SURF

CALTECH

Student-Faculty Programs

2020 Student Abstract Book

STUDENT-FACULTY PROGRAMS

2020 Abstract Book

This document contains the abstracts of the research projects conducted by students in all programs coordinated by Caltech's Student-Faculty Programs Office for the summer of 2020.

Table of Contents

Simulating and Analyzing the Hyperfine Structure of Erbium-167 Doped Yttrium Orthosilicate With Spectral Hole Burning Spectroscopy
Sébastien Abadi
Mentors: Andrei Faraon and Mi Lei

$^{167}Er^{3+}$-doped yttrium orthosilicate ($^{167}Er^{3+}:Y_2SiO_5$) is a promising system for the implementation of long-lived multimode quantum memory at telecommunication wavelengths, which is important for the realization of long-distance quantum communication. Optical quantum memory has already been implemented in this system. The next goal is to implement spin-wave quantum memory using hyperfine spin states, which will improve storage times and enable on-demand recall. It is thus of interest to determine the hyperfine energy splittings of $^{167}Er^{3+}:Y_2SiO_5$ at the magnetic field used by the Faraon group, which can be accomplished via spectral hole burning (SHB) spectroscopy. To that effect, I created simulations to reproduce SHB spectra of $^{167}Er^{3+}:Y_2SiO_5$ under a stronger magnetic field where the hyperfine structure is known. With those simulations, I improved my understanding of how the burned hole position and linewidth affects which energy splittings can be inferred. I then made predictions about what sets of SHB experiments are sufficient to determine the hyperfine structure of interest. Fitting procedures were also developed so that once that structure is known, relative oscillator strengths of the optical transitions can be determined. Future work will consist of acquiring SHB spectra and determining the hyperfine structure from those spectra.

Developing Digital Infrastructure for Synthesizing and Probing Quantum Materials
Adam Abbas
Mentor: Joseph Falson

As a new group at Caltech, the Falson Lab has been in the process of being constructed for the past year. To complement the soon to be completed physical lab, a digital infrastructure was developed to provide a seamless integration of lab equipment, computing devices, and all other aspects of the lab's network. The Raspberry Pi computing device was chosen as an optimal way to communicate with individual pieces of equipment, due to both its space and processor efficiency. A network of these was designed, connecting each piece of critical equipment to a Pi, allowing for easy communication and programming for equipment. A network-based digital backup solution was also identified to connect to this network and guarantee the safety of our data from device failures or ransomware.

Along with planning the network's infrastructure, it was important to ensure that it would be possible to control the physical lab equipment. This involved developing combined systems of software and hardware based on the requirements of each instrument. Software was developed to monitor and display the outputs of lab devices, such as the cryostat temperature monitors, over time. Circuits were designed to be able to successfully communicate with a vacuum Pirani gauge, as well as to control stepper motors for a measurement apparatus while balancing power constraints for both input and output. Finally, communication protocols were designed to be able to control the equipment for the superconducting magnet that will be used in the cryostat. All of the work done will be able to be implemented shortly after the opening of the lab and will allow for a strong beginning to its operations.

Imaging Simulated Exoplanets With the James Webb Space Telescope
Jea Adams
Mentors: Dimitri Mawet and Jason Wang

The launch of the James Webb Space Telescope (JWST) in 2021 will revolutionize direct imaging by providing the first characterizations of exoplanets within 3-13 microns. Our recent ability to empirically model and remove starlight from imaged extrasolar systems with post-processing algorithms like py has proven vital to the discovery of exoplanets with high contrast imaging using current telescopes. However, such software systems are not presently prepared to handle the incoming JWST data. We aim to develop and refine new and existing data analysis tools that use pyKLIP in order to have a functional data reduction pipeline prior to the telescope's launch. Using simulated data of an extrasolar system through the telescope's Near-Infrared Camera (NIRCAM), we confirmed pyKLIP's ability to fit and subtract the stellar PSF from an image and reveal exoplanets hidden underneath. We computed NIRCAM's sensitivity to exoplanets, and upgraded pyKLIP to correct for coronagraphic usage when determining the astrometry and photometry of exoplanets. As part of JWST's Early Release Programs, we share their goal of demonstrating JWST's capabilities and swiftly disseminating them to the community. We therefore documented our analysis steps as public notebooks to help astronomers undertake the most suitable observing programs during JWST's limited lifetime.

Developing a Simulator for Complex Quadrotor Maneuvers
Kasey Adams
Mentors: Soon-Jo Chung and Michael O'Connell

In the field of autonomous flight, simulations allow for testing to be done before expensive hardware needs to be risked in testing complex maneuvers. Thus, the development of a realistic simulator is imperative for testing the algorithmically generated, complex maneuvers that model systems being researched in the Center for Autonomous

Systems and Technology (CAST) lab. We developed a 6 degree of freedom quadrotor simulator in python to test these maneuvers and algorithms. The simulator incorporates a Rapidly-Exploring Random Tree (RRT) path planning algorithm, which will plan agile maneuvers between two states following environmental and dynamic constraints. The simulator emulates real hardware by having motor delay and by running the controller and simulator at different frequencies. The simulator will allow the simulation and testing of more complex algorithms and maneuvers than were able to be simulated and tested before, and in the future, it will allow for a framework of developing the simulator to support multiple bodies and swarms.

(1) Evaluating Hindsight During the COVID-19 Pandemic; (2) EEG Correlates of Pre-Clinical Alzheimer's Disease
Sara Adams
Mentors: Shinsuke Shimojo and Daw-An Wu

(1) Hindsight bias is the tendency to recall events as having been more predictable than they actually were. This topic is particularly relevant for COVID-19: the actions we take, both individually and as a society, are shaped by our predictions of what is to come. In later evaluating our response, it is important to understand the extent to which our recollection is biased. Thus, we are conducting a global survey of 500+ individuals in order to gauge their opinions at particular points in the pandemic and how they recall those opinions later on. Preliminary results have confirmed that individuals do exhibit hindsight bias regarding the pandemic, and ongoing surveys are evaluating contributing factors to this bias. (2) In addition to the COVID-19 project, we have been evaluating how EEG may be able to be used in order to aid the development of non-invasive early diagnoses of Alzheimer's disease. Pilot experiments involving EEG alongside a Simon task in individuals who have biomarkers suggesting pre-clinical Alzheimer's (CSF $A\beta42$/tau ratio > 2.7132) have demonstrated that participants with these markers may have a difference in readiness potential compared to a control group.

Simulating Scalar Vortex Fiber Nulling for Telescopes With Segmented Primary Mirrors
Maximilian Adang
Mentors: Dimitri Mawet and Daniel Echeverri

Direct imaging and characterizing exoplanets at close angular separations from their host stars is a challenge in modern astronomy that high contrast imaging is trying to solve. Vortex Fiber Nulling (VFN) is an interferometric technique combining the phase pattern of a vortex coronagraph mask with the light coupling properties of a single mode fiber, nulling the starlight in the image to improve contrast between the exoplanet and the star at small separations. A pupil-plane vector vortex-based version of VFN has been demonstrated in a laboratory at Caltech. This implementation works well but is sensitive to polarization effects in the light that reduce the achievable null. I am exploring the viability of scalar VFN as an alternative for telescopes with segmented primary mirrors by imprinting the phase pattern across the segments of the primary instead of using a vector vortex mask. This is being done by simulating the phase pattern for the Keck telescope and optimizing the ratio of planet light to starlight coupling into the single mode fiber and the system's integration time. This technique produces a null unlimited by polarization effects and enables immediate on-sky testing of VFN with telescopes such as Keck, and eventually TMT and GMT.

Follow-Up of Supernovae Using the Liverpool Telescope
Yasmin Afshar
Mentor: Daniel Perley

The Bright Transient Survey, a project within the Zwicky Transient Facility, aims to classify every bright transient object in the night sky. Reliable classification of supernovae and investigation of properties of their host galaxies may lead to a better understanding of the physics of supernovae and answer questions about the progenitors to the explosion and the mechanism of the collapse. Follow-up data of bright transients were processed with Python, using packages such as emcee and sncosmo to implement nonlinear methods of parametric modeling. Markov Chain Monte Carlo techniques can be used to fit light curves of supernovae, from which useful parameters may be extracted to improve measurements such as timescales and peak luminosities of supernovae, thereby aiding in the accuracy of their classification. We present parameters from fitting methods that were explored and examine the relationships between pairs of model parameters for various supernova subtypes.

Negative Electrothermal Feedback in Thermal Kinetic Inductance Detectors Over High Readout Power Regimes
Shubh Agrawal
Mentors: Jamie Bock and Bryan Steinbach

The BICEP Array is the latest generation of instruments of the BICEP Keck Collaboration for Cosmic Microwave Background polarimetric measurements. It comprises four 550mm telescopes observing between 30-270 GHz, aiming to generate sensitive CMB polarization maps for the detection of B-mode polarization signal, which would be the first direct evidence of inflationary gravitational waves.

Superconducting micro-resonators are micrometre-scale LCR circuits operating at millikelvin temperatures, created by the inductance and resistance of lithographically patterned superconducting films under alternating current. Thermal Kinetic Inductance Detectors (TKIDs) use several superconducting micro-resonators multiplexed in the radio-frequency domain on a single readout line and exhibit the superior noise performance of traditional Transition Edge Sensor (TES) bolometers, allowing for large detector counts needed in future millimetre-wave instrumentation. Here, TKIDs fabricated by the Caltech Observational Cosmology group are inspected at high readout powers for negative electrothermal feedback. This involved collecting such frequency sweep data under varying optical loading and numerically testing physical models for the observed hysteresis. Such observations would enable operating the detectors at higher loop gain and higher linearity in responsivity, required for environment-independent, fast-processable, and repeatable measurements.

Implications of analyzing negative electrothermal feedback extend to general superconducting micro-resonators that are employed in dark matter search, neutrino mass, quantum circuit readout, and n-body simulation.

NDI GUI: A Graphical Interface for Data Analysis in Neuroscience
Gabriel Aguiar
Mentor: Stephen Van Hooser

Neuroscience Data Interface (NDI) was designed to combat the lack of standardization in the neuroscience community. The interface provides tools that facilitate the exchange and analysis of data between research groups, without regard for the format and organization of the raw data. In its current form, NDI can be operated through a command-line interface, which requires the user to a) be comfortable with basic programming conventions, b) be familiar with the operations provided by NDI, and c) interpret the effects of calculations from simple text output. A graphical interface was developed to allow the user to view or manipulate the layout and contents of an experiment in an intuitive and user-friendly environment. Interactions allowed by the graphical interface include searching through documents, viewing pertinent information and related files, organizing elements of an experiment, and mapping the progression of data through those elements.

Plasma Effects of Ioffe-Pritchard Coils for Tritium Beta Decay
Nezir Alic
Mentors: Joseph Formaggio and Ryan B. Patterson

Tritium Beta Decay allows for the indirect determination of neutrino mass through the collection of data regarding electron energies. To carry out such an experiment, it is first necessary to capture tritium atoms. In the case of Project 8, this is accomplished magnetically through the use of Ioffe-Pritchard coils. A particular design of Ioffe-Pritchard coils with specific shape, dimensions and current value is simulated in Kassiopeia, a particle tracking software. This simulation will be used to study plasma effects that occur within the arrangement. The fluctuations of charge that exist within a plasma give rise to small potentials that may have an effect on the energy distributions and other properties of the particles of interest. Thus, to help with this problem, data concerning energy losses from scattering, among other phenomena and statistics, will be collected and analyzed.

Projectors for Evaluation of Observables With Two Dimensional Tensor Networks
Alessio Amaolo
Mentors: Garnet Chan and Matthew O'Rourke

A major challenge in modern physics is the accurate microscopic description of strongly-correlated quantum many-body systems. Tensor networks are a promising way of representing complicated quantum states and computing their observable properties. We characterize the accuracy of a numerical technique for speeding up observable computations and enhancing the stability of tensor network optimizations. This technique works by "recycling" computational operations that are approximately conserved between the evaluation of two different, yet related tensor networks.

Designing and Manufacturing an Autoinjector Testing Fixture Which Mimics the Human Arm
Eric Amaro
Mentors: Joseph Shepherd and Donner Schoeffler

Autoinjectors aid in reducing the uneasiness of self-administering necessary injections. Research performed on autoinjectors strives to further ease the hesitation of self-administering. However, testing has been done on static fixtures that do not mimic the motions of the human arm. Developing a testing device that can actuate the autoinjector and accurately respond to its violent recoil provides more accurate and reliable test results, furthering research on this topic. To design the testing fixture, various constraints must be met. These include the human arm's length, mass and center of mass, as well as the reactive forces from the joints. The fixture should be simple to manufacture. Therefore, it should be fabricated from commercially available material, have a low parts count and be straightforward to assemble and test. Given these constraints, I was able to design a human arm that could perform the tasks it was required to do. My continuation of the project will involve finalizing the drawings so that the fixture can be prototyped and manufactured for initial testing.

Scheduling Machine Learning Jobs in the Cloud: Performance and Energy Considerations
Vivek Anand
Mentor: Adam Wierman

Machine learning models are increasingly important in several domains with big data processing. Many of these models are computationally intensive and require days to run. Consequently, they are run on expensive state of the art hardware, usually in the cloud. Due to this expensive hardware, efficiently scheduling such jobs on limited hardware resources is paramount. While optimizing performance metrics is critical, equally important is the energy consumed due to their huge power requirements. Therefore, it is essential to not only schedule such jobs efficiently but also minimize energy consumption. Hence, we are developing an algorithm to schedule machine learning jobs on multiple machines with a dual objective of both performance and energy. So far, we have successfully examined various candidate heuristics to develop a concrete algorithm, built simulators to calculate empirical bounds for the objective and analyzed their results. Currently, we are proving theoretical bounds on these heuristics to develop an algorithm that schedules machine learning jobs with theoretical guarantees on energy consumption and performance. Subsequently, we plan on testing our algorithm on real life workloads to verify and fine tune its performance in practice.

Fast and Adaptive Simulations of Turbulent Vortex Rings
Rahul Arun
Mentors: Tim Colonius, Benedikt Dorschner, and Ke Yu

While the energy cascade associated with turbulent flow has been understood from a statistical point of view for many decades, there remain many open questions regarding the physical, nonlinear mechanisms by which turbulent motion is sustained in specific flows. As a prototypical example of this process, we study the dynamics of turbulent vortex rings by numerically solving the incompressible Navier-Stokes equations. We utilize a recently developed, second-order accurate fast lattice Green's function method to discretize the governing equations on a formally unbounded, staggered Cartesian grid. Computation time is minimized by adapting the finite computational region to coincide with the vortical flow region and by implementing a block-wise adaptive mesh refinement technique. To verify the accuracy of our solutions, we reproduce four integral quantities from previous solutions: the enstrophy, kinetic energy, impulse, and propagation velocity for a laminar ring at Re_{Γ_0} = 500 and an initially laminar ring at Re_{Γ_0} = 7,500. For the latter case, we examine the turbulent regime and study the formation and shedding of hairpin vortices and the corresponding step-wise decay of circulation. Simulations of a vortex ring at Re_{Γ_0} = 20,000 are currently in progress, and we aim to use these results to characterize the mechanisms by which turbulence is sustained, including vortex reconnection.

Comparative Analysis of Workflows for Processing Single Cell mRNA Sequencing Data
Joeyta Banerjee
Mentor: Lior Pachter

Single-cell mRNA sequencing allows for large amounts of gene expression data to be collected and analyzed at once. A major challenge of such a large amount of data is extracting meaningful information. Machine learning approaches to data analysis generally involve methods of dimensionality reduction that condense the number of variables of each cell while trying to preserve the information. However, some information is always lost in this process, and we are aiming to see how much can be kept without significantly increasing the algorithm time. Our approach focuses on using Neighborhood Component Analysis (NCA), an uncommonly used technique, and comparing the results to those of published papers using other methods. The main difference is that NCA preserves the local structure of the data, not just the global structure like many other methods do.

Quantum Imaginary Time Evolution for the Fermi-Hubbard Model
Anthony Bao
Mentors: Austin Minnich and Adrian Tan

The recent development of a quantum imaginary time evolution (QITE) algorithm has promising implications for quantum simulation on near-term quantum computers. We apply the QITE algorithm to the Fermi-Hubbard model to solve for the ground state of its Hamiltonian. The Fermi-Hubbard model is a cornerstone of condensed matter theory with wide-ranging applications in solid-state physics, but its simulation is classically difficult. We focus on implementing an efficient simulation with QITE on a noisy intermediate-scale quantum (NISQ) device lacking ancillae and complex circuits. Successful quantum simulation on a near-term quantum computer could pave the way for great advances across a broad range of technologies.

Modeling and Detection of Dendritic NMDA Spikes With Neuropixels Probes
Alex Bardon
Mentors: Michael Hausser, Arnd Roth, Brendan Bicknell, and Thanos G. Siapas

One of the fundamental goals in neuroscience research is to understand how neurons perform computations, which allow our brains to take in and process information and eventually lead to behavior. Although the traditional view states that dendrites simply pass signals on to the soma, a growing body of research points to just how important dendrites are for neural computations. The diversity of dendrites allows them to perform a wide variety of linear and nonlinear computations, which are an important component of the activity of networks of neurons. Many open questions remain about how these dendritic computations work, and how dendritic spikes relate to behavior.
In order to begin to answer these questions, it is necessary to have a way to easily record large amounts of dendritic spike data from neurons while an animal is performing a behavioral task. While recent research has shown that patch-clamp recording and extracellular tetrode recording can capture dendritic activity from single sites on a neuron, these techniques are not easily translatable to large scale recording, nor do they work well in animals completing behavioral tasks. Neuropixels probes combine high spatiotemporal resolution with large volume coverage to allow for the collection of large amounts of neuronal activity in unrestrained animals. These probes have been used in many studies to observe somatic action potentials; however, it has thus far not been possible to isolate dendritic spiking activity from Neuropixels recordings. We use a detailed biophysical model of a layer 2/3 pyramidal neuron and virtual Neuropixels probe to simulate recordings of dendritic spiking activity. With this simulated data, we train an algorithm to isolate dendritic NMDA spikes from Neuropixels recordings.

Machine Learning With Non-linear Fluid-Like Systems
Sarah L. Barrett
Mentors: Beverley McKeon and Simon Toedtli

Turbulence plays a role in everyday life, from an airplane's movement in flight to the lower performance of a ship experiencing biofouling. Our physical understanding of turbulent flows is limited. By gaining a better understanding of turbulent flows, better solutions will be found to reduce energy expenditure. The Navier-Stokes equations (NSE) describe fluid flows, yet they are difficult to solve computationally. The Kuramoto-Sivashinsky equation (KSE) describes phenomena similar to the NSE but is easier to solve, so the KSE was used as a model for turbulence and decomposed into a set of bases and coefficients using modal decomposition. Getting these coefficients requires a large dataset with modal decomposition, so machine learning was used to solve for these coefficients across different bases without needing these sets. Results of machine learning analysis will be discussed. In future works, the results of this investigation into the KSE will be applied to the NSE.

Inverse Design for Nano-Photonic Antennas
Hrishika Basava
Mentor: Ali Hajimiri

Nano-photonic antennas are a key component in many integrated photonic systems. For large-scale optical phased arrays as well as many emerging nano-photonic systems, antennas with larger effective aperture and a well-shaped far-field radiation pattern are needed. Such antennas can be created by expanding the design space beyond standard periodic patterns like the conventional grating coupler. We are investigating the possibility of designing high-performance nano-photonic antennas for integrated transmitter and receiver systems with a focus on multi-port and multi-mode antennas, utilizing the technique of inverse design. A search algorithm was designed that includes the fabrication design rules as well as the performance specifications in order to find the optimal design in the fabricatable nano-photonic design space. This algorithm defines etch boundaries using a Fourier series, and optimizes the coefficients in order to best meet the specified objective. In the future, we plan to continue to improve upon this search algorithm, as well as expand its scope so we can apply it to design a wide range of nanophotonic devices.

Predicting the Spread of COVID-19 Using Artificial Intelligence
Rahil Bathwal
Mentor: Yaser Abu-Mostafa

The COVID-19 pandemic has brought epidemiological models to the forefront in helping policymakers and public health officials make important decisions in curbing the spread of the virus. The objective of this study was to adopt data-driven models to inform the estimates of the pandemic's future trajectory in the US. A combination of data-augmented compartmental epidemic models (SIR models), neural network models, clustering techniques and model aggregation techniques were developed to predict the number of daily fatalities at a county-level. These data-driven models provided accuracy improvements over conventional epidemiological models. This project identifies machine learning as a promising tool in developing more sophisticated epidemiological models that can be informed by as well as simulate the real-time effects of policy decisions.

