Distinguishing Stellar Contamination in Exoplanet Transmission Spectra Using CSST/MCI (Post-print)

Authors:

Date: 2025-02-12T00:00:00+00:00

Abstract

The Chinese Survey Space Telescope (CSST) is anticipated to conduct atmospheric observation studies of exoplanets, presenting new opportunities for the field of exoplanet observations. Currently, exoplanet atmospheres are primarily investigated via transit transmission spectroscopy. However, signs of stellar activity have already been detected in actual observational data, and results indicate that it is challenging to accurately constrain stellar activity contamination. Therefore, accurately distinguishing and eliminating stellar contamination represents a major challenge for transmission spectroscopy. Consequently, when selecting targets for follow-up transmission spectroscopy observations, it is necessary to quantitatively evaluate potential stellar contamination and its distinguishability, which necessitates large-scale simulations to provide anticipated data. Employing multicolor photometry combined with machine learning to identify potential stellar contamination in the transmission spectra of gaseous exoplanets, and designing filter observation combinations for CSST' s Multi-Channel Imager (MCI) that can accurately predict the presence of stellar contamination in transmission spectra, can provide valuable references for formulating strategies for subsequent exoplanet atmospheric observation targets.

Full Text

Preamble

Vol. 66 No. 1
Jan., 2025

Acta Astronomica Sinica

Discrimination of Stellar Contamination in Exoplanet Transmission Spectra with CSST/MCI

DENG Ying-ping[1,2], JIANG Cheng-zi[1,2], CHEN Guo[1,3], YAN Zhao-jun[4]

[1] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023
[2] School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026
[3] Key Laboratory of Planetary Sciences, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023
[4] Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030

Abstract

The Chinese Survey Space Telescope (CSST) is expected to conduct atmospheric observations of exoplanets, providing new opportunities for exoplanet research. Currently, exoplanet atmospheres are primarily studied through transit transmission spectroscopy. However, evidence of stellar activity has already been detected in observational data, and results indicate that it is difficult to accurately constrain stellar contamination. Therefore, how to accurately distinguish and eliminate stellar activity contamination represents a major challenge in transmission spectroscopy. Consequently, when selecting targets for follow-up transmission spectroscopy observations, a quantitative assessment of potential stellar contamination and its distinguishability is required, necessitating large-scale simulations to provide expected data. This work employs multi-color photometry combined with machine learning to identify potential stellar contamination in the transmission spectra of exoplanet gas giants. We design filter observation combinations for the CSST Multi-Channel Imager (MCI) that can accurately predict whether transmission spectra contain stellar contamination, providing a reference for developing observation strategies for subsequent exoplanet atmospheric studies.

Keywords: methods: statistical, methods: data analysis, techniques: photometric, stars: activity, stars: starspots, planets and satellites: atmospheres

1. Introduction

Transit transmission spectroscopy is one of the most widely used methods in exoplanet atmospheric research. Retrieval analysis of transmission spectra can yield information about planetary atmospheric composition [?, ?, ?], and this approach has led to the detection of numerous atoms and molecules in dozens of exoplanet atmospheres [?, ?]. Although transmission spectroscopy has achieved substantial results, it also faces significant challenges. Since transmission spectra identify absorbing or scattering species in planetary atmospheres by measuring the difference between incident and transmitted light, accurate interpretation of transmission spectra directly depends on precise knowledge of the host star' s incident spectrum [?]. However, the presence and variability of heterogeneous regions in the stellar photosphere—such as cooler starspots and hotter faculae—cause time- and wavelength-dependent variations in the stellar spectrum. When these variations are superimposed on the planetary transmission spectrum, they affect measurements of transit depth and can even completely distort signals of atmospheric features, interfering with analysis of planetary atmospheric composition. This "stellar contamination," also known as the "transit light source effect [?]," arises because the non-uniform distribution of spots and faculae means that the intensity of the stellar surface traversed by the planet during transit (the transit chord) is no longer representative, and the normalization process in spectrophotometric light curve modeling propagates this spectral difference into the transmission spectrum.

