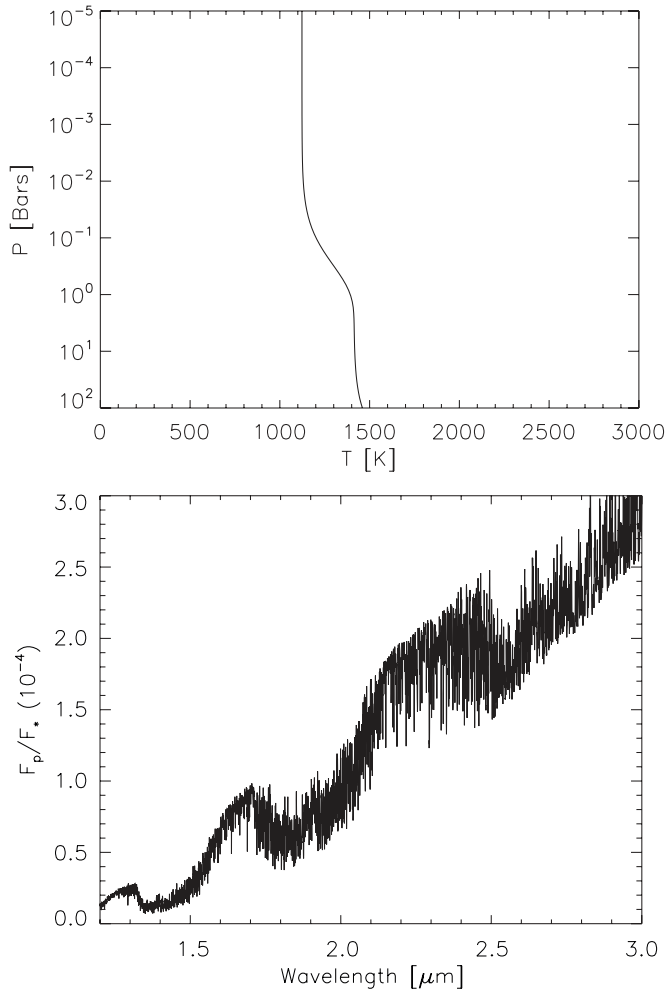

Figure 1. Synthetic spectrum (bottom) generated with the model atmosphere (top) with a spectral resolution of 1 cm^{-1} or $R \sim 5000$ at 2 μm. The model temperature profile is generated from Equations (19) and (20) with $\kappa_{v_1} = \kappa_{v_2} = 4 \times 10^{-3}$ cm^2 g^{-1}, $\kappa_{IR} = 1 \times 10^{-2}$ cm^2 g^{-1}, $\alpha = 0.5$, $T_{irr} = 1223$ K, and $T_{int} = 200$ K. The constant-with-altitude mixing ratios are $f_{H2O} = 5 \times 10^{-4}$, $f_{CH4} = 1 \times 10^{-6}$, $f_{CO} = 3 \times 10^{-4}$, and $f_{CO2} = 1 \times 10^{-7}$.