Stream Intersection in Tidal Disruption Events
Gauri Batra
Mentors: Sterl Phinney and Wenbin Lu

In a tidal disruption event (TDE), a star approaching a black hole is torn apart by the black hole's tidal force, resulting in the squeezing of the star to form a stream of tidally disrupted material. Some of this material is bound and keeps orbiting the black hole while some is unbound and escapes. To understand the outcome of TDEs, it is crucial to find where the bound stream intersects itself since intersection can lead to potentially observable shocks, accretion disks and secondary outflows. We compute the intersection points for a Schwarzschild (non-spinning) black hole and a Kerr (spinning) black hole, first for the equatorial case where the stream lies in a plane and then for the general case where the stream evolution is in three dimensions. We numerically integrate the geodesic equations of motion to find the path of the stream and discuss the algorithm used to find the region of self-intersection. This algorithm takes into account various aspects of the model including stream thickness and energy distribution of the stream material. As a result, we obtain the dependence of the self-intersection region on multiple parameters such as angular momentum, black hole spin, and stream thickness.

Feed-Forward Inhibition Can Strengthen Modeled Neural Responses to Narrow Odors in *Drosophila*
Matthew Bauer
Mentor: Betty Hong

The goal of this study is to expand upon efforts to implement an accurate, dynamic, and broadly applicable model of odor processing in the vinegar fly *Drosophila melanogaster*. Previous work suggests that the model is a poor predictor of neural responses to highly selective 'narrow' odors that activate only a small fraction of the olfactory input channels; in particular, the model predicts a much weaker mushroom body response to these odors than has been observed *in vivo*. In this paper we demonstrate that adding a feed-forward inhibition term to the model is sufficient to strengthen modeled mushroom body responses to these narrow odors. In addition, we profile properties of the model when configured with anatomically-accurate circuit connectivity data from the hemibrain connectome published by the Janelia FlyEM project.

Computational Design of Binding Pockets for Biosensor Proteins
Zoe Beatty
Mentors: Henry Lester and Anand Muthusamy

Biosensors are important tools for studying the presence and localization of target molecules in biological environments. In the Lester lab, our biosensors are fusion proteins made from a periplasmic binding protein (PBP) and a green fluorescent protein (GFP). When a target molecule binds to the binding pocket of the PBP, a conformational change occurs that results in increased GFP fluorescence. These biosensor proteins are engineered to bind molecules of interest through directed evolution, in which key amino acids are mutated based on the knowledge of solved crystal structures. While this process has yielded many effective biosensors, we can improve this experimental process further in tandem with computational approaches. Autodock is a program that simulates the binding of ligands to receptors. I use Autodock to generate ligand positions in the binding pockets of PBPs with prospective mutations. I also use Chimera to assess non-covalent interactions responsible for binding. This project will allow for a better understanding of the interactions responsible for strong binding in the PBP binding pocket, identification of additional key residues, and streamline the experimental process by identifying promising mutations.

Stellar Abundances of RR Lyrae Variables in the Outer Halo of the Milky Way Galaxy
Trinity Bento
Mentor: Judith Cohen

RR Lyrae Variables are easy to distinguish, allowing purity in data, and are commonly used as standard candles for their well defined period to luminosity relationship with only a small dependence on metallicity. Spectroscopic data of approximately 600 RR Lyrae in the outer halo of the Milky Way (40-90 kpc) has been collected. Using these spectra, their velocities were calculated. Since RR Lyrae are commonly found in globular clusters, the velocities could help find new globular clusters and stellar streams. After velocity corrections were performed on the spectra, they were compared with synthetic spectra produced by the program SYNTHE. Using synthetic spectra representing different metallicities and phases, the metallicities of the RR Lyrae were calculated by way of correlation analysis. Initial calculations suggest errors in the log of the metallicities are under 0.3 dex for approximately half of the RR Lyrae. These are some of the first measured metallicities for the most distant part of our galaxy. This information has many possible applications, including helping to constrain the ΛCDM cosmological model.

Understanding the Surface Reflectance of Pluto Using the 2S-ESS Model
Daniel W. Bi
Mentors: Yuk L. Yung, Vijay Natraj, and Zhao-Cheng Zeng

Understanding the surface reflectance of celestial bodies through hyper-spectral remote sensing measurements requires running RT (radiative transfer) simulations on a large scale, where two-stream approximations are often used to speed up calculations. The linearized 2S-ESS (two-stream exact single scattering) multiple-scatter model can be used with respect to any atmospheric or surface property for RT simulations. Serving as a data source, the *New Horizons* flyby of Pluto provides valuable hyper-spectral measurements and sheds light on Pluto's aerosol properties. With the aerosol properities and instrument geometry as input, the measured radiance as output, we determine the surface reflectance of Pluto using the 2S-ESS model. In the future, we can apply the same pipeline to infer the surface reflectance of other celestial bodies and gather valuable data (such as using Cassini observations of Titan). Additionally, the 2S-ESS code has much room for improvement from several seconds to sub-millisecond calculation of broad spectral regions (say 300 nm – 3 μm).

Automotive Autonomy: Sensing and Motion Planning
Chase Blagden
Mentors: Soon-Jo Chung and Anthony Fragoso

Autonomous technologies in automobiles are becoming more common and promise a reduction in accidents. Although there have been significant advances in self-driving car technology, significant obstacles remain. Sensors will fail in adverse weather conditions as measurements are degraded. Data capturing these conditions are scarce and thus existing sensor fusion techniques have not matured. To further sensor fusion models and progress for the CAST car, we present trajectory generation using polynomials and develop a collection pipeline for multimodal sensor data. Lidar, radar, and IMU sensors are configured and the incoming data processed in preparation for object detection. This work enables the construction of new multimodal datasets.

Design of Arbitration Circuitry for a Failsafe Mechanism in Unmanned Aerial Vehicles
Chase Blanchette
Mentors: Soon-Jo Chung and Matt Anderson

Caltech's Aerospace Robotics and Control Lab (ARCL) is developing next-generation flight controllers for unmanned aerial vehicles (UAVs), including the Autonomous Flying Ambulance (AFA). Experimental flight code is typically implemented as a flight mode so that, in the event of a software error, it can be quickly turned off and replaced with more reliable failsafe algorithms. In some instances, these failsafe algorithms may be affected by uncontained errors (such as memory leaks or infinite loops) in the experimental code, leading to full lockup of the flight controller and loss of the UAV as a result.

To address this problem, my project developed a mechanism that allows a commercial off-the-shelf (COTS) flight controller to ride along in standby, stepping in only if the experimental controller fails. Connecting one set of sensors and motors to two flight controllers requires switching outputs based on which is in control ("arbitration"). The most power- and space-efficient method of arbitration was found to be a combination of active components (transistors/multiplexers) and relays. This arbitration circuitry is a low-impact addition to UAVs; it allows for the testing of experimental flight controllers without requiring developers to rigorously account for all potential software or hardware failures.

Using Twitter to Predict Rates of Covid-19 Infections
Eve Blank
Mentor: K. Mani Chandy

Social media has been continually shown to be an accurate source of data for the prediction of epidemics. Long term analyses of Twitter data have shown correlations between the usage of health-related words and occurrences of illness. I aimed to identify a correlation between COVID-19-related tweets indicating infection and COVID-19 cases in order to predict case counts in real time. First, a 1% sample of tweets was extracted from a set of COVID-19-related tweet ids. Tweets were then filtered by words indicating personal connection in order to determine the frequency of tweets that indicated COVID-19 infections. Performing a cross-correlation between case counts and tweet frequencies yielded insignificant results. The lack of correlation was likely due to the large amount of political content contained in the dataset. More work needs to be done to be able to accurately filter out this content. Despite the insignificance of these results, code for real time analysis was developed using the IoTPy library. This program may be used in further research involving Twitter data streams.

Literature Curation of *Caenorhabditis elegans* Anatomy Functions
Fernando Bolio
Mentors: Paul Sternberg and Raymond Lee

Caenorhabditis elegans are a non-parasitic, transparent nematode who shares many essential biological functions and characteristics with humans. When placed in harsh environmental conditions they enter a period of suspended development known as dauer. How the animal decides to enter dauer has been extensively studied, with over 4200 articles appearing through a nematode-specific search. Due to this immense amount of information available, I decided to employ literature curation for input into Wormbase's Anatomical Ontology. Ontologies are used to take existing information and extract salient summaries from them making the information more accessible and digestible for readers. This paper describes the initial literature search and start of literature curation for anatomical function related to the dauer-decision process in *C. elegans*. Future research will continue this literature curation and submit information into Wormbase and potentially create a visual model to demonstrate found connections.

Development of Statistically Sound Workflows for Single-Cell RNA Sequencing
Mihir Borkar
Mentor: Lior Pachter

Single-cell RNA sequencing (scRNA-seq) allows for the transcriptomes of individual cells in a sample to be sequenced and analyzed. The transcripts from each cell in a sample are barcoded with a DNA sequence in order to identify which cell they come from. During the sequencing process, these cell barcodes can be accidentally modified. The Pachter Lab introduced a new file format for scRNA-seq data (BUS) and new software called bustools that makes possible the development of more efficient scRNA-seq workflows. We are examining methods for pre-processing scRNA-seq data, notably a new DNA sequence clustering algorithm called Starcode, which can be used for correcting cell barcode errors. Based on the results, we can decide whether to incorporate some of these methods into the lab's current software. We are also studying various downstream data analysis techniques, such as dimensionality reduction, to find the best workflows for downstream analysis. We found that the Starcode algorithm seems to include many more transcripts than the lab's current software. Next, we will study the effect of using Starcode on the downstream analysis of scRNA-seq data.

Deep Kernel Learning for Bayesian Optimization
James Bowden
Mentors: Yisong Yue and Jialin Song

Optimizing complex, expensive-to-evaluate functions with a limited budget is a big challenge in adaptive experiment design. Traditional Gaussian process (GP) models struggle to fit many such functions and are beset by difficult kernel design and poor scalability. We propose using deep kernel learning (DKL) with Bayesian Optimization (BO) for increased accuracy and scalability over current GP models. The DKL framework consists of training a deep neural network alongside a GP model to learn kernel parameters more effectively. The advantages of DKL arise from the flexibility of neural networks and their ability to extract underlying features from incredibly complex functions. Additionally, the DKL model can be made much more scalable than GP models by overlaying the base kernel with a grid interpolation kernel that makes quick but accurate approximations. Based on preliminary experiments, DKL outperforms standard GP models on complex objectives in terms of simple regret and train/test MAEs. We will apply DKL to real-world problems including protein engineering and nanophotonics filter design with input spaces of up to 2000 dimensions. Based on DKL's single-fidelity/task/objective success, we will also expand implementation to multi-fidelity/task/objective BO, where complex relationships between multiple functions pose a challenge to GP models.

Absolute Single Shot Picosecond-Precision Timing Photonic Integrated Circuit Enabling Ultrafast Clock Synchronization and TS/s Digitization
Cole Brabec
Mentor: Ali Hajimiri

A CMOS-compatible photonic integrated circuit capable of single-shot timing the arrival of optical signals with picosecond precision is demonstrated. The circuit achieves a minimum detectable power of -20dbm. The core of the circuit is a two-photon absorption photodiode implemented via a slow-light photonic crystal waveguide. Two applications are demonstrated. The first is a clock synchronization system. The second is a TS/s digitizer.

Leveraging Deep Learning to Strengthen the Security of Mobile Biometrics
Hernan Caceres
Mentors: Kiran Balagani and Adam Wierman

At the present moment, smartphones are the most popular and widespread personal devices, and offer the capability to perform multiple security sensitive activities, such as online banking and shopping, social networking and e-mailing. Simultaneously, they also pose security and privacy threats to the users' stored data, calling for a need for increased security beyond the current one-time authentication model standard for smartphones. Sensor data was previously collected from 100 smartphone users during 8 free text typing sessions, and accelerometer, gyroscope and magnetometer sensor readings, raw touch data from the touchscreen, key press, and key release latencies were all collected. The objective of this study was to see if features could be extracted from the swipe data of the users using a convolutional neural network to accurately discriminate between users. For future study, there is potential to use swipe data in conjunction with data from the other sensors to improve accuracy rates, as well as testing whether different portions of a swipe lead to higher accuracy rates compared to the entirety of a swipe.

Web-Based Spectral Visualization Tool for the Southern Stellar Stream Spectroscopic Survey (S5)
Joseph Cachaldora
Mentors: Ting Li and Brian M. Stoltz

S5 aims to map the kinematics and chemistry of stellar streams in the southern sky. Studying stellar streams provides useful information on the shape and radial profiles of their gravitational fields, as well as the history and distribution of matter of our galaxy. S5 has led to significant discovery, including fastest known star S5-HVS1 ejected from the galactic center. We develop a user friendly web-based spectral visualization tool for improved analysis of the large dataset to increase the likelihood of similar remarkable findings from S5. The spectral viewer allows the user to inspect spectra of specified characteristic from the S5 dataset through a query filtering form or inspect a set of spectra provided unique identification (e.g. sky coordinates or ID tag). The tool presents results in a customizable display containing data from the S5 catalog and a flux-wavelength spectral plot for each object for comparable spectral analysis. This web application is developed in Python 3.7.6 utilizing Python module Django 3.0.7 for web development and SQL database maintenance. The listed features afford a simplistic and organized spectral analysis experience for the S5 collaboration and eventually for the entire astronomy community with a potential S5 data release.

10 and 11 Dimensional Linearized Supergravity and the Holoraumy Tensor
Luis Camargo-Carlos
Mentors: Jim Gates and Sergei Gukov

In the search for a gravitational theory, one new development that may help is a tool called the holoraumy tensor. This can be used for imensional enhancement, where the goal is to build higher dimensional supersymmetric multiplets from lower dimensional supersym-metric multiplets. In using this tool, we want to focus on scalar gravitational theories, which, in analogy to those first thought of in the development of general relativity, might help to find the true nature of gravitational theories in 11 dimensions. This project seeks to use various supermultiplets, along with the various tools such as new Fierz identities in 11D, to study what information the holoraumy tensor gives when used in 11 dimensions. In doing so, we focus on scalar theories and examine possible analogies between the 11 dimensional and 4 dimensional cases.

Using Neural Networks to Investigate Advantages of the Discrete Consciousness Model
Isabella Camplisson
Mentors: Michael Herzog and Ralph Adolphs

The debate as to whether consciousness is a discrete or continuous process is ongoing, with the timescales involved often being too small for experimental results to definitively validate one model. Taking a different approach to this debate than conventional human experiments, we aimed to provide evidence showing the benefits of discrete consciousness to visual processing using neural networks, thus forming an argument for discrete consciousness surrounding its evolutionary advantages over the continuous model. Using the 20bn-jester dataset, a very large set of hand-gesture videos for which we determined the temporal aspect to be imperative for classification, we trained two neural networks: one representing discrete consciousness and one representing continuous consciousness. The continuous neural network contained a convolutional LSTM (Long Short-Term Memory) architecture, while the discrete neural network contained a convolutional LSTM with inputs from multiple LSTMs that analysed consecutive 450ms segments of the videos. Our results showed that the discrete model more accurately classified the videos than the continuous model, implying that discrete consciousness would allow humans to more accurately identify their surroundings and thus be evolutionarily advantageous.

Detection of BFB in Cancer Cells Using Human Genomic Data
Ann Caplin
Mentors: Vineet Bafna and David Van Valen

Breakage-fusion-bridges (BFBs) are a type of chromosomal rearrangement that are associated with certain cancers. For this reason, detection of BFB in cells could have applications in better understanding and targeting these cancers. Current widely-available sequencing data is not informative enough to reconstruct certain genomic sequences. Instead, the data is given in the form of “count vectors”, which count the number of occurrences of each genomic region and “foldback read vectors”, which provide information on adjacencies between a genomic region and its reverse. The Bafna lab has developed efficient algorithms that use count vectors to see whether or not a certain genomic region could have undergone BFB rearrangements. One goal of this project is to incorporate information from foldback read vectors as well as count vectors to get a more accurate prediction of BFB. The other goal is to combine this widely-available data with less available but more informative structural data to better predict BFB. So far, we have modified previous algorithms developed by the Bafna lab to incorporate both foldback read and count information into BFB prediction. Further work will involve testing these algorithms on real-world data to evaluate their performances and incorporating structural data to improve these performances.

Developing Feedback-Based Coding Schemes for Stochastic Bit Arrival Times
Ali Cataltepe
Mentors: Victoria Kostina and Nian Guo

Error-correcting codes (ECCs) that rely on noiseless feedback have been studied extensively for decades. However, the most theoretically rigorous methods in this area do not apply to a causal setting, where the encoder does not know the entirety of the message they wish to transmit beforehand. Of particular interest among feedback-based ECCs are those operating on the Small-Enough-Difference (SED) principle, where the entire vocabulary of possible messages is deterministically divided into two groups with probability mass difference less than that of a single message. The Information Theory Group at Caltech has developed a causal ECC which assumes predetermined bit arrival times. In this paper we propose two implementations of a causal ECC for a setting with uniformly random bit arrival times, performing a natural grouping of low-probability messages using the tree structure of binary strings to achieve linear-time performance. Even when it does not attempt to meet the SED criterion, this algorithm performs at 95% of channel capacity with fixed message lengths, and stabilizes the leading 90% of bits of streamed data from a random plant at bit arrival probabilities up to channel capacity.

Utilizing SPINS Inverse Design to Create On-Chip Accelerator Structures for MeV Dielectric Laser Accelerators
Dominic Catanzaro
Mentors: Jelena Vuckovic, Jean-Phillipe MacLean, and Axel Scherer

MeV particle accelerators are used in a wide variety of applications, from medical devices to manufacturing to basic research. Leveraging developments in silicon nanofabrication techniques, the creation of on-chip dielectric laser accelerators would result in an order of magnitude improvement in the size and cost of MeV accelerators. This project utilizes adjoint sensitivity gradient descent optimization through the SPINS inverse design package to design accelerator structures pumped by femto-second laser pulses. Because inverse design can create novel, non-intuitive structures, accelerator designs can be more effective than conventional structures. Designs are externally verified and analyzed through Lumerical simulations. The design of accelerator structures is advanced primarily through broadband functionality and investigation of the efficacy of various dielectrics, but also through the fine-tuning of fabrication constraints and structure sizing. In addition to providing optimal designs for experimental verification, this project seeks to provide comparisons between the efficacy of various dielectrics for accelerator structures designed for various input electron velocities.

A Novel Application of Polygenic Risk Scores to Transcription Phenotypes
Isha Chakraborty
Mentors: Lior Pachter and Ingileif Hallgrímsdóttir

In Genome Wide Association Studies (GWAS), researchers query millions of genetic variants across the whole genome in large samples of individuals to identify genetic variants that associate with certain traits or diseases. GWAS have been used to study hundreds of different traits or diseases and currently around 190,000 statistically significant associations between a trait and genetic variant have been reported. In many GWAS, tens or even hundreds of loci are reported from a single study. In order to combine the information across genetic loci it is commonplace to calculate a polygenic risk score (PRS) as a linear combination of the estimated risk associated with each variant. This allows researchers to calculate a PRS for an individual which depends on their genotype and the (risk) coefficients for each variant.

Expression quantitative trait loci (eQTL) analysis aims to identify genetic variants that affect the expression of one or more genes by running an association analysis where the trait is gene expression. cis-eQTL analysis is limited to looking for associations to loci in a region around the gene, whose expression we are considering. Some analyses are limited to identifying cis-eQTLs but we can also search for associations across the genome to identify trans-eQTLs, expression loci that act on distant genes. Trans-eQTLs tend to have much weaker effects than cis-eQTLs. We will study whether combining the signal across trans-eQTLs by defining a polygenic eQTL score akin to a PRS will provide new insights into the role of trans-eQTLs.

We build a pipeline to determine the use and interpretability of polygenic risk scores in this novel setting. The Pachter Lab has developed kallisto, a tool to quantify the abundances of transcripts from bulk and single-cell RNA-Seq data using high-throughput sequencing reads,we will apply kallisto to publicly available datasets, then use Plink to perform genomewide cis- and trans- eQTL analysis and a custom script to calculate polygenic scores. We will also perform a simulation study to study the utility of the polygenic trans-eQTL scores under different conditions.

