PhD topics
Statistical learning for systematics reduction on space coronagraphs

Laboratoires proposants : CNES, LAM, Lagrange
Direction de thèse : Elodie Choquet (LAM), David Mary (Lagrange), Orphée Faucoz (CNES)
Financement : CNES
Structure/Laboratoire d'accueil envisagé : CNES, Toulouse

Profil du candidat (préciser la spécialité du master) :
Applicants must have a MSc degree or equivalent degree, in Computer Science, Physics, or Optical Sciences. Applicants with a background or experience in any of the following topics will be given additional attention: machine learning, optical modeling, wavefront sensing.

Descriptif du sujet (5000 ca.) :
Context:
In three years of operation, coronagraphic imaging with JWST has transformed the field of exoplanet research, providing spectacular images and spectra of exoplanets and dusty disks in the near and mid-infrared. These results have demonstrated the telescope's excellent optical quality and stability, enabling it to achieve detection limits down to 10^-5 in planet/star contrast. As the science programs shift toward observations of fainter exoplanets, several factors become limiting: contamination by the light of the host star at the shortest angular separations and unavoidable instrumental defects, such as inconsistent or suboptimal pixels responses. Reaching the ultimate performances of the instrument requires detecting, understanding, and fixing all such optical and instrumental systematics. This ambitious endeavor requires advanced image processing methods. Several methods exist, with rapid development of AI techniques, for fine modeling of stellar contamination, for detecting the exoplanetary signal within the noise, and for flagging and debiasing detectors read out up-the-ramp. These methods are yet mostly developed for ground-based high-contrast imaging, with limited applications to space coronagraphy.

At the same time, the JWST archives have accumulated a significant amount of data: coronagraphic science images, which can be used to quantify the performances at the detector level, and telemetric measurements from the telescope's wavefront sensor, which can be used to monitor the stability of its optical quality. This database offers the opportunity to explore new image processing methods to better model and eliminate detector and optical systematics, which are critical for exoplanet searches. Developing such methods now is particularly timely to make the most of JWST scientific exploitation in the coming years. Furthermore, NASA's next space telescope, the Roman Space Telescope (RST, launch scheduled for Sept. 2026), for which CNES and ESA are partners, also includes an exoplanet imaging instrument that will be equipped with an active wavefront measurement and control system. The methods developed in this thesis will therefore be applicable to Roman's first data with unprecedented contrast limits.

Thesis program description:

The goal of this thesis is to develop high-contrast image processing methods for space coronagraphs, with a particular emphasis on the systematics limiting JWST and Roman coronagraphic observations.

In the first part of the program, the PhD student will explore an analytical approach, consisting of reconstructing the contaminating coronagraphic response of the star using an existing JWST propagation model (WebbPSF), fed by wavefront measurements recorded by the observatory since its commissioning at a frequency of 2-3 days. This will constitute an initial reference approach based solely on optical modeling (method 1). A second approach will be obtained using existing coronagraphic images of reference stars from the archives. This method, particularly effective on the Hubble archives, needs to be studied for JWST, in a much more stable optical environment (method 2). The reconstructor used here will be principal component analysis (PCA), state of the art in that class of approaches. The student will examine 1) how the inclusion of telemetry data in method 1 can improve the adjustment of the optical model to better remove the starlight contamination; 2) how to improve method 2 by considering alternative reconstructors augmented by AI (e.g. autoencoder).

In the second part of the program, the PhD student will develop and explore hybrid approaches, combining optical modeling with the joint consideration of telemetry data and small-scale archive data in order to efficiently model and remove systematics. The focus on these new AI-based methods will not detract from the attention paid to statistical methodology and, in particular, to accurate data modeling, from the sensors to the final scientific image. Indeed, experience shows that the more care is paid upstream to get information on the optical and statistical properties of the data, the higher the performances of the post-processing methods fed with this information. These methods will be applied on the existing JWST data collection assembled during the first part of the thesis. The detection limits achieved with the hybrid methods will be compared to the state-of-the-art baseline method. Provided a significant gain is demonstrated, they will be submitted for implementation in the L3 stage of the JWST data processing pipeline.

In the third part of the program, the student will collect the first data acquired by the Coronagraph Instrument on Roman, which will be publicly available mid-2027 / early 2028, and investigate the systematics limiting the detection limits of the instrument. They will then adapt the methods developed for JWST to these observations, using both the science images and wavefront sensor telemetry.

References

[1] A.-M. Lagrange, et al., 2025, Nature, 642, 8069
[2] H. Le Coroller et al., 2020, A&A, 639
[3] O. Flasseur et al., 2024, MNRAS, 527
[4] T. Bodrito et al., 2024, MNRAS, 534
[5] B.Y. Feng et al., 2025, IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING, VOL. 11
[6] D. Lafreniere et al. 2007, ApJ, 660, 770,
[7] R. Soummer et al., 2012, ApJL, 755, 2, L28
[8] O. Flasseur et al., 2018, A&A, 616, A138
[9] R. Soummer et al. 2011, ApJ, 741,55
[10] C. Gomez Gonzalez et al. 2018, A&A, 613, A71
[11] K.H. Yip et al. 2020, ECML PKDD 2019
[12] C. Gomez Gonzalez et al., 2016, A&A, 589, A54

[13] A. Kuznetsov et al. 2024, A&A, 687, A221
[14] M. Mâlin et al. 2024, A&A, 690, A316
[15] Timothy D. Brandt, PASP, 136:045004, 2024