AI for astronomical spectral line analysis

(Sub)millimeter spectroscopic observations are fundamental tools in astrophysics that allow us to explore the molecular content of the Universe, and provide key constraints on the physics and chemistry of the interstellar gas. In particular, sites where new generations of stars form represent active sites of chemical evolution. This allows the formation of complex organic molecules including prebiotic species in space that are pre-requisites for the emergence of life on Earth.

Molecules in the interstellar gas can be detected through spectral line observations from their rotational transitions. State-of-the-art (sub)millimeter astronomical observatories are capable to measure spectra with a wide frequency range instantaneously providing a vast amount of spectral lines (Fig. 1). However, the increase in sensitivity and band-width leads to data with extremely high complexity, and thus the efficient analysis of complex (sub)millimeter spectra becomes a serious challenge. Currently available tools severely underperform for large data volumes, and therefore we aim here to explore artificial intelligence (IA)/machine learning (ML) algorithms for extracting information from spectral lines. We have to tackle several challenges to adapt these methods for (sub)millimeter spectroscopic analysis of astrophysical data, such as severe line confusion, and varying source specific properties as well as low signal-to-noise. The aim of this project is to test various methods to perform a quick and efficient assessment of the molecular content of complex (sub)millimeter spectra (line identification). The project will explore both statistical methods, such as 1.) spectral alignment 2) ..? as well as 3) deep learning methods.

The expected result is to explore the applicability and the limitations of the above listed methods for astrophysical (sub)millimetric spectral line observations.

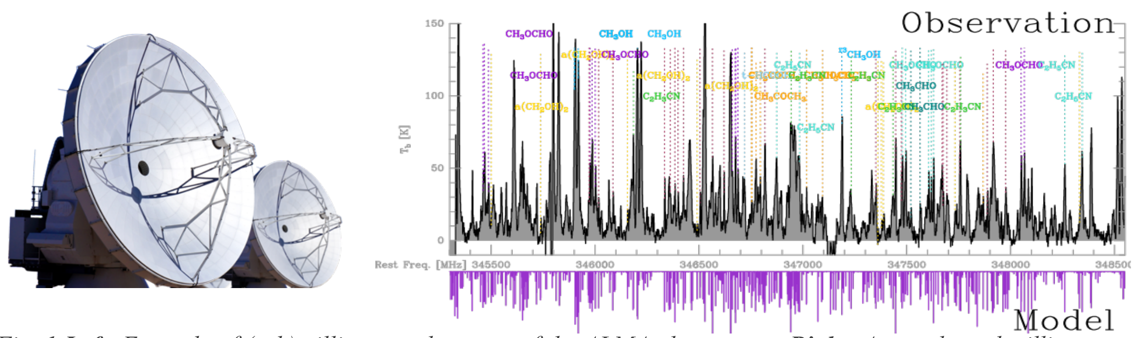

*Fig. 1 **Left**: Example of (sub)millimeter telescopes of the ALMA observatory. **Right:** A complex submillimeter spectrum from the ALMA observatory (gray) covering a frequency range of 3.5 GHz at 345GHz towards a chemically rich source about to form a star 50 times as massive as the Sun. Spectral lines from various COMs are labelled in color after manual identification using a model spectrum (purple). Adapted from Csengeri et al. (2019).*