

Atmospheric characterization with the *Twinkle Space Telescope* following advances from *JWST* observations

Tailong Zhang,¹ Benjamin Wilcock,¹ Sushuang Ma,² Giovanna Tinetti,^{1,2★} Lawrence Bradley,¹ Ian Stotesbury,¹ Marcell Tessenyi¹ and Jonathan Tennyson^{1,3}

¹Blue Skies Space Ltd., 69 Wilson Street, London EC2A 2BB, UK

²Department of Physics, King's College London, Strand, London WC2R 2LS, UK

³Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Accepted 2026 June 19. Received 2026 March 6; in original form 2025 August 11

ABSTRACT

The *Twinkle Space Telescope* is a satellite designed for spectroscopic observations of a wide range of extrasolar and Solar system objects. Equipped with a 0.45 m diameter telescope and a spectrometer covering from 0.5 to 4.5 μm simultaneously, *Twinkle* will be launched in a sun-synchronous, low-Earth orbit, and it is expected to operate for 7 yr. *Twinkle* is developed, managed, and operated by Blue Skies Space Ltd., a space science data company whose vision is to accelerate and expand the availability of new, high-quality data sets to researchers worldwide, complementing the space-observatories delivered by government space agencies. Over its lifetime, *Twinkle* will conduct large-scale survey programs. The scientific objectives and observational strategy of these surveys are defined by researchers who join the Science Team. Leveraging advances made possible by recent observations with the *James Webb Space Telescope*, we present here updated simulations evaluating *Twinkle*'s observational capabilities in the context of exoplanet atmospheres. Through retrieval analyses of HD 209458 b, WASP-107 b, GJ 3470 b, and 55 Cnc e, we demonstrate how increasing observational investment enhances the retrieval of atmospheric parameters and molecular abundances. Our sensitivity study highlights *Twinkle*'s capability to detect less abundant/detectable molecules depending on the observing strategies adopted. This work provides practical guidance for developing targeted observational strategies to maximize *Twinkle*'s scientific return.

Key words: instrumentation: spectrographs – techniques: spectroscopic – planets and satellites: atmospheres – planets and satellites: gaseous planets – planets and satellites: terrestrial planets.

1 INTRODUCTION

The field of exoplanetary science has experienced unprecedented growth, with the number of confirmed exoplanets now exceeding 5800. A significant proportion of these detections – approximately 4000 – has been made using transit photometry from space-based surveys such as *Kepler*, *K2*, and *TESS* (W. J. Borucki et al. 2010; S. B. Howell et al. 2014; G. R. Ricker et al. 2014). Since the retirement of the *Kepler/K2* mission in 2018, the *TESS* mission, operational since 2018, has significantly advanced planet discovery efforts. As of 2025 March 12, *TESS* has identified 7525 *TESS* Objects of Interest (TOIs). After excluding false positives, 5253 candidates remain confirmed by the *TESS* Follow-up Observing Program Working Group (TFOPWG) (R. Akesson & J. Christiansen 2019). This wealth of discoveries of transiting planets around bright stellar hosts has been a game changer for atmospheric characterization studies using transit, eclipse, and phase-curve spectroscopy.

The European Space Agency's PLATO, due for launch in 2026, will aim at detecting small transiting planets (above) around bright stars (≤ 11 mag), including terrestrial planets in the habitable zone of solar-like stars (H. Rauer et al. 2025). Similarly, the Earth 2.0 (ET) mission in transit mode will continuously monitor over 2 million FGKM dwarfs in the original *Kepler* field and its neighbouring fields for 4 yr to search for new planets, including terrestrial-like planets, across a wide range of orbital periods (J. Ge et al. 2022, 2024).

Since the first atomic, ionic, and molecular species were revealed in exoplanetary atmospheres (e.g. D. Charbonneau et al. 2002; A. Vidal-Madjar et al. 2003, 2004; G. Tinetti et al. 2007; C. J. Grillmair et al. 2008; S. Redfield et al. 2008; M. R. Swain et al. 2008, 2009), the interest in studying the chemistry and structure of exoplanet atmospheres has progressively become more widespread. As a result, tens of exoplanetary atmospheres have been observed with the *Spitzer* and *Hubble Space Telescopes* (*HST*) and with ground-based facilities in the past couple of decades. These measurements have guided our initial understanding of the atmospheric composition, dynamics,

★ E-mail: giovanna.tinetti@kcl.ac.uk

and thermal and scattering properties for a variety of atmospheres in very diverse environments (e.g. J. Harrington et al. 2006; H. A. Knutson et al. 2007; C. Majeau et al. 2012; K. B. Stevenson et al. 2014; B.-O. Demory et al. 2016).

In particular, the Wide Field Camera 3 (WFC3) G141 instrument on board *HST*, covering 1.1 to 1.7 μm , has been central to the hunt for key molecular species such as H_2O , CH_4 , TiO , VO , and clouds for a long list of very diverse exoplanets (e.g. R. A. Windhorst et al. 2011; Z. K. Berta et al. 2012; D. Deming et al. 2013; J. Fraine et al. 2014; A. Tsiaras et al. 2016; J. Arcangeli et al. 2018).

While this effort has been pursued through a number of complementary techniques, including direct imaging and high-resolution spectroscopy from the ground (J. L. Bean et al. 2010; I. A. G. Snellen et al. 2010; B. Macintosh et al. 2015; G. Chauvin et al. 2017; L. Nortmann et al. 2018; GRAVITY Collaboration 2019), transit, eclipse, phase-curve spectroscopy, and multiband photometry from space observatories have enabled large-scale atmospheric studies throughout the years (D. K. Sing et al. 2016; A. Tsiaras et al. 2018; Q. Changeat et al. 2022; B. Edwards et al. 2023; A. Saba et al. 2025; L. Dang et al. 2025).

The recent launch of the *James Webb Space Telescope (JWST)* has offered unprecedented insights into exoplanet atmospheres, spanning wavelengths from 0.5 to 14 μm (e.g. *JWST* Transiting Exoplanet Community Early Release Science Team 2023; N. Madhusudhan et al. 2023; Q. Xue et al. 2024; A. L. Carter et al. 2024; R. Hu et al. 2024; T. Bell et al. 2024). *JWST* has refined measurements of molecular abundances, thermal profiles, atmospheric dynamics, and cloud/haze characteristics. Despite these significant advances, *JWST*'s broad scientific remit – including galactic and extragalactic astrophysics and Solar system studies – limits its dedicated observation time for exoplanetary research.

Addressing this limitation, dedicated missions such as ESA's Ariel (scheduled for launch in 2029) have been developed. Ariel aims to conduct a comprehensive chemical survey of exoplanetary atmospheres using transit spectroscopy over the spectral range of 0.5–7.8 μm , thereby deepening our understanding of how atmospheric chemistry correlates with planetary characteristics (G. Tinetti et al. 2018, 2022).

The *Twinkle Space Telescope*, managed and operated by Blue Skies Space Ltd. (BSSL), will be dedicated to provide spectroscopic data for various science themes developed by its members (S. Jason et al. 2016; G. Savini et al. 2018; B. Edwards et al. 2019a, b; R. Archer et al. 2020; I. Stotesbury et al. 2022, 2024). Equipped with a 0.45-m telescope and a spectrometer covering simultaneously from 0.5 to 4.5 μm , *Twinkle* is capable of providing spectra for a wide range of targets such as brown dwarfs, stars, protoplanetary discs, Solar System objects, and exoplanets. Two large survey programmes focusing on extrasolar targets and Solar System objects are being planned during the first three years of *Twinkle*'s expected seven years operations (I. Stotesbury et al. 2022). While BSSL is responsible for the satellite, the data analysis tools and facilitating the survey planning process, researchers who join the survey program (Science Team) will decide the scientific objectives and observational strategy. *Twinkle* Science Team members have been tasked with delivering a comprehensive and scientifically optimized observation plan and list of targets.

Based on the existing scientific themes proposed by the *Twinkle* Science Team members, a large fraction of the survey time will be allocated to extrasolar objects and the atmospheric characterization of diverse types of exoplanets is expected to be a significant part of the survey. This paper presents new simulations to evaluate *Twinkle*'s potential for atmospheric characterization in the context of more recent findings by *JWST* and discusses possible approaches to optimize the observational strategy.

Twinkle's spectral range overlaps substantially with that of *JWST*, making *JWST*'s observations an invaluable reference for simulating *Twinkle*'s expected performance. The combination of *JWST*'s comprehensive spectral coverage and its recent discoveries offers essential benchmarks that drive our updated simulations of *Twinkle*'s observational capabilities. In this study, we incorporate the latest *JWST*-calibrated atmospheric models into *Twinkle*'s simulation framework, leading to more accurate predictions of signal-to-noise ratios (SNR) and a refined assessment of the detectability of key atmospheric features.

B. Edwards et al. (2019c) previously assessed *Twinkle*'s ability to characterize the atmospheric composition and structure of three iconic exoplanets: HD 209458 b, GJ 3470 b, and 55 Cnc e. However, due to lack of IR spectroscopic data pre-*JWST*, those analyses were based on theoretical models and simplified assumptions. Recent *JWST* observations have revealed complexities in atmospheric chemistry, thermal, and cloud structures that necessitate an update to these models (e.g. Q. Xue et al. 2024; T. G. Beatty et al. 2024; R. Hu et al. 2024). More specifically, recent *JWST* observations of HD 209458 b by Q. Xue et al. (2024) have provided a high-quality transmission spectrum from 2.3 to 5.1 μm , revealing pronounced absorption features from CO_2 and H_2O . Based on thermochemical equilibrium assumptions, these results suggest a supersolar metallicity and a low C/O ratio (~ 0.11). GJ 3470 b has emerged as a key target of interest due to its relatively transparent atmosphere and the detection of SO_2 in its atmosphere, reinforcing the evidence of disequilibrium chemistry in lower mass exoplanet atmospheres (T. G. Beatty et al. 2024). *JWST*'s eclipse observations of the super-Earth 55 Cnc e by R. Hu et al. (2024) have confirmed the presence of a volatile-rich atmosphere, potentially maintained by a magma ocean. In addition to these three exoplanets, we have added here the study of WASP-107 b: L. Welbanks et al. (2024) have identified the presence of several molecular species in the atmosphere of this planet, including H_2O , CH_4 , CO , CO_2 , SO_2 , and NH_3 when combining *HST* and *JWST* data.

This paper first reviews *Twinkle*'s instrumental capability, the models and methodology used in estimating *Twinkle*'s science performance in Section 2. It updates the radiometric model employed to simulate the noise characteristics (Section 2.1) and the list of potential target candidates within *Twinkle*'s field of view and details. Section 2.3 describes the atmospheric forward and inverse models obtained with the open-source exoplanet atmospheric retrieval framework TauREx 3. Simulations and retrieval results are presented in Section 3, highlighting *Twinkle*'s ability to detect molecular species and other atmospheric properties across a diverse exoplanetary

Table 1. Exoplanets and their properties considered in this work. HD 209458 b (K. G. Stassun et al. 2017; Q. Xue et al. 2024); WASP-107 b (D. R. Anderson et al. 2017; C. Piaulet et al. 2021); GJ 3470 b (S. Awiphan et al. 2016); 55 Cancri e (R. Hu et al. 2024).

	HD 209458 b	WASP-107 b	GJ 3470 b	55 Cancri e
Planet type	Hot Jupiter	Warm Neptune	Warm Neptune	Super Earth
Planet mass (M_J)	0.73 ± 0.04	0.096 ± 0.005	0.035 ± 0.005	0.0251 ± 0.001
Planet radius (R_J)	1.39 ± 0.02	0.94 ± 0.02	0.37 ± 0.016	0.1673 ± 0.0026
T_{eq} (K)	1140^{+71}_{-83}	770 ± 60	600 ± 96.5	1958 ± 15
Stellar magnitude (V/K)	7.65 ± 0.03 (V)	11.592 ± 0.208 (V)	12.332 ± 0.016 (V)	5.95 ± 0.023 (V)
	6.308 ± 0.026 (K)	8.637 ± 0.023 (K)	7.989 ± 0.023 (K)	4.015 ± 0.036 (K)
<i>JWST</i> availability (μm)	2.3–4.5	2.4–4.5	2.45–4.97	3.94–12
<i>HST</i> wavelength range (μm)	0.5–0.9; 1.1–1.7	0.8–1.6	0.524–1.027; 1.1–1.7	–

sample and observational strategies, based on recent *JWST* advances. Finally, Section 4 discusses the results, addresses study limitations, and proposes directions for future simulation work.

2 METHODS

2.1 Twinkle's radiometric model

Twinkle will observe spectra simultaneously across two channels covering the wavelength range from 0.5 to 4.5 μm . Channel 0 operates from 0.5 to 2.43 μm with a resolving power ($R = \frac{\lambda}{\Delta\lambda}$) of up to 70, and channel 1 operates from 2.43 to 4.5 μm with a resolving power of up to 50.

Twinkle's radiometric performance is assessed using the *Twinkle* Radiometric Tool (I. Stotesbury et al. 2022), adapted from the open-source radiometric software EXORAD2 (L. V. Mugnai et al. 2023). The radiometric model simulates observational and instrumental noise contributions, including photon noise, instrument emission, detector dark current, read noise, and zodiacal background. This model has been refined to incorporate recent advances in instrument design (as elaborated in I. Stotesbury et al. 2024). Detailed noise contributions for each target simulated in this study are presented in Appendix A.¹ It is worth noting that the scope of this study covers only radiometric simulation, without detailed time-domain simulation, which will be addressed in a forthcoming paper. However, an average jitter value derived from ground tests is applied. The assumed instrument parameters are summarized in Table A1 of Appendix A.

The radiometric tool calculates realistic noise estimates, which in turn allows determination of the number of observations necessary to reach a specific SNR through observation stacking.

The cumulative SNR achieved by combining multiple observations can be expressed as:

$$\text{SNR}_N = \sqrt{N} \times \text{SNR}_1, \quad (1)$$

where N is the number of observations (each corresponding to a single transit or eclipse event), and SNR_1 is the signal-to-noise ratio (SNR) of a single observation. The signal is derived from the amplitude (A_p) of spectral features in transit spectra, approximated by:

$$A_p = \frac{2R_p \times z}{R_s^2}. \quad (2)$$

Here, R_p and R_s represent the planetary and stellar radii, respectively, and z denotes atmospheric thickness. The atmospheric thickness is typically estimated as nH , with n ranging from 3 to 5 for hydrogen/helium (H_2/He)-dominated atmospheres. The scale height H is defined by:

$$H = \frac{k_B T_{eq}}{\mu g}, \quad (3)$$

where k_B is Boltzmann's constant, T_{eq} the equilibrium temperature of the planet, g the planet's surface gravity, and μ the mean molecular weight of the atmosphere (usually assumed as 2.3 for H_2/He atmospheres but refined using recent *JWST* data in this study).

Previously, due to limited observational constraints, atmospheric thickness (z) was set as 1, 3, or 5 scale heights, with 5 scale heights being a commonly used assumption (e.g. B. Edwards et al. 2019c; L. Booth et al. 2024). However, with the aid of recent data release from *JWST* and its overlapping wavelength coverage with *Twinkle*, constraints on atmospheric composition have improved, enabling more precise estimates of atmospheric scale height and thickness above the optically thick continuum. Thus, we can derive a more realistic SNR and resolution needed for *Twinkle* to better retrieve the atmospheric chemical composition, and consequently, determine the number of observations required to achieve a given SNR. It is worth noting that the radiometric model employed in this study calculates

¹Appendix A summarizes the random (photon, dark-current, and read) and systematic (pointing jitter and zodiacal background) noise terms included in the model, together with the additional systematic effects (gain noise, detector drift, thermal variations) deferred to a future time-domain simulation study.

Table 2. Prior distributions adopted in the free-chemistry retrievals. All molecular log-abundances share a common LogUniform prior; planet-specific parameters are listed separately. R_p is expressed as a multiplicative factor of the nominal planetary radius.

Parameter	Prior type	Range
<i>Common to all planets</i>		
All molecular log-VMRs	LogUniform	$[10^{-12}, 10^{-1}]$
<i>HD 209458 b</i>		
T (K)	Uniform	[100, 3000]
R_p factor	Uniform	[0.5, 2.0]
$\log P_{\text{cloud}}$ (Pa)	LogUniform	$[10^{-2}, 10^7]$
<i>WASP-107 b</i>		
T (K)	Uniform	[100, 2000]
R_p factor	Uniform	[0.2, 2.0]
$\log P_{\text{cloud}}$ (Pa)	LogUniform	$[10^{-2}, 10^5]$
<i>GJ 3470 b</i>		
T (K)	Uniform	[10, 1500]
R_p factor	Uniform	[0.2, 2.0]
$\log P_{\text{cloud}}$ (Pa)	LogUniform	$[10^0, 10^6]$
<i>55 Cnc e</i>		
T_{surface} (K)	Uniform	[1000, 4000]
T_{top} (K)	Uniform	[1000, 2000]
R_p factor	Uniform	[0.01, 0.5]
$\log P_{\text{surface}}$ (Pa)	LogUniform	$[10^3, 10^8]$
$\log P_{\text{top}}$ (Pa)	LogUniform	$[10^0, 10^5]$

Table 3. Predicted and realistic number of transits needed for *Twinkle* to achieve a specific SNR for various planets. Bold value column indicates the number of transits needed to achieve the updated SNR from *JWST*, while the *Italic* values indicate assumptions that is closest to the *JWST* observation. Real numbers are shown in the brackets and rounded integers are shown next to them. H in the brackets following Predicted Transits is the demonstrate scale height in each planet.