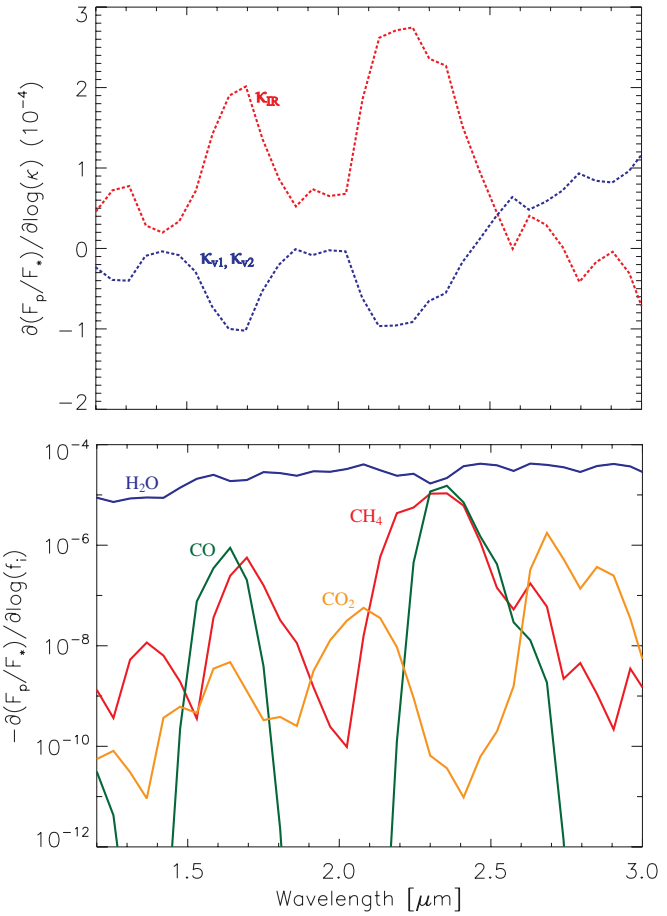

Figure 2. Columns of the Jacobian for the synthetic spectrum evaluated at the true state. This is the response of the flux as a function of wavelength due to a small positive perturbation in one of the parameters in **x**. The top panel is the flux response for the parameters that govern the temperature profile, κ_{v_1}, κ_{v_2}, and κ_{IR}. The bottom panel is the flux response to a small perturbation in the gas mixing ratios, f_{H2O}, f_{CH4}, f_{CO}, and f_{CO2}. The Jacobian is calculated as a change in the planet-to-star flux ratio, $\Delta(F_p/F_*)$, to a positive logarithmic perturbation in a given parameter, $\Delta \log(x_j)$. Note that in the bottom panel an increase in the gas mixing ratios always results in a decrease in F_p/F_*. In this particular case, the spectrum is equally sensitive to κ_{v_2} and κ_{v_1} because α is 0.5. If $\alpha = 0$ then the spectrum will have no sensitivity to κ_{v_2}, and if $\alpha = 1$ the spectrum will have no sensitivity to κ_{v_1}. Also, for this synthetic data set $\kappa_{v_2} = \kappa_{v_1}$, which results in no sensitivity to α.

(A color version of this figure is available in the online journal.)

0.055 μm ($R \simeq 40$ at 2 μm; Swain et al. 2009b), and reducing the measurement signal-to-noise ratio of each spectral channel to ~10. Rather than be guided by physical and chemical models, or some previous observation of the object, we arbitrarily chose an a priori state, $\mathbf{x}_a$, far from the true physical state. The remaining unspecified quantity is the a priori covariance matrix, $\mathbf{S}_a$. Once more, the diagonal elements of $\mathbf{S}_a$ are allowed a large range as we are dealing with a relatively novel type of observations and lack detailed prior information. We also assume that there are no cross correlations between different state parameters (e.g., f_{CO} and f_{CO_2}, even though from chemical models we know that such quantities have high correlations). Because the state parameters are logarithmic, the elements of $\mathbf{S}_a$ are also logarithmic (with the exception of α) so we set, somewhat arbitrarily, $\sigma_{\kappa_{v1}} = 2$, $\sigma_{\kappa_{v2}} = 2$, $\sigma_{\kappa_{IR}} = 2$, $\sigma_{\alpha} = 0.5$, $\sigma_{f_{H2O}} = 6$, $\sigma_{f_{CH4}} = 6$, $\sigma_{f_{CO}} = 6$, and $\sigma_{f_{CO2}} = 6$ meaning that the opacities are permitted to span 4 orders of magnitude centered around their a priori value and the mixing ratios are allowed to span 12 orders of magnitude. Such large a priori uncertainties lead to a flat a priori distribution, relative to the data, reducing the current problem to a maximum likelihood estimation (as opposed to Bayesian), with the option of using the a priori information if the data are sparse.

The entirety of the forward model can be summarized with the Jacobian. Figure 2 shows the columns of the Jacobian evaluated at the true state (response of the flux in each channel to a perturbation in each of the parameters in **x**) for the synthetic data (Figure 3). The spectrum is most sensitive to perturbations in the opacities that govern the temperature profile. The 1.7 μm and 2.2 μm channels are most sensitive to changes in the temperature profile. This is because there are not large absorption features at these wavelengths, meaning these channels are most sensitive to the flux from deeper layers (1–10 bar). This also partially explains why κ_{IR} and κ_{v1} have opposite responses. An increase in κ_{IR} results in an increase in flux due to an increase in temperature in the deep layers probed by these channels, as can be seen in Equation (19). An increase in κ_{v1} results in a decrease of flux in these channels due to a decrease in temperature in the deeper layers. From Equation (19), an increase in κ_{v1} increases the temperature above the ~0.1 bar level, and in order to maintain radiative equilibrium at the top of the atmosphere, a decrease in