Evidence of stellar activity contamination has already been found in observational data. For example, [?] used the Gran Telescopio CANARIAS (GTC) to conduct two transit spectroscopic observations of the hot Saturn HAT-P-12b orbiting a K-type star, finding systematic deviations between the two measured transmission spectra that could be attributed to contamination from time-varying starspots and faculae. Additionally, signs of stellar activity have been detected in the transit spectra of the warm Neptune GJ 3470b [?] and the warm sub-Neptune GJ 1214b [?] orbiting M-type stars. Therefore, the potential impact of stellar contamination must be considered in transmission spectroscopy studies. The key to efficiently utilizing existing observational resources for greater scientific value lies in how to promptly determine whether transit observations are affected by stellar contamination and how to maximize the extraction of planetary atmospheric properties when such contamination is present.

Photometric monitoring of stars at specific wavelengths can yield light curves produced by spots or faculae rotating with the star. Based on measured or assumed properties of spots or faculae and stellar spectral models, the correction factors for spots or faculae measured in the observed band can be extrapolated to other wavelengths, enabling preliminary correction of stellar contamination in spectrophotometric transit light curves [?, ?, ?]. However, the morphology of stellar rotational light curves depends on the contrast, size, and distribution of spots or faculae, and some spots or faculae may not contribute to rotational variability (e.g., polar spots), potentially leading to underestimation of stellar contamination corrections [?]. Another widely used approach is to model stellar contamination directly into planetary atmospheric retrieval, which has the advantage of propagating uncertainties from stellar contamination into the posterior distributions of planetary atmospheric parameters [?].

[?] proposed a stellar contamination modeling formula assuming no spots or faculae within the transit chord, which [?] subsequently extended to cases where spots or faculae are present within the transit chord. The upcoming 2-meter-class Chinese Survey Space Telescope (CSST) combines excellent wide-field capabilities with high image quality. Its Multi-Channel Imager (MCI) enables simultaneous high-precision photometric observations across three channels covering the full ultraviolet-to-optical band: NUV1 + u (0.255-0.43 m), g + r (0.43-0.7 m), and i + z (0.7-1.0 m), making it suitable for exoplanet transit observations. Therefore, this work attempts to combine multi-color photometry with machine learning methods to discriminate whether stellar contamination is present in exoplanet transmission spectra and to provide filter combination recommendations for MCI exoplanet multi-color photometric observations.

The structure of this paper is as follows: Section 2 describes the model construction process and the filters used for analysis, Section 3 analyzes the complete model sample to identify optimal filter combinations, and Section 4 presents our conclusions.

2. Model Construction

We construct a series of gas giant planet samples, simulate their atmospheric transmission spectra, and consider varying degrees of activity contamination from host stars of multiple spectral types. We then employ machine learning algorithms to investigate the statistical trends exhibited in the simulated multi-color photometry to determine whether parameter space exists that can distinguish stellar contamination. Similar methods were previously used by [**?**] to study whether classification of directly imaged planets could be performed based on multi-color photometry of their reflected spectra.

Model construction proceeds in three steps: First, we select gas giant planets with 0.6 Jupiter masses and 1.25 Jupiter radii as our study objects and generate theoretical planetary atmospheric transmission spectra across a grid of atmospheric parameters. Second, we choose F, G, K, and M spectral type stars as representatives and generate stellar activity contamination models for them. Finally, we generate planetary transmission spectrum models containing stellar contamination and produce corresponding photometric model data according to specific filter configurations. Since the color characteristics of transmission spectra we focus on are independent of host star radius, we set the host star radius to 1.1 solar radii for all models to simplify the parameter grid.

