Computational Summer Projects in the Rice Astroparticle Laboratory

My group is an interdisciplinary group focusing on data-intensive astroparticle physics. More information on the group is available at astroparticle.rice.edu. In the summer of 2020, we have up to 6 projects that can be done remotely, where depending on the level of interest and funding, these will be either full-time or part-time positions that can be done from anywhere.

Each project will interact directly with a PhD student, where you will be expected to meet through Slack and Zoom with them regularly, and ultimately supervised by PI Tunnell with weekly check-ins. Furthermore, you are expected to participate in weekly Zoom group meetings where you will present updates and plans for the coming week. These projects will run for 10 weeks, where at the end of them you will be expected to deliver a written report, presentation, and reproducible code repository. Contingent on the success of the project and dynamic, these may transition into research courses and/or senior theses.

1. Signal processing and data optimization (supervisor: Andaloro)
    a. The XENON experiments produce petabytes of time series data from flash analog-to-digital convertors. This data undergoes a series of signal processing steps. Our experiment depends on efficient and accurate processing of this data so as to not bias scientific measurements, reduce storage costs, and ensure data can be monitored live when operating the experiment. A range of topics in this realm exist ranging from normal filtering routines, optimal filtering, and compressive sensing. After achieving this primary project goal, you will have the background to then explore any machine learning techniques that can help us with compression and filtering of particle physics data.
    b. Preferred skills: Python, basic knowledge of waveform processing (though not required)
2. Microphysics simulations with the NEST collaboration (supervisor: Andaloro)
    a. The group is part of the NEST microphysics collaboration and maintains a package called nestpy. These models for particle interactions with xenon and argon serve as the basis of most neutrino and dark matter experiments. You will help develop a validation system that compares these simulations to data from various experimental sources. This project requires running simple simulations and comparing to known measurements to alert us when our model disagrees with data. This will develop software engineering skills as such an automated continuous integration system, and you will learn to perform functional testing of these simulations using TravisCI. Expansions of this project (after the primary goals) could include analyzing incoming xenon and argon detector data from

experiments which come online this summer against your simulations, and verifying our model for use in these next-generation dark matter experiments.
    b. Preferred skills: Python, some interest/background in software development, C++
3. Clustering of time-series data [Multiple spots] (supervisor: joint with Dr. Merenyi STAT)
    a. These detectors produce many different types of signals (scintillation, ionization, photoionization backgrounds) that can be measured independently of one another. Specifically, they manifest themselves in spatiotemporal data where these signals must be separately clustered with extreme efficiency. Millions of backgrounds signals can never be classified as a dark matter signal. This project is an unsupervised learning project with the Rice statistics department, where we form the physics part of the project in defining the problem and example classifiers, and they work on using neuromaps (a.k.a. self-organizing maps) to solve the clustering problem.
4. Front-end web development for time series visualization (supervisor: Liu)
    a. As the XENONnT Dark Matter experiment becomes operational, a critical task is being able to visualize parts of the time-series data with what we call 'waveform watching.' This is used to diagnose new types of problems with the detector and data that may not clearly be visible in higher-level reconstructed quantities about the data. This project will involve creating a front-end to visual and interact with such data. The technologies used are nominally a mix of web frameworks (Flask, NodeJS, Ajax) and data management tools providing random access to petabytes of data (GridFTP, Rucio, MongoDB).
5. Graphical neural networks for dark matter reconstruction (supervisor: Liang)
    a. The graphical nature of particle physics data makes it a natural candidate for graphical neural networks. As part of the DIDACTS project, we are looking for students interesting in solving the inverse problem using such regressions. This involves inferring underlying properties of particle interactions. The student would focus on validation of novel networks by analyzing data from the XENON1T experiment. You will be expected to work with both physics and ECE people.
6. Likelihood-free inference for dark matter reconstruction (supervisor: Liang)
    a. Related to the above, this is a different approach to a similar problem requiring a stronger mathematical background and willingness to learn Bayesian statistics.