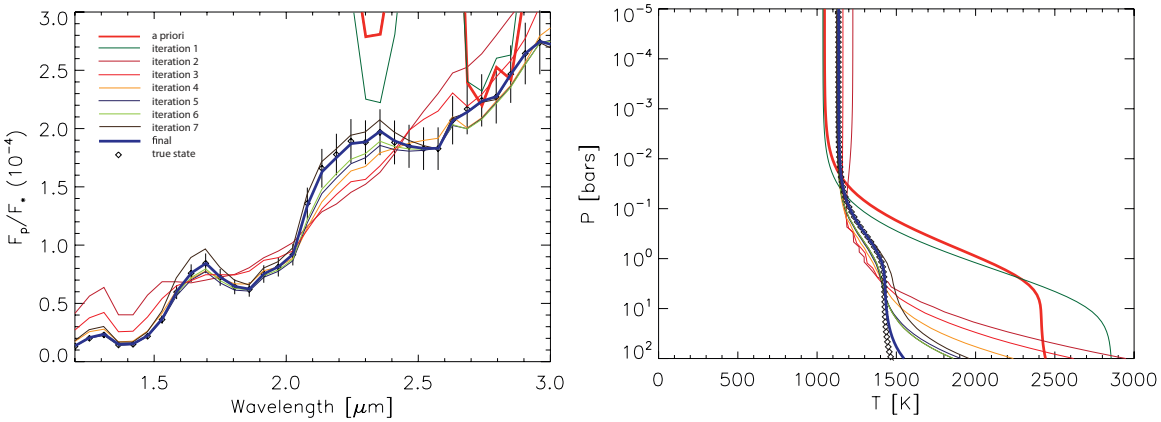

Figure 3. Synthetic spectrum retrieval. Left: iteration sequence of the model spectrum, $\mathbf{F}(\mathbf{x}_k)$. The diamonds with error bars are the synthetic data convolved down to a resolution of 0.055 μm ($R \sim 37$ at 2 μm) and a signal-to-noise ratio of 10. The thick red curve is the forward model spectrum generated from the a priori, $\mathbf{F}(\mathbf{x}_a)$. Note that it is a poor fit to the data. Each subsequent curve is the new model spectrum after each iteration of Equation (13). The thick solid blue curve is the final retrieved model spectrum. Right: evolution of the temperature profile with each iteration. The thick red curve is the a priori temperature profile. The thick blue curve is the retrieved temperature profile. The diamond symbol curve is the true temperature profile as in Figure 1. χ^2 converges to 0.007 after eight iterations of Equation (13).

(A color version of this figure is available in the online journal.)

temperature in the deeper layers must occur, and also a higher κ_v prevents the stellar flux from penetrating into the deeper atmosphere. The opposite is true near 2.9 μm, which is more sensitive to higher altitudes because of the large absorption; thus an increase in κ_{v1} will result in an increase in temperature, which in turn results in a flux increase. Also, in this particular case, $\alpha = 0.5$, meaning both κ_{v1} and κ_{v2} have identically the same results. Additionally, $\kappa_{v1} = \kappa_{v2}$, which causes the spectrum to have no sensitivity to changes in α.

The spectral response is most sensitive to the water abundance, more than any other gas across all wavelengths in this example (Figure 2). This makes the retrieval of water more precise than the other species. The greatest sensitivity to changes in the CO_2 abundance occurs at 2.1 and 2.8 μm, which both happen to be located near the sensitivity minima of CO and CH_4, though it still has to contend with water. Both CO and CH_4 have greatest sensitivity in the 2.3 μm band, making it difficult to simultaneously retrieve both.

Figure 3 shows the retrieval process for this initial synthetic test case. We determine the quality of the retrieval using the standard reduced χ^2 given by

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - F_i)^2}{\sigma_i^2}, \quad (23)$$

where N is the total number of data points, y_i, F_i, and σ_i, are defined in Section 2.1. If χ^2 is less than one, then the difference between the model fit and data is typically better than 1σ. We should stress, however, that a perfect fit ($\chi^2 = 0$) does not necessarily mean that the true state has been retrieved because of the degeneracies between some of the parameters. Table 1 compares the true state to the retrieval results along with the retrieval precision. The synthetic retrieval demonstrates the robustness of the retrieval to a poor a priori. The reason for this can be seen by inspecting the elements of the averaging kernel. From Table 1, all but κ_{v1} and methane are fairly well characterized by the data (A_{jj} is close to 1). Summing these values gives the total degrees of freedom, and thus the total number of useful retrievable parameters of ~6.

3.1. Resolution and Signal-to-noise Effects on the Degrees of Freedom and Information Content

The S/N and R are two important factors that influence the quality and usefulness of a spectrum. It is thus imperative to consider them when designing a spectrometer. In this section, we use our synthetic data set to explore how the degrees of freedom, both total and per atmospheric parameter, and the information content evolve with increasing S/N and R.