2.1. Planetary Atmospheric Transmission Spectrum Model

We use Platon [**?**, **?**] to generate planetary atmospheric transmission spectra with a spectral resolution of 1000. Platon calculates transmission spectra of equilibrium-chemistry, hydrostatic atmospheres through one-dimensional radiative transfer, accounting for gas absorption, collisional absorption, optically thick clouds, and scattering absorption. In its default configuration, Platon assumes an isothermal atmospheric structure at temperature T_{iso}, divided into 250 layers uniformly distributed on a logarithmic scale from 10^{-3} to 10^{-9} bar. Gas abundances are drawn from a pre-computed equilibrium chemistry grid generated by the GGChem code [**?**], which is controlled by parameters including atomic and molecular species, temperature (T), pressure (P), metallicity (Z), and carbon-to-oxygen ratio (C/O). Collisional absorption coefficients reference the HITRAN molecular spectroscopic database [**?**, **?**]. In Platon, clouds are modeled as an optically thick layer with cloud-top pressure P_{cloud}. Scattering absorption by atmospheric hazes and aerosols is parameterized as Rayleigh-like scattering with slope γ and amplitude A_{scatter}, where the scattering cross-section at wavelength λ is $\sigma(\lambda) =$ A_{scatter} $\sigma_0\lambda^{-4}$, with σ_0 being the H_2 Rayleigh scattering cross-section at a reference wavelength.

In constructing our transmission spectrum models, we grid the parameters T_{iso}, Z, C/O, P_{cloud}, and A_{scatter}. The T_{iso} range is 600-2400 K in steps of 200 K. Metallicity Z takes values of 0.1, 1, 10, and 100 times solar metallicity. P_{cloud} takes values of 10^{-3}, 10^{-2}, and 10^{3} bar, representing strong cloud influence, moderate cloud influence, and cloud-free conditions, respectively, where larger P_{cloud} indicates lower-altitude clouds (with 10^{3} bar located at the bottom of the planetary atmosphere in our models). A_{scatter} takes values of 1, 10, 100, and 1000 times the Rayleigh scattering strength, representing haze scattering characteristics (Rayleigh scattering corresponds to $\gamma = 4$ in $\sigma(\lambda)$). This parameter grid yields a total of $10 \times 4 \times 4 \times 4 \times 3 = 1920$ planetary atmospheric transmission spectrum models.

[Figure 1: see original paper] illustrates how the transmission spectrum varies with cloud-top pressure P_{cloud}. As P_{cloud} decreases, the absorption features in the transmission spectrum gradually become shallower and narrower because the presence of high-altitude optically thick clouds prevents stellar light from penetrating deeper into the atmosphere, weakening spectral features and reducing pressure-broadening effects. [Figure 2: see original paper] shows how the transmission spectrum varies with planetary atmospheric temperature T_{iso}. The low-temperature region exhibits prominent Rayleigh scattering and Na and K absorption lines, while the high-temperature region is dominated by complex band absorption features from molecules such as TiO and VO.

2.2. Stellar Activity Contamination Model

Quantitative correction of stellar activity contamination in planetary transmission spectra typically employs two methods. One involves photometric monitoring of the star to estimate contamination levels for correction in the transmission spectrum [**?**, **?**, **?**]. The other models stellar contamination directly into planetary atmospheric retrieval, which avoids underestimation of contamination and propagates uncertainties into the posterior distributions of planetary atmospheric parameters [**?**]. This work focuses on the latter method.

For this approach, a three-component model is commonly used, approximating the spectrum of an active star as a composite of the immaculate photosphere, spot spectrum, and facula spectrum [**?**]. Considering only the photosphere plus spots (hereafter called pure spot contamination), the formula is:

λ^s = Dλ^{obs} / D_λ = (S_λ^{phot} - F_{spot}(S_λ^{phot} - S_λ^{spot})) / (1 - F_{spot}) = 1 - F_{spot}(1 - S_λ^{spot} / S_λ^{phot})

where D_λ^{obs} represents the observed transit depth (the ratio of transit-induced flux variation to stellar flux), D_λ represents the theoretical transit depth (the square of the planet-to-star radius ratio), and _λ^s represents the contamination ratio at different wavelengths on the theoretical transmission spectrum due to spots. The model inputs include the stellar photosphere spectrum (S_λ^{phot}), spot spectrum (S_λ^{spot}), and the filling factor of spots not occulted by the planet (F_{spot}).

If all three components—photosphere, spots, and faculae—are considered (hereafter called mixed contamination), the formula generalizes to:

_λ^{sf} = (1 - F_{spot}(1 - S_λ^{spot} / S_λ^{phot}) - F_{facu}(1 - S_λ^{facu} / S_λ^{phot}))

where _λ^{sf} is the contamination ratio at different wavelengths due to both spots and faculae. This model adds the facula spectrum (S_λ^{facu}) and its filling factor (F_{facu}) to the previous equation.

