



Department of Chemistry

MSc Thesis in **Computational/Environmental Chemistry at EMPA**

Molecule Structure Reconstruction Algorithms

Description / Project:

Mass spectrometry is a valuable tool that allows scientists to break down molecules into smaller fragments and determine their structures. While traditional methods involve manually investigating possible arrangements of fragments, modern algorithms are being investigated to automate this process. These algorithms can be divided into two categories: combinatorial-based methods and learning-based methods. Combinatorial-based methods use puzzle-solving approaches, resulting in exponential running times as the problem size increases. Learning-based methods, such as deep learning, utilize large databases and can provide results in a timely manner. However, these methods may lack interpretability, making it difficult to understand how they generate the structure of molecules.

In this master thesis project, we aim to implement and improve upon various algorithms for molecule structure reconstruction. Our focus will be on testing these algorithms using measurements from a tandem chromatogram-mass spectrometer located at Jungfraujoch, operated by the laboratory of climate gases of Empa. This instrument measures halogenated compounds, which are small molecules responsible for greenhouse gas emissions and ozone layer depletion. The algorithms developed during this project will be integrated into the laboratory's existing program for identifying new pollutants. Currently, the program can predict the molecular formula of a compound, and our project aims to expand its capabilities to predict the structure of unknown compounds as well.

Methods:

Programming using Python to expand an existing program for non-target search for new organic trace gases in the atmosphere, related to climate change and air pollution.

What we expect from you:

We are looking for a person with good programming skills in Python and a basic understanding of analytical chemistry. Additionally, experience with machine learning and algorithms is a plus.

Starting date: Fall semester 2023

Research Groups: Dr. Stefan Reimann (Empa); Prof. Jürg Hutter (UZH)

Supervisor / Contact: Lionel Constantin (Empa), Dr. Stefan Reimann (Empa)

lionel.constantin@empa.ch, stefan.reimann@empa.ch

<https://www.empa.ch/web/s503/climate-gases>