We would intuitively expect d_s and H both to increase with increasing R and S/N. Figure 4 shows a contour plot of d_s and H calculated for the synthetic spectrum generated in Figure 1

Table 1
Synthetic Retrieval Results

Parameter	True State ($\mathbf{x}$)	A priori ($\mathbf{x}_a$)	Retrieved State ($\hat{\mathbf{x}}$)	Retrieval Precision	$\frac{\partial \hat{x}_i}{\partial x_j}$
κ_{v1}	4.00×10^{-3}	1.00×10^{-3}	3.59×10^{-3}	2.76×10^{-3}–4.68×10^{-3}	0.997
κ_{v2}	4.00×10^{-3}	1.00×10^{-2}	1.70×10^{-9}	1.70×10^{-11}–1.70×10^{-7}	0.0
κ_{IR}	1.00×10^{-2}	3.16×10^{-2}	8.93×10^{-3}	7.13×10^{-3}–1.12×10^{-2}	0.998
α	0.5	0.1	0.003	0.00–0.022	0.999
f_{H2O}	5.00×10^{-4}	1.00×10^{-6}	4.18×10^{-4}	2.58×10^{-4}–6.76×10^{-4}	0.999
f_{CH4}	1.00×10^{-6}	1.00×10^{-4}	3.43×10^{-7}	4.34×10^{-12}–2.70×10^{-2}	0.334
f_{CO}	3.00×10^{-4}	1.00×10^{-6}	1.96×10^{-4}	2.27×10^{-6}–1.69×10^{-2}	0.896
f_{CO2}	1.00×10^{-7}	1.00×10^{-4}	7.70×10^{-7}	9.95×10^{-10}–5.96×10^{-4}	0.768

Notes. κ_{v1}, κ_{v2}, and κ_{IR} are in units of cm^2 g^{-1}. f_i is the volume mixing ratio for species i. We also show the diagonal averaging kernel elements ($A_{jj} = (\partial \hat{x}_j / \partial x_j)$) for each parameter. The retrieval uncertainties are given as $\hat{x} - \hat{\sigma}$ to $\hat{x} + \hat{\sigma}$ for each parameter.

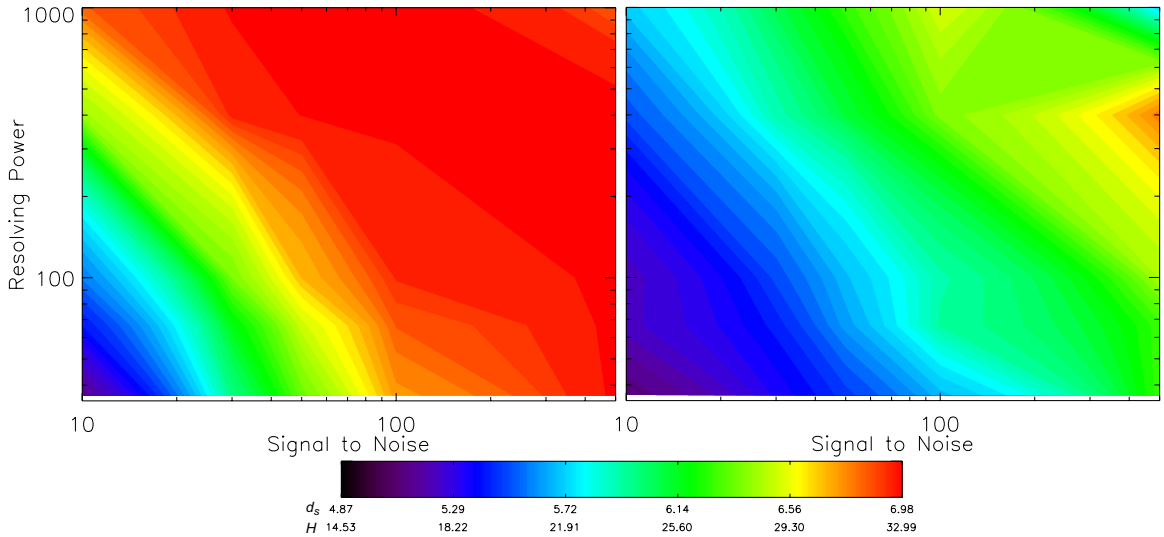

Figure 4. S/N and R effects on the total degrees of freedom (left) and the information content (right). In general, as S/N and R increase, the total number of degrees of freedom obtainable from the data and the information content increase. See Equations (24) and (25).

(A color version of this figure is available in the online journal.)

for a variety of S/Ns and Rs. The maximum increase in both occurs with a simultaneous[5] increase in S/N and R.

We point out that the contour plots in Figure 4 can only be taken in the context of the spectral window within which we are applying the retrieval and the number of parameters we are trying to retrieve. In other words, for the eight parameters we are retrieving here, there is no benefit to increasing R or S/N beyond a few hundred and ∼100, respectively. If we do happen to have a higher R and S/N, it is likely that we would be able to retrieve more forward model parameters such as the concentrations of other gases, or information on the vertical distributions of the gases. Current observations, like the *HST*/NICMOS observations of HD189733b, generally fall toward the bottom left corners in Figure 4. This suggests that S/N and Rs of such data are not high enough to fully constrain even our simple forward model, and thus even less constraining for more complicated models.