Planet	Median SNR	Predicted transits (5H)	Predicted transits (3H)	Predicted transits (2H)	Realistic transits
HD 209458 b	5	1 (0.18)	<i>1 (0.51)</i>	2 (1.21)	1 (0.85)
	7	1 (0.35)	<i>2 (1.04)</i>	3 (2.6)	2 (1.77)
	10	1 (0.75)	<i>3 (2.32)</i>	7 (6.79)	5 (4.24)
WASP-39 b	5	3 (2.26)	7 (6.31)	15 (14.31)	3 (2.11)
	7	5 (4.45)	13 (12.45)	29 (28.45)	5 (4.14)
	10	10 (9.12)	26 (25.74)	60 (59.84)	9 (8.48)
WASP-80 b	5	3 (2.01)	6 (5.7)	14 (13.41)	8 (7.49)
	7	4 (3.98)	12 (11.56)	29 (28.5)	16 (15.35)
	10	9 (8.33)	26 (25.44)	71 (70.86)	35 (34.67)
WASP-107 b	5	1 (0.16)	1 (0.45)	2 (1.02)	2 (1.31)
	7	1 (0.32)	1 (0.88)	2 (2.00)	3 (2.59)
	10	1 (0.65)	2 (1.81)	5 (4.15)	6 (5.4)
GJ 3470 b	3	1 (0.59)	2 (1.65)	4 (3.77)	4 (3.58)
	5	2 (1.64)	5 (4.69)	12 (11.09)	11 (10.49)
	7	4 (3.27)	10 (9.55)	24 (23.85)	23 (22.45)

Twinkle's observational errors assuming complete transit/eclipse coverage. However, observational availability of targets are influenced by Earth obstruction due to *Twinkle*'s orbital configuration, causing *Twinkle* to sometimes capture partial transits or eclipse events, similar to the situation encountered by CHEOPS (W. Benz et al. 2021). The extent of transit or eclipse coverage depends on the target's declination, the satellite's orbital phase, and the ratio of the planet's orbital period to *Twinkle*'s orbital period. Partial transits can still yield scientifically useful data but at reduced observing efficiency compared with full-coverage events. A detailed quantification of this efficiency factor, incorporating the effects of Earth obstruction and target-specific orbital geometry, will be addressed in a forthcoming time-domain simulation study.

While for some exoplanets, the number of required observations is high, it is important to note that lowering spectral resolution can increase the SNR per wavelength bin and thus achieve the desired SNR with fewer stacked observations. Although post-observation rebinning is widely used (e.g. L. Kreidberg et al. 2014; K. B. Stevenson et al. 2014), artificially reducing resolution in simulations introduces a loss of spectral information along the wavelength axis, particularly around spectral peaks. To avoid this, this study retrieves

spectra at *Twinkle*'s full native resolution. After launch, rebinning applied to real observational data – rather than to simulated spectra – will mitigate this limitation, as it preserves the full information content of the original observations.

2.2 *Twinkle*'s target candidates

Twinkle will be placed in a 1200 km, sun-synchronous low-Earth orbit, with a 6 am Local Time of Ascending Node (LTAN). The spacecraft's observational capabilities are restricted by the Field of Regard (FoR) (a 40 degree cone around the anti-Sun vector), solar and lunar exclusion constraints and Earth obstruction across different seasons (I. Stotesbury et al. 2022).

To systematically assess target observability, we employed the BSSL *Twinkle* Radiometric Tool to simulate observational noise for all confirmed exoplanets within *Twinkle*'s FoR possessing well-defined stellar and planetary parameters. Planetary and stellar data from the NASA Exoplanet Archive are taken as input into the radiometric tool to simulate noise levels for single transits or eclipses at *Twinkle*'s maximum resolving power. For *TESS* Objects of Interest (TOIs) lacking direct planetary mass measurements and thus not able to be processed with Radiometric Tool, masses were estimated using the *Forecaster* tool (J. Chen & D. Kipping 2017). Approximately 200 TOIs within *Twinkle*'s FoR were subsequently excluded due to unavailable stellar mass data. Additionally, in the giant-planet regime (radii $\gtrsim 8 R_{\oplus}$) the mass–radius relation is strongly degenerate, so *Forecaster*'s initial mass-radius relationships for gas giants proved inconsistent with current empirical data, with the predicted masses poorly constrained and not reproducing the observed distribution of known transiting giants. Hence, planetary masses for these TOIs were recalculated based on a lognormal distribution fitted to known transiting exoplanet populations, as detailed in Fig. C1 and further described in Appendix C. TOIs lacking sufficient stellar or planetary data for reliable mass estimation were excluded from the candidate list.

Our primary focus was on exoplanets with H_2/He dominated atmospheres; thus, we assumed a mean molecular weight of 2.3 for initial assessments. Atmospheric thicknesses equivalent to 3 atmospheric scale heights were adopted uniformly for all candidate targets. Exoplanets predicted to achieve median SNR levels of 3, 5, or 7 within 30 transits or eclipses at *Twinkle*'s native spectral resolution were considered suitable candidates for detailed atmospheric characterization. The threshold of 30 observations represents an assumed upper limit per target, determined by *Twinkle*'s total available observing time, the number of targets in the survey, and the accessibility constraints imposed by the satellite's orbit over the mission lifetime. This choice of upper limit is an operational planning decision rather than a constraint imposed by *Twinkle*'s instrumental capabilities.

2.3 Simulated exoplanets and spectral retrievals

The detection of molecular species and characterization of chemical compositions and abundances in exoplanet atmospheres remain central objectives of transit spectroscopy. Achieving these goals offers critical insights into planetary processes, formation histories, and potential habitability. *Twinkle*'s extensive and continuous spectral coverage from optical to near-infrared wavelengths includes absorption features for many previously detected and theoretically predicted molecules. In this study, we conduct atmospheric retrieval simulations on selected exoplanets (listed in Table 1), chosen from *Twinkle*'s candidate list based on the availability of *JWST* data. These planets represent a diverse range of types, aligning with *Twinkle*'s broader exoplanetary science themes, and facilitating a comprehensive evaluation of *Twinkle*'s capabilities informed by recent *JWST* findings.

We employed the open-source retrieval code `TauREx 3.2.2` (A. F. Al-Refaie et al. 2021) to simulate forward models for transmission and emission spectra and to perform spectral retrieval analyses. Stellar spectra for each target star were simulated using the PHOENIX stellar library (T. O. Husser et al. 2013). Chemical profiles were generated using the FastChem equilibrium chemistry plugin (J. W. Stock et al. 2018) for equilibrium assumptions, supplemented by TauREx's free chemistry retrieval mode. In this mode, the volume mixing ratio of each molecule is treated as a free parameter, constant with altitude, and retrieved independently without imposing chemical equilibrium constraints (A. F. Al-Refaie et al. 2021). While this approach does not explicitly model disequilibrium chemistry, it provides a model-agnostic framework that can capture departures from equilibrium, such as those arising from vertical quenching or photochemistry. All atmospheric models assumed a plane-parallel geometry comprising 100 layers, including molecular absorption, Rayleigh scattering, grey clouds, and collision-induced absorption (CIA). For all exoplanets, CIA processes involving H_2 - H_2 and H_2 -He were considered, with additional contributions from N_2 - N_2 and O_2 - O_2 specifically included for 55 Cnc e. Molecular opacity cross-sections were sourced from the ExoMol data base (J. Tennyson et al. 2024), while CIA data were taken from the HITRAN data base (T. Karman et al. 2019).

Retrieval analyses were performed using the Bayesian MultiNest algorithm (F. Feroz et al. 2009), employing 1000 live points and an evidence tolerance threshold of 0.5. Table 2 summarizes the prior distributions adopted for all retrieval parameters. Non-informative priors for molecular abundances were uniformly set on a log scale within the range of 10^{-12} to 10^{-1} , while other planetary parameters utilized priors constrained by previous literature of specific planets.

To quantify the detectability of individual molecular species, we employ the Likelihood Ratio Test (LRT) following the methodology described by M. Tesseney et al. (2013). To guard against unreliable maximum *a posteriori* (MAP) estimates for poorly constrained parameters, we perform the LRT using both the MAP and posterior median values; a molecule is classified as a 'Detection' only if both tests independently exceed the 3σ threshold. The full LRT procedure is detailed in Appendix D. The more conservative (lower) detection significance is reported in the results tables; molecules below the 3σ threshold are classified as 'Upper Limit'.

To incorporate recent *JWST* observational findings, chemical compositions and planetary parameters for each simulated exoplanet were based on best-fitting solutions reported in recent literature, alongside equilibrium chemistry predictions. Forward modelling and retrieval analyses were performed based on simulated transmission and emission spectra incorporating predicted *Twinkle* observational errors.

For HD 209458 b, recent *JWST* observations by Q. Xue et al. (2024) were combined with *HST* data from A. Saba et al. (2025). The simulated spectrum was aligned with the observed data by applying a vertical offset to reconcile differences in absolute transit depth calibration between the *HST* and *JWST* instruments at overlapping wavelengths. Molecular abundances and temperature profile parameters were then fine-tuned to optimize the fit to the combined *HST*+*JWST* data set. Similarly, to assess *Twinkle*'s capabilities for WASP-107 b, a simulated spectrum covering *Twinkle*'s wavelength coverage (0.5–4.5 μm) was constructed using *HST* and *JWST* NIRSpec data as benchmarks. The atmospheric model included known molecular constituents identified by D. K. Sing et al. (2024), specifically H_2O , CH_4 , NH_3 , CO , CO_2 , SO_2 , and H_2S , alongside potential SiO detection proposed by S. Ma et al. (2025), and optical absorbers Na and K, with uniform abundances assumed.

For GJ 3470 b, we simulated *Twinkle*'s performance by closely aligning spectral data with *JWST* and *HST* observations across the 0.5–4.5 μm range. The atmospheric composition model incorporated nine primary molecules (H_2O , CH_4 , CO_2 , CO , SiO, SO_2 , NH_3 , HCN, H_2S), with atmospheric temperature fixed at 600 K and a deep cloud deck placed at 3000 Pa to highlight molecular absorption signatures. For the emission spectra of 55 Cnc e, we employed atmospheric assumptions consistent with R. Hu et al. (2024), featuring predominant CO_2 and CO composition alongside trace species (C_2H_2 , HCN, O_2 , and N_2) with constant mixing ratios. Given *JWST*'s spectral coverage starting around 4 μm , only select *JWST* NIRCам data points were used as references for the simulated *Twinkle* spectra.

To evaluate the retrievability of minor atmospheric species whose spectral signatures are typically obscured by clouds or dominated by major molecules, we conducted a dedicated sensitivity study. Using WASP-107 b as the test case, we systematically amplified the abundances of minor species (CH_4 , SO_2 , NH_3 , and H_2S) to identify the enhancement factor necessary for their spectral features to become clearly observable above simulated observational uncertainties. The baseline scenario employed the previously established retrieval setup for WASP-107 b, stacking six transits to achieve a median SNR of 10. For each molecule, the abundance was scaled from its baseline value by amplification factors ranging from $\times 5$ to $\times 80$ (the specific factors tested for each species are listed in Table 8), while all other atmospheric parameters were held fixed at their baseline values. Independent retrieval analyses were then performed at both SNR = 5 and SNR = 10 to determine the minimum amplification factor required for a constrained retrieval of each species.

3 RESULTS

3.1 Model assumptions: lessons learned from *JWST*

With recent advances resulting from *JWST* observations significantly refining our understanding of exoplanetary atmospheres, previous assumptions regarding atmospheric thickness and composition require reassessment. *JWST* data have provided critical updates, particularly regarding atmospheric composition and the prevalence of clouds, which directly impact observational strategies and the SNR achievable by *Twinkle*. Table 3 summarizes our revised estimates of the number of transits required by *Twinkle* to reach specific SNR targets (3, 5, 7, and 10) at native resolution after incorporating recent *JWST* findings. We compare these updated estimates with earlier calculations that assumed atmospheric thicknesses of 2, 3, and 5 scale heights. The updated calculations assume ideal observational conditions – complete transit visibility without interruptions – thus serving as lower bound estimates. These preliminary outcomes, integrating *JWST* observational results, will continue to evolve as additional *JWST* data become available.

The realistic observation requirements based on recent *JWST* constraints are marked in bold in Table 3, while atmospheric thickness predictions closest to the *JWST*-derived reality are highlighted in italic. Initial analyses indicate that Jupiter-sized exoplanets (e.g. HD 209458 b, WASP-39 b, WASP-80 b) typically have observable atmospheric thicknesses ranging from approximately 2.5 to 5 scale heights. In contrast, Neptune-sized exoplanets (e.g. WASP-107 b, GJ 3470 b) align better with atmospheric thicknesses around two scale heights. These findings broadly confirm previous assumptions but now offer significantly tighter constraints, improving observational efficiency.

Furthermore, *JWST* data confirm that cloud coverage significantly reduces observable atmospheric thickness. Most H_2 /He-rich planets studied thus far display observable atmospheric features between 2 and 3 times their scale height. Notably, exceptions such as WASP-39 b exhibit significantly extended atmospheric signals, a result directly linked to its particularly low planetary density. These refined assumptions provide a robust foundation for the detailed retrieval studies and sensitivity analyses of individual exoplanets presented in subsequent sections.

3.2 Candidate list for atmospheric characterization survey

Among the confirmed exoplanets, approximately 1070 transiting exoplanets lie within *Twinkle*'s FoR. Of the TOIs, 2009 candidates fall within *Twinkle*'s FoR. Fig. 1 illustrates the right ascension (RA) and declination (Dec) of all confirmed exoplanets and TOIs distribution in the *Twinkle*'s FoR.

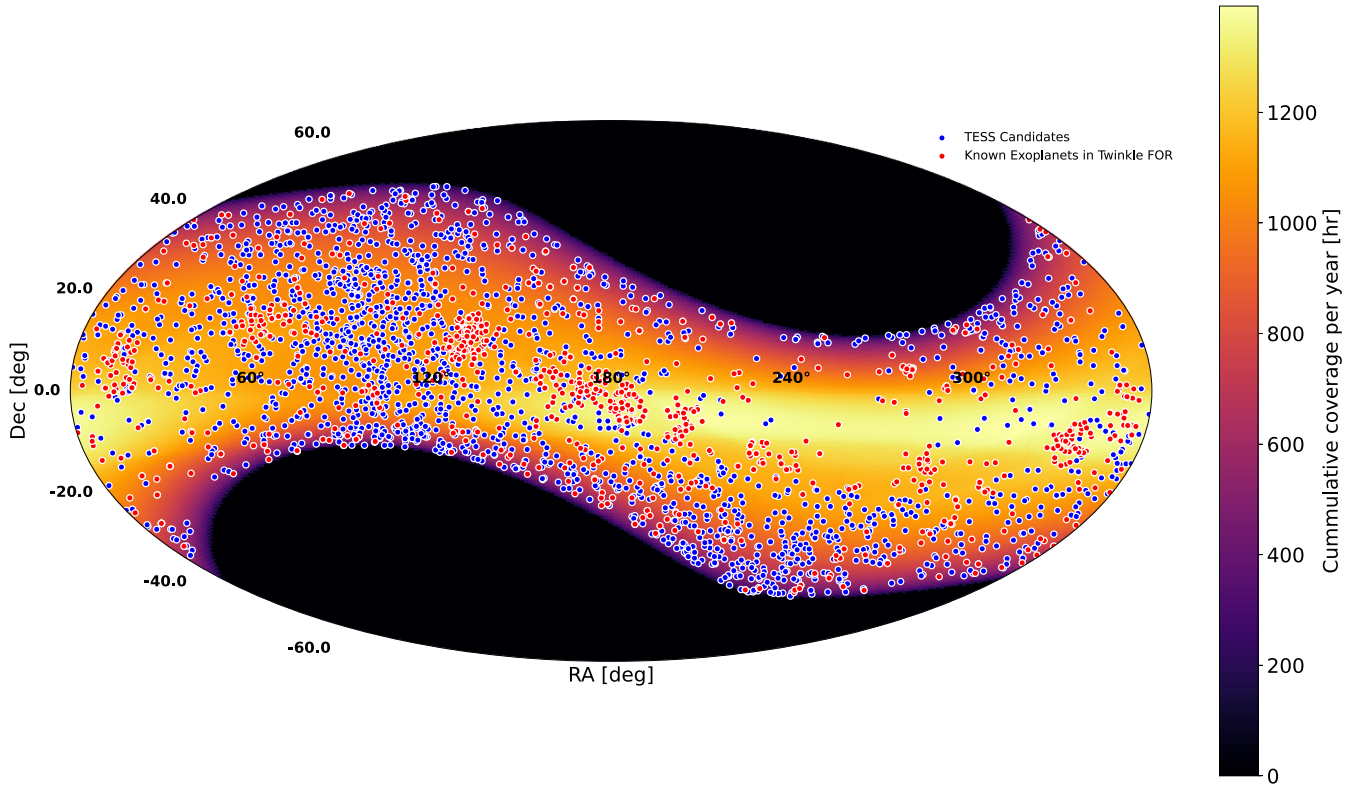


Figure 1. The sweep of *Twinkle*'s FoV and location of currently known exoplanets (red) and *TESS* objects of interest (blue).

After processing all targets through the radiometric model, a combined set of known exoplanets and TOIs capable of achieving median SNR thresholds (3, 5, or 7) within 30 observations at native resolution is presented in Fig. 2.

This initial candidate list used for the observability study for *Twinkle*'s atmospheric characterization survey is specifically tailored to H_2/He -rich planets, due to current uncertainties surrounding the atmospheric composition of smaller planets. Recent observations of terrestrial planets have highlighted significant diversity, emphasizing the need for further study with dedicated future missions. Consequently, terrestrial planets have not been extensively analysed in this observability study. Further refinement of the candidate list will be informed by complementary observations from *JWST*, *PLATO*, *CSST*, and *Earth 2.0*, enabling tailored observational strategies aligned with the scientific objectives of the *Twinkle* Science Team. The candidate list presented in this work will be made publicly available to support the broader exoplanet community in planning complementary observations.

