ML approaches for exoplanet atmosphere characterization and detrending methods in the Ariel space mission

A. Syty (IAP*), O. Faucoz (CNES), JP. Beaulieu (IAP), P. Drossart (IAP), J. Amiaux (CEA), T. Pichon (CEA), C. Cossou (CEA)

The Ariel space mission is a European Space Agency (ESA) mission that aims to study the atmospheres of a large and diverse catalog of transiting exoplanets (Tinetti et al. 2021). Scheduled for launch in 2029 to the L2 Lagrange point, Ariel will observe a diverse catalog of transiting exoplanets in visible and near-infrared wavelengths (0.5–7.8 μm) via low-resolution spectroscopy. By analyzing starlight filtered through exoplanet atmospheres during transits, the mission will detect wavelength-dependent variations in a planet's apparent radius caused by molecular absorption.

As machine learning methods are increasingly used in many astrophysics fields, these methods have recently arrived in the exoplanet community. The Ariel Consortium puts effort into the machine learning development from data processing to molecular retrievals. Applications include accelerating parameter space exploration, investigating the highly degenerate molecular composition of planet atmospheres in exoplanets' spectra, studying complex undesired artifacts in Ariel data, such as the effect of the jitter of the line of sight of the telescope during the observation, or even interpreting pre-launch calibration data.

In this talk, I will present the Ariel Data Challenge 2024, extended in 2025, a competition hosted by the NeurIPS conference, which gathered 23000 model submissions from almost 1500 participants (Yip et al. 2024). The task of this competition is to extract the atmospheric spectra from every observation, with an estimate of its level of uncertainty. To obtain such a spectrum, we required the participant to detrend many sequential 2D images of the spectral focal plane taken over several hours of observing the exoplanet as it transits in front of its host star. I will also introduce a project in collaboration with CNES and CEA on calibration data for Ariel, measured in laboratories, to flag the pixels having non-nominal behaviors ("bad pixels"), which must be either masked during the data processing step or parametrized to avoid introducing any bias in the scientific data to come. Machine learning is a powerful tool to cluster those pixels according to their behavior, on datacubes containing millions of time series of pixels' responses.

References:

Kai Hou Yip et al., NeurIPS - Ariel Data Challenge 2024. NeurIPS - Ariel Data Challenge 2024. https://kaggle.com/competitions/ariel-data-challenge-2024, 2024. Kaggle.
Giovanna Tinetti et al. "Ariel: Enabling planetary science across light-years". In: arXiv e-prints,arXiv:2104.04824 (Apr. 2021), arXiv:2104.04824. arXiv: 2104.04824

Acknowledgments: This work has received support from France 2030 through the project named Académie Spatiale d'Île-de-France (https://academiespatiale.fr/) managed by the National Research Agency under bearing the reference ANR-23-CMAS-0041, and from the Centre National d'Études Spatiales (CNES).

*Institut d'Astrophysique de Paris, Sorbonne Université, CNRS