Molecular Dynamics Simulations of Nanocrystalline Additively Manufacture Zinc Oxide
Antoine Chamoun-Farah
Mentors: Julia Greer, Zachary Aitken, Rebecca Gallivan, and Kai Narita

Zinc oxide (ZnO) is a metal-oxide ceramic that is important for photonic and electronic applications, notably as a key component of ceramic varistors. In addition to its low cost and high abundance, ZnO exhibits piezoelectricity, a coupling between stress and generated electric field, which makes the material highly desirable for devices and sensors including energy harvesting devices such as nanogenerators. Internal interfaces, such as grain boundaries (GB), play an important role in determining the electrical and mechanical properties of such materials. The complex microstructure that results from non-equilibrium additive manufacturing processes is often nanocrystalline, where internal interfaces are the dominant microstructural feature. Using Molecular Dynamic (MD) simulations, this project studies the role of an isolated (GB) in the piezoelectric response of ZnO. We conduct a series of compression, tension, and shear simulations to elucidate the effect of an isolated GB on the mechanical response and dominant atomistic deformation mechanisms. We relate and discuss the observed mechanical behavior to the atomic structure of the GB. These simulations can be used to inform ongoing experiments on nanocrystalline ZnO by showing the specific relationship between grain boundaries and both the electrical and mechanical behavior of nanocrystalline Zinc Oxide.

Stratification for Sets of Half-Average Nulls Generate Risk-Limiting Audits
Sophie Chan
Mentors: Philip B. Stark and R. Michael Alvarez

Risk-Limiting Audits (RLAs) limit the probability of confirming an election result if the declared winner differs from the result of a full hand count. Stratified audits allow ballots to be sampled independently from different collections of ballots, e.g., ballots cast in different counties, ballot cast by mail versus in person, or ballots tabulated using different equipment. Stratification is particularly useful to improve the efficiency of audits when some ballots are tabulated using equipment that tracks its interpretation of each ballot using cast vote records (CVRs) and some ballots are tabulated with equipment that does not produce a CVR that can be linked to the corresponding physical ballot.

Sets of Half-Average Nulls Generate Risk-Limiting Audits (SHANGRLA) assigns non-negative values to ballots based on the votes marked (and the CVRs, in some cases). It tests null hypotheses in the form "the average value of all ballots cast is less than or equal to 1/2." SHANGRLA can be used with stratified samples, but that has not yet been implemented in software nor tested in practice. This research project builds software suitable for pilot audits by election officials.

The underlying strategy is to examine all possible allocations of outcome-altering error across strata, rejecting the null only when the maximum *P*-value across all allocations is sufficiently small. Evidence from different strata is combined using Fisher's combining function for independent *P*-values. This approach was used in SUITE (Stratified Union-Intersection Tests of Elections). The resulting method is expected to yield more powerful audits (i.e., audits that require smaller sample sizes when the outcome is correct) than previous methods for stratified risk-limiting audits, including SUITE, because SHANGRLA tests sharper conditions.

Modelling Particle Movement Through Meshes With Implications for Cloth Masks
Anjini Chandra
Mentors: Lydia Bourouiba and Mikhail G. Shapiro

Personal protection equipment, including face masks and face shields, is essential to protecting people from COVID-19 and other respiratory diseases. Unfortunately, the repeated use of disposable face masks has a negative impact on the environment. In addition, the current demand for surgical face masks far outweighs the industrial production capacity. One alternative to surgical masks is cloth masks that are reusable and readily available to the public. In this project, we investigated the efficiency of cloth masks by studying the dynamics of particle meshing interactions in the masks. We implemented a theoretical model that simulated submicron particles moving in air and through a cloth mask at low Reynolds numbers. The model used the steady Kuwabara-Happel stream function to calculate the position of the particles in the air flow and accounted for effects of Brownian motion and electrostatic fields. For a single layer of a grid-like fabric, our model calculated the particle capture efficiency as 85%, and for staggered layers, the capture efficiency increased linearly with vertical staggering distance. In the future, we plan to validate our model by comparing its results to experimental measurements and creating an analogous computational fluid dynamics model to account for unsteady flows in three dimensions.

Impact of Forced Exploration on Habit Persistence
Katherine Chang
Mentors: Colin F. Camerer and Aniek Fransen

Human behavior is regularly influenced by habit tendencies causing a reluctance towards experimentation. When these habitual choices temporarily become unavailable, such as in a pandemic, humans are forced to explore new, alternative options. The effects of a period of forced exploration were examined with a free-operant coin paradigm run on Prolific. Participants were initially trained to associate two coin rewards with distinct visual stimuli and actions to develop stimulus-response and response-outcome associations. One coin was devalued and removed from selection, then a third coin with a distinct visual stimulus and action was introduced and trained. Finally, the removed coin was reintroduced and revalued, and the subject was given a choice period with all three coins. The coin paradigm was implemented in three experiments with varying training times and reward values for the removed and new coins. Across all experiments, differences in coin preference between habitual and goal-directed participants were generally insignificant. However, the distribution of coin preferences for habitual participants showed significantly less variability than goal-directed participants. Results suggest that the degree of habituation influences the predictability of human behavior after a period of forced exploration. In addition, bursty and metronomic response behaviors were observed.

Using Machine Learning to Predict the Spread of COVID-19
Nicholas Chang
Mentor: Yaser Abu-Mostafa

As COVID-19 is still prevalent throughout the U.S., governmental and healthcare institutions need to be able to predict how COVID-19 will spread to better inform the policies and practices put in place to minimize the damage. Machine learning models have shown to be effective in predicting the spread of COVID-19. However, there are still many difficulties in capturing the changes in this ever-evolving situation, as policies and public sentiments shift, and creating a model that makes meaningful predictions. This research aims to provide a model that can compete with the top COVID-19 prediction models in the country.

Digit Analysis: An *R* Package for Detecting Data Manipulation
Wenjun (Happy) Chang
Mentors: Jetson Leder-Luis and Jean Ensminger

When individuals steal money, they have incentives to cover their tracks to avoid detection. This inhibits traditional forms of monitoring, such as auditing, but provides another opportunity for detection: humans are bad at fabricating data. Ensminger and Leder-Luis (2020) present a set of tests based on digit analysis, which checks whether self-reported numbers appear to be fabricated by humans, indicative of fraud. These tests are applied to data from a World Bank development project, and the results are validated by a forensic audit conducted by the Bank. In this project, we develop a statistical software package in R to conduct similar statistical tests and facilitate their use across different data sets. The package allows users to customize their statistical testing with flexible input options. The digit tests include comparison to Benford's Law, binomial tests of high vs low digits, and a test for padding with Monte Carlo simulations. The package also produces interpretable 2D and 3D bar plots for visualization. The package is available on R CRAN and at https://github.com/jlederluis/digitanalysis. We hope that this open source package will contribute to monitoring corruption and data fabrication more efficiently in various fields of study.

Uncertainty Quantification Using a Game Theoretic Approach
Diego Chavez
Mentors: Houman Owhadi and Peyman Tavallali

Uncertainty quantification is an important concept when it comes to models and simulations for systems. The main focus of the project was using a game theoretic approach to uncertainty quantification to estimate the parameters of physical models. The method uses the idea that the matching of a model to data can be modeled by an adversarial game between nature and the researcher, and by doing so, the problem is simplified down to the use of a loss function to quantify how well the researcher is doing with estimating the behavior of nature. In researching this topic, we started with exploring using the decision theoretic approach on a fixed set of possible model parameters, applied this to use Bayesian optimization to be able to find probability distributions for a region of possible parameters, and set up sequential Bayesian optimization to use the iterations to perform optimizations on adjusted regions of interest. We showed how the algorithm works with one-dimensional parameters and then extended it to multi-dimensional parameters. This method can be compared to other methods of uncertainty quantification, such as the frequentist or Bayesian approaches that have been developed in the past.

Modeling the Dynamics and Control of Airborne Virus in Hospital Rooms

Peter Chea
Mentor: Richard Flagan

Considering the COVID-19 pandemic and its high probability of airborne transmission, understanding the dynamics of airborne virus in hospital rooms is vital. Through literature searching, it was found that the half-life of the virus was 1.1 hours and the peak concentration of virus aerosols was between 0.25 to 1 μm in diameter and larger than 2.5 μm. Through literature searching of indoor air quality models concerning bioaerosols, a well-mixed mass balance model was observed to be satisfactory. The model is a differential equation that accounts for the volume of the room, ventilation, emission rate, decay rate, and UV inactivation rate. UV-C light is used as a control measure for the virus since it has been previously implemented and would be the least harmful to humans. Modeling of the room was done in two ways: a single well-mixed room and as two compartments with a constant flow rate between them. With this model, graphs were produced of the concentration through solving the differential equations using Python. Realizing the over-estimation of the efficiency of UV-C light, more work is being done to better understand the effects of UV light in terms of a probabilistic standpoint.

Redeveloping the DARPA Quadcopter

Kristine Chelakkat
Mentors: Soon-Jo Chung and Matt Anderson

The Aerospace Robotics and Control Lab at Caltech is working on the development of various robotic assistants, including the Autonomous Flying Ambulance (AFA) which seeks to develop an autonomous helicopter-like vehicle capable of collecting scientific samples, transporting human passengers, or delivering packages. This project has focused on the mechanical and electrical design of a quadcopter interfacing with various payloads, making headway on systems similar to the Autonomous Flying Ambulance. Key design decisions included the geometry of the propeller arms, the frame of the body, and the body itself; the former K-shape frame was changed into a more stable H-frame configuration with a rectangular body to properly support the sensors, flight electronics, and payloads. The thin-walled circular carbon fiber tubes were kept for the arms as these were the most efficient in terms of strength-to-weight ratio, providing ample bending and torsional strength. After completing the preliminary mechanical design, work was completed on the electrical systems, including a board interfacing between a Teensy 4.1, simplertk2blite, and Pixhawk flight controller to communicate GPS data between the flight controller and the computing hardware. Ongoing work involves development of the electrical flight systems and the interfaces for payloads on the quadcopter.

Deep Learning for Image Artifact Correction in Non-Cartesian Imaging

Cynthia Chen
Mentors: Shreyas Vasanawala, Adam Bush, Frank Ong, and Adam Blank

Magnetic resonance imaging is a non-invasive imaging technique commonly used for disease diagnosis. Clinically, data is sampled along a Cartesian trajectory, but Cartesian imaging is corrupted by motion, resulting in reduced image diagnostic quality. In comparison, non-Cartesian imaging is more robust to motion. However, eddy currents due to changing magnetic fields cause deviations in the k-space sampling trajectory, leading to image artifacts. This work aims to correct for image artifacts due to eddy currents in non-Cartesian rosette imaging using a convolutional neural network (CNN). Our neural network architecture enables residual learning in the k-space domain and loss calculation on the image domain. To generate training data, we simulate the effect of eddy currents on rosette trajectories and resample Cartesian images using both nominal and corrupted rosette trajectories. The neural network is tested on resampled Cartesian data and data acquired using rosette trajectories. Our model successfully corrects for eddy current artifacts in resampled Cartesian data, obtaining predictions with an improved average structural similarity index (SSIM).

Integrating Single Cell Transcriptional and Proteomic Data With Deep Learning: An Application to Cancer Immunotherapy

Hannah X. Chen
Mentors: Vanessa Jonsson, Jonathan Hibbard, and Rachel Ng

Chimeric antigen receptor (CAR) T-cells are a promising immunotherapy for treatment-resistant cancers, though the marked spatial and genomic heterogeneity of solid tumors in particular pose many challenges in understanding how the composition, functional and phenotypic status of both cancer and immune cells change in response to therapy. We aim to build a deep learning model to perform an integrative analysis of that heterogeneous, multi-omic data, namely one that is better able to leverage spatial information than currently existing methods, with the long-term goal of identifying diverse CAR targetable tumor antigens and thereby enabling more data-driven, personalized treatment strategies. At present, the model is being developed using publicly available spatial transcriptomics (ST) and single-cell RNA-Sequencing data from a series of tumor and matched normal tissue samples. Preliminary analyses performed on the ST data involve manifold embedding and clustering based on transcriptional similarity, followed by visualization in spatial coordinates over the whole-slide images. Additional framework elements being examined for the model include pre-trained convolutional neural networks to extract latent features from the imaging data and coupled variational autoencoders to learn a lower-dimensional representation that captures information derived from both datasets.

Continuous Time Opinion Formation on a Graph
Haoxuan Chen
Mentors: Andrew M. Stuart, Bamdad Hosseini, and Franca Hoffmann

Due to the growing number of social media users these platforms have played an essential role in spreading information and impacting the formation of opinions in societies. In this project, we study a mathematical model for the formation of people's opinions both with and without the influence of media accounts. We introduce a dynamic network model for opinion formation based on graph Laplacians that encode the network geometry. We analyze this model theoretically and numerically using MATLAB. Our results reveal interesting properties of the model such as uniqueness and existence of solutions and characterization of steady states and formation of communities with similar opinions in the steady state.

Investigating Mutational Breakage in Spatial Patterning Circuits
Victoria Chen
Mentors: Richard Murray and John Marken

Synthetic biology relies heavily on the usage of synthetic circuits in order to understand biological design principles and to implement and test new cellular functions. However, such synthetic constructs impose an unnatural load on the host, leading to increased cellular stress and decreased overall function. Thus, cells that randomly mutate to inactivate the synthetic circuit propagate more successfully than engineered cells. This eventually leads to complete loss of circuit function in the population in any continuous-growth environment. Although circuit stability is very important for pattern-forming cellular systems, especially for applications that rely on the integrity of the pattern, it is less well-understood the extent to which mutants would overtake the population in a non-continuous-growth environment, such as in a spatial patterning circuit. We predict that it will be more difficult for a mutant to overtake the population when spatial dynamics are present. This is due to additional factors exclusive to spatially-explicit 2D environments, such as the ability of other cells to physically block mutants from spreading. In addition, nutrient diffusion is much more constrained in a 2D field such that a higher-fitness mutant emerging in a low-nutrient zone may have a lower overall growth rate than a non-mutant in a nutrient-rich zone. Thus, there are likely different developments when circuit breakage occurs in spatial systems versus non-spatial systems. The goal of our project is to investigate and document the effects of mutational breakage in spatial patterning circuits. We use an agent-based model to simulate a bulls-eye patterning circuit and implement various circuit breakage scenarios. We hope to use our findings to develop new strategies for ensuring mutational robustness for spatial circuits.

On-Chip Training for AI Diagnosis of Cardiac Arrhythmias on an Analog Interface
Kathleen Chiu
Mentors: Azita Emami and Lin Ma

Cardiovascular diseases can be fatal, resulting in the deaths of millions annually. Many of these deaths can be prevented by earlier and more accurate diagnosis. Cardiologists often confuse similar arrhythmias in ECGs, creating up to a 25% misdiagnosis rate. AI implementation of cardiac arrhythmia detection and classification is more accurate and allows for portable biomonitoring devices. These devices would be worn like a watch and detect cardiac arrhythmias, anytime and anywhere, without affecting the tester's regular life. This is both convenient and has potential to detect diseases early because the devices do not require the presence of a doctor and provide continual monitoring. However, portable devices have size, weight, and power constraints. Digital circuits and implementations of AI algorithms are computationally expensive and require too much power. Instead, our group is integrating AI into an analog interface for the first time, which requires less power. One phase of implementing AI is training to obtain weights for the classification algorithm, which can occur on-chip or off-chip. Differences in component values from their ideals creates an average misclassification rate of 43%[1] in off-chip training, rendering these algorithms practically useless. On-chip training, however, creates chip-specific weights and an average 8.4%[1] misclassification rate.

1. Gonugondla, S. K.; Kang, M.; Shanbhag, N. A 42pJ/Decision 3.12TOPS/W Robust in- Memory Machine Learning Classifier with on-Chip Training. In 2018 IEEE International Solid - State Circuits Conference - (ISSCC); IEEE: San Francisco, CA, 2018; pp 490–492. https://doi.org/10.1109/ISSCC.2018.8310398.

Implementation of RRT Motion Planning Algorithms for Stochastic Model Predictive Control in Spacecraft Simulators
Ellie Cho
Mentors: Soon-Jo Chung and Yashwanth Nakka

Motion planning in spacecrafts requires the spacecraft to navigate from an initial position to a final position by navigating towards the goal in a dynamic, high-dimensional environment. The RRT approach to motion planning involves searching through space by randomly building a space-filling tree and outputting a path. There exists sampling-based RRT implementations (such as RRT and RRT*), and dynamics-based implementations (such as LQR-RRT* and Kinodynamic RRT*). Planning techniques used for computing a feasible trajectory with dynamics constraints at high frequency generally require an initial collision free path. The output of RRT is used as an input into the stochastic model predictive control to produce a smooth trajectory and nominal control. My project is focused on implementing the RRT algorithms in Python.

Viscoelastic Characterization of Jellyfish Mesoglea Using Finite Element Modeling (FEM)
Chloe Choi
Mentors: Julia R. Greer and Jane Zhang

Mesoglea is a highly hydrated and collagen-fibrous extracellular matrix (ECM) hydrogel found between the epidermal and endodermal layers of a jellyfish. Since jellyfish comprises mainly of its mesoglea, it has been hypothesized that the mechanical properties of mesoglea are the origin of jellyfish's shape-retaining behavior (radial symmetrization). In this work, we aim to model the *Aurelia aurita* jellyfish response under flat punch indentation and to understand the symmetrization process using finite element analysis (FEA) simulations. Mesoglea was modeled as a homogeneous gel matrix embedded with evenly spaced collagen fiber elements. Experimental results and existing literature were used to determine material properties, expressed as Maxwell-Weichert parameters. Static tests using a tip size of 120 μm and different tip depths of 100 μm, 200 μm, and 300 μm were used to verify the model. Dynamic tests using varied frequencies of 1-10 Hz and dynamic displacement of 1 μm will be used to test the microstructure of the simulated mesoglea and mimic radial symmetrization.

Data Structure Visualizer for Introductory Computer Science Courses
Devin Chotzen-Hartzell
Mentor: Adam Blank

When students are tasked with implementing complex data structures in a course similar to Caltech's CS 2, they often struggle to debug their code, as traditional debugging methods do a poor job of displaying program state for multi-layered, abstract data structures. We assert that including a graphical visualization of the program's state in the Integrated Development Environment (IDE), the program students use to write code, would allow students to more efficiently debug their code, while honing their own developing skills and conceptual understanding. We have created such a visualizer for IntelliJ, a popular IDE, along with a framework for customizing logical visualizations of advanced data structures to help students find granular bugs in their code and solidify their understanding of algorithms associated with data structures. The tool will be used by students in Caltech's CS 2 course starting in the Winter 2021 term, and can be used in similar introductory-level computer science courses covering data structures.

Solving Moving Boundary Problems by Incorporating Explicit and Implicit Approximation Schemes
Jennie Chung
Mentors: Oscar Bruno and Daniel Leibovici

Moving boundary problems have been of great interest in a wide variety of contexts, such as simulating viscous flows around aerospace configurations, biomedical applications, and describing trajectories of projectile objects. Solving PDEs over domains with moving boundaries thus has many desirable consequences, as phenomena in numerous fields can be modeled using such explorations. This project focuses on approximating the solution of the heat equation on domains containing moving boundaries in MATLAB while maintaining stability and accuracy over time. Components of both explicit and implicit methods were used depending on whether new points were added to the domain over the last time step. For penultimate points to the moving boundary, explicit methods employing boundary conditions were used, while interpolation was used for newly added points. For the 1D heat equation, a moving boundary on the right was introduced to analyze stability and accuracy. Analogously, the moving boundary problem in 2D was approximating the heat equation on a rectangular domain excluding a moving ellipse, i.e. a punctured domain. Special considerations were made in planning to approximate over a mesh grid discretization of points in 2D, for which the moving ellipse introduced more complicated evolution schemes and stability considerations compared to the un-punctured domain.

Investigating the Application of Particle Image Velocimetry to Multiphase Flows
Norman Chung
Mentor: Melany Hunt

Particle image velocimetry (PIV) is a process by which one can calculate instantaneous velocity fields for a fluid flow. Traditionally, tracer particles, or particles which are assumed to not affect the fluid flow because of their size, are seeded into a fluid flow, which is then recorded and analyzed using PIV software. Because one can assume the velocities of these tracer particles accurately represent the fluid velocities at a given point, calculating an instantaneous velocity field for the tracer particles is akin to calculating the instantaneous velocity field for the fluid itself. My work focused on whether one could use PIV to calculate the velocities of larger particles of interest in multiphase flow, a phenomenon with many applications to industry and a phenomenon not typically analyzed using PIV. To answer this question, I developed a video pre-processing procedure using MATLAB to create images compatible with PIV software and analyzed the data outputs from said software. Post-processing of the data suggests that while there are imperfections in applying PIV to non-tracer particle flows, PIV is not entirely unsuitable for analyzing multiphase flows.