3.3 Spectral retrievals: Gas giants

3.3.1 HD 209458 b

Simulated spectra for HD 209458 b with *Twinkle*'s predicted uncertainties after fitting with combined set of *HST* WFC3 and *JWST* are shown in Fig. 3. Retrievals conducted under chemical equilibrium assumptions indicate that, even with a single transit, abundances of major molecules can be accurately recovered. However, minor molecular species such as CH_4 , SO_2 , and HCN exhibit weaker constraints, introducing uncertainties in metallicity and C/O ratio estimates (Fig. 4). Increasing observations to 10 transits significantly enhances the accuracy of CH_4 and HCN abundance constraints, leading to more robust metallicity and C/O ratio estimates (Fig. 4). This improvement mainly results from the distinct and abundant H_2O absorption due to the assumed low C/O ratio, prominent TiO features at optical wavelengths, and clearly defined CO_2 features at the limit of *Twinkle*'s coverage. Additional planetary parameters, including radius, atmospheric temperature, and cloud-top pressure, were also accurately recovered within 1σ .

To provide a more comprehensive assessment, we performed a free retrieval analysis without assuming chemical equilibrium by selecting the eleven most critical molecules identified in previous simulations, along with species suggested by Q. Xue et al. (2024). A constant vertical mixing ratio (VMR) profile based on equilibrium abundances and suggestions from Q. Xue et al. (2024) was used for these molecules in a free retrieval analysis. Results in Fig. B1 confirm robust retrieval of H_2O and TiO , whereas other molecular signatures remain masked by the dominant H_2O absorption and cloud opacity. *Twinkle*'s extensive spectral coverage from optical to infrared allows for clear identification of TiO absorption features, while VO remains below the detection threshold ($\text{LRT} \sim 2\sigma$; Table 4) despite showing a posterior shift from the prior. However, due to the relatively low assumed abundance of CO_2 (10^{-6}) and increased noise at longer wavelengths, the CO_2 feature was only partially detected and thus less reliably constrained.

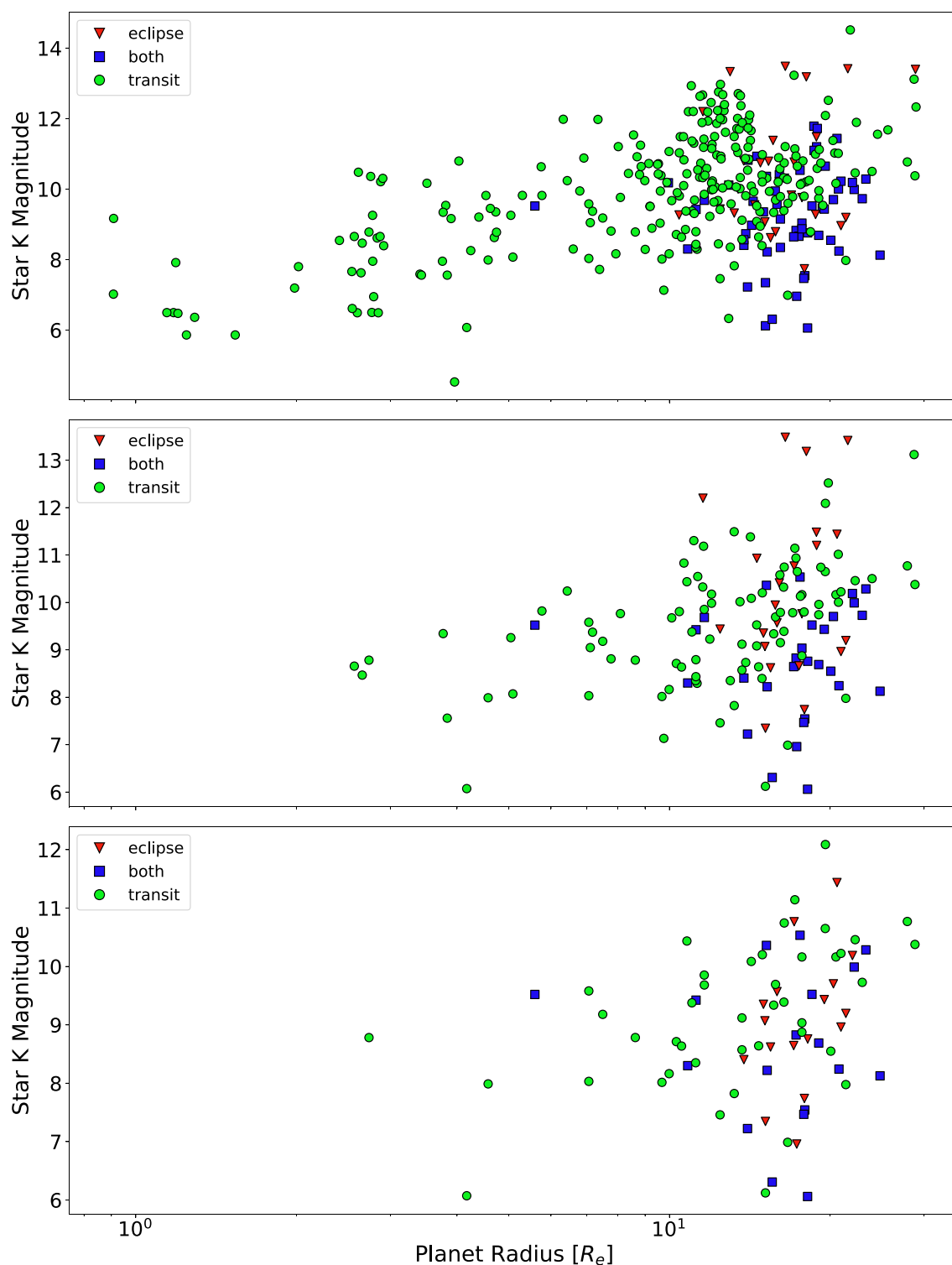


Figure 2. Known planets and TOIs capable of achieving median SNR thresholds within 30 observations at native resolution. They are distributed based on planet radius versus host star K magnitude. Different shapes demonstrate the preferred type of observation. Top Panel: $\text{SNR} \geq 3$; Middle Panel: $\text{SNR} \geq 5$; Bottom Panel: $\text{SNR} \geq 7$.

Retrieval comparisons at different SNR values (5, 7, and 10) for HD 209458 b, as illustrated in Fig. B1, indicate improved parameter constraints for strongly featured molecules such as H_2O and TiO at higher SNR. Conversely, parameters not adequately retrieved at an SNR of 5 exhibit limited improvement at higher SNR levels. This limitation of improving SNR and difficulty of retrieving minor species in HD 209458 b likely arises from spectral masking by prominent water absorption and simplified grey cloud assumptions, rather than reflecting *Twinkle*'s intrinsic capabilities.

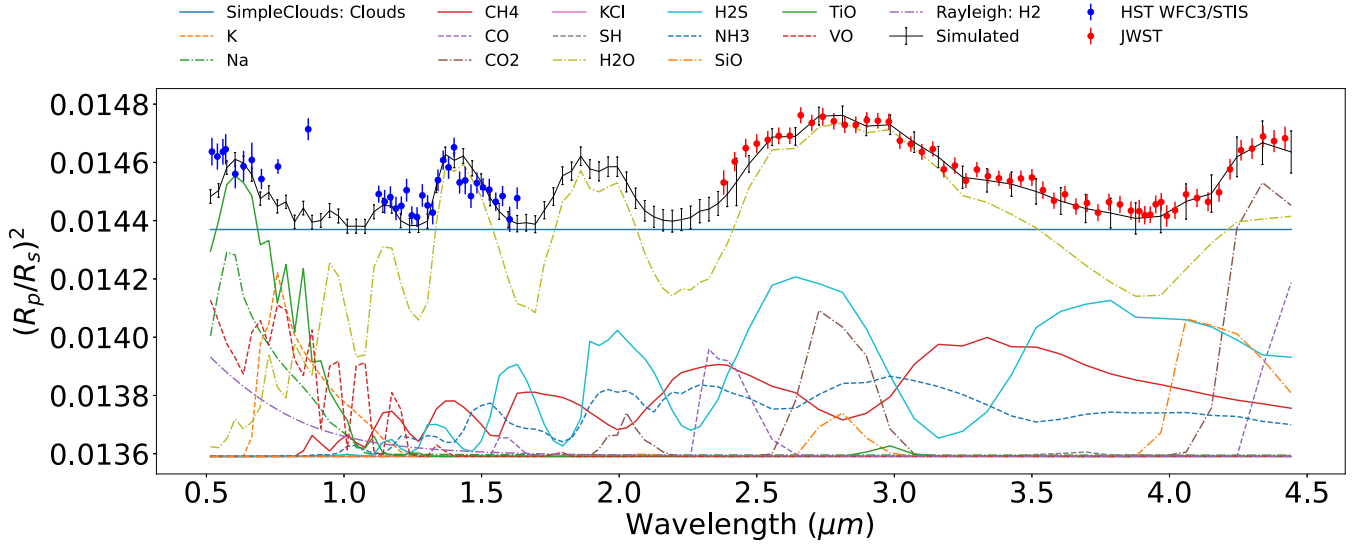


Figure 3. Comparison of the *JWST* and *Twinkle*-simulated spectrum for HD 209458 b, highlighting the spectral contributions of various molecular absorbers and scattering effects. The error bars on the spectrum are based on simulations of 10 transits, and the spectra are fitted under the assumption of chemical equilibrium. For clarity, spectral contributions from effects that are insignificant have been omitted. The *HST* data points have been shifted downwards by 100 ppm.

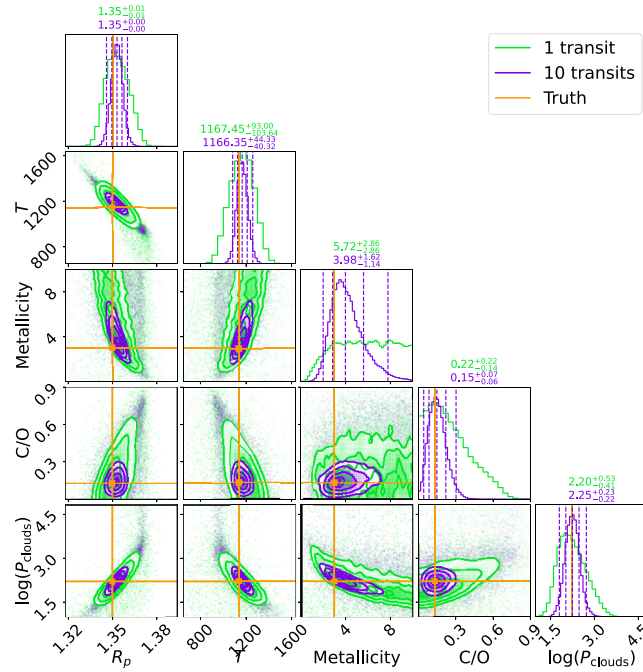


Figure 4. Posterior distributions of retrieved parameters from retrieval analysis of HD 209458 b assuming chemical equilibrium. Green: 1 transit; Purple: 10 transits.

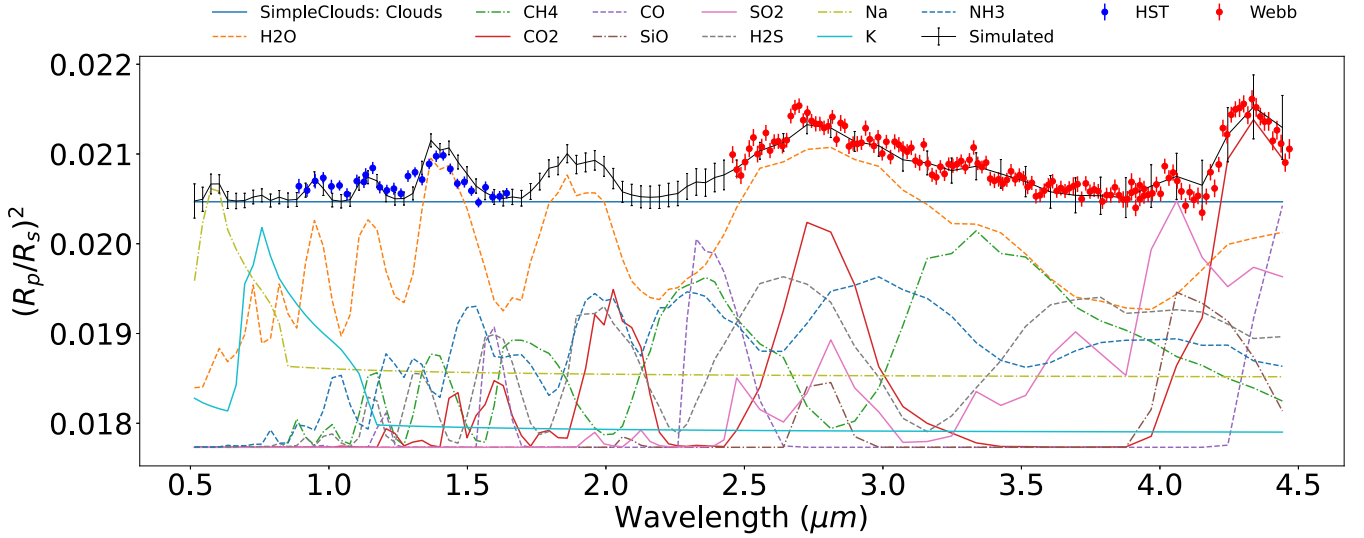
3.3.2 WASP-107 b

Simulated spectrum for WASP-107 b with *Twinkle*'s predicted error after fitting with available *HST* WFC3 and *JWST* NIRCcam from L. Welbanks et al. (2024) are shown in Fig. 5. Calculations shown in Table 3 suggest the atmospheric thickness of WASP-107 b is slightly less than two scale heights, likely due to high-altitude cloud cover. To reflect this observational constraint, our simulation adopted a high cloud altitude at 70 Pa.

Posterior distributions for WASP-107 b in Fig. B2 revealed significant degeneracies between the planetary radius and cloud-top height parameters, with the radius typically underestimated and cloud height overestimated in retrievals. Nevertheless, robust constraints were obtained for the equilibrium temperature and the abundances of H_2O , CO_2 , and Na. The retrieved parameter values are compared with the input truth values in Table 5. The abundances of NH_3 and CH_4 remain undetectable due to their lower contribution

Table 4. Comparison of truth values and retrieved parameters for HD 209458 b with 1 transit (SNR = 5).

Parameter	Truth	Retrieved (SNR = 5)	LRT (σ)
Planet radius (R_p)	1.35	$1.34^{+0.01}_{-0.01}$	
Temperature (K)	1140	1205^{+118}_{-122}	
Detection			
$\log(X_{\text{H}_2\text{O}})$	-2.7	$-2.25^{+0.57}_{-0.62}$	15.18
$\log(X_{\text{TIO}})$	-8	$-7.51^{+0.61}_{-0.63}$	8.99
Upper Limit			
$\log(X_{\text{VO}})$	-9	$-9.31^{+0.93}_{-1.42}$	2.05
$\log(X_{\text{CO}_2})$	-6	$-7.63^{+2.20}_{-2.80}$	1.20
$\log(X_{\text{CO}})$	-3.4	$-7.39^{+3.33}_{-2.96}$	< 0.01
$\log(X_{\text{H}_2\text{S}})$	-5	$-7.57^{+2.78}_{-2.82}$	< 0.01
$\log(X_{\text{NH}_3})$	-5.3	$-7.99^{+2.47}_{-2.60}$	< 0.01
$\log(X_{\text{HCN}})$	-6	$-8.28^{+2.52}_{-2.44}$	0.10
$\log(X_{\text{CH}_4})$	-8	$-8.57^{+2.25}_{-2.20}$	0.05
$\log(X_{\text{C}_2\text{H}_2})$	-8	$-8.30^{+2.25}_{-2.34}$	0.12
$\log(P_{\text{clouds}} \text{ (Pa)})$	2.11	$1.63^{+0.59}_{-0.53}$	

**Figure 5.** Comparison of *JWST* transmission spectrum and *Twinkle*-simulated spectrum, with spectral contributions of various molecule absorbers and scattering effects for WASP-107 b. Error bar on spectrum is simulated with observation of six transits (SNR = 10). The spectra is fitted assuming free chemistry. The *JWST* data points were offset upward by 300 ppm.

compared with cloud opacity within *Twinkle*'s spectral coverage as in Fig. 5. The primary absorption band for CO (4.5–5.0 μm) lies outside *Twinkle*'s spectral range, making CO undetectable. The spectral signature of SO₂ at 4.0 μm is too weak to be reliably retrieved, primarily due to interference from cloud opacity and observational noise.

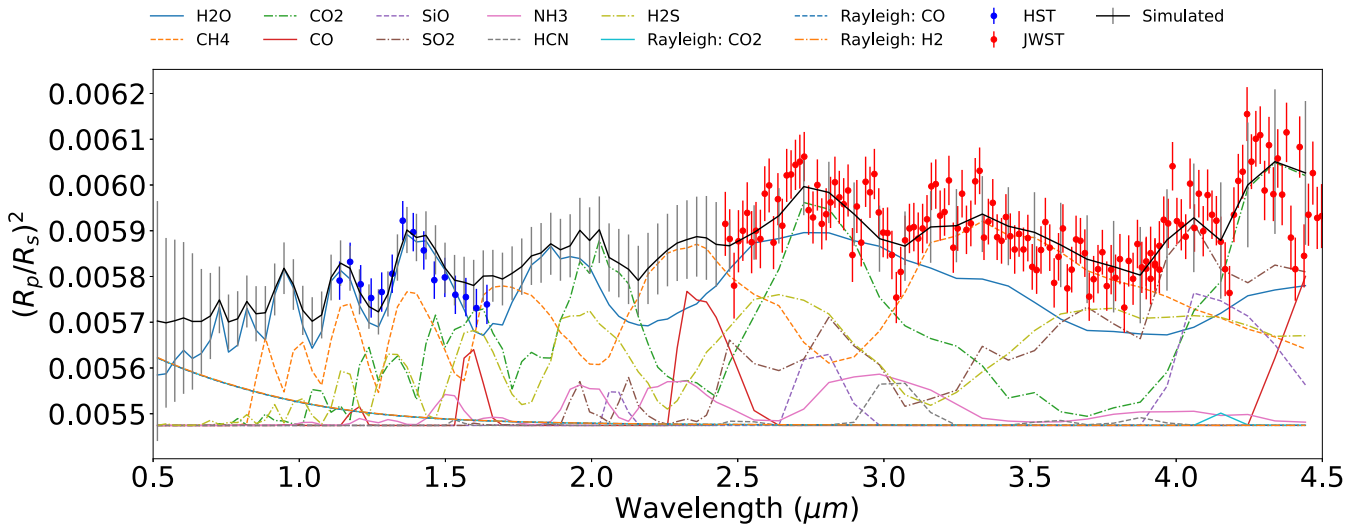
Comparative retrieval results at different SNR values showed that parameters clearly observable in the spectra, such as H₂O and CO₂, exhibited tighter constraints with increasing SNR. Parameters initially unconstrained at an SNR of 5 but exhibiting spectral contributions stronger than cloud opacity – such as Na – became retrievable at higher SNR values (SNR = 10). Furthermore, while CO₂ could be constrained at all simulated SNR levels, increasing the SNR to 7 or 10 significantly improved the accuracy and reliability of its retrieval.

