ATOMS Lab Rotation Plan

Prof. Tyler R. Josephson

Fall 2020

E-mail: tjo@umbc.edu
Web: **cbee.umbc.edu/josephson/**
WebEx Personal Room: umbc.webex.com/meet/tjo
Office Hours: By Appointment

Welcome to the ATOMS Lab!

I'm excited to guide you through a rotation in the ATOMS lab (AI & Theory-Oriented Molecular Science) over the next 6-8 weeks.

The ATOMS lab seeks to:

1. Predict adsorption and catalysis in nanoporous materials and in solution using computational chemistry tools
2. Develop artificial intelligence tools (especially symbolic regression and automated reasoning) to learn complex relationships in material performance and thermodynamic properties
3. Apply these tools to gain molecular-scale insights on processes in the environment to enable new solutions for cleaner air and water, in collaboration with experimental partners

To get a feel for what we do, your rotation will consist of a number of mini-projects that will expose you to quantum chemistry and molecular simulation software. You will write some code from scratch for data management, fitting, and visualization; transferable skills that will pay dividends whether you join this group or not. As you finish each mini-project, please record your time spent - I'll appreciate these measurements as I plan the future of this rotation program.

In the ATOMS lab, we focus on Python and Julia, popular and up-and-coming languages for scientific computing. We access HPC (high-performance computing) resources through a Linux environment, so bash commands are also helpful (though it's good practice to limit the scripting you do with bash; Python is much more flexible). Though not required for your rotation, typesetting reports and papers in LaTeXwill be required if you join the group.

Finally, I encourage you to spend *at least* 3 distraction-free hours every week reading textbooks or papers (see the final section of this document), no matter how excited you are about writing code and running calculations. Whether deep-diving into textbooks and theory or assessing review papers on a surface level, it's important to read, and read *consistently* to grow and benefit from the work of others.

Rotation Objectives

By the end of this rotation, successful students will be able to produce quality results for three of the following tasks (selected in consultation with Prof. Josephson). Students can pursue stretch goals of their interest during their rotation, keeping in mind their relevance to funded projects.

- Pick a dataset from the NIST Adsorption database, fit the Langmuir and dual-site Langmuir isotherms to it, and plot using Python or Julia
  - Why? Coding and using theories of adsorption is essential in the ATOMS group
  - Stretch goal: Rediscover Langmuir using symbolic regression
  - Future task (Fall 2021): Write down axioms for Langmuir and prove using Keymaera
- Using a chemical science database of your choice, produce a high-quality visualization with Python or Julia, write a descriptive caption, and store the data in tabular, machine-readable format (like JSON or CSV)
  - Why? Visualization is essential to understand the data we produce, and quality figures are essential for science communication. To promote data transparency and reproducibility, papers from the ATOMS group must include tabulated files for all results in the Supporting Information (if file sizes are reasonable for journals to host), and generating plots and tables using scripts is MUCH safer than using spreadsheets.
  - Stretch goal: Perform a meaningful regression or classification task on the dataset
- Simulate vapor-liquid equilibrium in the Gibbs ensemble using MCCCS-MN (Monte Carlo for Complex Chemical Systems-Minnesota)
  - Why? MC is a fundamental technique in the ATOMS group for simulating phase coexistence, gas- and liquid-phase adsorption, and predicting thermodynamic properties
  - Stretch goal: Calculate critical properties, and contribute to TraPPE validation data
- Optimize the unit cell parameters of zeolite A using plane-wave DFT (CP2K or ASE)
  - Why? Plane-wave DFT is great for quantum simulations of novel molecules and reactive phenomena in solids and liquids. Atomic simulation environment (ASE) is a great scripting platform for high-throughput calculations.
  - Stretch goals: High-throughput calculations in some chemical or materials space using GFN-xTB or GFN-FF set up using ASE (estimate comp resources, and get authorization from Prof. Josephson before using)
- Calculate the heat of combustion of methane using Gaussian
  - Why? Cluster-based electronic structure calculations are useful when we don't have accurate force fields, and are especially important for force field parameterization, chemical reactions, and solvation
  - Stretch goals: Fit the torsion potential for butane, find transition states for F–F–H and decarboxylation of butenoic acid, calculate solvation free energy and pKa of phenol