Achieving Near-Unity Absorbance in TMDC Monolayers at Room Temperature to Enable Monolayer Excitonic Solar Cells
Tyler Colenbrander
Mentors: Harry Atwater and Joeson Wong

Semiconducting transition metal dichalcogenide (TMDC) monolayers like tungsten disulfide (WS_2) are a promising group of atomically thin materials for use in solar cells because of their intrinsically passivated surfaces, high radiative efficiencies, and large absorption coefficients. However, their monolayer nature (~7 Å thick) limits their free-standing absorbance to approximately 10% or 20%. Efforts to design perfect absorbers have required cryogenic temperatures (~100 K) or used multilayer TMDCs which have an indirect bandgap and therefore substantially lower radiative efficiency than direct bandgap TMDC monolayers. Here, we show that with appropriate photonic design, we can theoretically achieve >85% and >99% monolayer excitonic absorbance at room temperature using a metallic or dielectric cavity, respectively. We show that this high absorbance stems from matching radiative and non-radiative decay channels. Hence, the Purcell factor which enhances the radiative decay rate of the excitonic dipole is an important figure of merit for the optical design of these systems. Our results demonstrating near-unity absorbance therefore enable monolayer excitonic solar cells with efficiencies approaching their detailed balance limit.

Searching for Pulsations in Ultraluminous X-Ray Source X-1 in Galaxy IC 342
Danny Collinson
Mentors: Fiona Harrison and Sean Pike

Ultraluminous x-ray sources (ULXs) are extragalactic x-ray sources with a luminosity of at least 10^{39} erg/s. We present a study of the ULX X-1 in the galaxy IC 342 using NuSTAR x-ray telescope observations from August 2012. It was previously believed that most ULXs are black hole binary systems, but the detection of pulsating x-ray flux from a ULX in 2014 indicated that the ULX was actually a neutron star binary, in which a neutron star accretes matter from a companion star and the neutron’s star rotating magnetic field produces pulsations in the observed brightness of the source. We perform a timing analysis of X-1 in order to search for such pulsations, using methods developed recently to detect other ULX pulsars. Identification of X-1 as a neutron star would add to the number of known neutron star ULXs. These objects are not well understood given that their observed luminosity exceeds the Eddington limit—determined by the balance of the radiation pressure of photons being emitted and the force of gravity inwards—and so their detection is significant to help better understand super-Eddington accretion in neutron stars.

An Investigation on Machine Teaching Schemes for Continued Fraction Regression
Jonathon Corrales de Oliveira
Mentor: Pablo Moscato

The goal of symbolic regression is to find the best model that fits a certain dataset, and approaches have largely depended on genetic programming and present polynomial models. The application of a Memetic Algorithm, taking advantage of heuristic and recombination methods to improve population diversity, to symbolic regression which presents models with recursive representations such as continued fractions or continued logarithms seem to generate more accurate models that are simpler to understand. Experimentation with a wide range of datasets through the generation of many models for each dataset using different configurations for each has indicated that the addition of other recursive representations better suited some datasets and the diversity arising from partial seeding of a population when initializing with a previous solution (as opposed to total seeding) resulted in generally better models. The implementation of various user configurations should allow this method to generate models for a wide variety of use-cases.

Modeling Planet-Planet Interactions in Planetary Astronomy
Sofia Covarrubias
Mentors: Dimitri Mawet, Sarah Blunt, and Jason Wang

Orbitize! is a Python package that streamlines the process of modeling the orbits of directly-imaged planets. The *orbitize!* development team aims to make planet-fitting faster and more accessible by combining multiple existing techniques into one code base. The VLTI-GRAVITY instrument has demonstrated that interferometry can achieve 10x better astrometric precision than existing methods, a level of precision that makes detection of planet-planet interactions possible. I have researched and developed my own N-body simulator in order to make the best implementations and updates to the package that I can. I have replaced the standard Keplerian orbit solver in orbitize! with a module using REBOUND, an N-body solver, in order to account for planet-planet interactions present in the data. Users have the option to use this module when modeling highly precise astrometric data. I will present this updated implementation in *orbitize!*, discuss validation tests I performed, and discuss future use cases.

A Search for Outbursting AM CVn Systems in Zwicky Transient Facility Data
Leah Creter
Mentors: James Fuller and Jan van Roestel

The Zwicky Transient Facility is located at the Palomar Observatory in California. ZTF is an astronomical survey of the northern sky that collects potential outbursts of transient objects with sensitive cameras, allowing us to observe stars with a 20.5 magnitude or brighter. The outbursting systems detected by ZTF have given us the opportunity to search for AM CVn binary systems. AM CVn binary systems are interesting to our research since these systems are known sources of gravitational waves, have short orbit periods of 5 minutes – 60 minutes, and are about 1 out of a million. These systems are in the stages of final binary evolution and both the host and donor star are helium dominated, while lacking all hydrogen from the spectra. Since these systems are helium based, we can further observe the spectra of these systems by using the Palomar Observatory telescope and the Keck Observatory telescopes. These observations will provide us with the tools to identify new AM CVn systems.

Applications of Detailed Spectroscopic Studies to the Quantification of H_2O in Rock
Molly Crotteau
Mentors: George Rossman and Rebecca Greenberger

The improvement of the understanding of the spectral proxies for hydration is becoming increasingly important when considering the future of planetary science and geoscience. Current and upcoming missions to Mars, Europa, and other planetary bodies will utilize spectroscopy to detect H_2O and OH based on spectroscopic absorptions. The precise interpretation of these spectral features will be paramount when studying the aqueous history and astrobiologic potential of Mars, water-rock interactions and aquatic chemistry on ocean worlds, and the hydrothermal alteration, hydration, and formation of the oceanic crust on Earth. While hydration has been well constrained via spectroscopy for single minerals, the relationship between spectral features and the water content in bulk rock has yet to be elucidated. This project combines data from spectral databases and microimaging spectroscopy measurements of basaltic/gabbroic core sections from the ICDP Oman Drilling Project to study the relationship between the H_2O/OH absorptions and geochemical measurements, which provide precise H_2O content. We utilize spectral data, mineralogy, and geochemistry to determine a relationship between spectral indices and hydration. Ultimately, this work will advance our understanding of ocean crust formation on Earth, while also having numerous applications to the interpretation of mafic lithologies on other planets.

Unlikely Newton Polygons Arising From Abelian Covers of the Projective Line
Miles Cua
Mentors: Elena Mantovan and Trung Cẫn

The characterization of the open Torelli locus has long been a subject of great interest. To explore this problem, we first study the Jacobians of abelian covers of $\mathbb{P}^1$ branched at three points; in particular, we prove that they have complex multiplication. Then, we apply the Shimura-Taniyama formula to compute Newton polygons of such covers. We also explore special cases arising from the m-cyclic case, such as when m is a Fermat prime. Furthermore, we utilize these Newton polygons to find unlikely intersections of Newton polygon strata and the open Torelli locus inside Siegel varieties of large genus.

Simulations of Low-Swirl Burner for H_2 Combustion
Robert Daigle
Mentors: Guillaume Blanquart and Joseph Ruan

In the last twenty years, Low Swirl Burners (LSB) have been considered as an alternative to the traditional High Swirl Burners (HSB) for power production from the combustion of hydrogen due to the LSB's ultra-low emissions and fuel efficiency. However, experimental costs of Low Swirl Injector (LSI) design and access to experimental diagnostics have prohibited systematic design. In the current work, a new process of simulating

LSI inside ANSYS was developed in order to characterize the injector's swirl, pressure gradient, and turbulence. This process centered on decomposing a complex LSI into separate parts: the outer pipe, the inner pipe, the perforated screen, and the vanes. These parts were modeled through a systematic progression of complexity such that each successive model included the geometry from the previous model. A variety of turbulence models k-epsilon, k-omega, Spalart-Allamaras) were used to test the robustness of the ANSYS simulations. The simulation results were validated against previous experimental work [Cheng 2009]. The overall process quantified each design addition's effect on the injector's overall pressure drop and turbulence intensities. The present work will help future research of LSIs by proving a standard methodology of simulating the flow field at a low computational cost.

A Geotechnical Approach to Early Landslide Detection
Aneel Damaraju
Mentors: Katherine Bouman and Aviad Levis

Landslides cause billions of dollars in damages per year in the US alone. Landslides often start as slow-moving "creep" until the mass reaches a critical failure point, in a process that can take several years. Determining regions of land that are potential landslide areas long before the landslide has occurred is necessary to mitigate infrastructural landslide damages. In this project, we create an inversion approach for the characterization of landslide regions from the surface displacement of the landslide slope. The geologic landslide regions are parameterized utilizing the Burgers creep model, consisting of multiple plastic, elastic, and viscosity parameters. These parameters dictate the rate of surface displacement which can be detected using aerial imagery. In this project we aim to estimate these subsurface material parameters using surface displacement measurements. To approximate these material and geometric properties of the landslide region, we use a three-step approach for determining these material parameters. We show results on linear test cases that are used to explore and validate the inversion algorithm and introduce a pipeline for applying this method to artificial landslide data using the FLAC finite difference simulation software. This approach is not limited to landslides and is useful for other inverse problems where the process is computationally complex, and traditional gradient-based methods are not applicable.

Are All Fast Radio Bursts the Same?
Lucca S. de Mello
Mentors: Charles L. Steinhardt and Melany L. Hunt

In 2007, highly-energetic, millisecond-long bursts in the radio spectrum were discovered. These Fast Radio Bursts (FRBs) are poorly understood: there is no convincing explanation for their origin yet, and attempts to classify them have not gone well. Using a novel machine learning technique, we have determined whether there are multiple classes of FRB. The technique consists of modeling the entire intensity-versus-time curve of these bursts as vectors in a high-dimensional space and clustering these vectors in two dimensions. To accomplish this, we have also compiled a standardized catalog of raw FRB data, as no such catalog currently exists. We hope that, in grouping FRBs by their physical properties, we have also grouped them by physical origin.

Towards a Physical Understanding of Water Molecular Dynamics Through Simulation
Lily DeBell
Mentors: Geoffrey Blake and Haw-Wei Lin

Water is a necessary molecule for life on Earth. Despite its essential role, the underlying physics and chemistry that govern intermolecular interactions between water molecules are still poorly understood. One way to probe the molecular dynamics of water is through Terahertz (THz) Kerr Effect (TKE) spectroscopy. TKE spectroscopy makes use of ultrafast laser pulses to orient molecules and excite intermolecular vibrational modes in water samples, which is then followed by a period of relaxation. The response is measured by the THz pulse induced birefringence, a change in the refractive index that is polarization dependent. In water, this is driven by the preferential orientation of molecules that follow the polarization characteristics of the applied electric field. During my SURF project, I used the molecular dynamics package GROMACS to simulate ultrafast TKE experiments on liquid water. I wrote a Python script to compute the orientation and birefringence of molecules from GROMACS output files. This script was used to probe the relationship between pulse main carrier frequency and sample response, as well as the role of ions in preferential orientation. Together, these experiments helped to clarify the intermolecular dynamics of water and will serve to guide future work.

Towards Learning Representations for Automatic Protein Classification and Design Using Neural Networks
William Dembski
Mentors: Matt Thomson and Emanuel Flores Bautista

While DNA holds the code for life, it relies on the proteins it encodes to build the complex macromolecules that make up organisms. Understanding these proteins and their functions is an essential task for biologists. With the rise of modern sequencing technologies, the amount of sequenced proteins in databases is growing exponentially. However, the capacity for experimentally assessing protein function and understanding its biological role inside cells is still relatively slow. Even in model organisms, about ⅓ of the proteome is still functionally unannotated, which hampers our ability to completely understand and engineer even these well studied living systems. In this work, we exploit advances in deep learning to classify proteins based on their amino acid sequence alone, achieving better results than current methods like BLASTp for protein classification. We also present initial results towards learning representations for protein design using generative models.

PAC-Bayes Generalization Guarantees for a Robot Learning to Pour

Shrikeshav Deshmukh
Mentors: Anirudha Majumdar, Allen Z. Ren, and Richard M. Murray

Using imitation learning to develop control policies in robotic applications does not guarantee success when the policies are implemented in novel environments due to imperfect demonstrations or imprecise policies. To improve the generalization of policies, the Probably Approximately Correct (PAC)-Bayes framework is used to guarantee a lower bound of success for such control policies. Here, the framework is implemented on a Franka PANDA arm completing the task of pouring spheres (meant to represent liquid) into a bowl, all done in simulation via PyBullet. The robot learns a distribution of stochastic policies from a collection of expert demonstrations via behavior cloning with a conditional variational auto-encoder; PAC-Bayes is applied upon this distribution to improve generalization of those policies. Exact results have not yet been accomplished due to the complexity of the task. Subsequent work involves improving the process of embedding the trajectories into a latent variable (this will improve the initial distribution of policies), increasing the diversity of the expert demonstrations, and increasing the variety of environments. The lower bound of success for this task is predicted to be 60-70%.

Metal–Organic Materials as Barocaloric Materials for Pressure-Induced Solid-State Cooling

Vidhya Dev
Mentors: Jarad Mason, Jinyoung Seo, and Theodor Agapie

With growing concerns over the devastating impact of fluidic refrigerants on global warming, scalable alternatives to the vapor compression cycle are being investigated, including refrigeration based on solid-state caloric effects. Caloric effects are thermal changes of the material due to application of an external stimuli such as pressure (Barocaloric), magnetic (Magnetocaloric) or electric fields (Electrocaloric). Previously, the Mason Group has worked on two-dimensional metal–organic materials with long alkyl chains that undergo thermally-induced order–to-disorder phase transitions (known as "chain melting" transitions) that offer large volume and entropy changes beneficial for barocaloric cooling. To further explore these materials, SURF 2020 has focused on designing a new class of metal–organic barocaloric materials that undergo a spin crossover (SCO) in the metals. The large volume changes associated with SCO transitions coupled with chain melting may increase the sensitivity to pressure, inducing large barocaloric effects under easily accessible pressures. After surveying the structural/thermal properties of various 1-D and 2-D iron(II)-based SCO compounds in the literature, we have compiled a library of compounds that may exhibit a SCO triggered by the chain melting transition. We hope to eventually synthesize and characterize the compounds via high-pressure calorimetry to test the SCO effect on barocaloric cooling.

X-ray Pulsation Searches With Improved NuSTAR Clock Corrections

Audrey DeVault
Mentors: Fiona Harrison and Amruta Jaodand

NASA's NuSTAR (Nuclear Spectroscopic Telescope Array) satellite is a telescope extending up through the hard X-ray range (5-80 keV) used to observe many types of X-ray sources, including pulsars. Pulsars are neutron stars emitting coherent, pulsed emissions. Tracking these pulsations allows us to understand the intrinsic spin period and its evolution, which may help answer a fundamental question as to how, sometimes, pulsars are recycled to spin frequencies as rapid as 700 Hz. Extreme timing precision (~microsecond) is critical to clearly interpret and fit neutron star pulsations. NuSTAR clock corrections introduced in 2020 reach timing precisions of <100 μs, enabling novel pulsar timing analysis in the previously unexplored hard x-ray range for the transitional millisecond pulsar J1023+0038. We have successfully isolated millisecond pulsations and are now modelling the torques acting on the system, affecting its spin. This pipeline, once tested on PSR J1023+0038, can be applied to a wide range of millisecond pulsars and used to construct long term spin down solutions. This would help us understand how the torques from soft and hard X-ray emission processes evolve differently in the accretion phase. Importantly, this study also allows us an absolute comparison of NuSTAR's timing abilities with other X-ray telescopes.

Development of a Low-Cost Optical Particle Counter for Aerosols

Schuyler Dick
Mentor: Richard Flagan

Currently, atmospheric aerosols are measured via large and expensive equipment, limiting the amount of data that is able to be gathered. The ongoing increase in atmospheric aerosols in our world calls for a more accessible measuring device that can give us more information about the changing climate and how it is affecting the entire planet, including remote regions and areas higher in the atmosphere. We seek to replicate the small, low-cost optical particle counter presented in Gao et al. 2016. To count particles optically, we use a 605nm laser diode with two cylindrical lenses and four apertures to focus the laser, while aerosol particles are pumped into the device. The aerosol particles scatter the laser light, which is detected with a photodiode and Arduino. The Arduino collects the scattering data as voltages, which can then be translated into intensities to determine the count of the aerosol particles. The completed device has the potential to have its components upgraded to ensure even more accurate results. Overall, we hope this device can be used in the future as an easy and accessible measuring tool for atmospheric aerosols.

Developing Durable Biocomposites Through Synthesis of *Chlamydomonas* Algae and Agricultural Waste
Evan Dicker
Mentors: Chiara Daraio and Helen Wexler

Prior investigations into plant-based materials have focused on the use of polymer or petroleum-based plastic matrices to compensate for the low mechanical strength of standalone biocomposites. However, these biocomposites are produced by unsustainable manufacturing practices and are unrecyclable, characteristics not typically associated with environmentally conscious materials. This work aims to both develop and characterize biocomposite materials composed of *Chlamydomonas* algae matrices and agricultural green wastes (i.e. pistachio shells and avocado pits) in order to rectify prior issues with carbon neutrality. In the past, SEM and confocal imaging of fabricated biocomposite algae samples have confirmed proper integration of the algal matrices and agricultural wastes; however, it was unknown whether the lamellated fibrous layers produced from the compression-based manufacturing process would result in orthotropicity. A parametrized search was conducted using finite element analysis (FEA) to determine the angled orientation of the multi-lamellated and globular structures within the biocomposite. By applying an inverse analysis process and dimensionality reduction to this FEA data, the directional mechanical properties were extracted. It was ultimately discovered that despite the initial compression, the biocomposite's overall orthotropicity was not a direct result of fibrous layer lamellation, but rather a potential effect of localized *Chlamydomonas* cell conglomerates generating orthotropic behavior.

Growing Functional Neural Networks on Neuromorphic Chips
Rachel Ding
Mentors: Matt Thomson, Guru Raghavan, and Cong Lin

Neuromorphic engineering utilizes systems that implement the same basic operations as those of the nervous system. Neuromorphic hardware, such as Intel's chip Loihi, allows for simulating large spiking neural networks over long periods of time more efficiently than on conventional CPUs. Currently, spiking neural networks implemented in chips like Loihi are sensitive to the initial wiring architecture and spiking input signal decreases in later layers. To address this, the Thomson Lab recently created a modular tool-kit in the form of a dynamical system to self-organize multi-layer neural networks where waves in a layer trigger autonomous waves in subsequent layers. We would like extend this method and apply the system on neuromorphic chips to autonomously grow functional 'brains' in-silico. For this project, we create a python script implementing a three-layered network using the modular tool-kit and Nengo, a Python package for building neural networks that includes a backend for simulating Nengo models on Loihi boards. Within the Nengo architecture, we implement a custom Leaky-Integrate-and-Fire (LIF) neuron type, a local spike-timing dependent plasticity (STDP) learning rule, and competition rules on layer inputs and outputs. The result is a multilayered spiking neural network that can grow and self-organize without human intervention.

Study of Low Mass LLP Decaying in the Muon System
Gabby Dituri
Mentors: Maria Spiropulu, Christina Wang, Cristian Pena, and Si Xie

The goal of this project is to determine how sensitive the Cathode Strip Chamber (CSC), the endcap muon system in the Compact Muon Solenoid (CMS), is in detecting long lived particles (LLPs, predicted by the Twin Higgs Theory) that decay in the CSC. Simulated signal samples of LLPs with different masses, lifetimes, and decay modes are studied by analyzing the cluster of reconstructed hits (RecHits) deposited by the showering of the LLP decays. Furthermore, we wanted to understand how the RecHit shower shapes are related to mass and the decay mode of the LLPs, and how the distribution was different from that of background (particles produced at the interaction point, not from the displaced decays). Preliminary study shows the muon system is sensitive to several LLP decay modes, including electrons, neutral pions, charged pions, down quarks, bottom quarks, charged kaons, and tau leptons, and to LLPs with mass as low as 1 GeV. This study enables us to set competitive limits on light LLPs which has not been done before at the CMS, making this the first study to show the sensitivity of low mass LLPs.

LAX-Ray (LAteral X-Ray)
Marcus Dominguez-Kuhne
Mentors: Kenneth Goldberg and Joel Tropp

This summer in the UC Berkeley AUTOLab, I have been on a new project LAX-Ray (LAteral X-Ray), where given an occluded object on a shelf, the robot uses pushes to move occluding objects out of the way (not removing any objects) until the target object is exposed and can be grasped. For this project, we extend the results and applications of the X-Ray paper [1]. X-Ray is an algorithm for estimating occupancy distributions of an occluded target object using a set of neural networks given a depth image of the environment and a segmentation mask. However, this algorithm does not consider varying object sizes (large near camera, small far from camera) as a property of the novel shelf environment. Additionally, we introduce a new policy that uses solely pushes to move objects out of the way so we can grab the target object.