3.3.3 GJ 3470 b

Simulated spectra of GJ 3470 b with *Twinkle*'s predicted error after fitting with available *HST* WFC3 and *JWST* NIRCcam from A. Saba et al. (2025) and T. G. Beatty et al. (2024) are shown in Fig. 6. The contribution plot illustrates distinct spectral features from key molecules. In addition to strong absorption by H₂O, significant CO₂ features at 2.0, 3.8, and 4.4 μm allow robust constraints on its

Table 5. Comparison of truth values and retrieved parameters (RP) for WASP-107 b stacking 2 transits (SNR = 5).

Parameter	Truth	RP (SNR = 5)	LRT (σ)
Planet radius (R_p)	0.933	0.90 ± 0.02	
Temperature (K)	700	689^{+91}_{-88}	
Detection			
$\log(X_{\text{H}_2\text{O}})$	-2	$-2.07^{+0.51}_{-0.64}$	13.39
Upper Limit			
$\log(X_{\text{CO}})$	-1.4	$-7.35^{+3.08}_{-2.81}$	< 0.01
$\log(X_{\text{CO}_2})$	-4.52	$-5.30^{+1.17}_{-2.52}$	1.35
$\log(X_{\text{Na}})$	-4	$-7.03^{+2.65}_{-2.92}$	0.07
$\log(X_{\text{CH}_4})$	-5.7	$-8.35^{+2.12}_{-2.24}$	< 0.01
$\log(X_{\text{H}_2\text{S}})$	-4.3	$-7.62^{+2.70}_{-2.69}$	< 0.01
$\log(X_{\text{K}})$	-6	$-8.46^{+2.31}_{-2.19}$	0.08
$\log(X_{\text{NH}_3})$	-5.7	$-8.09^{+2.25}_{-2.41}$	< 0.01
$\log(X_{\text{SO}_2})$	-5	$-8.13^{+2.47}_{-2.32}$	< 0.01
$\log(X_{\text{SiO}})$	-4	$-7.18^{+3.21}_{-3.05}$	0.01
$\log(P_{\text{clouds}} \text{ (Pa)})$	1.85	$1.49^{+0.62}_{-0.47}$	

**Figure 6.** Comparison of *JWST* and *Twinkle*-simulated spectrum and spectral contributions of various molecule absorbers and scattering effects for GJ 3470 b. Error bar on spectrum is simulated with observation of 11 transits (SNR = 5). The spectra is fitted assuming free chemistry. The *HST* data points were offset downward by 200 ppm.

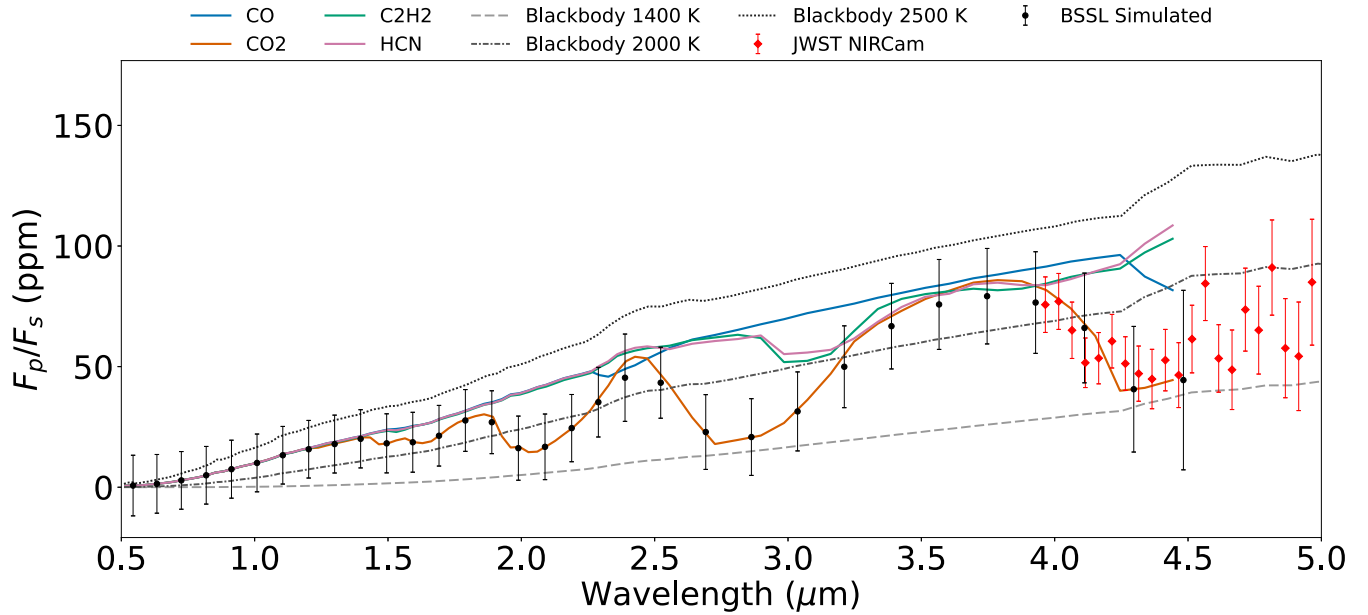
abundance. Prominent CH_4 absorption features at 1.6, 2.4, and 3.4 μm are also clearly discernible. A detectable but weaker absorption peak of SO_2 at 4.0 μm provides crucial insight into atmospheric disequilibrium chemistry.

Considering the observational challenges posed by GJ 3470 b's low transit depth and relatively flat spectrum, we evaluated retrieval outcomes at SNR levels of 3, 5, and 7, approximately 23 transits are required to reach an SNR of 7. The retrieval results, depicted in Fig. B3, demonstrate accurate constraints for parameters such as planetary radius and equilibrium temperature. Conversely, the cloud deck height remains unconstrained due to its depth, rendering it indistinguishable from other atmospheric components. The abundances of H_2O , CH_4 , and CO_2 were successfully retrieved at all SNRs considered. The retrieved values for the four-transit (SNR = 3) case are summarized in Table 6. However, SO_2 could only be effectively constrained at the highest SNR (SNR = 7). Minor molecular species such as CO , SiO , NH_3 , HCN , and H_2S were undetectable even at high SNR due to their weak absorption features within *Twinkle*'s spectral range.

Comparative retrieval results across different SNR scenarios yielded consistent conclusions with previous exoplanet cases. Parameters with strong spectral signatures above cloud opacity, such as H_2O , CO_2 , and CH_4 , benefit significantly from higher SNR, resulting in tighter constraints. Conversely, molecules like SO_2 , with a single prominent absorption feature near the spectral range boundary, require high SNR levels for accurate abundance retrieval.

Table 6. Comparison of truth values and retrieved parameters (RP) for GJ 3470 b stacking four transits (SNR = 3).

Parameter	Truth	RP (SNR = 3)	LRT (σ)
Planet radius (R_J)	0.36	0.36 ± 0.002	
Temperature (K)	600	650^{+235}_{-192}	
Detection			
$\log(X_{\text{H}_2\text{O}})$	−1	$-0.92^{+0.51}_{-0.89}$	5.03
$\log(X_{\text{CH}_4})$	−3.3	$-3.48^{+0.80}_{-1.75}$	3.37
$\log(X_{\text{CO}_2})$	−1.3	$-1.48^{+0.66}_{-2.98}$	9.52
Upper Limit			
$\log(X_{\text{SiO}})$	−2	$-4.62^{+3.29}_{-4.25}$	0.02
$\log(X_{\text{SO}_2})$	−3.3	$-3.84^{+2.44}_{-4.35}$	0.39
$\log(X_{\text{CO}})$	−1	$-5.03^{+3.48}_{-3.96}$	< 0.01
$\log(X_{\text{NH}_3})$	−7	$-7.46^{+2.61}_{-2.59}$	< 0.01
$\log(X_{\text{HCN}})$	−8	$-5.91^{+2.85}_{-3.50}$	< 0.01
$\log(X_{\text{H}_2\text{S}})$	−3	$-5.71^{+3.20}_{-3.43}$	< 0.01
$\log(P_{\text{clouds}} \text{ (Pa)})$	3.48	$4.27^{+0.95}_{-0.99}$	

**Figure 7.** Simulated emission spectrum of 55 Cnc e (black circles) with spectral contributions of individual molecular absorbers (coloured lines), three blackbody reference curves at 1400, 2000, and 2500 K (grey lines), and *JWST* NIRCcam observations (red diamonds; R. Hu et al. 2024). Error bars on the simulated data are for 10 eclipses with ch0 binned to one-third and ch1 binned to half of the native resolution; native resolution with 10 eclipses was used in retrieval.

3.4 Spectral retrievals: super-Earths

3.4.1 55 Cancri e

Fig. 7 shows a simulated spectrum for 55 Cnc e, along with *Twinkle*'s predicted error after fitting with available *JWST* NIRCcam data from R. Hu et al. (2024). This baseline scenario uses three blackbody curves representing temperatures of 1400, 2000, and 2500 K. Comparing these baselines with the simulated and observed data points reveals distinctive absorption features present in the emission spectrum. *Twinkle*'s expected performance, calculated by reducing the resolution of channel 0 to one-third and channel 1 to half their native resolutions and stacking ten eclipse observations, demonstrates comparable sensitivity to *JWST* within its spectral range. Fig. 7 reveals significant spectral divergence from blackbody emission beginning around 2.0 μm . Although the higher resolution of channel 0 below this wavelength is underutilized, the spectral range above 2.0 μm effectively captures prominent CO_2 absorption features. Additional CO_2 features are clearly identifiable within *Twinkle*'s channel 1 coverage. The contribution curves further highlight that CO_2 dominates the spectral signature throughout *Twinkle*'s wavelength range, overshadowing contributions from minor species such as C_2H_2 , CO, and

Table 7. Comparison of truth values and retrieved parameters (RP) for 55 Cnc e stacking 10 eclipses.

Parameter	Truth	RP (10 eclipses)	LRT (σ)
Planet radius (R_J)	0.167	$0.150^{+0.01}_{-0.02}$	
T_{surface} (K)	2300	2487^{+204}_{-169}	
T_{top} (K)	1500	1585^{+232}_{-273}	
$\log(P_{\text{surface}})$ (Pa)	5	$5.57^{+1.50}_{-1.01}$	
$\log(P_{\text{top}})$ (Pa)	2	$1.56^{+1.29}_{-0.98}$	
Detection			
$\log(X_{\text{CO}_2})$	−2	$-2.38^{+1.01}_{-1.19}$	5.38
Upper Limit			
$\log(X_{\text{CO}})$	−3	$-6.32^{+3.34}_{-3.43}$	0.01
$\log(X_{\text{HCN}})$	−5	$-7.92^{+2.49}_{-2.53}$	0.01
$\log(X_{\text{C}_2\text{H}_2})$	−5	$-8.17^{+2.53}_{-2.47}$	< 0.01

HCN. Notably, an increase in CO contribution is evident beyond 4.2 μm , suggesting enhanced detectability if *Twinkle*'s coverage could extend slightly further.

Using this simulated spectrum as input, we performed emission retrievals to evaluate *Twinkle*'s capability in constraining atmospheric properties. To avoid the loss of spectral information introduced by post-processing wavelength binning, retrievals were conducted at *Twinkle*'s native resolution, despite larger resultant uncertainties. Retrieval outcomes for scenarios involving 10, 20, and 30 stacked eclipse observations are presented in Fig. B4. The retrieved parameters for the 10-eclipse case are listed in Table 7. The results illustrate that, with just 10 eclipses, *Twinkle* effectively constrains the temperature structure and reliably retrieves the abundance of the dominant atmospheric species, CO_2 . Additional observations enhance parameter constraints incrementally but do not significantly improve the detection of dominant atmospheric constituents. This demonstrates *Twinkle*'s robust performance, particularly for bright super-Earth targets such as 55 Cnc e.

3.5 Minor species sensitivity study on WASP-107b

Contribution plots from simulations (Figs 5, 6, and 7) revealed that several minor species, such as CH_4 , SO_2 , NH_3 , and H_2S , typically possess abundances too low to yield distinguishable spectral features, especially in atmospheres dominated by water vapour absorption or obscured by clouds. Direct retrieval of these species from featureless continuum regions was thus found unreliable, as illustrated in posterior plots provided in the Appendix. Such limitations reflect realistic observational challenges faced when attempting to detect minor constituents even at higher SNR levels.

To better quantify these observational constraints, we assessed the amplification factors necessary for each minor species' spectral features to surpass the noise threshold of *Twinkle* observations at $\text{SNR} = 10$ (Comparison of the original and amplified spectra in Fig. 8). Our results show that CH_4 and SO_2 features become clearly observable when their abundances increase by approximately fivefold, corresponding to abundances of 1×10^{-5} and 5×10^{-5} , respectively. For NH_3 and H_2S , which have broader and weaker absorption bands overlapping significantly with water vapour, spectral signatures required approximately a twentyfold abundance increase (to 4×10^{-5} and 1×10^{-3}) before becoming discernible.

Individual retrieval analyses for each amplified abundance scenario confirmed that minor species, initially undetectable at baseline abundances, can indeed be reliably retrieved when their spectral signatures exceed noise thresholds (posterior distributions in Fig. 9). At *Twinkle*'s native resolution with $\text{SNR} = 10$, CH_4 and SO_2 could be constrained down to abundances of 1×10^{-5} and 5×10^{-5} , respectively. Lowering the SNR to 5 increased these retrieval thresholds to approximately 2×10^{-5} for CH_4 and 9×10^{-5} for SO_2 . NH_3 could be reliably constrained at an abundance of 4×10^{-5} at $\text{SNR} = 10$, requiring further amplification to 8×10^{-5} (fortyfold increase from baseline) at $\text{SNR} = 5$. H_2S retrieval was challenging at $\text{SNR} = 5$ due to its weaker, indistinct absorption features compared to water.

While the amplified spectra may appear visually consistent with the baseline within individual spectral bin error bars (Fig. 8), the LRT analysis captures the cumulative effect of systematic spectral differences across multiple bins simultaneously. The retrieval posteriors in Fig. 9 confirm that the amplified species are indeed constrained, validating the sensitivity thresholds identified by our analysis.

Table 8 reports the LRT detection significance for each amplification scenario. SO_2 achieves robust detection ($> 3\sigma$) at $\times 5$ amplification with SNR 10 and at $\times 9$ amplification (6.20σ) with SNR 5. NH_3 requires $\geq \times 10$ amplification at SNR 10, while H_2S requires $\geq \times 20$. CH_4 is not included in the table because it remains undetected by the relative LRT (full model versus model without CH_4) at all tested amplification levels. This non-detection arises from strong spectral degeneracy with H_2O , whose retrieved abundance ($\log \text{VMR} \approx -1.8$) produces broad near-infrared absorption that overlaps and masks the CH_4 features; the retriever can compensate for the removal of CH_4 by adjusting H_2O , yielding a low LRT statistic even when the CH_4 posterior itself is constrained (Fig. 9). However,

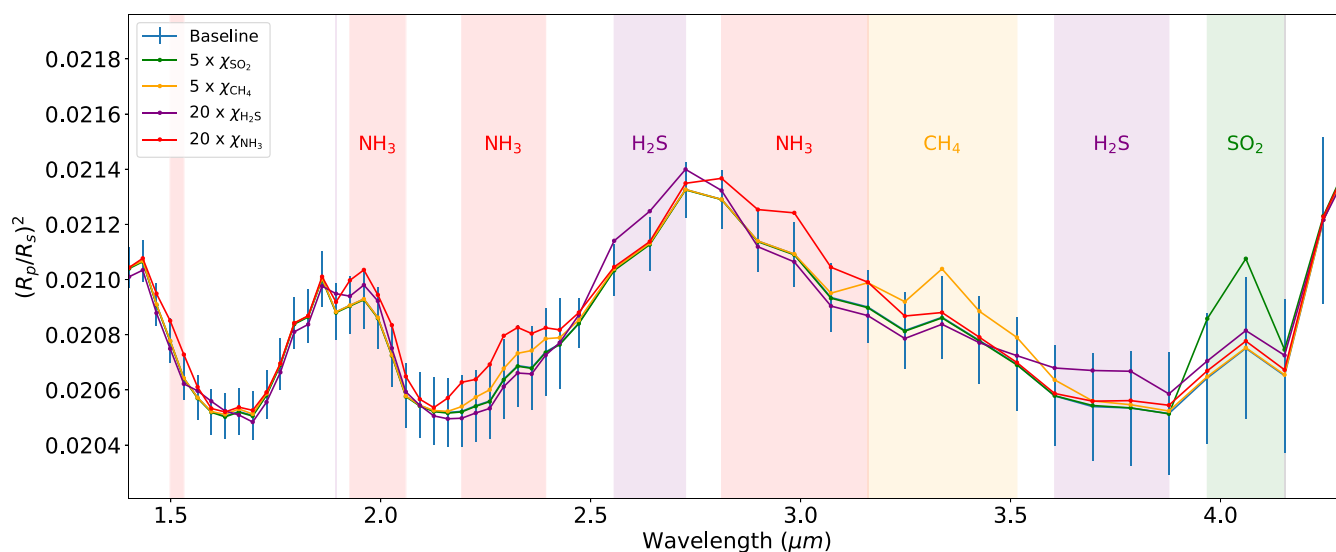


Figure 8. The resultant spectrum with amplifying abundance for CH_4 , SO_2 , NH_3 , and H_2S at $\text{SNR} = 10$. Magnifying factor are shown in the legend, assuring visibility in the spectrum. Errorbar on baseline spectrum is estimated with observation of six transits. The wavelength ranges where the molecular spectra contribute significantly after abundance amplification have been shaded for visual ease.

