Efficient and fast deep learning approaches to denoise large radioastronomy line cubes and to emulate sophisticated astrophysical models

L. Einig, J. Pety, P. Palud, J. Chanussot, A. Roueff, M. Gerin, E. Bron, F. Le Petit and the ORION-B consortium

The interstellar medium (ISM) is an important actor in the evolution of galaxies and provides key diagnostics of their activity, masses and evolutionary state. However, surveys of the atomic and molecular gas, both in the Milky Way and in external galaxies, produce huge position-position-velocity data cubes over wide fields of view with varying signal-to-noise ratios. Besides, inferring the physicals conditions of the ISM from these data requires complex and often slow astrophysical codes.

The overall challenge is to reduce the amount of human supervision required to analyze and interpret these data. I will describe two applications of deep learning to tackle this challenge.

1/ I will first introduce a self-supervised denoising method adapted to molecular line data cubes (Einig et al. 2023). The proposed autoencoder architecture compensates for the lack of redundancy between channels in line data cubes compared to hyperspectral Earth remote sensing data. When applied to a typical data cube of about 10ˆ7 voxels, this method allows to recover the low SNR emission without affecting the signals with high SNR. The proposed method surpasses current state of the art denoising tools, such as ROHSA and GaussPY+, which are based on multiple Gaussian fitting of line profiles.

2/ Numerical simulations are usually too slow to be used in Bayesian inference framework, as it requires numerous model evaluations. Here, I will present a supervised method to derive fast and light neural-network based emulations of a model from a grid of precomputed outputs (Palud et al. 2023). This emulator is compared with four standard classes of interpolation methods used to emulate the Meudon PDR code, a characteristic ISM numerical model. The proposed strategies yield networks that outperform all interpolation methods in terms of accuracy on outputs that have not been used during training. Moreover, these networks are 1,000 times faster than accurate interpolation methods, and require at most 10 times less memory. This paves the way to efficient inferences using wide-field multi-line observations of the ISM. The proposed strategies can easily be adapted to other astrophysical models.

References:

Einig et al. 2023, A&A, 677, A158

Palud et al. 2023, A&A, 678, A198