Material Testing in Autonomous Labs
Phillip Donnelly
Mentors: Harry Atwater and John Gregoire

As material testing continues to be a crucial part of making technological advances (e.g., more efficient solar panels), the development of autonomous laboratories offers a fast and resource-friendly method of making discoveries without the need for human oversight. Autonomous labs such as ChemOS have already shown that artificial intelligence paired with robots can carry out lab-based experiments and interpret the results to improve its approach before carrying out the test again; artificial intelligence allows the experimental procedures to improve with each iteration. The number of robotic parts compatible with the software and the variety of experiments that the AI can handle are both limitations. We are working to improve this situation by developing a software hierarchy that allows the lowest level of software (drivers) to control individual tasks done by the robotic parts. Drivers never have to be reprogrammed between experiments. Instead, a lab configuration is fed into higher-level software that can understand the new experiment and properly control the drivers. This allows for more flexibility in terms of adding new robotic parts and changing the experiment being carried out.

Disjointedness in the 1-4-6-4-1 Adinkra
Ismail M. Elmengad
Mentors: Sylvester Gates and David Hsieh

Adinkras are a graphical representation of supersymmetric systems of bosons and fermions. In these representations, information may be encoded through colored links, open or closed nodes, link grading and height of nodes. From an adinkra with nodes at five distinct heights we will use bosonic and fermionic operators to reach a valise adinkra with two distinct heights. This valise Adinkra will be studied to find how decomposition of its links into groups may result in disjointedness. A disjoint adinkra suggests that the supermultiplet is composed of multiple supermultiplets. Based on previous work, potential disjointedness may indicate that the 1-4-6-4-1 adinkra may be represented as a more compact system.

Simulations of Particle Flow During Shear
Olivia Ernst
Mentors: Melany Hunt and Han Hsin-Lin

LAMMPS is a molecular dynamics simulator that has been adopted for granular flows. Hunt’s lab previously analyzed the transient and steady-state behavior of granular material through experimentation with a rheometer, since granular flows are less understood than single-phase Newtonian flows such as water or air. This project focuses on understanding how granular material interacts when shear is present in different geometries through simulations, and applying the results to what was analyzed previously through experimentation. This will be useful in developing reliable continuum models for granular flows.

Modeling Photonic QIP Experiments Towards a Quantum Internet
Sergio Escobar
Mentors: Maria Spiropulu, Nikolai Lauk, and Raju Valivarthi

In this research we studied and used a modern technique in Quantum Information Processing (QIP) known as the characteristic function formalism. This formalism, usually used in continuous variable formulation of QIP, was used in novel ways to model discrete variable experiments. In particular, we have used it to model coincidence counts of a heralded single-photon source based on spontaneous parametric down conversion (SPDC) in a way that is independent of previous calculations. The model found using this formalism was shown to be equivalent to previous models for this experiment, suggesting that this new technique is widely applicable. This experiment allows us to measure the quality of heralded single-photon sources, and in our case gave quantitative bounds on the practical use of SPDC as a single-photon pair source through calculation of the $g^{(2)}$ coefficient. Looking forward, we will next use this formalism to compute the degree of indistinguishability for photons that are used in quantum teleportation experiments. Following that, we will develop a model for the teleportation and entanglement swapping experiments that will allow us to include many experimental imperfections.

Predicting Vertical Accretion in Wax Lake Delta, Louisiana
Sarah Feil
Mentors: Michael Lamb and Gerard Salter

Deltas are naturally low-lying and therefore susceptible to inundation from sea-level rise. For example, the Mississippi River Delta (MRD) in Louisiana is undergoing rapid land loss: since the 1930's 5200km^2 of wetland has been lost due to subsidence and sediment supply reduction. Unlike the rest of the MRD, Wax Lake Delta (WLD), a sub-delta of the Mississippi, is actively building land. Here, I develop a new technique for inferring sedimentation rates from sediment concentration data. I used sediment concentration previously derived from AVIRIS-NG spectrographic images and apply the method to WLD. Firstly, I extracted streaklines from the AVIRIS data and tracked sediment concentration along them. Next, by inverting the steady state 2D advection-diffusion-settling equation for suspended sediment and supplementing the AVIRIS data with depth and velocity data from the ANUGA hydrodynamic model, I estimated sedimentation rates throughout the WLD. I found that some regions of the delta had active sedimentation, while others were net-erosional. This method can be applied to deltas worldwide to estimate sedimentation rate to better predict land area change due to sea-level rise.

Development of Predictive Maintenance Methods for Anomaly Detection and an Application for Mobile Monitoring and Repair
Joshua Field
Mentors: Michael R. Hoffmann and Clement Cid

The Seva Project has developed, and is currently in the process of field testing, sanitation units that treat and repurpose wastewater. Through the course of this project, the significant problem of unit maintenance has arisen. When a unit breaks down it is challenging to diagnose and subsequently fix the malfunction. The goal of this project is to: 1) Apply Predictive Maintenance machine learning techniques in an unsupervised form in order to detect and diagnose unit malfunctions. 2) Develop a full-stack solution, in the form of an Android Application and AWS software, that connects maintenance operators to the unit, informs them when malfunctions occurs, and guides them step-by-step through the repair process.

A Generalization of a Generalized Turán Problem
Sara Fish
Mentor: David Conlon

Let ex(n, T, H, s) denote the maximum number of copies of T in a graph on n vertices which is the union of at most s H-free graphs. The problem of finding ex(n, T, H, s) can be viewed as a generalization of the problem of finding ex(n, T, H), a question posed by Alon and Shikhelman. Previously, Fischer and Matoušek have studied the case ex(n, K_3, C_4, 3). We find various upper and lower bounds of ex(n, T, H, s). In particular, we determine ex(n, K_2, H, s) when H is a complete graph or a complete bipartite graph.

An Interactive Visualization Tool for Transmission Clustering Across All TN93 Thresholds
Hannah Fisher
Mentors: Niema Moshiri and Claire Ralph

Epidemiologists use transmission clustering based on genetic sequences to analyze the spread of disease. As the collection of viral sequences has become easier and less expensive, molecular epidemiology efforts are rapidly increasing in scale. HIV-TRACE, a standard tool in viral molecular epidemiology, performs transmission clustering by estimating pairwise distances between viral sequences under the TN93 model of evolution and linking individuals whose pairwise distances are below a chosen distance threshold. The choice of distance threshold is challenging; it varies between viruses, populations, and regions of the viral genome. As a result, epidemiologists often perform transmission clustering using multiple distance thresholds, but it is unclear what specific thresholds should be selected. Here, we present TreeN93, a web-based tool that enables studying transmission clustering results across all possible thresholds. Given pairwise distances, TreeN93 constructs an ultrametric single-linkage hierarchical clustering tree; cutting the resulting tree *d* distance above the leaves yields the transmission clusters obtained using threshold *d*. TreeN93 allows the user to visualize and interact with the clustering tree, and users can easily view and export clusters at any selected distance threshold. TreeN93 is available at https://moshiri-lab.github.io/TreeN93/, and the code is available at https://github.com/Moshiri-Lab/TreeN93.

Sculpting Sustainable Bioresin: Castable and Optically Clear Biocomposite From Chitosan and Hagfish Slime Fibers
Sasha Fishman
Mentors: Julie Kornfield, Priya Chittur, Tiziana Dilucco, and Lealia Xiong

The majority of artists' materials contain chemically toxic, non-recyclable plastics with damaging effects on human health and the environment. For example, petroleum resins, used for casting and laminating, break down into microplastics containing endocrine disruptors and carcinogens. Our goal is to engineer a novel bioresin suitable for sculptural applications as an alternative to petroleum resin by employing chitosan and hagfish slime fibers. The target material should be optically clear, strong, hydrophobic and versatile while being sustainable; notably it should be non-toxic, bio-based, and compostable in marine environments. Working from standard tests to perform in a studio setting without expensive instruments, we first optimize protocols to quantitatively measure viscosity, optical characteristics, and mechanical material properties. Through this process, we also establish methods of data analysis adapted to the chitosan-based materials under study. In the second stage, we use these tests to assess the effects of chitosan suppliers, multiple concentrations, different crosslinkers, and hagfish slime fibers to understand which variables improve properties for our bioresin. The outcomes of this project will not only serve as a model for artists to innovate with environmentally conscious materials, but also contribute to the larger goal of creating sustainable, compostable, and bio-derived materials.

Using Density Functional Theory to Afford Multivariable Analysis of the Influence of Molecular Additives on Cu Catalyzed CO_2 Reduction
Alex Fontani Herreros
Mentors: Jonas C. Peters and Alonso Rosas-Hernández

Electrochemical CO_2 reduction, powered by renewable electricity, is a promising strategy to decrease global CO_2 emissions and fossil fuel dependence while producing valuable chemical feedstocks and carbon-neutral fuels. Highly reduced, multi-carbon (C_{2+}) products from this reaction are particularly desirable due to their greater energy density and wider applicability. However, even the best current copper-based catalysts have achieved only modest activities and efficiencies for multi-carbon products at levels far from commercial viability. Our group has previously identified a class of *N*-substituted aryl pyridium additives that improve multi-carbon selectivity on copper electrodes through the formation of an organic surface film. While these additives share a common aryl pyridinium scaffold, the addition and substitution of different functional groups results in significantly varied performance. In this work, density functional theory (DFT) calculations were employed to compute a series of molecular properties for a select library of *N*-substituted aryl pyridium additives, with the goal of identifying parameters that showed promising statistical correlations with multi-carbon CO_2 reduction performance. Future work may involve employing more rigorous computational models to obtain stronger, more definitive trends, which could potentially be used to inform the design of improved molecular additives.

Applying Computational Chemistry and Machine Learning Techniques to Optimize Antibody Neutralization for SARS-CoV-2
Tea Freedman-Susskind
Mentors: Vanessa Jonsson and Natalie Dullerud

To develop effective strategies to combat the current pandemic caused by the SARS-CoV-2 virus, a thorough understanding of possible viral mutations as well as avenues to enhance antibody neutralization is crucial. We propose robust computational techniques to evaluate the impact of potential mutations of the SARS-CoV-2 virus, and for neutralizing antibody design. The proposed algorithm predicts the effect of mutations on antibody neutralization of SARS-CoV-2 and optimizes antibody design utilizing concepts from computational chemistry and machine learning techniques. Experimental neutralization assay data is processed with variable selection/regression methods to select sites that detract and enhance binding. For each of these sites, mutation stability as change in Gibbs Energy ($\Delta\Delta$ G) is calculated using FoldX software and analyzed in Python scripts to identify stabilizing mutations for neutralization in the antibody and potential viral escape mutations. Generated data will inform optimizing antibody design. Preliminary pipeline testing on the C105 antibody has yielded mutations that enhance neutralization and binding stability with SARS-CoV-2. Results will inform the search for optimized antibody-based approaches to combating SARS-CoV-2, and validate a relatively fast computational strategy to examine viral fitness after mutations.

Generation of Discrete Pedestrian Occupancy Probability Distributions
Bruno Freeman
Mentors: Richard Murray and Francesca Baldini

As we seek to introduce an increasing robotic presence into our day-to-day lives, developing autonomous agents that can safely navigate human crowds becomes ever more critical. In this work, we develop a low-overhead pedestrian crowd simulator that employs the social force model to realistically guide pedestrians, represented as two-dimensional particles, through arbitrary static environments. Our simulator records chronological sequences of "occupancy grids" that encode continuous agent position data into a discrete form. We train a recurrent neural network on these sequences to the end of generating occupancy probability predictions for each grid cell at future time steps. These probabilities form heat maps over the environment that provide insight into future areas of high and low collision risk. Such information is integral to assessing the safety of potential routes during autonomous navigation.

Investigating Multiphase Flow in a Space-Based Catalytic Reactor
Diana Frias Franco
Mentor: Melany Hunt

As the future of long duration human-space exploration quickly approaches, packed bed reactors have been investigated as a system that can potentially convert liquid waste into useful products, such as potable water, because of their compact design and minimal operating power. NASA's Packed Bed Reactor Experiment (PBRE) that is currently operating on the International Space Station studies gas-liquid flow through porous media under microgravity conditions, providing pressure transducer data and images of the flow. Results from the PBRE show pressure variations within the tube, which is critical to understanding the effects of microgravity on the flow. Variations, such as those featured in pulse flow conditions, are particularly important to understand since an unsteady flow regime is optimal for catalytic reactor design, because an unsteady flow is accompanied by enhanced mixing between the gas and liquid phases. To fulfill this need, an analysis that featured additional standard deviation data was performed using the PBRE raw data to better understand the unsteadiness in some flows. Paired with the creation of high-speed videos, the current work will enable better designs for future space-based catalytic reactors.

Design of a Cryogenic Test System for Low Light Testing of Large Format Infrared Detectors
Anderson Furlanetto
Mentor: Roger Smith

The TMT Infrared Imaging Spectrograph (IRIS) is the latest large format infrared detector soon to undergo testing. A cryogenic test system has been developed for the IRIS detector and its associated hardware from a dewar with an annular LN2 tank and deployable test articles employing active and passive mechanisms. To expedite the testing process the test system has been designed for consistent, rapid, and precise deployment, configurable hardware to install additional test articles, and fewer thermal cycles between test configurations. Handling and installation guides have also been developed to minimize risk of damaging the detector during its assembly.

Spin–Phonon Coupling in Optically Addressable Molecular Qubit Candidates
Elisabeth Gallmeier
Mentors: Ryan Hadt and Alexandra Barth

Quantum computing and sensing are rapidly developing fields that have enthralled the scientific community with their potential to solve currently intractable problems ranging from structural biology to simulations of quantum phenomena. The overall realization of this hinges on the development of robust quantum bits (qubits). Recently, chemists have been investigating electron spin-based molecular qubits as promising candidates for the implementation of quantum computing and sensing. In addition to microwave addressability, the potential for optical addressability opens the door to additional diverse applications. Beyond the ground state phenomena of relevance for microwave addressable qubits, optical addressability relies also on light-induced spin-selective intersystem crossing, which defines particular ground and excited state criteria for applicable transition metal complexes. This research uses density functional theory (DFT), time-dependent DFT (TDDFT), and multireference *ab initio* calculations to explore d^2 tetrahedral Cr(IV) complexes as potential optically addressable qubit candidates. In doing so, we develop detailed magneto-structural correlations over the entire ground and excited state manifolds, including a spin-phonon coupling model to describe dynamic structural contributions to ground state zero-field splittings (ZFSs) and the mechanism of spin-selective intersystem crossings. These studies thus guide the syntheses and ground and excited state spectroscopic characterizations of new optically addressable qubit candidates.

Thermal Simulations for the Sensors for the Barrel Timing Layer of the MIP Timing Detector
Madeline Gardner
Mentor: Maria Spiropulu

CERN's particle accelerator (the LHC) is becoming the High-Luminosity Large Hadron Collider to study Beyond the Standard Model Physics. Increased luminosity means that bunched collisions will increase in frequency, so the MIP Timing Detector is being developed by CMS. The Barrel portion (BTL) of the MTD utilizes SiPM sensors whose operating temperature must be maintained below -30°C to reduce dark current. This project conducted thermal simulations on the cooling system and environment of these sensors. These simulations were conducted with ANSYS, and due to limitations of the student license, were conducted on 2 dimensional sections of the system. The system was simplified to include the following components in ¼ of an RU: cooling plate, copper brackets, SiPMs, LYSO crystals, Airex, Nomex piece, the I-beam etc. The SiPMs temperature should operate within one degree celsius of the cooling system, however various environmental factors will affect their temperature.

Properties of Pulsars in the Rapid ASKAP Continuum Survey
Mohit Garg
Mentors: David Kaplan and Tara Murphy

The Australian Square Kilometer Array Pathfinder (ASKAP), an array of 36 12-meter dish antennas, was used to form the Rapid ASKAP Continuum Survey (RACS). The large-area survey incorporates multiple properties including Stokes I and V. Stokes I measures the intensity of the radiation, while Stokes V measures the circular polarization. These properties will be used to crossmatch RACS observations with prior data of pulsars stored in the Variable Slow and Transients (VAST) Survey. The Australian Telescope National Facility (ATNF) Pulsar Catalog was used to identify all known pulsars. The properties recorded will include the brightness (mJy) and distance (arcseconds) of the closest RACS source. The paper will also present an analysis comparing against pulsar population models to estimate their visibility using different survey strategies.

Simulation of Deformation and Stress Distributions in Lattice Structures
Lauren Garriques
Mentors: Guruswami Ravichandran and Jack Weeks

The study of the deformation of lattice structures is important in furthering the understanding of the exceptional impact mitigation properties of cellular solid. Data from compression testing of lattice structures can advance the creation of strong, lightweight structures which may be used for energy management, thermal insulation, and filtration applications. Using finite element analysis in Ansys Mechanical software, deformation mechanisms and stress distributions are analyzed in different lattice specimens. The results are compared for 4x4x4 lattices with similar relative densities: an octet truss, Kelvin cell, and cubic lattice. Each of the lattices are tested using different strut geometries to analyze the effect of the cross sectional shape of the beams on the material properties of the lattice. Expected results are a greater understanding of how lattice strength is optimized through strut and lattice geometries.

The Effects of Brain Pulsation During the Cardiac Cycle on Detecting and Sorting Neural Spike Trains
Mahideremariyam Gessesse
Mentors: Ueli Rutishauser and Clayton Mosher

To understand how the brain supports cognition and behavior, neurophysiologists identify neurons by sorting their extracellular action potentials (EAP) (or spikes). When the heart beats, the brain pulses and we recently demonstrated that this impacts the spike waveform. Currently, no neural sorting software accounts for such cardiac-related motion, making current sorting protocols potentially prone to erroneous detection and sorting of spikes.

We developed an analysis pipeline to characterize how cardiac motion impacts sorting. We applied this pipeline to spikes recorded from human patients with electrodes implanted for monitoring the source of intractable epilepsy. We quantify how the spike waveform, spike sorting quality, firing rate, and signal to noise ratio vary with the cardiac cycle and created visual tools to examine how these features are correlated. We observed that most spikes belonging to a neuron are not mis-sorted to another neuron. Rather, some spikes may fail to be detected during the cardiac cycle or are misclassified into noise clusters.

Further research will use this pipeline to establish protocols to improve spike sorting. We will benchmark these protocols against neural simulations. Ultimately these protocols will allow us to differentiate neural responses to the heartbeat from those that are biophysical recording artifacts.

Determining Levels of Immune and Neurotransmitter Gene Expression in the Gut of a Mouse Model of Autism Spectrum Disorder
Allison Glynn
Mentors: Sarkis Mazmanian and Jessica Griffiths

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder which affects approximately one in 59 children in the United States. ASD is characterized by impaired social communication and repetitive stereotyped behaviors; it is often comorbid with chronic gastrointestinal (GI) issues such as frequent abdominal pain, diarrhea, and constipation. The *Shank3*$^{\Delta 13\text{-}16}$ mouse model for ASD is representative of the mutations of the *Shank3* gene found in approximately 1% of human ASD cases; these mice exhibit ASD-like behaviors in addition to an altered gut microbiome and GI abnormalities. This project seeks to identify genes differentially expressed in the *Shank3*$^{\Delta 13\text{-}16}$ mice which might contribute to their GI dysfunction phenotype by using Reverse Transcription quantitative Polymerase Chain Reaction (RT-qPCR) to quantify gene expression in the distal small intestine, the proximal colon, and the distal colon. The expression levels of fourteen immune and neurotransmitter genes related to ASD and GI dysfunction were quantified in *Shank3*$^{\Delta 13\text{-}16}$ mice relative to wildtype (WT) mice. Trends of differential expression of the immune related genes IFN-γ, IL-6, IL-17A, and Nos2 were found, indicating a possible role for inflammation in *Shank3*$^{\Delta 13\text{-}16}$ gut dysfunction. This requires further gene expression studies of related genes and pathways.

Disconnection-Mediated Grain Boundary Motion Resulting From Thermal Fluctuations
Mahi Gokuli
Mentors: Brandon Runnels, Marco Bernardi, and Julia Greer

In polycrystalline materials such as metals and ceramics, grain boundary migration influences mechanical properties including fracture toughness and ductility, playing a critical role in processes such as plastic deformation. A particular example of this is shear coupling, in which the grain boundary moves orthogonally to an applied shear load. Shear coupled grain boundary motion has been observed to occur through the formation of disconnections, step-like line defects described by a step height and burgers vector (together taken to be a disconnection mode). Disconnections have been shown to arise naturally due to the shear coupling factor, the nonconvex nature of grain boundary energy, and minimization of the dissipation potential (the energy loss per unit transformed volume as a result of boundary motion). If shear coupling is disregarded, thermal fluctuations would cause boundary migration via random walk. However, building on the aforementioned model, we show that thermal fluctuations can actually nucleate disconnection pairs from a flat Σ13 copper symmetric tilt grain boundary in a bicrystal under shear loading. These disconnection pairs are observed to aid boundary migration. We illustrate this behavior with symmetric tilt grain boundaries of various shear coupling factors, plotting stress-strain curves to reveal stick–slip motion of the boundaries.