Table 8. LRT detection significance (σ) for minor species with amplified abundances in WASP-107 b. Values exceeding the 3σ detection threshold are shown in bold. ‘–’ indicates no retrieval was performed for that combination.

Molecule	Amplification	SNR 10	SNR 5
SO_2	baseline	< 0.01	< 0.01
	$\times 5$	4.69	0.12
	$\times 9$	–	6.20
NH_3	baseline	0.01	–
	$\times 5$	0.03	–
	$\times 10$	8.23	–
	$\times 20$	13.58	–
	$\times 40$	–	11.99
H_2S	$\times 50$	24.10	–
	baseline	< 0.01	–
	$\times 5$	0.17	–
	$\times 10$	2.42	–
	$\times 20$	3.87	–
	$\times 50$	24.44	–
	$\times 80$	–	< 0.01

an alternative LRT formulation using a featureless (flat-line) null hypothesis confirms that the CH_4 spectral signature is intrinsically detectable by BSSL, exceeding 3σ at $\geq \times 5$ amplification, consistent with the spectral amplification results shown in Fig. 8. SiO , whose spectral contribution is negligible in all scenarios, is likewise omitted from the table.

The goal of this sensitivity analysis was to isolate the impact of instrumental capabilities (resolution, SNR) from atmospheric factors such as inherently low abundances and optically thick cloud cover. By establishing clear thresholds for the detectability and retrievability of minor atmospheric constituents at varying levels of observational sensitivity, our results provide valuable guidelines for mission planners.

4 DISCUSSION

JWST’s observations provide a critical empirical foundation for simulating expected outcomes with future telescopes, particularly for targets already characterized by *JWST*. For targets yet to be observed, empirical relationships derived from *JWST* data – such as those relating to atmospheric mean molecular weight or number of scale heights (see Table 3) – offer valuable guidance for initial observational planning. These preliminary simulations are essential, as they enable mission teams to refine their observational strategies and allocate resources efficiently to best achieve their scientific goals. In this context, our study evaluating *Twinkle*’s capabilities for atmospheric characterization offers important initial insights into the mission’s anticipated performance in exoplanet surveys, leveraging the most recent discoveries from *JWST*. Our results demonstrate that, through the use of low-resolution spectroscopy and strategies involving

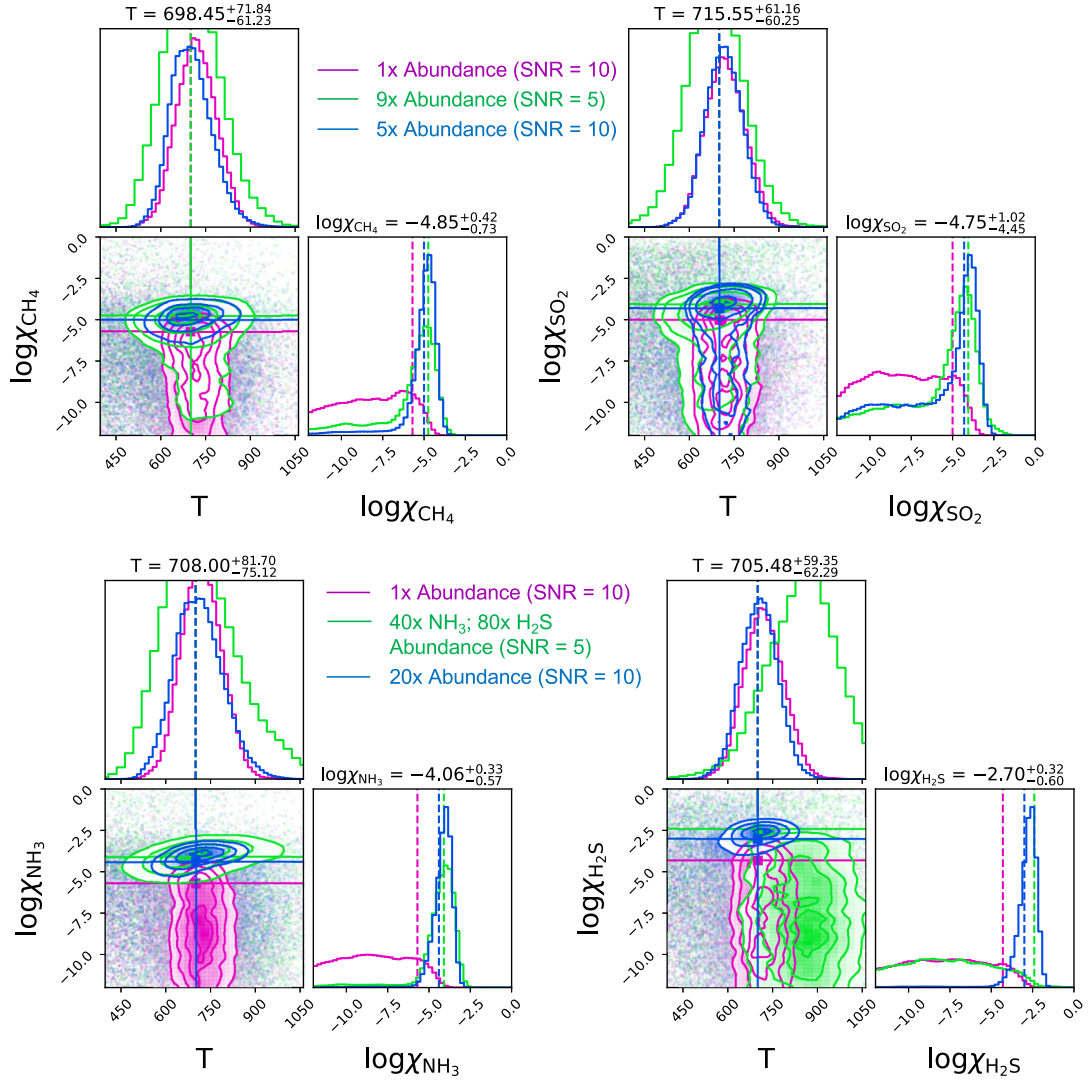


Figure 9. Posteriors for retrievals with amplified abundance for CH_4 , SO_2 , NH_3 , and H_2S at different SNR levels for WASP-107 b. Baseline (Pink) is same with SNR = 10 case in Fig. B2. Magnifying factor are shown in the legend, assuring visibility in the spectrum.

the stacking of multiple observations, *Twinkle*'s observations will effectively allow to determine planetary parameters and key chemical components for hundreds of exoplanets.

Focusing on transit spectroscopy of H_2/He -dominated atmospheres, we found that when a median SNR of 3 is achieved, key planetary parameters such as radius, atmospheric temperature, and cloud heights can be constrained within meaningful ranges (e.g. GJ 3470 b shown in Fig. B3). For chemical species identifiable in the spectra, retrievability improves significantly as SNR increases. However, higher SNR does not always enhance the detectability of spectral features from molecules that are intrinsically weak or masked by clouds or other more dominant chemical species, emphasizing the intrinsic limitations imposed by planetary atmospheres themselves or technique adopted or *Twinkle*'s spectral coverage and/or resolving power. In eclipse spectroscopy, *Twinkle* also exhibits promising capabilities, as demonstrated in our simulations for the super-Earth 55 Cnc e (more details in Appendix 3.4.1). While characterizing detailed chemical compositions remains challenging, our simulations confirm *Twinkle*'s ability to extract fundamental planetary parameters, such as the atmospheric thermal structure, critical for investigating internal/surface properties and heat transport processes. Moreover, with thousands of hours of observing time available to *Twinkle*'s science team, continuous phase-curve observations are feasible, particularly for short-period exoplanets such as 55 Cnc e. Such phase-curve data will offer insights into the 3D structure of planetary atmospheres, complementing other observational approaches (H. A. Knutson et al. 2007; K. B. Stevenson et al. 2014; B.-O. Demory et al. 2016; L. Dang et al. 2025). Additionally, *Twinkle*'s ability to observe multiple transits and its high-precision timing capabilities will enhance transit ephemeris accuracy and establish more robust baselines for transit timing variation (TTV) analyses. *Twinkle*'s participation in TTV and transit duration variation (TDV) studies will further refine orbital parameters, improve planet mass determinations in multiplanet systems, and aid in detecting potential new exoplanets.

4.1 H₂/He dominated exoplanets

Simulations presented for gas giants and Neptunes have significantly benefited from recent *JWST* observations, allowing updated assessments of their observability with *Twinkle* through improved understanding of their atmospheric compositions. For bright hot Jupiters, such as HD 209458 b, *Twinkle* can achieve an SNR of 10 by stacking just 5 transits at native resolution. As illustrated in Fig. B1, the improvement in retrievability for planets with prominent spectral features becomes less significant at higher SNRs. This diminished improvement for HD 209458 b arises because the error associated with a single transit observation by *Twinkle* is already relatively small. Consequently, while further enhancements through additional observations yield limited benefits, bright targets offer an excellent opportunity to study stellar and/or atmospheric variability. For bright H₂/He-dominant exoplanets without prior data, an efficient strategy would be to initially perform a single transit and/or eclipse observation, conduct preliminary analysis, and then decide whether further observations are required based on initial findings or existing data from other instruments.

In contrast, for warm H₂/He-dominant exoplanets orbiting fainter stars, such as WASP-107 b, lower intrinsic observability leads to larger errors in simulated observations. Thus, increased stacking of transits and corresponding SNR improvements markedly enhance retrievability. Retrieval results for WASP-107 b demonstrate that weaker species like Na become detectable only at higher SNR (e.g. SNR = 10). For this type of targets, an initial estimate of the required number of transits based on prior knowledge and scientific objectives would be advisable. If there is no previous knowledge, a good approach could be to start with a limited number of transits and then performing an initial analysis on the data acquired to design the follow-up observational plan.

Neptunes and sub-Neptune-sized planets are often cloudy and/or their atmospheres are heavier than hydrogen. For these reasons, typically ten to twenty transits are required for adequate SNR, the lower intrinsic signal results in greater gains when improving from lower to moderate SNR. For instance, GJ 3470 b parameters were weakly constrained at SNR = 3 but became well constrained at SNR = 5 for molecules with multiple distinct spectral features. However, retrieval of species such as SO₂, having limited spectral signatures, required higher SNR (SNR = 7). Increasing SNR beyond this point yielded diminishing returns. Similar observations were made for WASP-107 b, where increasing observations from SNR = 5 (2 transits) to SNR = 7 (3 transits) provided more improvement than further increasing to SNR = 10 (6 transits). The diminishing marginal returns, as described by equation (1), suggest an optimal observational strategy for smaller radius planets targeting initially median SNR levels (3 or 5). Comparing retrieval improvements at these levels and balancing against the observational investment may help optimize follow-up observations.

In summary, *Twinkle*'s capability to characterize exoplanet atmospheres depends primarily on stellar brightness in optical and infrared wavelengths, planetary and stellar radii, main atmospheric component and cloud coverage. While stellar brightness and planetary dimensions are usually well-characterized by previous observations, cloud properties and, in the case of small planets, atmospheric mean molecular weight may require dedicated characterization efforts. Cloud contributions may significantly impact parameters' retrieval, especially for minor species. For instance, in our simulated retrievals of WASP-107 b, SO₂ spectral features were fully obscured by the assumed cloud opacity (see Fig. 5), and degeneracies between cloud-top height and planetary radius were observed. Incorporating detailed cloud microphysical and radiative models into future simulations might allow more realistic analyses.

4.2 Planets with non-H₂/He dominated atmosphere

Atmospheric characterization of super-Earths presents significant challenges due to their relatively small radii and heavier atmospheres, resulting in weaker observable signals. Eclipse simulations for 55 Cnc e suggest that the primary information expected from such planets is their temperature structure and constraints in atmospheric composition, distinguishable through comparisons with blackbody emission curves. If an atmosphere exists and is dominated by detectable molecules such as CO₂, our simulations indicate that *Twinkle* can well constrain those. For bright targets such as 55 Cnc e, increasing the number of eclipses beyond an initial set (10 eclipses for 55 Cnc e) shows diminishing returns in parameters' retrievability. An alternative approach for gaining deeper insights into super-Earths is phase-curve spectroscopy. The short orbital period of 55 Cnc e makes continuous phase-curve monitoring feasible. We are currently conducting simulations to assess *Twinkle*'s potential in phase-curve observations, and these findings will be detailed in future studies.

4.3 Stellar activity effects, targets observability, and future works

While stacking multiple observations to achieve higher SNR per wavelength bin is an effective strategy to enhance target observability, variability between observations introduces potential challenges. Variations in planetary atmospheric conditions between different transits or eclipses can lead to differences in the collected spectra. More significantly, stellar activity may significantly contaminate atmospheric parameters such as molecular abundances and temperature retrievals. A population study of 20 exoplanets conducted with *HST* by A. Saba et al. (2025) demonstrated that stellar contamination could alter molecular abundance estimates by up to six orders of magnitude and atmospheric temperatures by up to 145 per cent. Stellar contamination predominantly affects the optical wavelengths, which are partially covered by *Twinkle*. Hence, the potential influence of stellar activity must be carefully considered, especially when multiple observations are stacked. Investigating stellar activity effects on exoplanet atmospheres is one of *Twinkle*'s core scientific themes, and a future study will explore *Twinkle*'s capability to identify and quantify these stellar variations. In addition to astrophysical variability, instrumental systematics such as detector drift, thermal variations, and pointing jitter may affect the consistency of stacked observations. A detailed assessment of these instrumental effects will be presented in a companion study focusing on time-domain simulations.

Another significant constraint is observational availability, influenced by Earth obstruction due to *Twinkle*'s orbital configuration, planetary orbital periods, and transit durations. To address this, the *Twinkle* team has developed an orbital planning tool, allowing members to identify available observational windows over a year. Integrating this availability with *Twinkle*'s simulated performance enables the development of optimized, well-informed observation strategies for target exoplanets.

However, one remaining problem is that Earth obstruction also causing *Twinkle* to sometimes capture partial transit or eclipse events, as described in Section 2.1. Consequently, a more realistic estimate of the number of required observations should incorporate an observing efficiency factor, as demonstrated by L. Booth et al. (2024). Meanwhile, the *Twinkle* team is actively developing methodologies for effectively stacking multiple partial transits or eclipses to maximize SNR. Comparable techniques have already been demonstrated for combining multiple nights of data with the same instrumental setup in both transmission and emission studies from ground-based instruments (e.g. P. Giacobbe et al. 2021; A. Y. Kesseli et al. 2022; G. Scandariato et al. 2023). A simulation by L. Boldt-Christmas et al. (2024) specifically addresses the compromises required when maximizing SNR through stacked observations for transiting exoplanets.

Several studies underline the importance of combining multiple observations for the future characterization of Earth-like and super-Earth exoplanets, even for powerful observatories such as *JWST* (F. Wunderlich et al. 2019). Indeed, recent analysis by J. Lustig-Yaeger et al. (2023) involved combining two *JWST* transit observations for improved results. Historically, enhancing SNR through stacking multiple observations has often been impractical due to limited telescope availability. However, as a commercially driven, rapidly deployable mission, *Twinkle* can offer substantial, on-demand telescope time, potentially thousands of hours for each scientific theme. Therefore, stacking multiple observations as availability permits is a viable and efficient strategy for the *Twinkle* science community, provided that stacking does not introduce significant data distortions.

5 CONCLUSIONS

Here we present updates on the *Twinkle* space missions simulated performance for characterizing exoplanetary atmospheres, given the latest design choices of the spacecraft and payload and new insight into the atmospheric composition and structure of select exoplanets provided by *JWST*'s observations. In addition, we have updated the list of exoplanet candidates suitable for characterization based on currently confirmed planets and TOIs, and incorporating updated planetary and stellar parameters as available in the literature.

More specifically, leveraging recent discoveries from *JWST*, we have provided more realistic spectral simulations across various exoplanet types and assessed the retrievability of their atmospheric parameters and chemical abundances at different achievable SNR levels. Retrieval analyses for HD 209458 b, WASP-107 b, and GJ 3470 b have indicated how increased observational investment to reach higher SNR improves the retrievability of the atmospheric parameters and chemical species for different planetary categories. At *Twinkle*'s native resolution, already at low median SNR values of 3 or 5, key atmospheric properties, such as temperature and cloud altitudes, are well constrained, alongside major molecular species. Increasing the median SNR to 7 or 10 further improves constraints on minor species, exemplified by the successful retrieval of SO₂ in GJ 3470 b. For small planets with atmospheres heavier than hydrogen, such as 55 Cnc e, emission retrieval studies revealed that after stacking ten eclipses, temperature structure and dominant atmospheric constituents could be constrained. Additional eclipse observations showed limited improvement, however 55 Cnc e would be an excellent candidate for phase-curve observations with *Twinkle*. This example showcases that tailored observational strategies are necessary for different exoplanet types to ensure an efficient use of the telescope time.

Our analysis indicates that approximately 100 exoplanets are suitable for in depth atmospheric studies and over 300 exoplanets are good candidates for characterization with *Twinkle*.

Future studies will build upon this work to better inform the observational planning of the *Twinkle*'s exoplanet survey. In particular we plan to include more realistic assumptions of target observability due to Earth obstruction and quantify the impact of stellar activity in our simulations. We also plan to refine our simulations by integrating more advanced cloud formation models and to assess the detectability of other atmospheric features. Additionally, the radiometric model will be continuously updated alongside instrumental or operational developments to ensure precise noise estimates. We will also broaden the scope of simulations to include phase-curve observations, exploring expected outcomes for other scientific themes and their mutual refinements.

