The Effect of Binning Transmission Spectra for Exoplanet Atmospheric Retrievals

Jack Davey (jack.davey.22@ucl.ac.uk)

Ingo Waldmann, Kai Hou Yip & Ahmed Al-Refaie

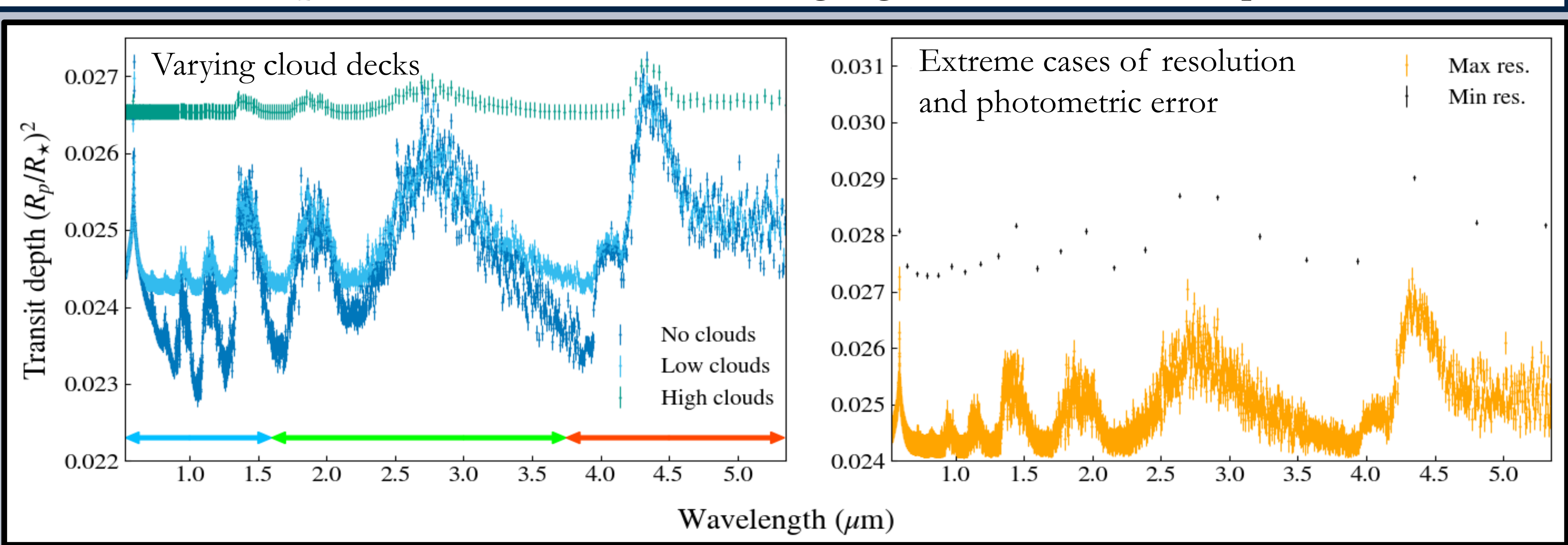
UCL

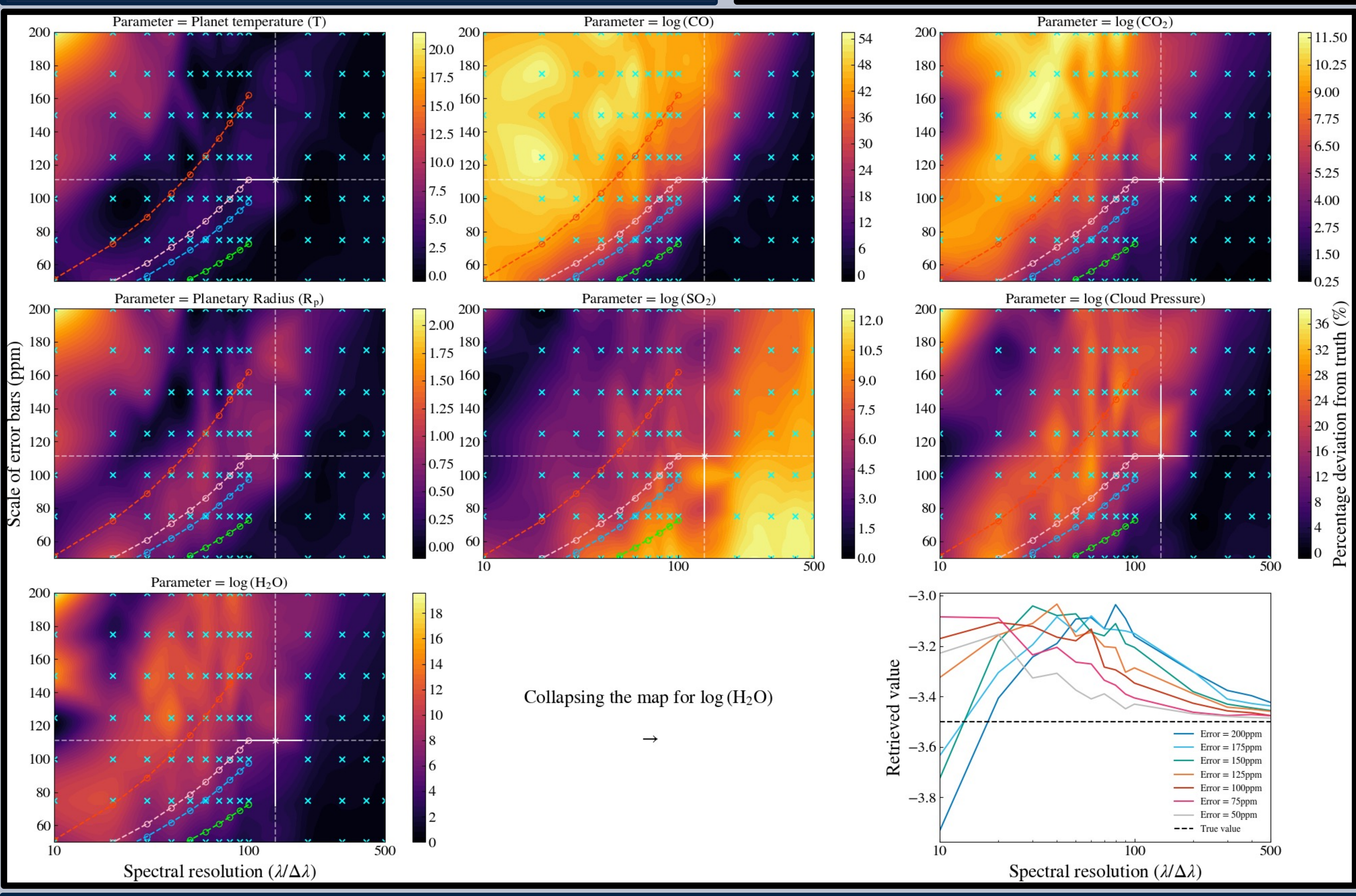
Abstract: With higher quality space-based data now available for transit observations, there is an increasing demand to maximise the efficiency and accuracy of our analysis methods and one processing step which may aid in this is binning spectra before retrievals

- We find **clear boundaries** in resolution-error space which separate well and poorly constrained retrieval predictions
- The **JWST NIRSpec data** [1] falls into a well retrieved region for low or no cloud cases but high clouds present a challenge
- Above R ≈ 500, retrievals become insensitive to error since the information in the spectra becomes **resolution dominated**
- **Correlations** with the CO abundance are resolution-dependent and higher resolutions can cause stronger degeneracies
- We don't find any dependence on the **model complexity** but the chosen wavelength grid can influence the predictions

Figure 1 (Right): Shown here is a subset of the complete grid of simulated spectra, demonstrating each cloud case and the extreme examples of our resolution-error grid. Coloured arrows mark three wavelength regions used in computing 'binning paths'.

Figure 2 (Below): The sensitivity map for retrievals on the high cloud case. Blue crosses mark the positions of retrievals, the white markers show the position of the JWST ERS data [1] and dashed coloured lines provide approximate 'binning paths'.