[Figure 3: see original paper] shows a schematic of a synthesized active stellar spectrum based on PHOENIX stellar atmosphere models [?], composed of three components: immaculate photosphere, spot spectrum, and facula spectrum. When a planet transits such a star, the observed transit depth can be expressed as the theoretical planetary atmospheric transit depth multiplied by _λ^{sf} from the equation above. We note that this formula assumes no spots or faculae within the transit chord; if spots or faculae are present within the chord, the formula from [?] should be used, though its higher degrees of freedom are typically difficult to constrain with current data precision. In summary, we parameterize stellar activity-induced contamination as a weighted combination of photosphere, spots, and faculae, characterized by wavelength-dependent scaling coefficients _λ^s and _λ^{sf}.

We use the Python package pysynphot to call PHOENIX stellar atmosphere models [?] to generate stellar photosphere, spot, and facula spectra, considering only solar-metallicity models without α-element enhancement ([α/Fe] = 0, [Fe/H] = 0). Following the empirical results in Tables 1, 3, and 4 of [?] and Tables 2 and 4 of [?], we select ten stellar spectral types as grid points: F5V, F8V, G2V, G5V, G8V, K2V, K5V, K8V, M2V, and M5V. We adopt the corresponding stellar physical parameters as prior inputs, with specific values listed in . We consider two types of stellar contamination: pure spot contamination and mixed spot-facula contamination, calculating _λ^s and _λ^{sf} using the respective equations. This yields a total of $10 \times 2 = 20$ contamination models.

2.3. Planetary Atmospheric Transmission Spectrum Models with Stellar Activity Contamination

After completing the planetary atmospheric transmission spectrum models and stellar activity contamination models separately, we multiply each type of stellar contamination coefficient (_λ^s, _λ^{sf}) with the planetary atmospheric transmission spectrum models to obtain contaminated transmission spectra. Including the uncontaminated planetary atmospheric transmission spectra, we have a total of 1920 × (20 + 1) = 40,320 models. We analyze these models as our sample in subsequent sections.

3. Sample Analysis

The MCI covers a wavelength range of 250–1050 nm and can observe three channels simultaneously, equipped with 30 filters. The narrowband filters are expected to capture spectral features in the transmission spectrum models. [Figure 4: see original paper] shows the throughput curves of ten narrowband filters with FWHM $\leq$ 20 nm. The F280N filter has relatively low throughput, and the corresponding stellar flux is also lower, potentially resulting in lower photometric signal-to-noise ratios, so this filter is excluded from subsequent analysis. Convolving our calculated transmission spectra with the remaining nine filters yields nine photometric measurements. Since the continuum levels of planetary atmospheric transmission spectra differ across parameters, only photometric differences between different bandpasses can reflect potential spectral features. Given that each channel can use only one filter at a time, the nine selected filters can be combined pairwise to produce 24 colors, meaning a single transmission spectrum yields 24 feature values. Therefore, our analysis can be framed as a classification problem with 24-dimensional color features.

3.1. Correlation of Transmission Spectrum Color Features

To visually demonstrate the distribution of various sample types in color feature space, we combine the 24 color features pairwise into 276 two-color spaces and qualitatively analyze sample distributions. We use subsamples with F5V and M5V host stars as examples to qualitatively analyze whether transmission spectrum color features correlate with stellar contamination.

[Figure 5: see original paper] shows partial two-color distributions for the F5V and M5V subsamples, representing early- and late-type host stars, respectively. The distribution morphology varies across different two-color spaces, indicating that color features capture differences between transmission spectra, which benefits classification. The distributions exhibit two general patterns: "quasi-correlated" and "T-shaped," suggesting that not all 24 color features provide strong discriminatory power—many capture similar transmission spectrum information.