Parameterized Simulation of Archeological Ceremonial Sites
Shir Goldfinger
Mentors: Norman Badler and Adam Blank

Due to the complexity of archeological ceremonial site usage, virtual simulations are extremely useful in visualizing activities and human movement in these spaces. Our aim is to develop a computational system that will be used to model and parameterize crowd movements based on a specific ceremonial site and event. Currently, we are specifically focusing on the Pachacamac site in Peru, working to create a large model of the site in Autodesk Maya that can be navigated by a heterogeneous crowd in Unreal Engine 4. Through the implementation of user interfaces, the system will allow for user variations on the large scale (e.g. crowd movement throughout the space) and on the small scale (e.g. individual emotion and performance), together providing a realistic and meaningful simulation for a specific event of a user’s choosing. Ultimately, by building the 3D application in WebGL (Web Graphics Library), users will be able to parameterize and run the crowd simulation regardless of their access to the authoring tools.

Neuroscience Data Interface (NDI)
Matthew Gonzalgo
Mentor: Stephen Van Hooser

In the field of neuroscience, sharing data across different labs has presented a problem with the absence of a standard data-sharing format. In the past, excessive effort was required between file conversion and organization to analyze data effectively, taking away from the efforts towards calculations and research. The Neuroscience Data Interface (NDI) seeks to fill this void as a platform-independent standard. Utilizing a simple vocabulary specifically geared for use in neuroscience, NDI is designed to allow users to easily upload and retrieve relevant information from raw data without needing to change the data’s original format. Currently there are two NDI implementations, one in Matlab and one in Python. The overarching goal of NDI is to create a multiplatform database system that is independent of the user and will allow separate labs to share and exchange information much more easily than the alternatives that are currently available.

Assessing the Deuterium Kinetic Isotope Effect of Hydrogen Radical Reactions on Photochemistry in the Mars Atmosphere
Aikaterini Gorou
Mentors: Mitchio Okumura and Danica Adams

Abundant geological evidence suggests the presence of surface liquid water at early Mars which would have required a thicker, warmer atmosphere than there is today. The present-day D/H enrichment in the Martian atmosphere provides understanding of the history of Martian water and its escape to space. In general, the mass difference facilitates H to escape faster than D, leading to an increase in the atmospheric D/H ratio over time, but this ratio is also affected by the deuterium kinetic isotope effect (^{D}KIE). The influence of the ^{D}KIE on the Mars D/H ratio is poorly constrained, primarily because many deuterated reaction rates are not well-known. This study investigates the influence of the ^{D}KIE on the Martian D/H ratio by (1) identifying which HO_x reactions the Martian atmosphere is most sensitive to in order to (2) improve constraints on their rates and ^{D}KIEs. We use KINETICS, the Caltech/JPL coupled photochemistry and transport code, to model the chemistry of 35 species linked by

193 reactions. As a sensitivity test, we vary the rates of HO_x reactions and find that a few reactions lead to substantial changes in hydrogen concentration. We then use semi-classical transition state theory to better estimate rates and ^{D}KIEs.

Predicting the Spread of COVID-19 Using Artificial Intelligence
Jethin Gowda
Mentor: Yaser Abu-Mostafa

Having accurate, local forecasts of the spread of the COVID-19 virus is critical for healthcare and political officials. To help satisfy this need, I am a member of the Caltech research team developing models, aggregation methods, and what-if scenarios for the spread of COVID-19 in the US. Specifically, my contribution to the team has been twofold; the first designing and implementing a robust epidemiological model to predict Covid-related deaths in each US county. This model, called PECAIQR-D, is an extension of the SIR epidemiological model, incorporating the effects of social distancing, asymptomatic infections, testing, and self-quarantine. The model's predictions were used in the team's aggregate model, marginally improving overall performance. My second contribution was taking "business-as-normal" model predictions that assume no future Non-Pharmaceutical-Intervention (NPI) policies to be implemented/removed, and creating a post-processing technique to apply a correction factor to these predictions based on real NPI policy changes. The technique can allow one to see the effect of enacting or revoking specific NPI policies on specific counties, giving a tangible measure of the importance of each individual NPI policy in different locales, critical information for political leaders trying to decide the best course of action regarding the reopening of the US economy.

Building Archival Query and Data Visualization Tools for Data-Driven Sample Exploration
Akshay Gowrishankar
Mentors: Vivian U, George Privon, and Andrew W. Howard

Telescope time is a valuable resource to the astronomy community. As more and morelarge data surveys are coming online, archival data science is becoming increasingly important inthe field, and the need for an efficient archival query system is imperative. To this end, we havecreated an archival query system and interactive data visualization tool to display the availabilityof observational data on user-selected sources. First, the archival query tool takes as input a listof sources, specified by coordinates, and runs queries in a number of astronomical archives,including the Keck Observatory Archive. The data visualization tool then takes the output ofthese queries and displays the observations with a number of useful attributes, such as angularresolution and frequency coverage. The user can interactively manipulate the selection criteriaand filter the data across these variables. The code, to be distributed publicly, will assist theastronomy community in utilizing various data archives and maximize the scientific value ofarchival science.

Nanoparticle Transport by Thermophoresis and Turbulence
Forrest Graham
Mentors: Guillaume Blanquart, Chandru Dhandapani, and Joseph Ruan

Soot can cause respiratory damage and environmental degradation and so simulating soot particle transport and deposition is important for pollution management and prediction. Because soot forms in rich regions of the flame ($T < T_{max}$) but oxidizes at stoichiometry ($T \approx T_{max}$), its particle density with respect to temperature is critical for understanding soot behavior. Unlike most passive tracers, soot particles transport is also highly sensitive to temperature gradients through a phenomenon known as thermophoresis. Modeling and isolating particle thermophoresis is therefore crucial. To this end, homogeneous, isotropic turbulent (HIT) simulations were conducted under different temperature fields with passive particles in the free-molecular regime ($K_n >> 1$). A "flame" was emulated by an imposed 1D sinusoid in temperature and a sweep of cases were run varying both the severity of the temperature gradient and the width of the "flame". First, two base cases with established asymptotic solutions were run: an isothermal HIT simulation and a quiescent flow with a "flame". Afterwards, the HIT simulations with a "flame" were run, and conditional means of the particle density with respect to temperature were extracted for comparison against general soot particle behavior.

GPS-Vision Fusion for Relative Spacecraft Navigation
Hannah Grauer
Mentors: Soon-Jo Chung and Vincenzo Capuano

The Aerospace Robotics and Control laboratory at Caltech proposed a novel method for relative spacecraft navigation that integrates GNSS and vision. This architecture relies on a tight fusion of carrier-phase GNSS observations and monocular vision measurements collected on board two space vehicles in earth orbits. This project focuses on the use of such algorithm for relative navigation between space vehicles of a formation in an on-orbit inspection mission. In particular, in this project, first, the GNSS visibility and of the field of views for each space vehicles and pair of space vehicles are analyzed. The second part of the project focuses on the preprocessing of the raw data provided by the GNSS receivers and their analysis, for definition of the system requirements. Finally, the project includes the implementation of modules for testing and validation of the pose estimation architecture.

Supernova Classification With Deep Learning Using Zwicky Transient Facility Alert System
Leah Griffith
Mentors: Matthew Graham and Dmitry Duev

In the developing field of time-domain astronomy, areas of the sky are observed every night and compared to the previous image of that area, creating a difference image sent out as an alert. One of these facilities specialized in this field is the Zwicky Transient Facility, which produces over a million alerts a night. Due to the sheer quantity of incoming data, humans can no longer individually analyze every alert.

Thus, the implementation of machine learning is vital to the further development of time-domain astronomy. Our research used a previously trained convolutional neural network that could detect if an alert was bogus - had some issue to render it unusable - or if it was real - could be used for analysis. Using this 2-class classifier, we developed a 3-class classifier to add the classification of “Supernova-Real”. This was in an attempt to eventually create an autonomous alert classification system that could classify real data into specific categories. Our trained network was able to accurately classify 93% of incoming data, yielding promising data for further development.

Building and Implementing Biologically-Inspired Spiking Neural Networks for Neuromorphic Computing
Nora Griffith
Mentors: Matt Thomson, Guruprasad Raghavan, and Cong Lin

Living neural networks grow and self organize from a single precursor cell. In the mammalian visual system, neurons in the retina are wired to neurons in higher layers through spontaneous spatiotemporal activity waves before the animal’s eyes open for the first time. Inspired by these biological neural networks, we have built a developmental algorithm for growing and self-organizing spiking neural networks (SNNs) that can be implemented on neuromorphic hardware. To begin implementing our network on neuromorphic chips like Intel’s Loihi, we translated the original network in MATLAB to Nengo, a Python library for building and simulating large-scale neural models. Future directions of this work include building networks of varying architectures for specific computing tasks and testing their functionality on-chip, as well as implementing varying wave dynamics.

Understanding Differences in Error for Separate Formulations of Chi Squared Minimization in the Presence of Systematic Uncertainties
Tomas Grossmark
Mentor: Ryan Patterson

The Deep Underground Neutrino Experiment (DUNE) aims to address problems in particle physics such as CP (charge parity) asymmetry. DUNE will collect vast amounts of data, and it is important to be able to understand the nuances in how this data will be analyzed. A chi squared minimization is a common tool for understanding the data, and previous error measurements can be accounted for using one of two methods: the addition of a penalty term or reformulating the entire chi squared equation with a covariance matrix. These methods should generate identical results. However, a notable difference between the error on the parameters generated by the two methods has been observed. This project aims to completely understand the differences between the penalty term and covariance methods, and then create a writeup to assist researchers in data analysis for DUNE. To do so, pseudo data was generated computationally based on a variety of different equations. The errors in both methods were then found and compared

Development of Caltech Curricula for Interactive Theorem Proving
Joshua Grosso
Mentor: Michael Vanier

Interactive theorem proving is the subfield of computer science and mathematics concerned with automated verification of mathematical proofs and computer programs. Despite the significant potential of interactive theorem proving, there is a surprising lack of emphasis on interactive theorem proving in undergraduate computer science curricula. The initial success of Caltech’s CS 101-2 (Interactive Theorem Proving) prompted this current work, which explores next steps in expanding this area of the Caltech curriculum.

Specifically, we studied *Programming Language Foundations* (Pierce, et al.) with the aim of developing course materials for a follow-up class to CS 101-2. We found the material to be mostly suitable for integration into the Caltech curriculum, and we developed several recommendations for how to effectively incorporate the material into a Caltech-focused academic setting.

Our project has pivoted into expanding the corpus of formalized computer science and mathematics. We have recently begun formalizing type systems in Coq beyond what *Programming Language Foundations* currently covers. Using *Types and Programming Languages* (Pierce, 2002) as a guide, we are formalizing definitions, theorems, and proofs about polymorphic type systems, higher-order type systems, and type systems with recursive types (and, possibly, Featherweight Java).

Barkhausen Noise in a Model Quantum Ferromagnet
Wenyan Guan
Mentors: Thomas F. Rosenbaum, Daniel M. Silevitch, and Chris Tang

When a time-varying magnetic field is applied to a ferromagnet, its magnetization does not change continuously, but rather in discrete steps corresponding to domain reversals. These discrete changes in magnetization can be observed as voltage spikes when measured with an inductive pickup coil known as Barkhausen noise. Here, we study Barkhausen noise in the $Li(Ho,Y)F_4$ family of model quantum ferromagnets. These materials can be tuned by applying a magnetic field perpendicular to the Ising axis and hence varying the relative strength of quantum fluctuations and random field pinning, changing the temperature, and varying the chemical doping. Each crystal is mounted in a helium dilution refrigerator to reach the mK-scale temperatures required to enter the ferromagnetic state, and run through a series of magnetic hysteresis loops to gather a large set of events. By analyzing both the statistics of various metrics (duration, area, energy, etc.) and the scaled shapes of the events, we can better understand the mechanisms driving domain wall motion, including the roles played by thermal fluctuations, quantum fluctuations, and structural disorder. In particular, we use the evolution of the average scaled shape as a probe of dissipation and drag in the domain dynamics.

Ternary Diagrams for Supernova Neutrino Emission Visualization
Rishi Gundakaram
Mentors: Kate Scholberg and Nick Hutzler

Over the course of a supernova event, neutrinos are emitted. The types, or flavors, of the emitted neutrinos evolve over the course of the event. Using simulated data for neutrino fluxes, we visualized the flavor evolution through a ternary diagram. Ternary diagrams plot 3D data on a 2D plane by taking advantage of the fact that the three pieces of data add up to a constant, in our case 100%. Our data comprises of three flavors of neutrinos: electron neutrino, electron anti-neutrino, and "x-flavor" neutrino (the other flavors combined). Our analysis confirmed that flavor evolution occurs most rapidly in the initial stages of the supernova. Once the supernova starts to cool, the flavors evolve less rapidly. Currently, our analysis is on theoretical fluxes, but we actually detect neutrino interactions. Future ternary diagram visualization on detected neutrino interactions may yield insights into inferring the underlying flux during a supernova event.

Genomic Analysis of Cyanobacterial Gas Vesicle Assembly
Bilge Gungoren
Mentor: Mikhail Shapiro and Bill Ling

Abstract withheld from publication at mentor's request.

p-adic Analogues of Exponential Sums
Arushi Gupta
Mentor: Xinwen Zhu

Over a finite field, an exponential sum is of the form $\sum_{x\in\mathbb{F}_q} \psi\left(Tr_{\mathbb{F}_q/\mathbb{F}_p}(f(x))\right)\chi\left(Norm_{\mathbb{F}_q/\mathbb{F}_p}(g(x))\right)$ for some additive character ψ and multiplicative character χ of $\mathbb{F}_p$. We consider a p-adic analogue of exponential sums, replacing the sum over a finite field with an integral over a p-adic field, and the characters of $\mathbb{F}_p$ with those of the p-adic rationals.

Investigation of Challenges in Deep Learning Framework for Diagnostic Image Quality Assessment
Matthew Hajjar
Mentors: Shreyas Vasanawala, Ukash Nakarmi, and Julian Tyszka

Magnetic Resonance Imaging (MRI) suffers from motion artifacts because of its inherent slow data acquisition procedure, and these artifacts can lead to misdiagnosis and necessitate patient revisits for rescan. Such inefficient use of scarce physical and expert resources could be avoided by developing an automatic machine learning pipeline that can evaluate medical image quality for the scans. However, the development of deep learning-based automated image quality assessment techniques suffers from several challenges such as inter-rater subjectivity, unbalanced/mislabeled data, and lack of proper deep learning modalities. In this project, we focus on three key aspects to study and address these challenges. i) We develop a simulated motion artifact dataset with different degrees of motion artifacts, ii) analyze the correlation between image quality scores from simulated motion artifacts and expert image quality ratings and iii) study a neighborhood-aware loss function.

Blind Domain Adaptation for Automotive Thermal Image Classification
Thomas Hayama
Mentors: Soon-Jo Chung and Anthony Fragoso

Currently, the most common approach to object detection in the automotive industry is with visible light cameras due to their availability, the abundance of information they provide, and their low cost, but this sensor fails in low light conditions or in extreme weather conditions that obscure the image or video feed. Using thermal imagery in tandem, though more expensive, would enable the additional detection of blackbody radiation, allowing people and vehicles to be seen even in complete darkness or obscured by precipitation.

The problem with training a thermal model is that there is very little thermal automotive data to train with. We are planning to make up for this deficiency by using visible light automotive data (of which there is an abundance) to train a model meant for thermal sensors.

We implemented principal component analysis to decorrelate the RGB color channels in visible wavelength images, randomly combined the new color channels to create a 1 channel image, and added minimal Gaussian blur to simulate thermal data. We were able to achieve a .972 accuracy classifying people from vehicles in thermal data which means using visible data is a viable strategy for training thermal models.

Automated Voltammogram Analysis for a Potentiostatic Electrochemical Sensor Capsule
Nicole Heflin
Mentors: Wei Gao and Jihong Min

The analysis of the contents of gastric secretion is important in detecting various stomach and duodenal diseases. Some older methods of analyzing electrolytes and gases in gastric juices consist of invasive, inconvenient, and sometimes unreliable methods. The development of a wireless ingestible capsule with potentiostatic electrochemical sensors that can sense various electrolyte concentrations directly in real time in the stomach can minimize the need for these invasive procedures and can provide an easier way to diagnose stomach-related disorders. This paper describes an automatic voltammogram analysis program that can be incorporated in a mobile and desktop application user-interface to analyze the electrochemical measurements produced by the sensors of the capsule. The program conducts automatic baseline correction through an iterative polynomial fitting algorithm, and automatic peak detection of input voltammogram data. The program is developed in both the Dart programming language, which can be built into an existing mobile application that uses Google's user interface, Flutter, and in Python programming language, which can be included in a desktop application that uses the Tkinter GUI.

Is Star Formation in Early Galaxies the Same as in the Milky Way?
Hagan Hensley
Mentors: Charles Steinhardt and David J. Stevenson

Observations of galaxies in the early Universe present a puzzle: there are so many massive galaxies in the early Universe that, under standard cosmology, they could not have possibly had time to form. But our measurement of the mass of a galaxy relies on the assumption that the distribution of its stars by mass, or the initial mass function (IMF), is the same as in the Milky Way. However, the IMF should theoretically depend on the temperature of star-forming gas clouds, which several lines of evidence suggest should be warmer in early galaxies than they are in the Milky Way. We fit galaxy observations to a model of this temperature dependence in order to find whether other galaxies' IMF temperature is the same as the Milky Way. We find that IMF temperature was higher in the early Universe, which may imply that the masses of early galaxies are lower than we thought.

Classical Simulation of Quantum Time Evolution Using the *qDrift* Algorithm
Valerie Hetherington
Mentors: *John Preskill and Ashley Milsted*

In our classical implementation of Earl Campbell's *qDrift*, we hope to use traditional computing resources as efficiently as possible to simulate the evolution of quantum systems through time. This may provide insight into the difficult task of quantum simulation, specifically it may identify systems that can be efficiently simulated either classically with Matrix Product States (MPS) or on a quantum computer. To accomplish this task, we will leverage MPS over *qDrift* and target local systems that simulate well under MPS. We will first simulate simple systems for the Trotter Decomposition and for *qDrift* that approximate the unitary matrix describing the system's evolution through time. In our initial *qDrift* implementation, for which we've chosen to simulate the Ising Hamiltonian, we sequentially evolved the state of the quantum system by randomly selecting a series of unitary gates. This process was repeated and its expectation values averaged for a number of samples. We generated scatter plots for the number of samples plotted against the error, showing that the error converges over time. We will move forward with simulation of other systems and begin to incorporate Matrix Product States into our current algorithms.

Opinion Formation on Networks
Sujai Hiremath
Mentor: Franca Hofffmann

As many people rely primarily on social media for sources of news and information, the content that is shared on social media can greatly impact people's opinions. Bounded-confidence models of opinion dynamics allow one to model the dynamics of how opinions spread online as a function of both homogeneous media ideology and non-media account opinions. In such a model, agents start with some opinion which takes a value in a continuum and evolves when agents interact with other agents whose opinion is sufficiently similar to their own. We build on recent work by adding heterogeneous media influence to such bounded-confidence models. In addition, as the number of the agents in such models increases the model becomes very computationally costly with respect to processing power and time. Therefore, we derive a macroscopic equation for an adjusted bounded-confidence model that more efficiently finds how the distribution of opinions on a social-media network evolves over time. In ongoing research, we are examining how closely this macroscopic equation approximates our original bounded-confidence model, investigating the possible stationary states of the macroscopic equation, and exploring other properties of the macroscopic model.

A Novel Proposal for Feynman Rules for Conformal Blocks
Sarah Hoback
Mentor: Sarthak Parikh

The AdS/CFT correspondence is a conjectural duality between d-dimensional conformal field theories (CFTs) and (d+1)-dimensional theories of gravity in Anti-de Sitter (AdS) space. It has significant implications for theoretical physics, because this correspondence implies an exact, though mysterious equivalence between CFTs, special Quantum Field Theories without gravity, and quantum theories of gravity. Conformal blocks provide a non-perturbative, theory independent bases for correlation functions within CFTs, thus, it is critical to study them to gain an understanding of both CFTs and quantum theories of gravity via the AdS/CFT correspondence. However, finding exact representations of conformal blocks is notoriously difficult. In this project, we propose a remarkably simple set of "Feynman-like" rules which can be used to write down explicit series representations of global scalar conformal blocks of arbitrary topology in d-dimensional CFTs. We prove this conjectural prescription for several infinite families of examples. The main tool we use in these proofs is the Mellin space formalism of CFTs. Finally, we explore various possible approaches to prove these Feynman rules in full generality.