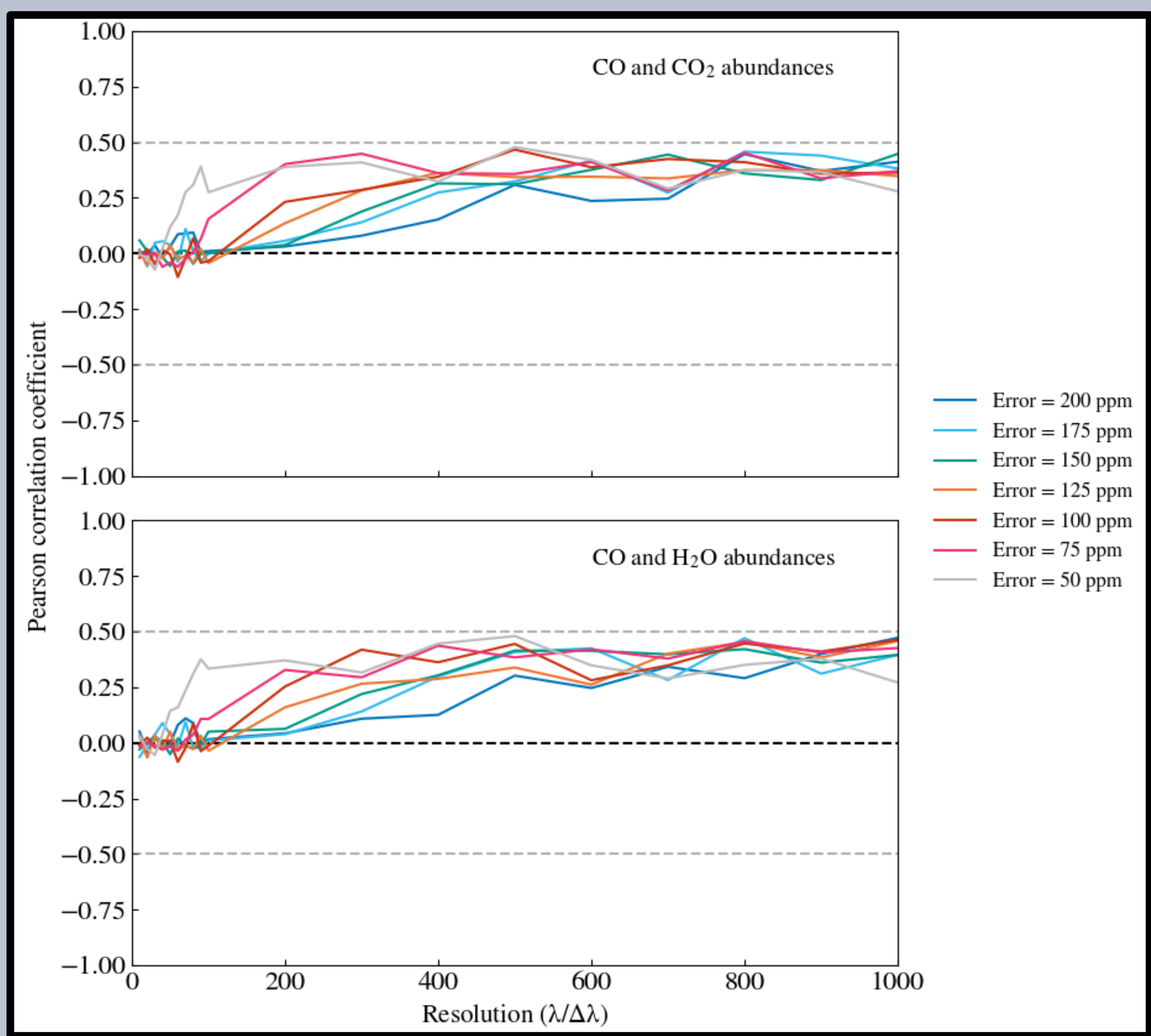
Simulated Spectra & Retrievals:

- We produced a simulated spectrum similar to that of WASP-39b [2] across the wavelength range of the NIRSpec PRISM instrument on JWST [1]
- Retrievals were run on a set of spectra with varying spectral resolution, photometric error and cloud levels (see figure 1) using TauREx3 [3] and MultiNest [4]
  - We quantify the accuracy of the retrieval by plotting the percentage deviation of the retrieved value from the input value of each parameter (see figure 2 for the high cloud case)
- We find the expected boundary between well and poorly retrieved cases is clearly visible and in the high cloud case, the ERS data's position is close to this limit

Correlation Analysis:

- In each case we also quantify the correlation between joint posterior distributions of retrieved parameters using the Pearson correlation coefficient
  - Focusing on cases where the correlation changes as a function of resolution, we find that the log-abundance of CO is the most variable parameter in the high cloud case (see figure 3)

Figure 3 (Right): The Pearson correlation coefficient for the log-abundance of CO with the log-abundances of CO_2 (top) and H_2O (bottom). Note that different lines show retrievals on spectra with different error scales. In both cases, the correlation becomes stronger at higher resolution.



Discussion:

- We advise caution when retrieving atmospheres suspected to have high cloud decks but find that the JWST ERS data [1] is within a well retrieved region
  - Analysis of low cloud and cloud-free scenarios show similar structures but the boundary is pushed towards lower resolutions
- This analysis is specific to a planet like WASP-39b but we will look to generalise this framework in future investigations
- Future studies may also consider whether certain features within a spectrum are more important to retain and how best to bin data given this information

References:

[1] Z. Rustamkulov et al., Nature **614**, 659-663 (2023)

[2] F. Faedi et al., A&A **531**, A40 (2011)

[3] A. F. Al-Refaie, Q. Changeat, I. P. Waldmann and G. Tinetti, The Astrophysical Journal **917**, 37 (2021)

[4] F. Feroz, M. P. Hobson and M. Bridges, MNRAS **398**, 1601 (2009)