Examining the distribution of uncontaminated (w/o Cont.), pure spot contamination (Spot), and mixed contamination (Sp+Fa) categories in each two-color space reveals that for the F5V subsample, the three categories almost completely overlap and are difficult to distinguish visually, with distribution morphology varying significantly across two-color spaces. For the M5V subsample, however, the three categories show clear separation, though the distribution for any single category (e.g., mixed contamination alone) maintains similar "quasi-correlated" and "T-shaped" patterns as seen in F5V. The relative offsets between the three categories introduce large-scale "quasi-correlated" patterns on top of the original

"T-shaped" morphology.

We calculate correlations between the 24 color features for both F5V and M5V subsamples, displayed as heatmaps in [Figure 6: see original paper]. If two colors are highly correlated (correlation coefficient > 0.75), they capture similar transmission spectrum features. Combining them does not help distinguish transmission spectrum differences, whereas combinations of less-correlated colors may be more sensitive to different spectral features and provide more discriminatory information. [Figure 6: see original paper] shows that 44.2% of color pairs have correlations above 0.75 for the F5V subsample, while this fraction is significantly higher for M5V at 99.3%, indicating stronger correlations among its colors. This aligns with our qualitative observations of distribution morphology. The stronger overall correlation in the M5V subsample arises from the dispersion of samples across the three contamination categories, introducing new large-scale linear correlations. When considering only a single contamination category, the fraction of highly correlated colors decreases substantially.

Among our ten spectral type subsamples, only M5V shows such strong overall correlations; the other nine spectral types have fractions of color pairs with correlations above 0.75 ranging from 42.0% to 44.6%. This is partly because M5V adopts higher spot and facula filling factors, and partly because when stellar effective temperature drops below 3000 K, molecular absorption bands dominate the optical band, making photospheric-active region spectral differences per unit temperature interval much stronger than in the >3000 K regime.

3.2. Discriminating Stellar Contamination Spectral Features with XGBoost

To quantitatively determine the relationship between these color features and stellar activity contamination in transmission spectra, we employ the supervised machine learning method XGBoost (eXtreme Gradient Boosting). XGBoost is a gradient boosting tree algorithm that supports regularization to control model complexity and reduce overfitting, offering high scalability and excellent performance across various machine learning tasks. In our classification problem, the stellar spectral type is prior information—i.e., the host star' s spectral type is typically known before observation. Therefore, our full sample can be divided into ten groups by spectral type: F5V, F8V, G2V, G5V, G8V, K2V, K5V, K8V, M2V, and M5V. Each group contains three categories of transmission spectra: uncontaminated (w/o Cont.), pure spot contamination (Spots), and mixed contamination (Spots+Faculae). Our goal is to use XGBoost to distinguish whether transmission spectra contain stellar contamination for each of these ten groups, making this a three-class classification problem.

We implement XGBoost using the Python package scikit-learn. First, each group is randomly split into 70% training and 30% testing sets. We then create and train XGBoost models after hyperparameter tuning, settling on 100 base decision trees, a learning rate of 0.05, and maximum depth of 5. After training, we evaluate performance on the test sets using confusion matrices, which display the counts of correct and incorrect predictions. We quantify classification accuracy as the proportion of correctly classified samples.

[Figure 7: see original paper] shows the confusion matrices for all ten test groups. (Column 2) presents the prediction accuracies. The results show a gradual decrease in accuracy from late-type to early-type stars. For M5V and M2V groups, discrimination accuracies reach 99.7% and 99.9%, respectively, enabling excellent separation of uncontaminated, pure spot, and mixed contamination categories. As spectral types transition from K to G, accuracy decreases gradually from 97.9% for K8V to 91.6% for G2V. For G2V, some misclassification occurs among the three categories. In contrast, accuracies for F8V and F5V groups are only 79.0% and 66.8%, respectively, indicating that discriminating stellar contamination for F-type stars is more challenging, though this also correlates with weaker prior input stellar activity contamination. Early-type stars are generally more quiescent, and stellar activity contamination can typically be neglected in exoplanet photometric observations, while late-type stars are relatively more active, and their contamination cannot be ignored. Active stellar activity causes stellar contamination to dominate the observed transmission spectrum, making color features primarily capture spectral differences between the stellar photosphere and active regions.