Adding a Robotic Telescope to the ZTF Time-Domain Survey
Andrew Hoffman
Mentor: Shrinivas Kulkarni

The Zwicky Transient Facility (ZTF) is a three-day cadence survey of the northern sky being conducted with the 48-inch Oschin Schmidt telescope at Palomar Obser-vatory. ZTF discovers a majority of the new supernovae and fast transients present in the northern hemisphere each night, but following up on these targets is difficult due to the number of objects. There is potential for smaller and under-utilized telescopes to contribute to time-domain astronomy by providing timely photometry of ZTF targets. The Katzman Automatic Imaging Telescope (KAIT) isa 30-inch robotic telescope located at Lick Observatoryon Mt. Hamilton, CA. While very successful for over a decade after it's first light in the 1990s, KAIT struggles to compete with large-scale sky surveys at novel detections. We show our method for implementing KAIT into the ZTF program, the scientific results, and discuss the potential of small telescopes to contribute to transient science.

Mathematical Modeling of Natural Selection in Cancer
Alexandria Hong
Mentor: Paul K. Newton

Cancer is the chaotic perversion of delicate human machinery, wreaking havoc on orderly biological processes. Efforts to mitigate and eliminate it systemically primarily center around chemotherapy. However, the progression of cancer is an evolutionary process subject to the laws of natural selection. Neoplastic lesions are comprised of populations of heterogeneous subclones, engaging in competitive processes analogous to those on the ecological and behavioral levels. Chemotherapy exerts selective pressure on tumors, driving evolution towards chemoresistance. Attempts to subvert chemoresistance include various methods of combination therapy and selection inversion, as well as radical pathogen elimination and intermittent dosage. Using evolutionary game theory and the replicator equations governing competing subpopulations of cells, a mathematical model is built to explain competitive release in tumors and design chemotherapy schedules to avoid chemoresistance.

Extracting Radial Velocities of Cool, Low Mass Stars Using Forward Modeling
Katelyn Horstman
Mentors: Dimitri Mawet and JB Ruffio

Observations have shown exoplanets are abundant around type K & M stars. To search for companions around these cooler, low mass stars, we benefit from observing targets in near infrared (NIR) wavelengths (758 -1388 nm) where the star has its peak emission. We use observations from the Palomar Radial Velocity Instrument (PARVI) to measure radial velocities of type K & M stars. We present an addition to the existing data reduction pipeline created for PARVI that obtains stellar radial velocities from the one-dimensional spectra of the object. Using a forward modeling technique, we create a model spectrum to mimic the reduced spectrum taken from instrument observations. We focus on three factors when creating our model: a stellar template represented by high resolution PHOENIX models, the telluric absorption of the atmosphere derived from standard star exposures, and line broadening as a result of the resolution of the instrument and stellar spin. We apply our radial velocity data reduction pipeline to observations of GJ229 and recover a precision on the order of 100 m/s. Our results are a first step toward showing PARVI can obtain the dynamical masses of candidate planets orbiting K & M stars detected by the NASA TESS mission.

High-Mass Stars Stripped in Binaries Missing at Low Metallicity: Tests of Stellar Astrophysics and Gravitational Wave Progenitors
Beryl Hovis-Afflerbach
Mentors: Ylva Götberg and James Fuller

Recent detections of gravitational waves from merging binary black holes (BBH) have sparked interest in binary stellar evolution. A common envelope phase, initiated via binary stellar interaction, is predicted to be crucial in bringing two progenitor stars close enough for a merger to occur. To form a tight binary, the common envelope must be ejected, leaving behind a stripped star orbiting a black hole. Using models computed with the stellar evolution code MESA, however, we predict that low metallicity stars with masses 20-50 $M_\odot$ expand late in evolution, preventing the formation of long-lived binary products. The most common binary products are stars stripped of their hydrogen envelopes when they expand. Since stripped stars are the exposed helium cores of their progenitors, those with masses 7-22 $M_\odot$ should be missing in the low metallicity Small Magellanic Cloud. If this is not the case, about 50 such stripped stars should be present, making this prediction observationally testable. Since the most massive stripped stars are the brightest, they should be the easiest to detect, but they remain elusive, supporting the theory that few massive stripped stars are created at low metallicity. This brings to question the common envelope evolutionary pathway towards BBH mergers.

Quantum Simulations for XUV Absorbance Spectroscopy
Jennifer Hritz
Mentor: Scott Cushing

Table-top generation of extreme ultraviolet light (XUV) by high harmonic generation (HHG) can be used to probe the electronic and structural dynamics through core-level transitions. This can be used to gain more understanding of coupled carrier-phonon dynamics, eventually enabling control of ultrafast carrier thermalization and transport processes. In order to better understand these spectrum theoretically, we can generate a Silicon spectrum through quantum many-body calculations, as modeled with density-function theory (DFT) through Quantum ESPRESSO. To account for the effects of excitons, an electron-hole paired quasiparticle, we used OCEAN code (Obtaining Corelevel Excitations using Ab initio methods and the NIST Bethel-Salpeter (BSE) solver). These models are being improved to better match experimental spectrum results.

Linear Regression Methods for Fitting a Continued Fraction Regression Model
Kevin Huang
Mentors: Pablo Moscato and Adam Blank

Continued fractions as a representation for symbolic regression is a novel method proposed by Moscato et al., in which a memetic algorithm is used as the learning algorithm. This project expands upon this technique with the aim of using linear regression to create a faster version of the algorithm. By viewing the continued fraction representation as a hierarchical model, each depth can be iteratively fitted with linear regression. As linear regression encompasses a versatile class of well studied and computationally efficient techniques, it is shown that a linear regression based algorithm is a good candidate for improving the continued fraction approach.

Investigating the Role of Malat1 in Co-Transcriptional Splicing
Wesley Huang
Mentors: Mitchell Guttman and Prashant Bhat

Long noncoding RNAs (lncRNAs) have long been implicated in serving a diverse set of functions such as gene expression regulation. The metastasis-associated lung adenocarcinoma transcript 1 (Malat1) is one such lncRNA that is evolutionarily conserved and abundantly expressed in mammalian species. Malat1 has been shown to be a critical driver of many cancers, as its overexpression correlates to splicing defects in various carcinomas. However, the mechanisms by which Malat1 drives disease are unknown due to a lack of understanding of its cellular functions. Here, we report on the effects on splicing in Malat1 knockout (KO) cells. We observe significant splicing defects across different RNA-seq data sets and find that Malat1 likely plays an important role in coordinating co-transcriptional splicing.

Creating a Set of Virtual Decision-Making Tasks
Alexandra Hummel
Mentor: Dean Mobbs

Humans and other animals display different characteristic decision-making behaviors when faced with different levels of threat, as per Fanselow and Lester's Threat Imminence Continuum Model. A typical model used to experimentally evaluate such decision making is a "two-step task," in which a human participant is presented with two stages in every trial of a task, the first of which asks them to make a decision that will determine the state of the second. Using jsPsych, we created a set of four virtual two-step tasks which subjects will be able to complete remotely. Of these, two have randomized stakes, with one asking participants to maximize a reward and the other asking participants to minimize a threat. The other two feature an extra decision-making step, in which the participant must choose whether to use or save protection resources for later. Both of these are threat-minimizing, with one using stable high and low stakes and the other using volatile stakes.

Computational Modeling of Gut Butyrate Dynamics: Towards Enabling Multi-Modal Butyrate Detection Using Noninvasive Yeast-Based Sensors
Isabella Hurvitz
Mentors: Arnab Mukherjee, Emily Sun, and Mikhail G. Shapiro

Butyrate is a short chain fatty acid (SCFA) that is produced primarily in the gut and which modulates several major biological functions. A key challenge in addressing how gut microbial topography affects butyrate distribution is the inability to report on chemical profiles from within the gut of intact animals. A proposed solution are cell based biosensors that utilize a butyrate responsive G-protein coupled receptor to activate expression of aquaporin AQP1 reporters (for *in situ* MRI detection) as well as GFP-based reporters (for *ex vivo* detection in fecal material) in non-pathogenic ingestible yeast cells. The proposed sensor framework would enable spatially resolved multi-modal imaging of butyrate dynamics, providing measurements for how gut chemical profiles change as microbial biogeography varies. An *in silico* pharmacokinetic model has been developed to investigate the metabolism reactions of butyrate in the intestines of mice. The approach treats the gastrointestinal tract as a multi-compartment cylindrical reactor in order to model the transit of yeast cells and the fluctuation of butyrate concentration through the gut. Developing a quantitative modeling and analysis framework for the sensor-gut system could help elucidate the pharmacokinetics, sensor kinetics, spatial offset, and sensor decay that enable the development of the multi-modal butyrate detection system.

Modification of Y-Factor Method for Measuring Low Noise Amplifier Performance
Joren Husic
Mentors: Austin Minnich and Alex Choi

In any application where an analog measurement is made, an understanding of how much noise is present in the signal is crucial in quantifying the precision of the measurement. A method known as the Y-factor method is a common technique for quantifying the noise of an amplifier device. For low-noise devices, a number of challenges arise in the application of this technique. In particular, the noise performance of the device in question is quantified in terms of a number of parameters of the measurement system, several of which are difficult to measure directly or with high precision. Here we develop a modified version of the Y-factor method that addresses some of these issues. By measuring the Y-factor a number of times while varying certain parameters of the measurement, we are able to eliminate the need for some absolute measurements of parameters in favor of less difficult differential measurements, or in some cases eliminate the need for measurement of certain parameters entirely.

Quadruped Robot for Extreme Environments
Saehui Hwang
Mentors: Yong-lae Park, Taekyung Kim, and Sudong Lee

Abstract withheld from publication at mentor's request.

"Team Flow" Real-Time Monitoring and Modulation System
Justin Hyon
Mentors: Shinsuke Shimojo and Mohammed Shehata

During an extended mission in an isolated space environment, the ability of the crew to maintain optimal synergy and group cognitive function has been identified as an outstanding concern by NASA's Human Research Program (HRP). The flow state, or "getting into the zone" is a positive psychological phenomenon which has been shown to result in elevated creativity, productivity, and positive emotions during task completion. This state can be synchronously experienced among members of a group who reach optimal synergy in completion of a common goal, in a scenario dubbed "Team Flow". The quantification of Team flow is a necessary step towards the improved training and monitoring of crew operation. In a previous study (Shehata et al, 2020), an analysis pipeline, which utilized electroencephalogram (EEG) analysis of two participants playing a music rhythm game, was developed in MatLab to study the existence and extent of Team Flow post-experiment. This project aimed to develop from scratch in Python, utilizing the open-source MNE-Python EEG analysis library, a pipeline to mirror and expand upon the results of the MatLab analogue in real-time, providing immediate feedback for the teamwork performance of participants. This allows for the generation of a real-time Team Flow score based on Auditory-Evoked Potential (AEP), intra- and inter-brain Phase Locking Value (PLV) synchrony, and Power Spectrum Density (PSD), which allows for instantaneous and objective monitoring of the Team Flow state. Though further experimentation using this system into the conditions for optimizing participant synergy was impossible to to COVID-19 related constraints, this neurofeedback will allow for future exploration into maximizing induction of the Team Flow state.

Regime Change With Analytical Focus on Models of Overlapping Generations
Michelle Hyun
Mentors: *Jonathan Bendor and R. Michael Alvarez*

When regime change occurs, one may question the cause of the change and the sustainability of the new regime. Economic and political crisis, consideration of future allocation of power, and the role of globalization in a nation are some of many factors considered when analyzing regime change. Considering the fundamental difference between the creation of a democracy versus the creation of a nondemocracy, and the difference between democracies that survive and those that do not, we will analyze models of overlapping generations of agents with finite planning horizons. Understanding regime change is important to avoid unwanted transitions and to instigate enduring change – as an example, the United States has involved itself in the regime change of many nations, such as Nicaragua and Panama, but has not always succeeded in establishing a lasting democracy. By understanding why democracies form and fail, we can hopefully avoid fatal mistakes and help sustain democracy.

Application of Holoraumy to 10 and 11D Supergravity Theories
Erik Imathiu-Jones
Mentors: Sylvester James Gates and Justin Campbell

Holoraumy is a mathematical tool arising from calculating the commutator of supersymmetric derivatives. Our current goal is to investigate the use of holoraumy in 10 and 11-dimensional supergravity theories. We can do this by finding a holoraumy formulation in preexisting supergravity theories. To do this, we have worked through the 11D Clifford algebra representations and developed Mathematica code to perform calculations in this algebra. Future work will be to continue working through the 11D representations of preexisting supergravity theories to search for connection to holoraumy tensors.

Exoplanet Discovery From Microlensing Event OGLE-2015-BLG-1726
Nissia Indradjaja
Mentors: Jim Fuller and Przemek Mróz

Gravitational microlensing is a phenomenon produced by gravity's effect on light. It is often used to find types of exoplanets that other methods, such as transit photometry and radial velocity, are unable to detect. In particular, it is sensitive to planets located around a planetary disk's snow line, the distance past which volatile compounds can condense into ices and where planets are believed to form. Initial observation of the photometric data taken by the OGLE and MOA surveys for event OGLE-2015-BLG-1726 shows the presence of a secondary object; analysis of the data was conducted to determine if this object is an exoplanet. Current approximations for separation between the host star and the object in terms of Einstein radius ($s = 1.09$) and mass ratio between the host star and the object ($q = 0.00027$) were obtained, suggesting that this object is indeed an exoplanet. Further analysis is needed to obtain more precise parameter sets and approximations for the information regarding the host star.

GPU Accelerated Variability Detection
Ethan Jaszewski
Mentor: Matthew Graham

Variability detection is an important task in the analysis of astronomical time series, often involving determining the periodicity of the magnitude of a given astronomical object. Many algorithms for variability detection exist, but available implementations often do not take advantage of GPU acceleration, and those that do can be improved upon. However, with the advent of large databases of astronomical time series, current implementations are prohibitively slow for comprehensive analyses of the available data. This project focuses on developing optimized, GPU-accelerated implementations of popular variability detection algorithms to allow for complete analyses of large time series datasets. The final GPU implementations are between 50 and 100 times faster than equivalent CPU implementations, and between 1.2 and 15 times faster than existing GPU implementations.

Quantitative Analysis of the Human Small Intestinal Microbiome
Jenny Ji
Mentors: Rustem Ismagilov and Jacob Barlow

The small intestine is home to a variety of microscopic residents that interact both with each other and the host. Relatively little is known about the how these interactions relate to abdominal pain and disease. Using a quantitative sequencing approach to analyze the microbiota of human duodenal samples, we were able to draw correlations between various microbial taxa, pain scores, weight, overall microbial load, diversity, and other relevant factors. Analysis of 250 samples allowed us to generate potential hypotheses relating specific taxa as risk factors for abdominal pain. Of the 286 detectable bacterial taxa, a few were of especial focus due to their disrupting nature of the ecosystem when present: Enterobacteriaceae (family), Escherichia-Shigella (genus), Lactobacillus (genus), Clostridium sensu stricto (genus), Enterococcus (genus), Aeromonas (genus), Romboutsia (genus), Pseudomonas (genus), Staphylococcus (genus), and Scardovia (genus). Analysis of the complex relationships between these disruptive species has helped elucidate part of the relationship between the small intestinal microbiota and gastrointestinal pain.

Developing Physical Lab Infrastructure for Thin Film Material Growth and Evaluation
Abigail Jiang
Mentor: Joseph Falson

The Falson Lab aims to grow high quality thin films using molecular beam epitaxy (MBE), and subsequently evaluate these materials at low temperature and high magnetic field for observation of quantum and topological phases. In order to fabricate and measure these materials, the physical lab infrastructure composed of mechanical, vacuum, and electrical equipment must be created. Three major physical tools essential to the infrastructure are: 1) a chip carrier and mounting socket to hold material samples for electrical measurements, 2) a variety of breakout and pomona boxes to connect samples with external measurement devices, and 3) a mobile pumping station to pump MBE equipment down to ultra-high vacuum, with dual purpose as a leak detector. All three projects and their corresponding components were designed, fabricated, and ordered, utilizing software for both CAD and PCB design. A more thorough understanding of vacuum and electrical system operations and their component manufacturers was also developed. The completion of these physical tools awaits on-site assembly.

Active Domain Randomization for Robust Control
Qixuan (Alice) Jin
Mentors: Yisong Yue, Anqi Liu, and Guanya Shi

The design of accurate and robust controllers for regulating robotic systems can be difficult due to the complexity of real world systems. Robust control strives to ensure desirable policy performance under the assumption that the real world parameters are bounded. Training controllers on real-world data can be expensive and inefficient. In recent years, simulations are used increasingly more often to train controller policies. However, discrepancies between simulation and reality can cause the failure of simulation-trained policies to generalize to real world applications. Domain randomization (DR) seeks to minimize this Sim2Real gap. The classic DR technique uniformly samples simulation parameters and trains the controller on the collection of these sampled simulations. The primary objective of our research is to improve the sampling efficiency of existing DR methods through active learning. We investigated the performance of model-based, active DR frameworks such as uncertainty sampling with Gaussian processes and meta-learning with neural networks. We observe notable faster empirical training convergence of these active frameworks for simulation environments such as the inverted pendulum and 2D quadrotor. In future work, we will attempt to expand these active DR frameworks to model-free policy learning.

Using Stars and Polarized Radio Emission to Map the Galactic Magnetic Field
Samir Johnson
Mentors: Anthony Readhead, Georgia Panopoulou, and Clive Dickinson

In order to detect the predicted B-mode signal imprinted the Cosmic Microwave Background (CMB) polarization which is a critical test of the Inflationary Cosmology Theory, the foreground emissions from our Milky Way Galaxy need to be modeled and subtracted from the CMB polarization map. With new sensitive data from synchrotron emission (WMAP + Planck surveys combination) and thermal dust emission (Planck 353 GHz survey) as well as optical polarimetry data with reliable distance measurements (Gaia Archive), a three dimensional model of the Interstellar Medium (ISM) and its effect on the galactic magnetic field can be determined. Here we show the structure of the ISM using cross-correlation studies of the synchrotron and thermal dust emission polarization angle data with stellar emission polarization angle with distance measurements. Preliminary results within the region bounded by -90° to 60° galactic longitude and 30° to 90° galactic latitude show better correlation between the stellar polarization angle data and the thermal dust emission polarization angle data (0.6342 difference statistic) than with the synchrotron emission polarization angle data (1.1683 difference statistic). This is to be expected, as the thermal dust emission data originates from the same phase of the ISM as the optical polarimetry data (thermal dust absorption) so the polarization angle data from both surveys should be tracing the same magnetic field orientation in the galaxy. The synchrotron emission data is detected from the relativistic phase of the ISM as a result of supernovae shockwaves; thus, the synchrotron emission data does not necessarily trace the same magnetic field orientation as the optical polarimetry data. The quantification of the correlation between the synchrotron and thermal dust emission data with the optical polarimetry data within the regions of the North Polar Spur will provide a more accurate determination of its distance. Several regions of the Galactic Plane suitable for similar studies have been identified.

Detecting Election Fraud With Machine Learning in Bolivia
Benjamin Juarez
Mentor: R. Michael Alvarez

With the increased demand of ensuring the legitimacy of democratic elections, the prominence of election forensics has grown as it has been applied in several circumstances in recent years. Moreover, machine learning algorithms have been utilized in order to detect instances of various types of voter fraud, from ballot box stuffing to vote stealing. In particular, Bolivia's recent presidential election in 2019 turned controversial abruptly as late-counted votes gave the impression of fraud in favor of the incumbent, resulting in a period of political instability. However, recent studies of this election are suggesting that no such fraud had occurred. With this project, once the data sets for the official and preliminary counts of the election were acquired, further validation was necessary in order to become familiar with the data and prepare it for modeling. The next step involves constructing the appropriate training data set from a combination of historical data, demographic data, and synthetic data through regression analysis. Future work in this project involves the construction, training, and implementation of a supervised machine learning model that classifies mesas as clean or at-risk through consideration of elements such as turnout and vote shares.

Determination of the Turbulent Flame Speed Scaling for Hydrogen Combustion
Calle Junker
Mentor: Guillaume Blanquart

Due to its high heat of combustion, hydrogen generates more energy per kilogram than any hydrocarbon fuel. Burning cleanly with only water as a byproduct, hydrogen combustion provides a potential solution to address growing energy needs without leading to greater emission of pollutants. The objective of this study is to determine the burning efficiency of hydrogen to further characterize hydrogen combustion. This is made challenging due to the presence of thermo-diffusive instabilities in hydrogen flames and because most practical combustion devices operate under turbulent conditions. The combustion simulations examined utilized lean premixed hydrogen flames with an equivalence ratio of 0.4 where flow occurred through a cuboidal domain with cross-section area A. Previously performed simulations of varying complexities were examined. The turbulent flame speed, S_T, and turbulent flame area, A_T, are typically related through the formula $\frac{S_T}{S_L} = \frac{A_T}{A} I_0$, where S_L represents the laminar flame speed, and I_0 is the burning efficiency constant. In this work, the following formula is proposed for the calculation of I_0: $\int \langle \frac{\dot{\omega}_{H_2O}}{|\nabla T|} \mid T \rangle / \int \frac{\dot{\omega}_{H_2O\ 1D}}{|\nabla T|}\, dT$, where T represents temperature and $\dot{\omega}_{H_2O}$ represents the production rate of water. The burning efficiency can be used to predict hydrogen combustion behavior in highly turbulent flames.

