Number of Regions Created by Random Chords in Circle
Shi Feng, Senior, Mathematics
Mentor: Soumik Pal, Mathematics

Take a circular sheet of paper. Pick two random points on the edge of the circle and draw a line segment between them. Repeat this many times. Then you take scissors and cut the circle along those line segments. How many pieces of paper do you get? Such questions are a part of a subject called geometric probability. Examples include the famous 18th century Buffon’s needle problem that lets one estimate the value of pi by running a random experiment on the computer. In modern times geometric probability is used to estimate the coverage of cell phone towers and communicating drones. I review historical research in geometric probability and graph theory dating between the 18th and 20th centuries to find the expectation and the variance of the number of pieces in the circular sheet. Then I prove a Central Limit-type behavior (converges to normal distribution) of the number of pieces that is conjectured from experiments. The proof is from a modern result of probability theory, called Stein’s method for the Central Limit Theorem. Several open questions remain to be answered, including how likely one is to get a piece with a very large or small area. The results of this research enrich the content of geometric probability with possible applications in geology, image processing, transportation, and so on.

Kronecker Products of Perron Similarities
Janelle Marie Dockter, Recent Graduate, Computer Science & Software Engineering, Mathematics (Bothell Campus)
Jonathan D Tu, Non-Matriculated, Bothell Non Matriculated Student
Robert Perry, Non-Matriculated, Bothell Non Matriculated Student
Mentor: Pietro Paparella, Engineering & Mathematics, University of Washington Bothell

An invertible matrix is called a Perron similarity if one of its columns and the corresponding row of its inverse are both nonnegative or both nonpositive, a definition introduced in a previous work by Johnson and Paparella (2016). For a given matrix S, the spectracone is a polyhedral cone formed by the vectors that produce nonnegative matrices when multiplied on the left by S and the right by S-inverse. The spectratope is a set defined similarly, with the added condition that the included vectors have an infinity norm of one. This work identifies previously unknown relationships between Perron similarities and their Kronecker products. First, if two matrices are Perron similarities, their Kronecker product is also a Perron similarity. Second, the Kronecker product of spectracones (spectratopes) is a subset of the spectracone (spectratope) of the Kronecker product. In addition, if two matrices are Perron similarities with dimensions greater than 1, then the Kronecker product of the spectracones (spectratopes) is properly contained in the spectracone (spectratope) of the Kronecker product. Another significant result is if S is a Perron similarity, then the ray through the all-ones vector is properly contained in the spectracone of S, but the converse is not true. One reason this work is relevant is that Perron similarities were previously defined to have a necessary and sufficient condition that the ray through the all-ones vector is properly contained in the spectracone of S, but the converse is not true. One reason this work is relevant is that Perron similarities were previously defined to have a necessary and sufficient condition that the ray through the all-ones vector is properly contained in the spectracone. With the publication of this article, the condition was found to only be necessary, and the definition was corrected. Perron similarities are important when examining the nonnegative inverse eigenvalue problem, or NIEP, which is concerned with the spectrum of nonnegative matrices. This research expands on previous work related to the NIEP by examining relationships between Perron similarities and their Kronecker products.

Improving Respondent Privacy and Addressing Lack of Trust in RRT Models
Joia W Zhang, Senior, Statistics: Data Science
Mentor: Sat Gupta, Math Stats, UNC Greensboro

When conducting surveys containing sensitive questions, Social Desirability Bias (SDB), people’s tendency to provide
socially acceptable answers rather than truthful ones, often leads to low response rate or worse, untruthful responding. Randomized Response Techniques (RRT) combat SDB by allowing respondents to provide scrambled responses. However, if a respondent does not trust the RRT model, data accuracy will be compromised. Lack of trust in binary RRT models has been shown to lead to untruthfulness, and thus unreliable data and unreliable estimates of the sensitive trait. Yet, no quantitative RRT model currently accounts for respondent lack of trust. We propose an Optional Enhanced Trust (OET) RRT model that extends Warner’s Additive Model, a quantitative RRT model with additive noise, by allowing additional multiplicative noise if the respondent does not trust the Warner Additive Model. Using the Unified Measure, a combined metric of respondent privacy and model efficiency, we demonstrate both theoretically and empirically that the proposed OET model is superior to existing models in terms of respondent privacy and estimator accuracy.

Photometric Classification of Evolved Massive Stars: High-resolution Spectroscopic Validation
Ishan Coutinho, Junior, Pre-Sciences
Mentor: Trevor Dorn-Wallenstein, Astronomy
Mentor: Emily Levesque, Astronomy

In order to compare observations of massive stars with theoretical predictions, stars must be classified accurately. Classification traditionally relies upon expensive new telescope observations. As the field of astronomy enters a new era of Big Data, Modern computational techniques may be used in place of these methods; however the results still require rigorous validation in order to be trusted. Recently, Dorn-Wallenstein et al. (2021) utilized a novel machine learning technique to classify a large sample of massive stars. This resulted in putative classifications for ~2550 stars. Our project serves as a follow-up to validate the results of Dorn-Wallenstein et al. and identify stars with rare evolutionary phases that are most useful for probing stellar evolution. We test the hypothesis that these classifications are reliable by obtaining new observations with the ARCES instrument mounted on the Apache Point Observatory 3.5-meter telescope. Using these data, we assigned classifications to the stars in our sample, focusing on identifying rare objects and evolved supergiants. To this end, we have developed custom software designed to navigate through the key features in our observations and allow for easy identification of an object’s spectral type (i.e., its evolutionary stage). We find that our observations support the classifications made by Dorn-Wallenstein et al. Our future work will focus on expanding our sample with further observations in the Southern Hemisphere. This work is critical in order to prepare for the age of Big Data in astronomy.