XGBoost provides a feature importance function that outputs each feature’ s contribution to the model. We use this to identify the filter combinations that best predict stellar contamination. [Figure 8: see original paper] shows the importance values for each color feature across the ten groups, with the three most important colors for each group listed in .

3.3. Optimal Filter Combinations for Discriminating Stellar Contamination

The quantitative XGBoost analysis demonstrates that for observation targets with significant stellar activity, narrowband filter combinations can be used to assess whether stellar activity contamination is present in transmission spectra. This section analyzes which filter combinations maximize discrimination accuracy and efficiency.

Based on [Figure 8: see original paper] and , F656N-F815N is the most frequently occurring color feature, indicating that the photometric difference between filters F656N and F815N is most sensitive to stellar contamination features. Filter F656N covers the stellar Hα line, a common indicator of stellar activity [**?**, **?**], while our planetary atmospheric models contain no Hα absorption features. In contrast, filter F815N covers a wavelength range with no significant stellar activity-sensitive lines and no prominent atomic or molecular opacity sources in our planetary atmospheric models, providing a good reference baseline for F656N. Their difference thus reflects stellar activity.

Further testing shows that using only the top three color features for model training yields accuracy variations that are not substantial compared to using all 24 color features, with results shown in (Column 3). This likely arises because multiple colors are correlated or even strongly correlated, so additional colors do not provide significant extra information. This aligns with the qualitative analysis in [Figure 6: see original paper], where strongly correlated feature combinations provide limited additional spectral detail for classification.

In practice, since some color pairs involve filters from the same MCI channel that cannot be used simultaneously, the three most important color features for each spectral type in require multiple observing epochs to cover completely, with the minimum number of required observations listed in the final column of . If the spectral type is unknown and only one transit observation is possible, using filters F373N, F656N, and F815N across the three channels provides the most efficient preliminary screening and estimation of stellar contamination.

4. Summary and Discussion

By constructing a sample of planetary atmospheric transmission spectra covering various types and degrees of stellar contamination and multiple atmospheric parameters, simulating CSST/MCI multi-color photometric observations of exoplanet transits, and investigating the feasibility of distinguishing potential stellar contamination using machine learning methods, we reach two main conclusions:

(1) For gas giants orbiting host stars later than G2V, MCI narrowband multi-wavelength transit observations can effectively distinguish potential stellar activity contamination and differentiate between pure spot contamination and mixed spot-facula contamination. In contrast, transmission spectra of F-type stars are difficult to classify. XGBoost achieves classification accuracies above 90% for subsample groups later than G2V, while performance is poor for F-type (F8V, F5V) groups. This correlates with our empirical stellar activity priors—early-type stars are generally quiet, and stellar activity contamination can typically be neglected in exoplanet photometric observations, while late-type stars are more active, and their contamination cannot be ignored. Active stellar activity causes contamination to dominate the observed transmission spectrum, making color features primarily capture spectral differences between the photosphere and active regions.

(2) We find that using MCI for narrowband transit photometry, selecting three filter combinations can constrain the contribution of stellar activity contamination to transmission spectra, thereby improving the accuracy of planetary atmospheric parameter retrieval. For different host star spectral types, we have identified optimal filter combinations (). If the spectral type is unknown and only one transit observation is possible, using filters F373N, F656N, and F815N in the three channels provides efficient preliminary screening and estimation of stellar contamination.

Stellar contamination intensity is a crucial factor affecting classification accuracy. This work adopts the median spot and facula filling factors from [?, ?] based on actual observations as prior inputs. If we instead use the 1σ upper limits as inputs in similar classification tests, enhanced stellar activity contamination makes it easier to distinguish whether transmission spectra contain stellar contamination, with improved discrimination accuracy for all groups. The accuracy improvement primarily comes from enhanced ability to differentiate between pure spot contamination and uncontaminated categories. If actual stellar activity is much stronger than empirical expectations, accumulated observational data will help improve studies of activity levels for different spectral type host stars.

Our MCI photometric simulations do not include photometric errors, as our primary goal is to determine, from a forward-modeling perspective, whether stellar activity contamination can be distinguished and which filter combinations best extract it. Quantitative calculation of stellar contamination in real data can be obtained through retrieval analysis of individual multi-color photometric spectra, without requiring machine learning classification of large samples of actual photometric data.