The increasing behavior in d_s with increasing S/N can be seen through the use of Equations (11), (12), and (15). As S/N goes to infinity, the elements of $\mathbf{S}_e$ go to zero, causing $\mathbf{G}$ to approach $\mathbf{K}^{-1}$, in turn causing $\mathbf{A}$ to approach the identity matrix, meaning the diagonal elements are all ones with a trace equal to the total number of parameters and thus the maximum number of degrees of freedom. The relationship between d_s and S/N can be seen in a one-parameter one-channel model, where $d_s = A$. Upon reducing the matrix equations, the one-element averaging kernel becomes

$$d_s = A = \frac{K^2\sigma_a^2}{K^2\sigma_a^2 + (F/(\mathrm{S/N}))^2} = \frac{(\mathrm{S/N})^2}{(\mathrm{S/N})^2 + \frac{F^2}{K^2\sigma_a^2}}, \quad (24)$$

and the relation of these parameters to the information content is

$$H = \ln\left[1 + \frac{\sigma_a^2}{F^2}K^2(\mathrm{S/N})^2\right], \quad (25)$$

where K, σ_a, and F are the one-dimensional analogs for $\mathbf{K}$, $\mathbf{S}_a$, and $\mathbf{F}(\mathbf{x})$, respectively. We also have assumed that σ_e, the one-dimensional analog for $\mathbf{S}_e$, is the flux, F, divided by S/N. In this case, d_s approaches unity as S/N goes to infinity, and zero, if S/N is zero. H approaches infinity as S/N goes to infinity, and approaches zero when S/N goes to zero. One important thing to note from these relations is that increasing S/N will matter only if the Jacobian, K, is non-zero, meaning that there must be some sensitivity of the flux to a perturbation in the desired parameter. Otherwise, no amount of S/N increase will improve our knowledge of the atmospheric state. Increasing R or adding more spectral channels can also contribute to an increase in d_s and H. If channels are chosen such that the K is large, meaning large sensitivity to a given parameter, then d_s and H will both increase. As K approaches infinity (infinite sensitivity), d_s will approach unity and H will approach infinity.

From this simple analysis, though it may intuitively obvious, we can readily see that if we want to improve the characterization of a particular atmospheric property, it is best to design an instrument whose spectral regions offer the greatest sensitivity to that parameter, and to have a high S/N within those spectral regions.

4. TEST ON REAL DATA: HD189733b DAYSIDE EMISSION

Now that we have demonstrated that this retrieval procedure works and provides useful information about the quality of a data set through the degrees of freedom and information content, we wish to apply it to the dayside emission spectra of one of the best-studied exoplanet atmospheres, HD189733b. We assume the same forward model and a priori covariances as in the synthetic work.

The dayside emission spectrum of HD189733b has been subject to much investigation (Swain et al. 2009b; Grillmair et al. 2007; Madhusudhan & Seager 2009, and many others), and often times different analyses come up with different solutions for its composition and temperature structure. For simplicity, we investigate only the near-IR spectrum from Swain et al. (2009b). As an a priori atmospheric state we use the "Fortney 2π" (Fortney et al. 2010) temperature profile from Figure 2 of Moses et al. (2011) approximated with Equation (19) and the 0.1 bar mixing ratios for H_2O, CH_4, CO, and CO_2 from their Table 2 but assumed to be constant with altitude within

[5] This is true if R and S/N are independent of each other. In most cases, S/N decreases with increasing R because of the smaller spectral bins.