In the next 5 yr, numerous space- and ground-based observatories dedicated to exoplanetary science research will become operational. BSSL-*Twinkle* uniquely complements these efforts by offering the science team full flexibility in the selection of the list of targets and observational strategy. To support informed decision-making by the *Twinkle* science team, BSSL is committed to clearly showcasing the telescope's anticipated scientific outcomes.

ACKNOWLEDGEMENTS

We gratefully acknowledge insightful discussions and inputs from members of the *Twinkle* Science Team, and their participation in *Twinkle* Science Conference. This work made use of data and results from observations with the NASA/ESA/CSA *James Webb Space Telescope*, as reported in the published literature. The *JWST* data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for *JWST*.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- Akeson R., Christiansen J., 2019, American Astronomical Society Meeting Abstracts #233, 140.09
- Al-Refaie A. F., Changeat Q., Waldmann I. P., Tinetti G., 2021, *ApJ*, 917, 37
- Anderson D. R. et al., 2017, *A&A*, 604, A110
- Arcangeli J. et al., 2018, *ApJ*, 855, L30
- Archer R. et al., 2020, *Nat. Astron.*, 4, 1017
- Awiphan S. et al., 2016, *MNRAS*, 463, 2574
- Bean J. L., Miller-Ricci Kempton E., Homeier D., 2010, *Nature*, 468, 669
- Beatty T. G. et al., 2024, *ApJ*, 970, L10
- Bell T. et al., 2024, American Astronomical Society Meeting Abstracts, #343.04
- Benz W. et al., 2021, *Exp. Astron.*, 51, 109
- Berta Z. K. et al., 2012, *ApJ*, 747, 35
- Boldt-Christmas L., Lesjak F., Wehrhahn A., Piskunov N., Rains A. D., Nortmann L., Kochukhov O., 2024, *A&A*, 683, A244
- Booth L., Sarkar S., Griffin M., Edwards B., 2024, *MNRAS*, 530, 2166
- Borucki W. J. et al., 2010, *Science*, 327, 977
- Carter A. L. et al., 2024, *Nat. Astron.*, 8, 1008
- Changeat Q. et al., 2022, *ApJS*, 260, 3
- Charbonneau D., Brown T. M., Noyes R. W., Gilliland R. L., 2002, *ApJ*, 568, 377
- Chauvin G. et al., 2017, *A&A*, 605, L9
- Chen J., Kipping D., 2017, *ApJ*, 834, 17
- Dang L. et al., 2025, *AJ*, 169, 32
- Deming D. et al., 2013, *ApJ*, 774, 95
- Demory B.-O. et al., 2016, *Nature*, 532, 207
- Edwards B., Savini G., Tinetti G., Tessenyi M., Arena C., Lindsay S., Bowles N., 2019a, *J. Astron. Tel. Instr. Syst.*, 5, 014006
- Edwards B., Lindsay S., Savini G., Tinetti G., Arena C., Bowles N., Tessenyi M., 2019b, *J. Astron. Tel. Instr. Syst.*, 5, 034004
- Edwards B. et al., 2019c, *Exp. Astron.*, 47, 29
- Edwards B. et al., 2023, *ApJS*, 269, 31
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
- Fraine J. et al., 2014, *Nature*, 513, 526
- Ge J. et al., 2022, preprint (arXiv:2206.06693)
- Ge J. et al., 2024, in Coyle L. E., Matsuura S., Perrin M. D., eds, *Proc. SPIE Conf. Ser. Vol. 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave*. SPIE, Bellingham, p. 1309218
- Giacobbe P. et al., 2021, *Nature*, 592, 205
- GRAVITY Collaboration, 2019, *A&A*, 623, L11
- Grillmair C. J. et al., 2008, *Nature*, 456, 767
- Harrington J., Hansen B. M., Luszcz S. H., Seager S., Deming D., Menou K., Cho J. Y. K., Richardson L. J., 2006, *Science*, 314, 623
- Howell S. B. et al., 2014, *PASP*, 126, 398
- Hu R. et al., 2024, *Nature*, 630, 609
- Husser T. O., Wende-von Berg S., Dreizler S., Homeier D., Reiners A., Barman T., Hauschildt P. H., 2013, *A&A*, 553, A6
- Jason S. et al., 2016, in 4S, 4S Symp
- JWST Transiting Exoplanet Community Early Release Science Team 2023, *Nature*, 614, 649
- Karman T. et al., 2019, *Icarus*, 328, 160
- Kesseli A. Y., Snellen I. A. G., Casasayas-Barris N., Mollière P., Sánchez-López A., 2022, *AJ*, 163, 107
- Knutson H. A. et al., 2007, *Nature*, 447, 183
- Kreidberg L. et al., 2014, *ApJ*, 793, L27
- Lustig-Yaeger J. et al., 2023, *Nat. Astron.*, 7, 1317
- Ma S., Saba A., Faris Al-Refaie A., Tinetti G., Yurchenko S. N., Tennyson J., Cecchi Pestellini C., 2026, *ApJ*, 1002, 208
- Macintosh B. et al., 2015, *Science*, 350, 64
- Madhusudhan N., Sarkar S., Constantinou S., Holmberg M., Piette A. A. A., Moses J. I., 2023, *ApJ*, 956, L13
- Majeau C., Agol E., Cowan N. B., 2012, *ApJ*, 747, L20
- Mugnai L. V., Bocchieri A., Pascale E., 2023, *J. Open Source Softw.*, 8, 5348
- Nortmann L. et al., 2018, *Science*, 362, 1388
- Piaulet C. et al., 2021, *AJ*, 161, 70
- Rauer H. et al., 2025, *Exp. Astron.*, 59, 26
- Redfield S., Endl M., Cochran W. D., Koesterke L., 2008, *ApJ*, 673, L87
- Ricker G. R. et al., 2014, in Oschmann J. M. Jr, Clampin M., Fazio G. G., MacEwen H. A., eds, *Proc. SPIE Conf. Ser. Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. SPIE, Bellingham, p. 914320
- Saba A., Thompson A., Yip K. H., Ma S., Tsiaras A., Al-Refaie A. F., Tinetti G., 2025, *ApJS*, 276, 70

- Savini G. et al., 2018, in MacEwen, H. A., et al., eds, Proc. SPIE Conf. Ser. Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 1514
- Scandariato G. et al., 2023, *A&A*, 674, A58
- Sing D. K. et al., 2016, *Nature*, 529, 59
- Sing D. K. et al., 2024, *Nature*, 630, 831
- Snellen I. A. G., de Kok R. J., de Mooij E. J. W., Albrecht S., 2010, *Nature*, 465, 1049
- Stassun K. G., Collins K. A., Gaudi B. S., 2017, *AJ*, 153, 136
- Stevenson K. B. et al., 2014, *Science*, 346, 838
- Stock J. W., Kitzmann D., Patzer A. B. C., Sedlmayr E., 2018, *MNRAS*, 479, 865
- Stotesbury I. et al., 2022, in Coyle L. E., Matsuura S., Perrin M. D., eds, Proc. SPIE Conf. Ser. Vol. 12180, Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 1218033
- Stotesbury I., Bradley L., Wilcock B. J., Tinetti G., Tessenyi M., Savini G., Windred P., Tennyson J., 2024, in Coyle L. E., Matsuura S., Perrin M. D., eds, Proc. SPIE Conf. Ser. Vol. 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 1309213
- Swain M. R., Vasisht G., Tinetti G., 2008, *Nature*, 452, 329
- Swain M. R. et al., 2009, *ApJ*, 704, 1616
- Tennyson J. et al., 2024, *J. Quant. Spec. Radiat. Transfer*, 326, 109083
- Tessenyi M. et al., 2013, *MNRAS*, 436, 2974
- Tinetti G. et al., 2007, *Nature*, 448, 169
- Tinetti G. et al., 2018, *Exp. Astron.*, 46, 135
- Tinetti G., Eccleston P., Luefing T., Salvignol J.-C., Fahmy S., Alves de Oliveira C., 2022, in *European Planetary Science Congress Abstracts Vol. 16*. Granada, Spain, p. EPSC2022–1114
- Tsiaras A. et al., 2016, *ApJ*, 820, 99
- Tsiaras A. et al., 2018, *AJ*, 155, 156
- Vidal-Madjar A., Lecavelier des Etangs A., Désert J. M., Ballester G. E., Ferlet R., Hébrard G., Mayor M., 2003, *Nature*, 422, 143
- Vidal-Madjar A. et al., 2004, *ApJ*, 604, L69
- Welbanks L. et al., 2024, *Nature*, 630, 836
- Windhorst R. A. et al., 2011, *ApJS*, 193, 27
- Wunderlich F. et al., 2019, *A&A*, 624, A49
- Xue Q., Bean J. L., Zhang M., Welbanks L., Lunine J., August P., 2024, *ApJ*, 963, L5

APPENDIX A: NOISE CONTRIBUTIONS

The noise component plots presented in this section illustrate how the different noise sources considered by the Radiometric Tool vary across each target. For all targets simulated in this work, the noise budget remains in the photon-noise-dominated regime, with all noise contributions increasing towards the edges of each spectral channel. For relatively bright targets such as HD 209458 b and 55 Cnc e, read noise begins to exceed photon noise at the longer wavelengths covered by channel 1, as the detector requires more reads to avoid saturation.

The radiometric simulations presented in this study adopt the instrument parameters summarized in Table A1. In its current design, the telescope has a 0.45 m diameter primary aperture, corresponding to a geometric collecting area of $\sim 0.159 \text{ m}^2$. After subtracting the physical obscuration from the secondary mirror and its support structure, the effective collecting area is $A_{\text{tel}} = 0.146 \text{ m}^2$. The optical train consists of four mirror surfaces and a dichroic beam splitter in the common path, followed by a dispersing prism and three fold mirrors in each channel. The reflectivity of the mirror surfaces (98 per cent per surface), the transmission of the dichroic (~ 90 per cent) and prism (~ 80 per cent), and the wavelength-dependent detector quantum efficiency (75 per cent–95 per cent across the $0.5\text{--}4.5 \mu\text{m}$ bandpass) are then applied as multiplicative throughput factors to obtain the total optical throughput.

Table A1. Summary of *Twinkle* instrument parameters assumed in the radiometric simulations.

Parameter	Value
Collecting area (A_{tel})	0.146 m^2
Channel 0 range	$0.5\text{--}2.43 \mu\text{m}$ ($R \leq 70$)
Channel 1 range	$2.43\text{--}4.5 \mu\text{m}$ ($R \leq 50$)
Mirror reflectivity (per surface)	98 per cent
Dichroic transmission	90 per cent
Prism transmission	80 per cent
Detector QE	75–95 per cent
Read noise	25 counts
Dark current	35 counts s^{-1}
Well depth	72 000 counts
Detector temperature	90 K
Optics temperature	100–160 K
PSF (Ch 0)	Diffraction-limited
PSF (Ch 1)	$\text{WFE}_{\text{rms}} = 0.28 \mu\text{m}$

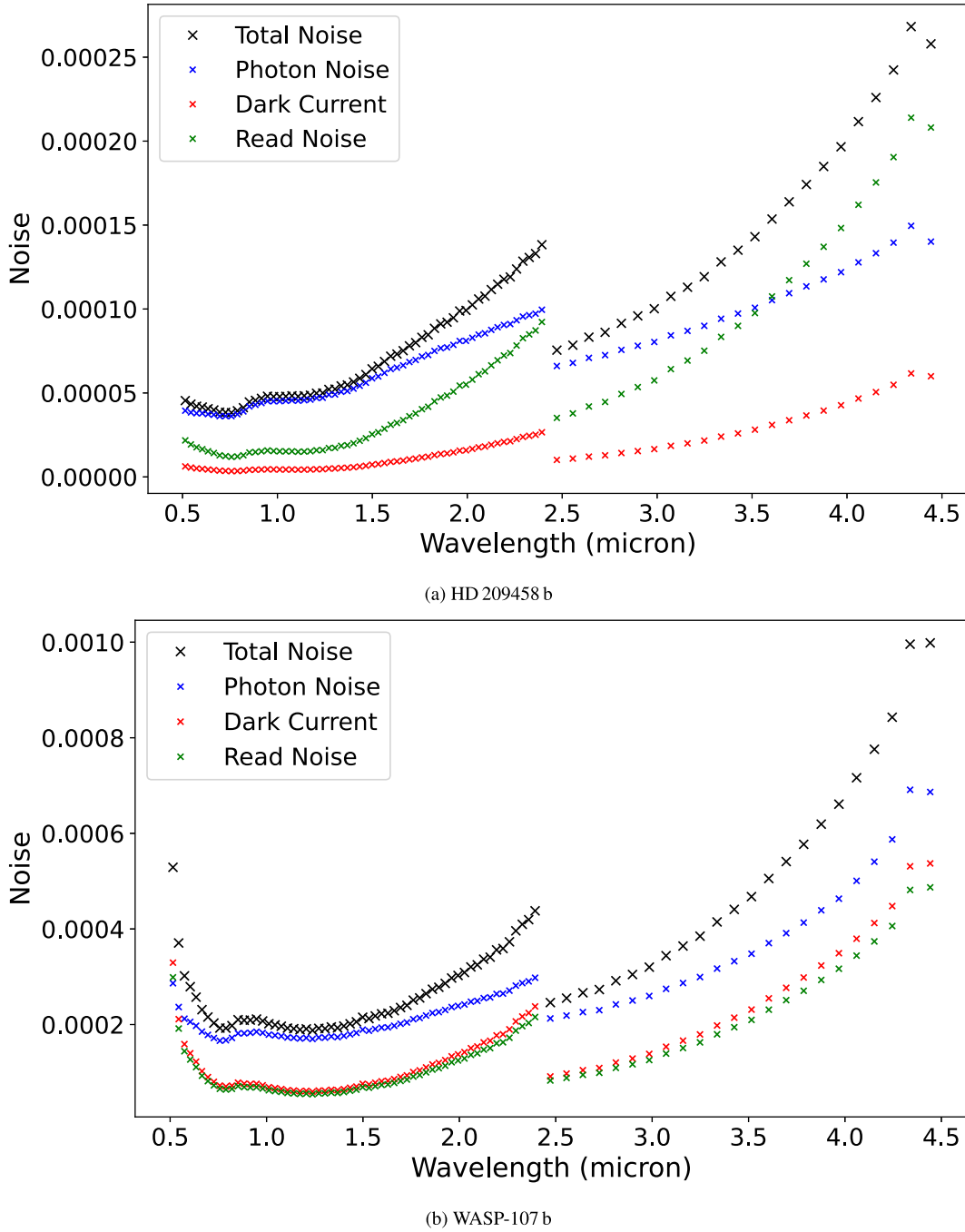
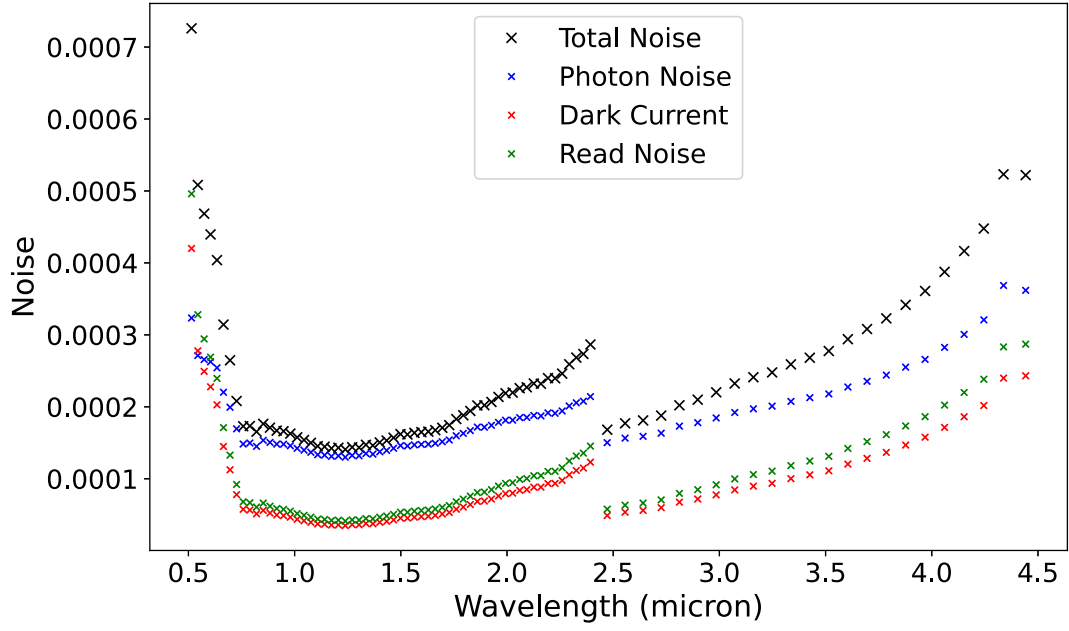


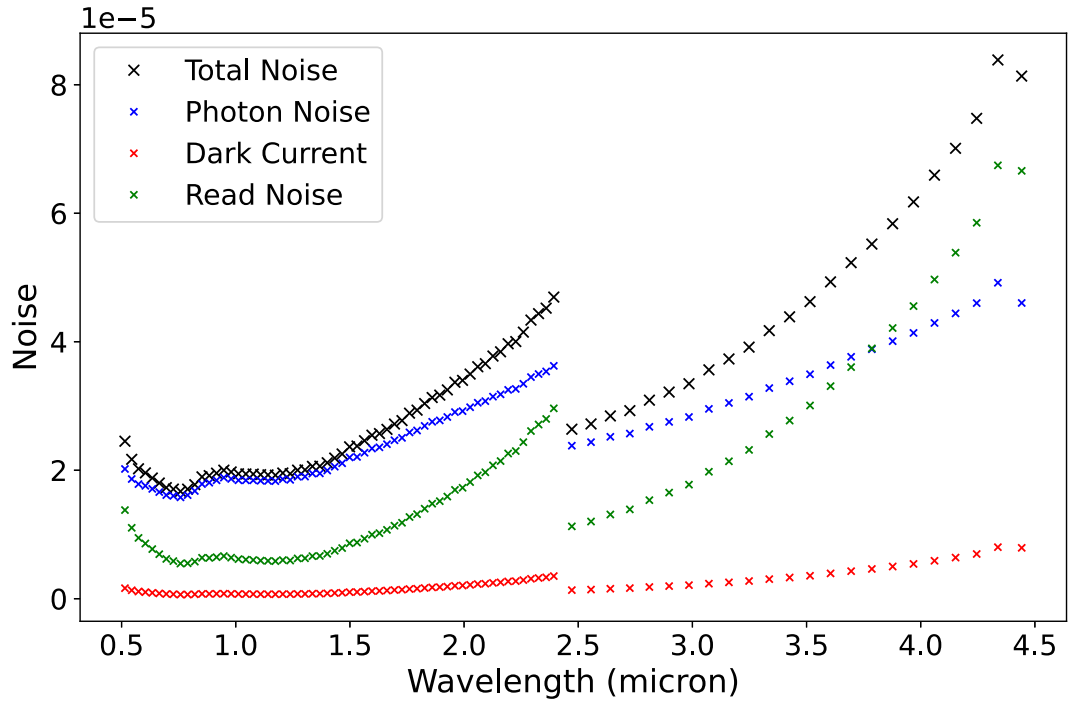
Figure A1. Noise component breakdown for a single transit observation at *Twinkle*'s native spectral resolution. The total noise (black) is decomposed into photon noise (blue), dark current (red), and read noise (green). The discontinuity at $2.43\ \mu\text{m}$ marks the boundary between channel 0 and channel 1. (a) For the bright host star of HD 209458 b ($K_{\text{mag}} \approx 6.3$), read noise becomes the dominant contributor in channel 1 because the detector requires more frequent reads to avoid well saturation. (b) For the fainter WASP-107 b ($K_{\text{mag}} \approx 9.0$), longer integration times are feasible and photon noise dominates across both channels; the overall noise level is correspondingly higher.