We assumed a planetary mass of 0.6 Jupiter masses when generating atmospheric models. Changing planetary mass in similar sample analyses shows that decreasing mass reduces classification accuracy while increasing mass improves it. This occurs because, at fixed radius, lower planetary mass corresponds to a more puffy atmosphere, making planetary atmospheric features dominate the transmission spectrum under equivalent stellar contamination. However, some atmospheric components (e.g., TiO) have spectral features similar to stellar contamination, increasing discrimination difficulty. For more massive planets, atmospheric spectral features are weaker and less similar to stellar contamination features.

Additionally, our chemical-equilibrium, hydrostatic planetary atmospheric models do not include Hα absorption features. However, recent high-resolution observations have detected Hα absorption in the upper atmospheres of many hot Jupiters heated by their host stars [?, ?, ?, ?, ?, ?]. In future work, we will investigate the impact of Hα absorption in planetary upper atmospheres on stellar contamination discrimination and study the feasibility and applicable conditions of constraining planetary upper atmospheric heating processes through narrowband photometry of transmission spectra.

Acknowledgments

We thank the referee for valuable suggestions that improved the quality of this paper.

References

[1] Seager S., Sasselov D. D. 2000, ApJ, 537, 916
[2] Brown T. M. 2001, ApJ, 553, 1006
[3] Hubbard W. B., Fortney J. J., Lunine J. I. 2001, ApJ, 560, 413
[4] Madhusudhan N. 2019, ARA&A, 57, 617
[5] Guillot T., Fletcher L. N., Helled R. 2022, arXiv:2205.04100
[6] Rackham B. V., Apai D., Giampapa M. S. 2018, ApJ, 853, 122
[7] Jiang C., Chen G., Pallé E. 2021, A&A, 656, A114
[8] Oshagh M., Santos N. C., Ehrenreich D. 2014, A&A, 568, A99
[9] Rackham B., Espinoza N., Apai D. 2017, ApJ, 834, 151
[10] Pont F., Knutson H., Gilliland R. L. 2008, MNRAS, 385, 109
[11] Berta Z. K., Charbonneau D., Bean J. 2011, ApJ, 736, 12
[12] Sing D. K., Pont F., Aigrain S. 2011, MNRAS, 416, 1443
[13] McCullough P. R., Crouzet N., Deming D. 2014, ApJ, 791, 55
[14] Zhang Z., Zhou Y., Rackham B. V. 2018, AJ, 156, 178
[15] Batalha N. E., Smith A. J. R. W., Lewis N. K. 2018, AJ, 156, 158
[16] Zhang M., Chachan Y., Kempton E. M. R. 2019, PASP, 131, 034501
[17] Zhang M., Chachan Y., Kempton E. M. R. 2020, ApJ, 899, 27
[18] Woitke P., Helling C., Hunter G. H. 2018, A&A, 614, A1
[19] Richard C., Gordon I. E., Rothman L. S. 2012, JQSRT, 113, 1276
[20] Karman T., Gordon I. E., van der Avoird A. 2019, Icarus, 328, 160
[21] Husser T. O., Wende-von Berg S., Dreizler S. 2013, A&A, 553, A6
[22] Rackham B. V., Apai D., Giampapa M. S. 2019, AJ, 157, 96
[23] Pasquini L., Pallavicini R. 1991, A&A, 251, 199
[24] Cincunegui C., Díaz R. F., Mauas P. J. D. 2007, A&A, 469, 309
[25] Yan F., Henning T. 2018, NatAs, 2, 714
[26] Casasayas-Barris N., Pallé E., Yan F. 2018, A&A, 616, A151
[27] Cabot S. H. C., Madhusudhan N., Welbanks L. 2020, MNRAS, 494, 363
[28] Chen G., Casasayas-Barris N., Pallé E. 2020, A&A, 635, A171
[29] Yan F., Wyttenbach A., Casasayas-Barris N. 2021, A&A, 645, A22
[30] Czesla S., Lampón M., Sanz-Forcada J. 2022, A&A, 657, A6

Note: Figure translations are in progress. See original paper for figures.