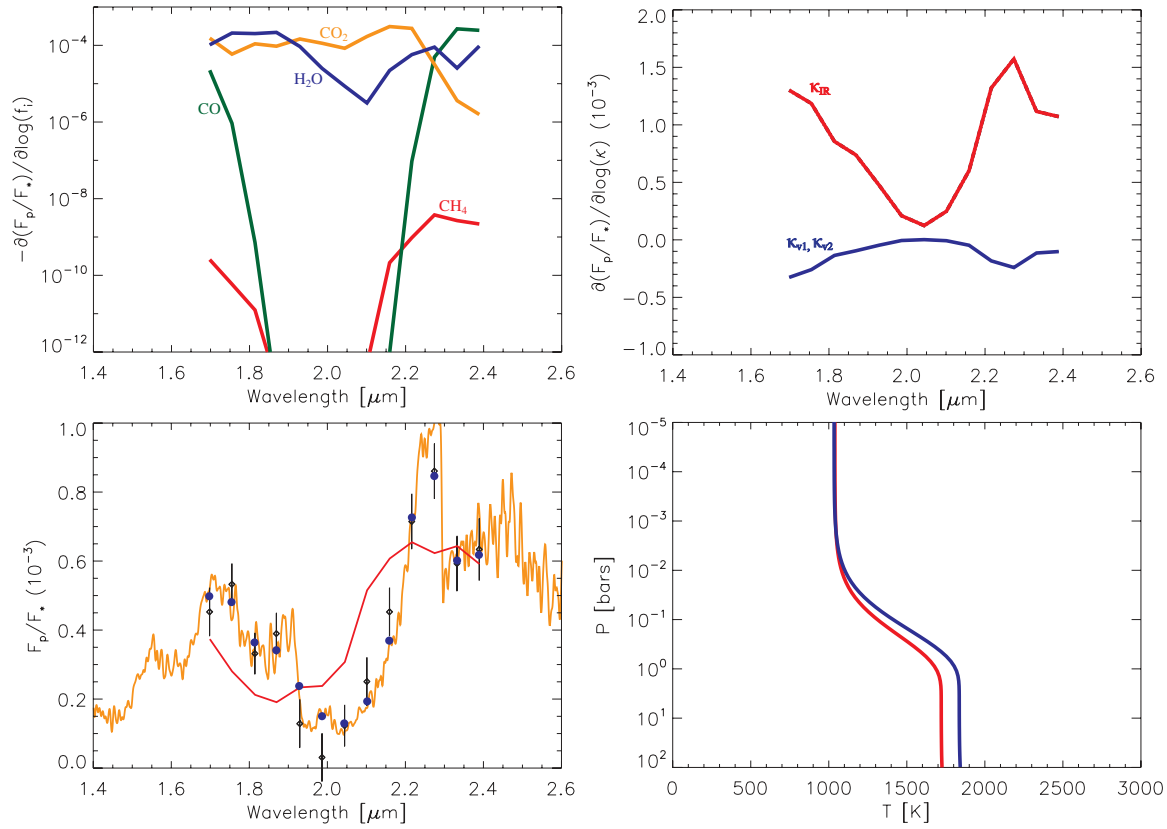

Figure 5. Retrieval results for the NICMOS dayside emission spectra of HD189733b from Swain et al. (2009b). Top left: the sensitivity of the planet-to-star flux ratio to a perturbation in the mixing ratios of H_2O, CO_2, CO, and CH_4 at each channel in the NICMOS data set. Top right: the sensitivity of the planet-to-star flux ratio to a perturbation in the parameters governing the temperature profile. Bottom left: the retrieved spectrum. The black diamonds with error bars are the Swain et al. (2009b) dayside emission data. The red curve is the a priori spectrum convolved with the instrumental broadening profile and sampled at the data wavelengths. The orange curve is retrieved spectrum at high resolution. The blue dots are the retrieved spectrum convolved with the instrumental broadening function and sampled at the data wavelengths. This optimal solution gives $\chi^2 = 0.76$. Bottom right: the a priori (red) and retrieved (blue) temperature profiles.

(A color version of this figure is available in the online journal.)

Table 2
Retrieval Results for HD189733b

Parameter	A Priori ($\mathbf{x}_a$)	Retrieved State ($\hat{\mathbf{x}}$)	Retrieval Precision	$\frac{\partial \hat{x}_i}{\partial x_j}$	MS10	S09a
κ_{v1}	4.00×10^{-3}	4.71×10^{-3}	1.67×10^{-4}–1.32×10^{-1}	0.475	...	...
κ_{v2}	4.00×10^{-3}	4.71×10^{-3}	1.67×10^{-4}–1.32×10^{-1}	0.475	...	...
κ_{IR}	3.00×10^{-2}	4.70×10^{-2}	3.00×10^{-2}–7.36×10^{-2}	0.990	...	...
α	0.5	0.5	0.00–1.00	0.00	...	...
f_{H2O}	4.00×10^{-4}	1.19×10^{-4}	5.29×10^{-5}–2.67×10^{-4}	0.997	$\sim 10^{-4}$	1×10^{-5}–1×10^{-4}
f_{CH4}	1.00×10^{-6}	9.78×10^{-9}	9.79×10^{-15}–9.77×10^{-3}	0.00	$\leqslant 6 \times 10^{-6}$	$\leqslant 1 \times 10^{-7}$
f_{CO}	5.00×10^{-4}	1.15×10^{-2}	3.60×10^{-3}–3.64×10^{-2}	0.993	2×10^{-4}–2×10^{-2}	1×10^{-4}–3×10^{-4}
f_{CO2}	1.00×10^{-7}	3.37×10^{-3}	1.69×10^{-3}–6.72×10^{-3}	0.998	7×10^{-4}	1×10^{-7}–1×10^{-6}

Notes. κ_{v1}, κ_{v2}, and κ_{IR} are in units of cm^2 g^{-1}. f_i is the volume mixing ratio for species i. We also show the diagonal averaging kernel elements ($A_{jj} = (\partial \hat{x}_j / \partial x_j)$) for each parameter. The total number of degrees of freedom for this spectrum is ~5. The retrieval precisions are given as $\hat{x} - \hat{\sigma}$ to $\hat{x} + \hat{\sigma}$ for each parameter. We also show for comparison the abundances derived by Madhusudhan & Seager (2009, MS10) and Swain et al. (2009b, S09a).

the IR photosphere sampled by the observations (because of quenching arguments). Figure 5 and Table 2 show the results of the retrieval. The Jacobian in Figure 5 demonstrates the high sensitivity of the spectrum to water and carbon dioxide, some sensitivity to CO near 2.3 μm, and very little sensitivity to methane at all wavelengths. The 1.7 and 2.2 μm channels are sensitive to the deep temperatures (effected by κ_{IR}) due to the higher transmittance at those wavelengths. The strong CO_2 absorption feature at 2.1 μm has less sensitivity to the deep temperatures and more sensitivity to temperatures higher up (controlled by κ_{v1} and κ_{v2}).