A Model for the Momentum Spectrum of Energetic Particles at Interplanetary Shocks Including Acceleration and Escape
Thomas Minh Do, Senior, Astronomy, Physics: Comprehensive Physics
UW Honors Program
Mentor: Federico Fraschetti
Mentor: Manpreet Singh, Department of Planetary Sciences, Lunar and Planetary Laboratory, University of Arizona

Charged particles in the heliosphere can be continuously accelerated by interplanetary shocks and eventually escape from these shocks without returning to it. Acceleration and escape are highly intertwined and both contribute to the shaping of the particles’ momentum spectrum at the shock. The simplest model which describes this phenomenon is called Diffusive Shock Acceleration (DSA). DSA has been very successful in describing several observations. However, DSA does not include an energy-dependent escape from the foreshock region. We expand upon DSA by presenting a model for interplanetary shock acceleration which includes this energy-dependent particle escape. We analytically solve a one-dimensional transport equation with a diffusion coefficient and an escape time that describes both the turbulence self-generated by the shock and the far upstream pre-existing turbulence. We consider the case where a shock encounters a population of pre-existing charged particles with a power law energy distribution as measured by spacecraft. We find that at lower energies our solution is concave, whereas at higher energies it asymptotically approaches a power law whose slope depends on the original energy spectrum’s power law index and shock parameters. We fitted the solution obtained from this transport equation to shock data measured from multiple shock events collected by ACE/EPAM (the Electron, Proton, and Alpha Monitor aboard the Advanced Composition Explorer spacecraft). We find that for the shock events considered, our model’s best fit parameters match very well with the predicted values, obtained by using the measured shock parameters. From this model, we can better understand the mechanism of interplanetary shock acceleration and how this phenomenon energizes charged particles near the Sun and around other objects (for example, blazars and supernova remnants).

Identifying and Estimating the Faint Light Emitted From the Diffuse Gas Around Galaxies
Erik Solhaug, Senior, Astronomy, Physics: Comprehensive Physics
Mary Gates Scholar, UW Honors Program
Mentor: Matthew McQuinn, Astronomy

The circumgalactic medium (CGM) is the extended gaseous halo that typically surrounds the visible parts of a galaxy - the beautiful spiraling disks that one may recognize from an image captured by the Hubble Space Telescope. Although
the CGM harbors no stars, making it only faintly illuminated by the interior galaxy and background sources, it is believed to contain far more mass than the galaxy itself. Understanding the CGM is thus key to understanding galaxy formation and address questions such as: Why are some galaxies red and some blue? Why are we only observing ~20% of the baryons (regular matter, not dark matter) we should observe in galaxies? How do galaxies like our own sustain star formation, enabling the particle diversity we see everywhere around us? The way we currently observe the CGM’s properties is through absorption spectroscopy by using a bright background light source (quasar) and observing how the light is absorbed as it travels through the CGM before reaching our telescopes. However, this limits our observations to only a “pinhole” view of the CGM’s properties. With the advent of instruments sensitive enough to observe the light emitted from the CGM, we will be able to create so-called “emission-maps” of the full state of the gas - including temperature, density, and element abundances. Our project has identified what wavelengths (ionic lines) are most observable with more sensitive telescopes, some of which are already underway. I have run computer simulations in the Cloudy program to identify these emission lines and developed models estimating their intensity. The culmination of our work shows that many of these CGM emission lines are detectable with feasible instruments in the near future, laying the groundwork to justify future missions targeting these specific lines in order to investigate the most pressing questions of galaxy evolution.

An Application of Convolutional Neural Networks with Solutions to the Black-Scholes Equation for Stock Option Price Prediction

James Zheng (James) Cao, Senior, Mathematics
Duncan Du, Senior, Computer Science
Mentor: Kirill Golubnichiy, Math

Mentored by Kirill Golubnichiy, this research project aims to apply mathematical finance and machine learning (ML) to forecast stock option prices. We create and evaluate new empirical mathematical models for the Black-Scholes equation to analyze data for 177,000 companies. For each company, we have 13 elements including stock and options’ daily prices, volatility, minimizer, etc. Because the market is so complicated that there exists no perfect model, we apply ML to train algorithms to make the best prediction. We first analyze several existing stock and option prediction models: Quasi-Reversibility Method (QRM), Binary Classification, and Regression Neural Network (RNN) ML. QRM is an analytical and analytical approach to find the minimizer by solving the Black-Scholes equation as an ill-posed problem; whereas the latter two combine QRM with ML. The current stage of research attempts to combine QRM with Convolutional Neural Networks (CNN), which learns information across a large number of data points simultaneously. Our current focus is to apply CNN to generate new results by programming, implementing, testing, and validating sample data. We will compare our CNN model with previous models to see if it is possible to achieve a higher profitable rate. If our CNN model successfully forecasts prices for a majority of stock options, it might be possible to deploy the model in the real world and help investors make better investment decisions.