The point spread function (PSF) in channel 0 is treated as diffraction-limited, while channel 1 assumes a wavefront error of $\text{WFE}_{\text{rms}} = 0.28\ \mu\text{m}$. Pixel-based PSF models are used for both channels. The detectors operate at 90 K with a read noise of 25 counts per read, a dark current of $35\ \text{counts s}^{-1}$, and a well depth of 72 000 counts (80 per cent of which is usable to minimize non-linear detector effects). For bright targets such as HD 209458 b and 55 Cnc e, shorter integration times are required to avoid well saturation, leading to a higher number of detector reads per exposure and consequently elevated read noise contributions (Figs A1a and A2b). For fainter targets (WASP-107 b, GJ 3470 b), longer integrations are feasible, and photon noise dominates across the full spectral range.

Line-of-sight pointing jitter is included in the noise model using a pre-computed jitter profile derived from ground-test pointing performance data. The out-of-transit noise overhead is accounted for by a factor $\gamma_{\text{noise}} = 1.67$, corresponding to an out-of-transit baseline



(a) GJ 3470 b



(b) 55 Cnc e

Figure A2. Same as Fig. A1 but for GJ 3470 b and 55 Cnc e. (a) For GJ 3470 b ($K_{\text{mag}} \approx 7.99$), photon noise dominates across both channels, with a noise level intermediate between the bright and faint targets. (b) 55 Cnc e is the brightest target in our sample ($K_{\text{mag}} \approx 4.0$); the detector saturates rapidly, requiring very short integration times and a large number of reads per exposure. Consequently, read noise dominates much of channel 1, more pronounced than for HD 209458 b, while the absolute noise level is the lowest among all four targets.

equal to 1.5 times the transit duration. The zodiacal background is modelled using a position-dependent sky map with a nominal scaling factor of 2.5. Additional systematic noise contributions (e.g. gain noise, detector drift, thermal variations) are not included in the present radiometric model but will be addressed in a forthcoming time-domain simulation study.

The observation stacking equation (equation 1) assumes that the noise contributions from successive transits or eclipses are uncorrelated. While this assumption is appropriate for photon noise and random detector noise, systematic effects such as instrumental drift, thermal cycling, or stellar variability may introduce correlated noise components that do not average down as $\sqrt{N}$. Experience

with *HST*/WFC3 transit spectroscopy suggests that systematic noise floors of $\sim 20\text{--}30$ ppm can be reached for bright targets after careful detrending (D. Deming et al. 2013). A detailed assessment of *Twinkle*'s systematic noise floor, incorporating time-domain instrumental effects, will be presented in a companion study.

APPENDIX B: RETRIEVAL POSTERIOR PLOTS

The posterior plots presented in this section of Appendix illustrate the statistical distributions of retrieved atmospheric parameters obtained from our retrieval analyses. Each posterior plot summarizes the likelihood distributions of key atmospheric properties, including planetary radius, equilibrium temperature, cloud-top pressure, and molecular abundances, under various observational conditions (e.g. different SNR levels). The contours indicate parameter uncertainties and correlations. Narrower, well-defined distributions imply robust retrievals and tightly constrained parameters, whereas broader, elongated distributions reflect higher uncertainties or parameter degeneracies. These visualizations highlight how observational strategies, such as varying SNRs or the number of stacked transits, impact the precision and reliability of atmospheric characterization.

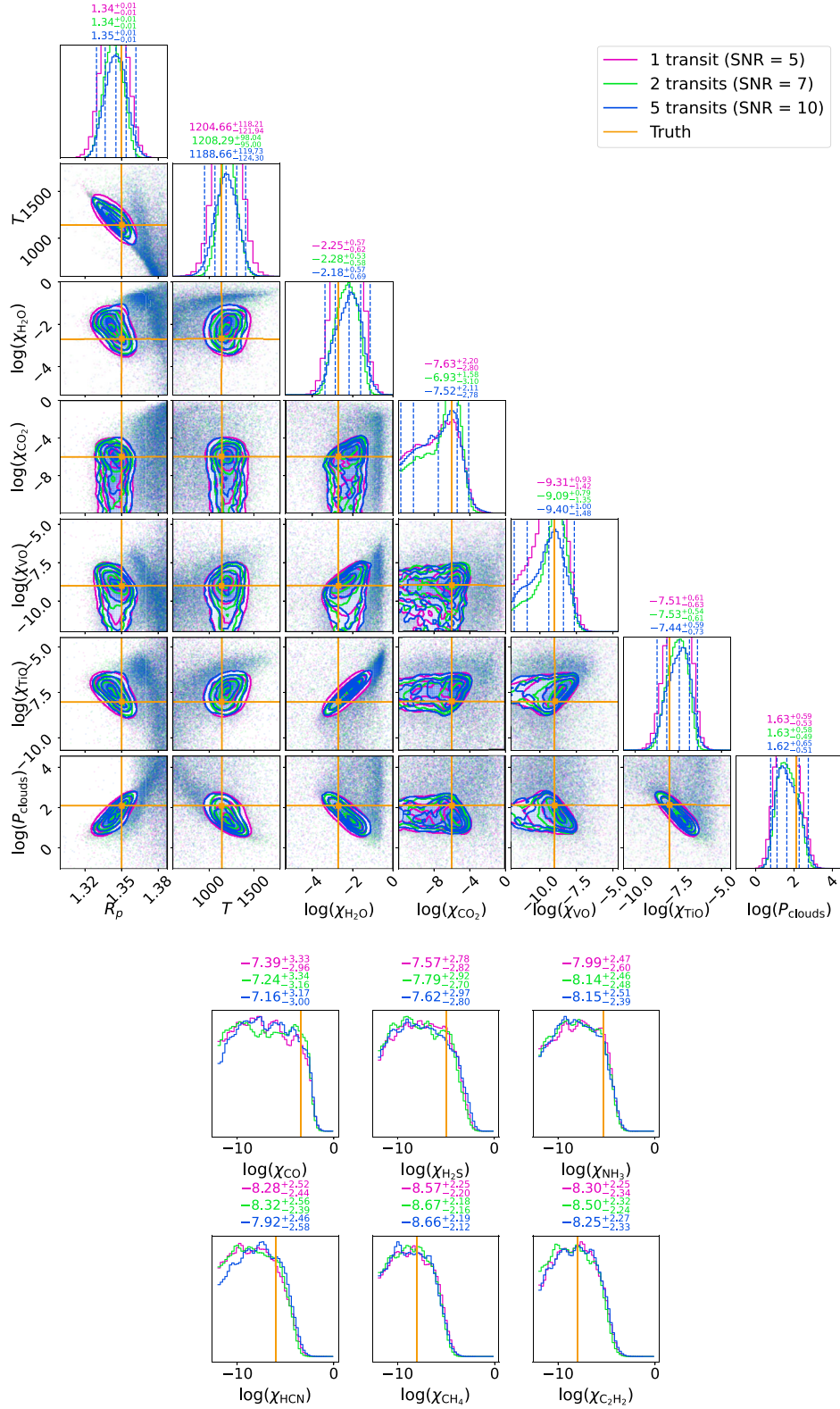


Figure B1. Posterior distributions of retrieved parameters from free retrieval analysis of HD 209458 b. Planet parameters and molecules visible in the spectrum are shown in the corner plot. Invisible molecules with retrieved upper limits are shown in the histograms at the bottom. The numbers above each block represent the median value of the posteriors.

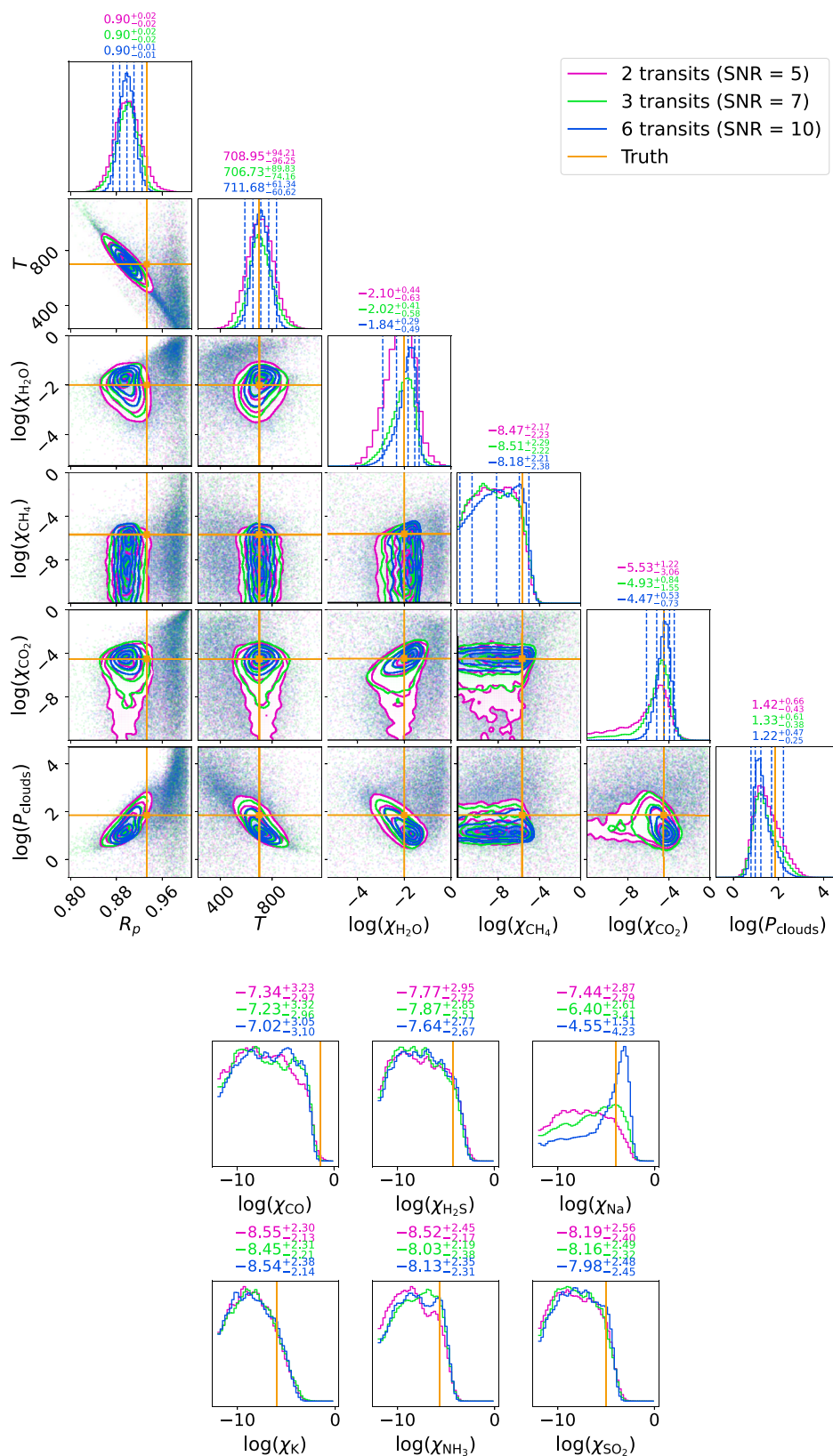


Figure B2. Posterior distributions of retrieved parameters from free retrieval analysis of WASP 107 b. Planet parameters and molecules visible in the spectrum are shown in the corner plot. Invisible molecules with retrieved upper limits are shown in the histograms at the bottom.

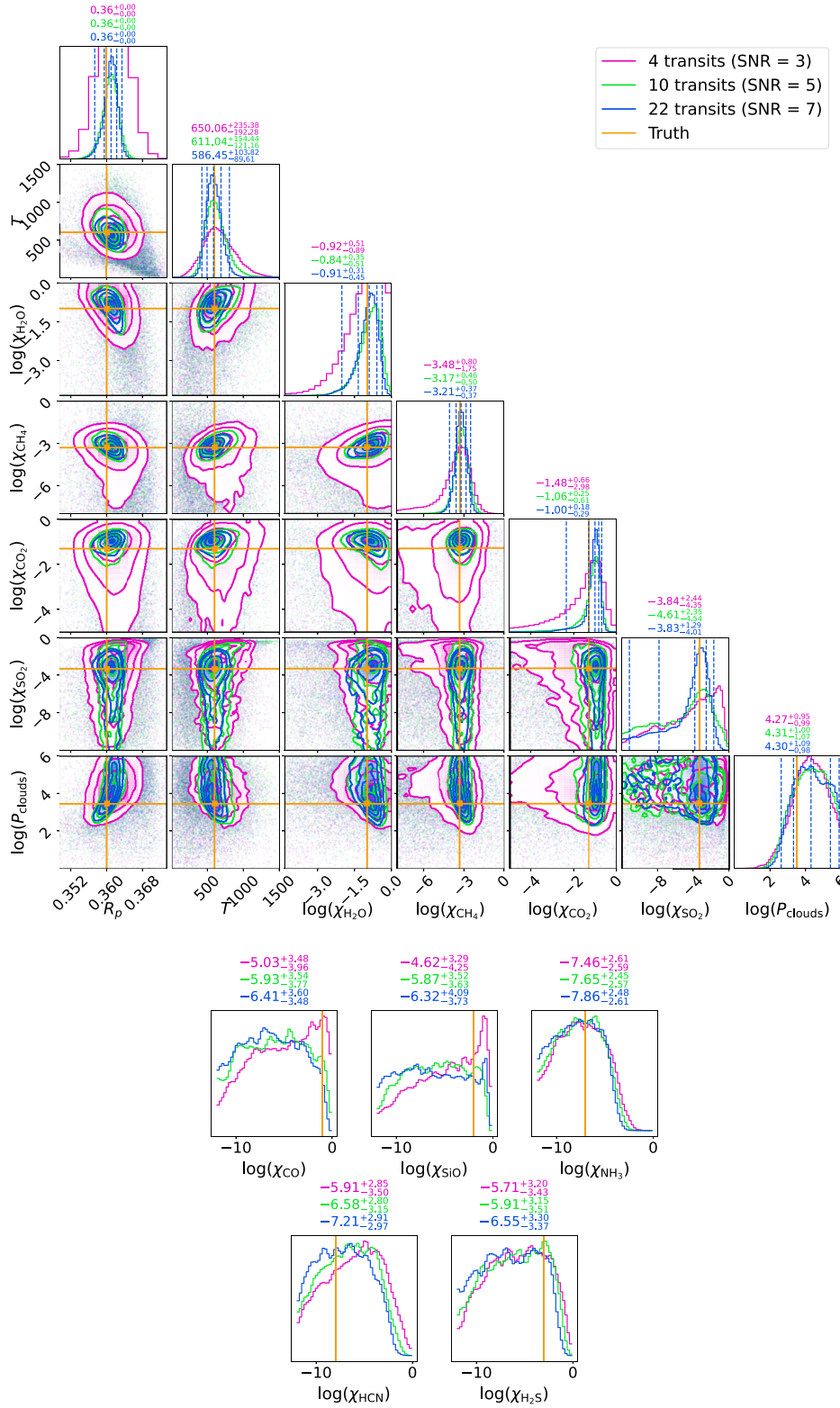


Figure B3. Posterior distributions of retrieved parameters from free retrieval analysis of GJ 3470 b. Planet parameters and molecules visible in the spectrum are shown in the corner plot. Invisible molecules with retrieved upper limits are shown in the histograms at the bottom.