The diagonal elements of the averaging kernel in Table 2 quantitatively tell us which parameters we can and cannot retrieve from the dayside emission spectra. Again, H_2O, CO, and CO_2 have averaging kernel elements that are near unity and are therefore well constrained by the data, as is also reflected in the retrieval uncertainty, which is smaller than the assumed a priori uncertainty. CH_4 is completely unconstrained. The retrieval

uncertainty is the same as the a priori uncertainty, suggesting that the observations contribute no information about its abundance. The trace of the averaging kernel gives the total number of degrees of freedom, and thus the total number of retrievable parameters, to be ∼5.

Our results compare quite well with those of Madhusudhan & Seager (2009) and with Swain et al. (2009b) with the exception of CO_2 (Table 2) which appears to be underestimated by three orders of magnitude in Swain et al. (2009b). Our derived temperature profile (Figure 5, bottom right) also appears to fall within the spread given in Figure 5 of Madhusudhan & Seager (2009).

5. DISCUSSION AND CONCLUSIONS

We demonstrate retrieval by inverse modeling of extrasolar planetary spectra. We first apply the technique to a synthetic model spectrum of a solar metallicity $T \simeq 1200$ K hot Jupiter, and then to a previously published *HST*/NICMOS spectrum of HD 189733b, showing results that are consistent with previous studies. The approach herein is much more efficient than other methods such as a gridded parameter search, or Monte Carlo techniques, as it only requires $\sim 10^2$ forward model computations as opposed to millions. The formalism also allows robust estimation of the retrieval uncertainties.

We have also investigated the information theory aspects of the problem, in order to assess the quality and usefulness of a spectral data set in constraining atmospheric properties. First, we discuss how the Jacobian matrix can be used to determine which spectral channels are most sensitive to chosen atmospheric parameters. Second, we show the use of the averaging kernel as a diagnostic tool to guide us to which parameters can be usefully retrieved from the spectrum in question. Third, we calculated the number of available degrees of freedom and often found that, given the current limited observational capabilities, the number of retrievable parameters was less than the number of parameters in our forward model. Fourth, using simple expressions for the degrees of freedom and information content, we showed semi-quantitatively how S/N and R affect our knowledge of the atmospheric state. These tools can be particularly useful in aiding the design of future instruments such that they can be optimized for observations of transiting exoplanets.

A recent paper (Lee et al. 2012), using the optimal estimation approach as applied to HD 189733b, was published while this article was in preparation. The details of the methodology in that paper are somewhat different from ours, i.e., in the parameterization of the atmospheric models and in the use of the correlated-K opacities (we use line-by-line radiative transfer). In addition, Lee et al. use multi-band (i.e., from various instruments inclusive of *HST*/NICMOS, *Spitzer*/IRAC, IRS, and MIPS), multi-epoch measurements of HD 189733b as a representative snapshot of the planetary dayside. We restrict our retrieval to a single epoch, 13 spectral-channel NICMOS observation spanning less than one octave of total spectral coverage between 1.45 μm and 2.5 μm. Our retrievals agree for the most part with those of Lee et al. in that H_2O and CO_2 are retrieved with confidence but neither retrieval can say much about the abundance of methane (a trace species in HD 189733b). One clear discrepancy is that we are able to retrieve CO whereas they cannot. Also, Lee et al. do not discuss the information content aspects of the atmospheric retrieval formulation presented in both of these papers.

In follow-on investigations, we plan to use the information content analyses to study aspects of combining *Spitzer* broad-band photometry with prior notions about the atmospheric state to constrain atmospheric properties such as CH_4/CO and C/O ratios. A powerful use of these methods is in optimizing the design of instruments that could be flown in NASA's *FINESSE* and ESA's *Exoplanet Characterization Observatory*, or in studying the potential of already designed instruments such as the *James Webb Space Telescope* NIRCAM that offer various observing modes, bandpasses, and spectral resolving power.

We thank Zhan Su, Aaron Wolf, Konstantin Batygin, Alejandro Soto, Run-Lie Shia, Leigh Fletcher, Kuai Le, Heather Knutson, Mimi Gerstell, Linda Brown, and the Yuk Yung group for reading the article and many useful discussions. M. Line is supported by the JPL Graduate Fellowship funded by the JPL Research and Technology Development Program. X.Z. and Y.L.Y. are supported by a grant from the PATM program of NASA to the California Institute of Technology. P. Chen and G. Vasisht are supported by the JPL Research & Technology Development Program, and contributions herein were supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

Borysow, A. 2002, A&A, 390, 779
Borysow, A., Jørgensen, U. G., & Fu, Y. 2001, J. Quant. Spectrosc. Radiat. Transfer, 68, 235
Chahine, M. T. 1968, J. Opt. Soc. Am., 58, 1634
Charbonneau, D., Allen, L. E., Megeath, S. T., et al. 2005, ApJ, 626, 523
Charbonneau, D., Knutson, H. A., Barman, T., et al. 2008, ApJ, 686, 1341
Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, Nature, 434, 740
Fortney, J. J., Shabram, M., Showman, A. P., et al. 2010, ApJ, 709, 1396
Grillmair, C. J., Burrows, A., Charbonneau, D., et al. 2008, Nature, 456, 767
Grillmair, C. J., Charbonneau, D., Burrows, A., et al. 2007, ApJ, 658, L115
Guillot, T. 2010, A&A, 520, A27
Hansen, B. M. S. 2008, ApJS, 179, 484
Harrington, J., Hansen, B. M., Luszcz, S. H., et al. 2006, Science, 314, 623
Harrington, J., Luszcz, S., Seager, S., Deming, D., & Richardson, L. J. 2007, Nature, 447, 691
Jacob, D. 2007, Inverse Modeling of Atmospheric Composition Data (ARCNESS Winter School), http://acmg.seas.harvard.edu/education/
Jørgensen, U. G., Hammer, D., Borysow, A., & Falkesgaard, J. 2000, A&A, 361, 283
Knutson, H. A., Charbonneau, D., Allen, L. E., Burrows, A., & Megeath, S. T. 2008, ApJ, 673, 526
Knutson, H. A., Charbonneau, D., Allen, L. E., et al. 2007, Nature, 447, 183
Lee, J.-M., Fletcher, L. N., & Irwin, P. G. J. 2012, MNRAS, 420, 170
Line, M. R., Liang, M. C., & Yung, Y. L. 2010, ApJ, 717, 496
Line, M. R., Vasisht, G., Chen, P., Angerhausen, D., & Yung, Y. L. 2011, ApJ, 738, 32
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Mandell, A. M., Drake Deming, L., Blake, G. A., et al. 2011, ApJ, 728, 18
Moses, J. I., Visscher, C., Fortney, J. J., et al. 2011, ApJ, 737, 15
Nixon, C. A., Achterberg, R. K., Conrath, B. J., et al. 2007, Icarus, 188, 47
Press, W. H., Teukolksy, S., Vetterling, W. T., & Fannery, B. 1995, Numerical Recipes: The Art of Scientific Computing (2nd ed.; Cambridge: Cambridge Univ. Press)
Redfield, S., Endl, M., Cochran, W. D., & Koesterke, L. 2007, BAAS, 38, 134.10
Rodgers, C. D. 1976, Rev. Geophys. Space Phys., 14, 609
Rodgers, C. D. 2000, Inverse Methods for Atmospheric Sounding: Theory and Practice (Singapore: World Scientific Publishing)
Rothman, L. S., Gordon, I. E., Barbe, A., et al. 2009, J. Quant. Spectrosc. Radiat. Transfer, 110, 533
Rothman, L. S., Gordon, I. E., Barber, R. J., et al. 2010, J. Quant. Spectrosc. Radiat. Transfer, 111, 2139
Shannon, C. E., & Weaver, W. 1962, A Mathematical Theory of Communication (Paperback ed.; Urbana, IL: Univ. Illinois Press)
Sharp, C. M., & Burrows, A. 2007, ApJS, 168, 140
Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., & Le Poole, R. S. 2008, A&A, 487, 357

Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, Nature, 465, 1049
Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, Nature, 464, 1161
Swain, M. R., Deroo, P., Griffith, C. A., et al. 2010, Nature, 463, 637
Swain, M. R., Tinetti, G., Vasisht, G., et al. 2009a, ApJ, 704, 1616
Swain, M. R., Vasisht, G., Tinetti, G., et al. 2009b, ApJ, 690, L114
Tinetti, G., Deroo, P., Swain, M. R., et al. 2010a, ApJ, 712, L139
Tinetti, G., Vidal-Madjar, A., Liang, M. C., et al. 2007, Nature, 448, 169
Twomey, S., Herman, B., & Rabinoff, R. 1977, J. Atmos. Sci., 34, 1085
Waldmann, I. P., Tinetti, G., Drossart, P., et al. 2012, ApJ, 744, 35