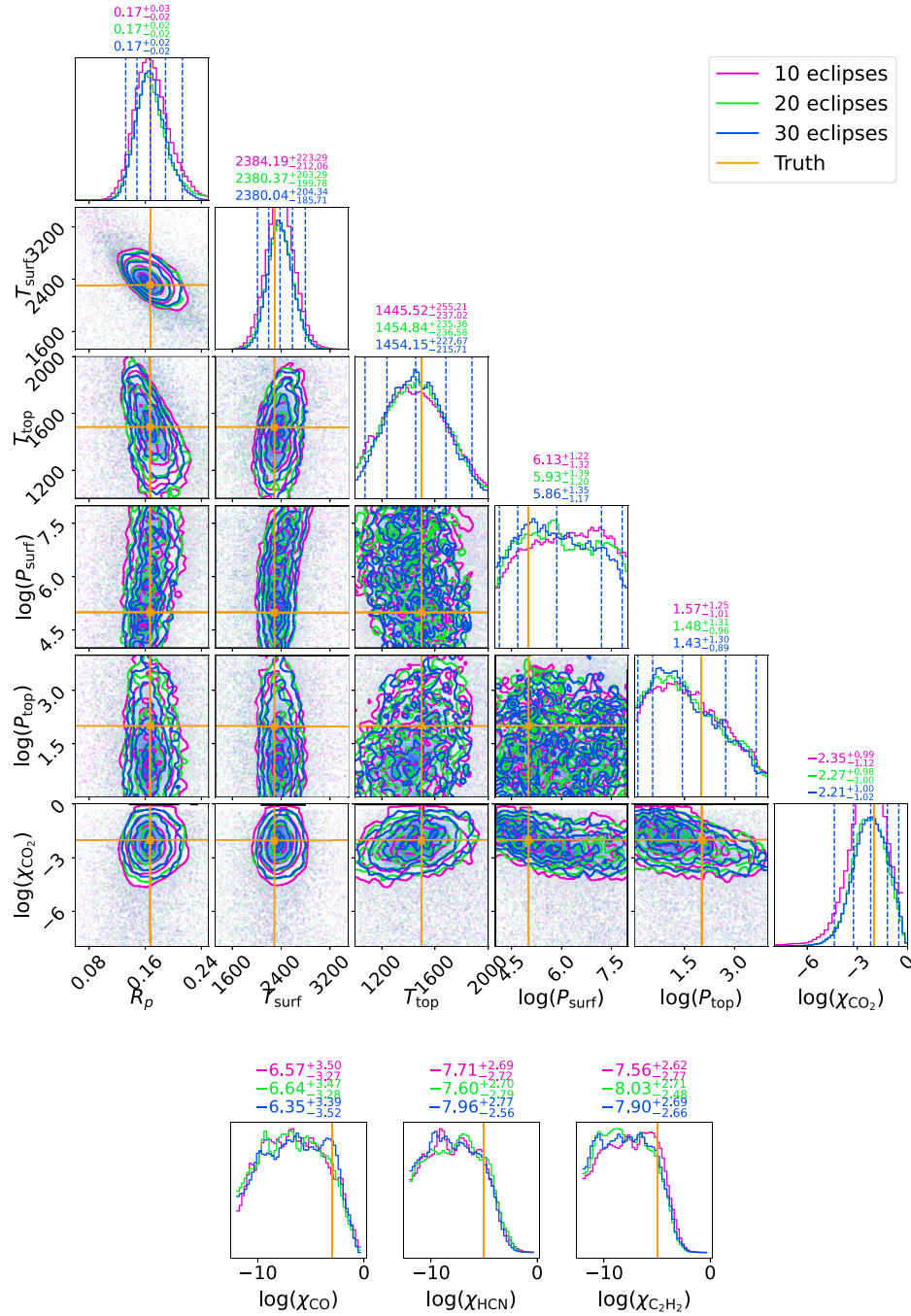


Figure B4. Posterior distributions of retrieved parameters from free retrieval analysis of 55 Cnc e. Planet parameters and molecules visible in the spectrum are shown in the corner plot. Invisible molecules with retrieved upper limits are shown in the histograms at the bottom.

APPENDIX C: MASS ESTIMATES FOR TESS OBJECT OF INTERESTS

In order to estimate masses for TOIs with radius available, we utilized *Forecaster* tool developed by J. Chen & D. Kipping (2017), a code that predict mass from radius for objects. Although *Forecaster* is calibrated across a wide radius range (~ 0.1 – $100 R_{\oplus}$), in the giant-planet regime the mass–radius relation is strongly degenerate, so its predicted masses there are poorly constrained and do not reproduce the observed distribution of known transiting giants. Mass estimates from the *Forecaster* tool for TOIs with radius less than eight Earth radii are adopted, as they align with the known exoplanet distribution shown in Fig. C1. For TOIs with higher radius, we assume an asymmetric lognormal distribution to randomly set the mass, with the mean value set as the median value of known exoplanets with radius larger than eight Earth radii. Resultant mass distribution for TOIs in *Twinkle*’s FoR is shown in Fig. C1.

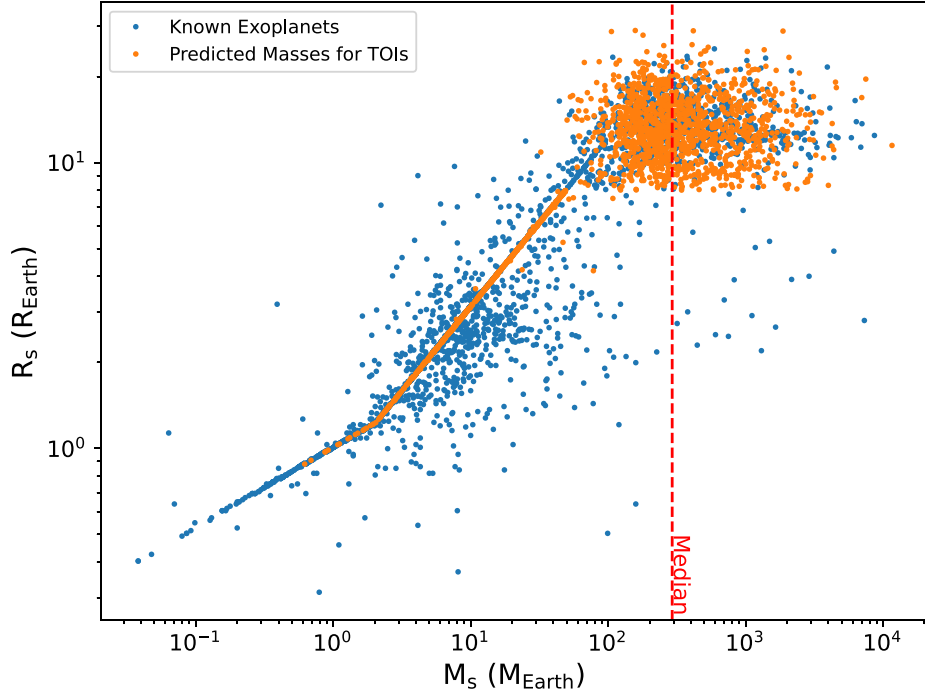


Figure C1. Mass–radius relationship comparison between all transiting known exoplanets and TOIs in Twinkle’s FoR with mass estimate based on radius. The red dashed vertical line in the plot represents the median mass of known exoplanets with radii larger than 8 Earth radii ($R_{\oplus}$), which is counted as the mean value for the lognormal distribution of TOIs mass estimate.

APPENDIX D: LIKELIHOOD RATIO TEST FOR MOLECULAR DETECTABILITY

To quantify the statistical significance of molecular detections, we employ the Likelihood Ratio Test (LRT) following the methodology described by M. Tessenyi et al. (2013). For each molecule under consideration, the alternative hypothesis H_1 corresponds to the full forward model containing all retrieved species, while the null hypothesis H_0 corresponds to the model with the test molecule removed. For each hypothesis, Monte Carlo simulations with 100 000 iterations generate synthetic observations by adding Gaussian noise (drawn from the per-bin radiometric uncertainties) to the corresponding model spectrum. For each realization, the likelihood ratio statistic $\Delta = -2 \ln \mathcal{L}_0 + 2 \ln \mathcal{L}_1$ is computed, yielding distributions of Δ (data generated under H_0) and Δ' (data generated under H_1). A molecule is classified as a ‘Detection’ when $\text{median}(\Delta') > \text{mean}(\Delta) + 3\sigma(\Delta)$, where $\sigma(\Delta)$ is the standard deviation of the null distribution. The detection significance, expressed in units of σ , is reported in the results tables. Molecules falling below the 3σ threshold are classified as ‘Upper Limit’, indicating that their abundances are either prior-dominated or only weakly constrained by the data.

For emission spectra (55 Cnc e), we adopt the same molecule-removal framework: H_1 is the full emission model containing all retrieved species, while H_0 is the full model with the test molecule removed. This approach is more conservative than the original M. Tessenyi et al. (2013) emission formulation (which uses a featureless blackbody as H_0), because it tests whether a molecule’s spectral signature is distinguishable in the presence of all other atmospheric constituents, rather than merely deviating from a smooth continuum.

To ensure robustness against unreliable MAP estimates for poorly constrained parameters, we perform the LRT using both the MAP and posterior median values for each molecule. A molecule is classified as a ‘Detection’ only if both tests independently exceed the 3σ threshold. The more conservative (lower) of the two significance values is reported in the results tables.

Figs D1–D4 show the resulting Δ (black) and Δ' (green) distributions for each molecule tested. The red dashed line indicates the 3σ detection threshold, while the blue solid line marks the median of Δ' . A molecule is detected when the blue line exceeds the red line.

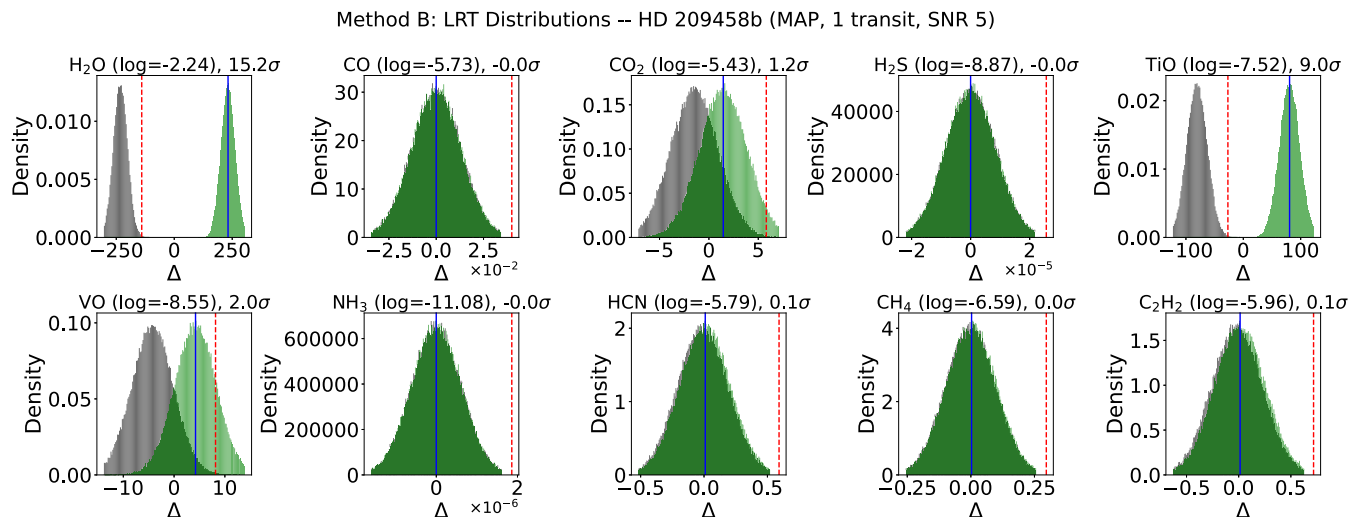


Figure D1. LRT Δ and Δ' distributions for HD 209458 b (1 transit, SNR 5). Each panel shows the null distribution Δ (black) and alternative distribution Δ' (green) for a single molecule removal test. The red dashed line marks the 3σ detection threshold; the blue solid line marks the median of Δ' . Detection significance in units of σ is indicated in each panel title.

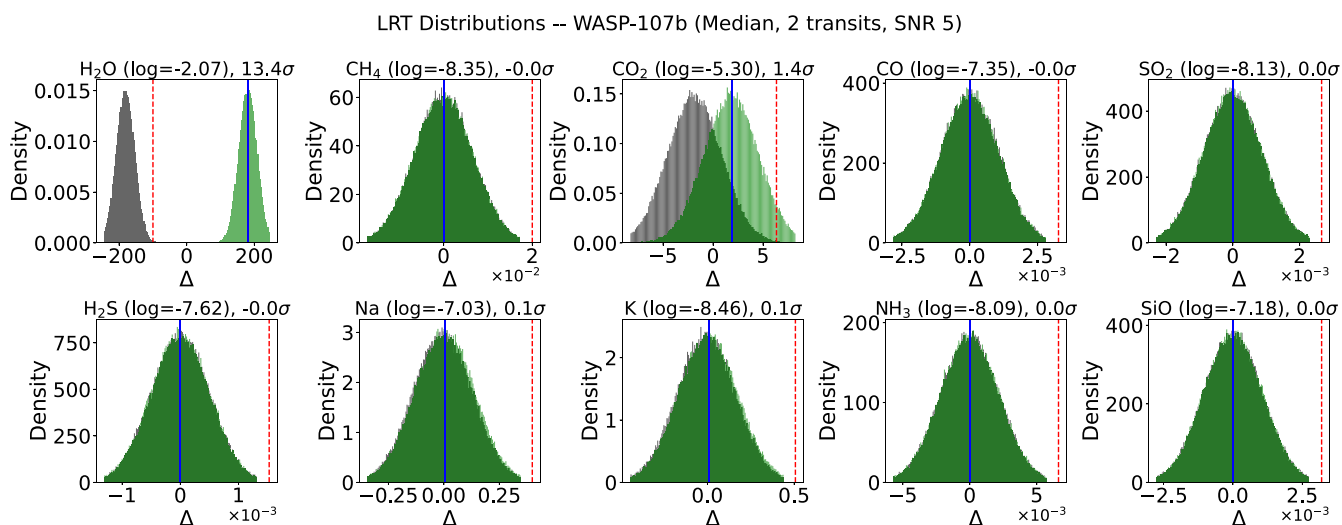


Figure D2. LRT Δ and Δ' distributions for WASP-107 b (2 transits, SNR 5). Panel layout and colour scheme as in Fig. D1.

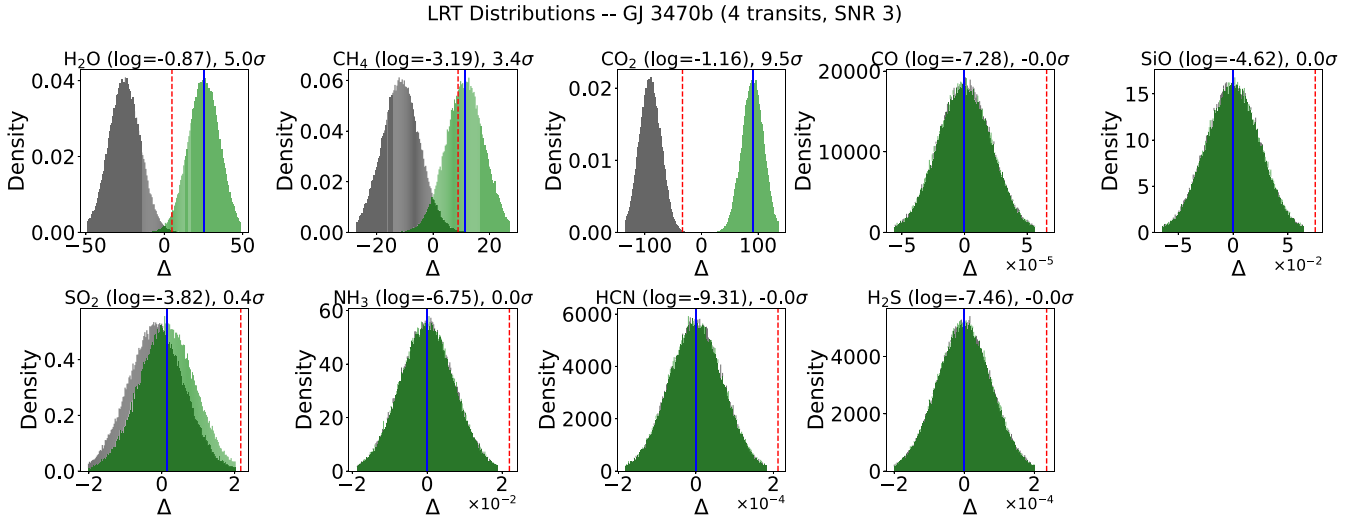


Figure D3. LRT Δ and Δ' distributions for GJ 3470 b (4 transits, SNR 3). Panel layout and colour scheme as in Fig. D1.

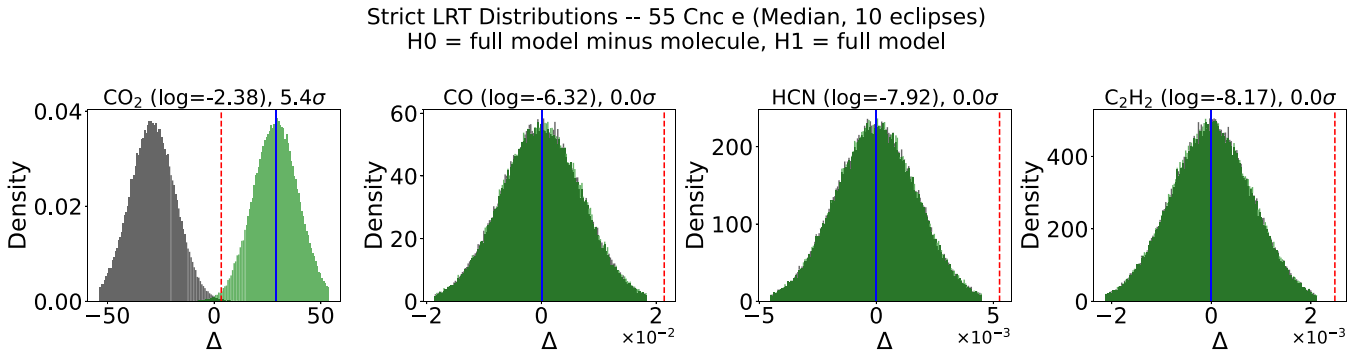


Figure D4. LRT Δ and Δ' distributions for 55 Cnc e (10 eclipses). For this emission target, H_1 is the full emission model and H_0 is the full model with the test molecule removed. Panel layout and colour scheme as in Fig. D1.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.