A Predictive Model for the Lactate-Pyruvate Ratio in Cerebral Edema Cases Using Gradient-Boosting Techniques
Jaeyoung Kang
Mentors: Jefferson Che, Shane Shahrestani, and Yu-Chong Tai

Cerebral edema, a common symptom of stroke and traumatic brain injury, describes swelling caused by excess fluid in the intracellular and extracellular parts of the brain. This phenomenon often results in dangerous secondary injuries such as hypoxia, neuroinflammation, and increases in intracranial pressure. The progression of these secondary injuries can be measured using multimodal monitoring, which captures data from several parameters simultaneously.

Machine learning has previously proven effective in utilizing multimodal monitoring data to predict patient outcome. In this project, a predictive model was designed for the lactate-pyruvate ratio, a significant indicator of mortality, via the implementation of gradient-boosting techniques, which construct a model using a summation of weak learners.

A patient data set was selected to represent common cases of cerebral edema, including strokes, intracranial hemorrhages and subdural hematomas. Since the data set was continuous, multivariate, and longitudinal, the R package boostmtree was used during the training process. As the sample size of the training data set expands, the model will be compared to existing statistical approaches, such as Cox regression analyses, to ascertain value in a clinical setting. Moving forward, CT scan data will also be collected to measure the impact of catheter location on observed readings.

Understanding the Evolution of the #MeToo Movement Over Time Using Topic Models

Sara Kangaslahti
Mentors: Anima Anandkumar, R. Michael Alvarez, and Anqi Liu

We seek to analyze the change in topics over time in a corpus. To do so, it is necessary to separate the data by arbitrary time intervals, such as months or weeks. In the current standard, Latent Dirichlet Allocation (LDA) modeling, it is difficult to analyze the differences across these time cut-offs, and thus the data from each month must be considered separately. Furthermore, traditional LDA has structural limitations, such as incorporating arbitrary topic correlations. Therefore, we instead utilize a novel topic modeling technique; we propose the use of a tensor-based LDA method, which currently lacks in-depth application to social science. As tensor LDA models have been shown to reduce the structural issues, including correlation, that are present in traditional LDA, this method improves on the existing standard. Furthermore, by using tensor operations between models, we can understand the change in topics between time periods and generate a continuous analysis of the topics over time. We apply this novel tensor technique to an important dataset, the Twitter discussion in the #MeToo movement. Particularly, we observe that the #MeToo movement has a fast evolution of topics, so especially in these data, our approach is beneficial to capture rapid changes over time.

Identification of Signaling States During Mammalian Development From Single-Cell RNA Sequencing Data

Nivedita Kanrar
Mentors: Michael B. Elowitz, Alejandro A. Granados, and James L. Linton

Abstract withheld from publication at mentor's request.

Investigating the In-Vitro Dynamics of High Intensity Focused Ultrasound Mediated Interactions Between Bacterial Gas Vesicles and Mammalian Tumor Cells Using Ultra High Frame Rate Photography

Rohit Kantipudi
Mentors: Mikhail Shapiro and Avinoam Bar-Zion

Abstract withheld from publication at mentor's request.

Graph-Based Semi-Supervised Learning

Ishani Karmarkar
Mentors: Andrew Stuart and Bamdad Hosseini

Graph-based semi-supervised learning is the problem of assigning labels to a data set given observed labels on a small subset of the data points. We focus on graph-based semi-supervised learning for binary classification and study the consistency of the probit optimization method in the case where the underlying data belongs to a single cluster. We assume that the labeled data are noisily observed in that some of the observed labels may be incorrect with respect to the ground truth. We develop theoretical conditions involving the parameters of the optimization problem and the number of mislabeled versus correctly labeled data points. We show that when these conditions are met, the probit method consistently assigns the correct labels to the entire data set. We provide empirical evidence to support these theoretical results in the form of numerical simulations of synthetic problems.

Comparing Regularized Maximum Likelihood and Deep Neural Networks Methods for Black Hole Imaging

Johanna Karras
Mentors: Katie Bouman and He Sun

Black hole image reconstruction is a challenging problem that relies on a sparse array of telescope measurements, which are subject to atmospheric, thermal, and systemic noise. Traditional regularized maximum likelihood methods for black hole imaging tackle the ill-posedness of the problem by introducing hand-crafted priors and regularizers, which aid gradient descent to avoid poor local minima while better minima may still exist for the predicted image. My objective is to compare these existing maximum likelihood methods with deep learning based reconstruction techniques for black hole imaging, as well as investigate the robustness to noise of images produced by both techniques. Upon completion of my implementations of the deep learning reconstruction algorithms, I plan to investigate bayesian neural network methods for uncertainty quantification and data analysis.

Identifying Protein Products and Inhibitors to Target the 2019 Novel Coronavirus

Catherine Kauber
Mentors: Ashish Mahabal and Nitin Singh

The outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent behind the COVID-19 pandemic, has infected over 17 million people, killed more than 670,000, and affected the lives of billions. Unfortunately, there are no antiviral drugs or vaccines with proven efficacy against the virus. Peptide based solutions have emerged as a promising area for COVID-19 mitigation strategies due to their short time of

production, targeted application, prolonged stability as compared to conventional medicine, and versatile applications. To help create these solutions, genomic analysis was used to reveal evolutionary history and highly conserved regions in the SARS-CoV-2 genome. Then, the protein products of conserved sequences were identified. By finding these conserved sequences and their associated protein products, it may be possible to identify antiviral drugs or other peptide based solutions that could prevent SARS-CoV-2 infection or help to treat it.

Brownian Dynamics Simulation at the Interface of Phase Separation and Active Matters
Tiernan Kennedy
Mentors: Paul Rothemund and Tyler Ross

While microtubule active matter and nucleic acid nanotechnology are both flourishing fields in biological self-assembly, intriguing phenomena at their intersection have gone unappreciated. In particular, the combination of microtubule swimmers and motor-functionalized DNA condensates both provides a novel framework of swimmers in active solvents and method for investigating unexplored interplays between phase-separation and active matter. Here we use coarse-grained Brownian dynamic simulations to explore emergent properties rigid-rod swimmers in an active solvent. In particular, the interaction of these swimmers with viscosity discontinuities is hypothesized to produce a mechanical equivalent of the index of refraction of light and may even be used in physical lensing phenomenon. To date we have developed a robust and tunable simulation framework to examine microtubule motion at the interface between fluids of differing viscosities. The computational power of this system allows for expedient identification of experimentally feasible regions of phase space that possess interesting phenomena. Using our simulation framework to focus and tune experiment may lead provide novel methods of swimmer control for bioengineering and lead to fundamental insight into biological self-assembly in energy dissipating systems.

Novel Immune Checkpoint Candidates and Their Ability to Inhibit T-Cells
Esther Kim
Mentor: Hansoo Park

Activating the immune system is an important component of eliciting an antitumor activity. However, the existence of inhibitory pathways that regulate the function of T lymphocytes often poses challenges to immunotherapy. These inhibitory pathways are referred to as the immune checkpoints (ICP) and are necessary for the control of excessive immune activation. While the blockade of the most commonly known ICPs, CTLA-4 and PD-1, by inhibitors was a therapeutic success, the majority of the patients were not able to benefit from such therapy. Due to these concerns, the identification of a new ICP that does not confer a negative therapeutic side effect is an important contribution to the field of immune-oncology. Hence, the objective of the study was to determine whether the novel ICP candidate effectively inhibits T-cells. To achieve this goal, my approach was to 1) clone cDNAs (KIRREL2 from human and mouse), 2) purify protein, and 3) perform the T-cell assay by CFSE staining. Fluorescence-activated cell sorting (FACS) was implemented to confirm that the novel immune checkpoint candidate has an inhibitory effect towards T-cells.

Individual Differences in Implicit Emotion Processing on Human Faces
Lily Kitagawa
Mentor: Shinsuke Shimojo and Shao-Min (Sean) Hung

Humans are fundamentally social, with a need to form emotional connections with one another. Socialization requires the identification of one's own emotions and the ability to judge the emotions of others. This is hindered by alexithymia, a subclinical personal trait associated with an inability to describe the emotions experienced by oneself or by others. Our experimental design provides insight into implicit emotional processing before assessing an index of alexithymia. Such a design could elucidate why alexithymics experience such difficulties describing emotion: they struggle to use the information from implicit emotional processing. Furthermore, as alexithymia is a personality trait, such a design also allows insight into the individual differences' role in perceptual processes. We hypothesize that people who have alexithymia trait have difficulty utilizing implicit emotional information in responding to human faces. A series of behavioral experiments are designed to directly examine this hypothesis.

Disruption of the Animal Magnetic Sense: Radio Wave Transduction by Biogenic Magnetite via Anisotropic Hysteresis
Alexandra Klipfel
Mentors: Joe Kirschvink and Isaac Hilburn

Many animals use Earth's geomagnetic field to navigate. A subset of these magnetoreceptive organisms have been observed to lose their ability to orient correctly when exposed to low intensity radio frequency (RF) radiation in the 0.1-10 MHz band. Additionally, many such organisms contain biologically precipitated magnetite in their brain tissues. We propose that these biogenic magnetite nanoparticles are responsible for both the magnetic sense and RF sensitivity of these organisms. Specifically, magnetosomes transduce energy from the incident RF into acoustic energy in the form of ultrasound via anisotropic hysteresis. This project models two possible mechanisms by which the increased internal energy of the crystal due to anisotropic hysteresis can be converted to ultrasound: heat dissipation and electron cloud interactions. For each mechanism, the transduced energy is compared to the background thermal noise and their efficiency is compared. A general result is that the energy available to a magnetosome over one cycle of the incident RF is lower than the thermal noise, implying that many magnetosomes are required to generate a signal capable of triggering a biological response. Additionally, our model allows the design of an experiment to test whether heat dissipation is the dominant ultrasound generation mechanism.

Investigating the Spin Evolution of Neutrons Under a Magnetic Field and Confining Geometry
Viktor S. Koehlin Lovfors
Mentors: Brad Filippone and Chris Swank

We analyze the spin evolution on neutrons in a confining geometry and magnetic field using simulations and analytic methods taking into account diffuse collisions of the neutrons with the walls. We derive an analytic form for the average spin polarization of an ensemble of neutrons as a function of time and compare this to results from simulations numerically solving the Bloch equations.

Elucidating Catalysis With the "Gold Standard" of Quantum Chemistry
Patryk Kozlowski
Mentors: Garnet K. Chan and Yang Gao

To develop enhanced catalysts, it is necessary to be able to accurately compute surface and adsorption energies, quantities that govern the chemistry occurring at catalyst surfaces. Though density functional theory (DFT), the workhorse computational method of surface science, has proven inadequate to this end, wavefunction-based methods that treat correlations between electrons have shown promise as a viable alternative. Here, we have used one such framework, coupled-cluster theory (CC), established as the “gold standard” of quantum chemistry at the molecular level, to study the adsorption of carbon monoxide on a platinum (111) surface, a bellwether system in surface chemistry for benchmarking the performance of computational methods. The performance of a periodic CC method including perturbative singles and doubles (CCSD) for computing accurate surface and adsorption energies has been compared with DFT and second order Moller-Plesset perturbation theory (MP2).

Income Targeting and the Labor Supply of Rideshare Driver-Partners
Anthony J. Kukavica
Mentors: Colin F. Camerer and Matthew S. Shum

Neoclassical economic theory indicates that when presented with higher wages, taxicab drivers are expected to drive longer hours on shifts as a means to optimize earnings. In recent decades, the possibility of a reference-dependent income targeting model has been raised as an alternative to the neoclassical model. Under an income targeting framework, taxicab drivers set an income target and make stopping decisions based on whether or not this target is reached. We extend the existing analysis in the taxicab literature to a dataset from a popular ridesharing platform. This setting shares similarities with the taxicab industry in the services it provides to riders but differs markedly in structure, most noticeably in the flexible work hours it offers driver-partners as well as a dynamic surge pricing mechanism for setting fares. Our investigation aims to shed light on the behavioral similarities and differences between the taxicab and rideshare markets, contributing to the existing debate between the neoclassical and reference-dependent models. We also investigate whether driver-partner experience or other state variables play a role, and we explore some alternative frameworks for decision-making.

Engineering Hepatitis C Virus Immunogens for Increased Binding Affinity for Broadly Neutralizing Antibody Precursors
Sanjana Kulkarni
Mentors: Pamela Björkman and Andrew Flyak

There is an urgent need for a vaccine for hepatitis C virus (HCV), which caused approximately 400,000 deaths in 2016 and is responsible for chronic disease in more than 70 million people today. The high genetic variability of the surface glycoproteins, E1 and E2, contributes to viral escape from neutralizing antibodies and hampers vaccine development. Many known broadly neutralizing antibodies (bNAbs) bind to conserved regions of E2 to prevent viral entry into host cells, and they have been found in the ~20% of infected individuals who spontaneously clear HCV without treatment. Our goal was to engineer an E2 immunogen capable of eliciting production of bNAbs towards the development of a vaccine. Based on amino acid polymorphisms found in naturally occurring HCV strains, we made three unique mutations in the E2 ectodomains of four HCV strains using site-directed mutagenesis. Single and double mutants were evaluated in enzyme-linked immunosorbent assay (ELISA) to measure binding affinity for HCV-specific bNAbs previously isolated from patients. Binding between some E2 variants and bNAb germline precursors was modulated by these mutations, but most were unaffected. In future work, we will randomize positions in E2 to produce and screen libraries of mutants with greater sequence variability.

Pseudobinomiality of the Sticky Random Walk
Vinayak Kumar
Mentors: Venkatesan Guruswami and Christopher Umans

Expander graphs are extremely useful as an efficient source of randomness. Simulating random walks on explicit constructions of these graphs allow for efficient constructions of pseudorandom generators, extractors, and error-reduction protocols. In this project, we analyze a random walk model believed to be simpler but approximateto the expander random walk, referred to as the"sticky random walk". This model is like standard coin-flipping, butinstead of each flip being independentlyand equally heads or tails,the current coin flip sticks to the previous flip with a parametrizedprobability that is at least ½.We analyze how much this sticky walk resembles the standard binomial distributionand attempt tofind any implications these results have for the general expander walk.

Exploring the Early Death of Galaxies With Cosmological Simulations
Shalini Kurinchi-Vendhan
Mentors: Michaela M. Hirschmann, Charles L. Steinhardt, and Philip F. Hopkins

Recent studies have revealed that for some galaxies, star-formation has been suppressed or *quenched* surprisingly early in the history of the Universe, causing them to become prematurely "dead" or quiescent. New observations point to the energy release of their central black holes, called *active galactic nuclei* (AGN), as a possible explanation for the early death of these galaxies. However, theoretical models repeatedly show that AGN regulate star-formation in only massive galaxies. This introduces a tension between observations and simulations which cannot be resolved with new observations, as only the most massive galaxies are bright enough to be seen at early times. Instead, we take advantage of the state-of-the-art cosmological simulation IllustrisTNG-100 with a particular focus on isolated quiescent dwarfs, some of the smallest but oldest objects in our Universe. We predict that the simulated isolated quiescent dwarfs, and therefore small galaxies in the early Universe, were mainly quenched by stellar feedback, with AGN as a contributing factor.

Towards Formalization of a Distributed Computing Language
Thomas Kwa
Mentors: Alex Aiken, Elliott Slaughter, and Chris Umans

Legion is a programming language and runtime for distributed computing; it can scale computations to hundreds of nodes with much less programmer effort than other such frameworks. It is uncommon to use proof assistants to prove the soundness of languages of this complexity; however, it is also useful, because a bug in the type system could cause bugs in thousands of scientific programs. We formalize the type system and execution rules for a simplified version of Legion in the proof assistant Coq, mechanize some existing proofs, and create a roadmap for mechanization of the remaining proofs.

Spectroscopy of Nova Shells
Rebecca Kyer
Mentors: Matt Darnley and Éamonn Harvey

Cataclysmic variables consist of a white dwarf (WD) primary and a Sun-like secondary star which are undergoing mass transfer. Material accretes onto the WD and a shell of hydrogen-rich material accumulates on its surface until it undergoes thermonuclear runaway and this material is ejected. The size, duration and shape of the eruption can occur in a variety of ways which define different classes of novae. In this project we constructed a centralized database of nova observations from the Liverpool Telescope (LT). The LT has operated robotically for about 15 years, taking photometry, spectroscopy and polarimetry, all of which are automatically fully reduced and stored in a searchable data archive online. We created a database of spectroscopic nova observations within the LT archive which will be used by the Astrophysics Research Institute nova group at LJMU to feed models of nova shell ejecta. The final product includes a MySQL database of spectra as well as Python scripts for calculating emission line characteristics and displaying spectra and line plots. We applied the PyCross pipeline to observations of a sample nova to demonstrate how the database and scripts will be used to create pseudo-3D photoionization models of nova shells.

Modelling Reactor Dynamics for Oceanic Carbon Dioxide Sequestration With Limestone Weathering
Albert Kyi
Mentor: Jess Adkins and Sijia Dong

Myriad climate models have demonstrated the necessity of effective carbon dioxide capture and sequestration to reach 1.5 and 2 degree celsius global warming targets. Using the ocean as a carbon sink has emerged as a potential method for millenia-scale carbon dioxide engineering. Weathering limestone by reacting carbon dioxide emitted in flue gas streams with water to generate dissolved inorganic carbon fixes anthropogenic carbon as bicarbonate ion. Maritime shipping vessels present a clear opportunity to apply such a sequestration method, due to their high emissions and easy access to both seawater and concentrated carbon dioxide streams. These factors, in addition to calcite properties, were first modelled as a two-staged reactor, separating equilibration of seawater with flue gas in a bubble column, and calcite dissolution in a packed bed reactor. The model was then modified to represent a three-phase fluidized bed reactor, which accounts for carbon dioxide bubbling and mass transfer. Both models were analyzed for reaction time scales, effluent dissolved inorganic carbon and alkalinity, storage efficiency, and reaction optimization. This analysis will enable large-scale reactor development, with the potential to halve maritime shipping emissions.

Integers Representable as Sums of Tetrahedral Numbers
Charley LaFayette
Mentor: Dinakar Ramakrishnan

We consider representations of positive integers as sums of binomial coefficients of the form $\binom{x}{3}$ where $x \in \mathbb{Z}$. Using programs we wrote to explicitly determine representations efficiently by searching over integer lattice points, we first look for positive integers $n < 5000$ as $n = \binom{x}{3} + \binom{y}{3} + \binom{z}{3}$. Then, we search for integers represented by a sum of two, $n = \binom{x}{3} + \binom{y}{3}$, and explore patterns in the results to attempt to determine what conditions allow for such a representation. Finally, inspired by the famous "Taxicab numbers" we consider a subset of integers as a sum of two where we search for integers than can be represented as a sum of two positive tetrahedral numbers, two nontrivially different ways (e.g. integers $n > 0$ such that, $n = \binom{a}{3} + \binom{b}{3} = \binom{c}{3} + \binom{d}{3}$ where $a, b, c, d \in \mathbb{N}$); we dub these numbers "Scooter numbers". We then delve into patterns and relations in the set produced to propose potential properties of scooter numbers.

Particle-Laden Flow
Joshua Lee
Mentors: Guillaume Blanquart and Joseph Ruan

Particle transport in rivers, blood vessels, and the atmosphere demonstrate fluid particle interactions. We introduced particles uniformly everywhere, near the wall, and outside the TNT interface in a computer-simulated flow field. The particles are generally constrained in entrainment structures or eddies as they experience the TNT interface and the viscous sublayer very near the wall. This project focuses on the boundary layer's effects on the particles' trajectory. We tracked the particle movement by using Paraview to display the particles trajectory through the field. The behavior of the particles near the wall was compared to the transport of turbulence. The behavior of the particles injected in the free stream was compared to the location of the TNT interface as well as the boundary layer thickness. While all the particles followed a one-dimensional trajectory, those around the TNT interface experienced instantaneously random displacement. To investigate the mixing ability of turbulence, pairs of particles were then injected an infinitesimal distance apart throughout the boundary layer. From these pairs, a time scale was extracted to quantify the time before the two particle paths become completely decorrelated. To fully illustrate these results, we created multiple videos to see this.

Connectomics-Based Investigation of Chemical Information Integration in Fly Olfactory Neurons
Katelyn Lee
Mentor: Elizabeth Hong

Olfaction is an important and complex system in *Drosophila melanogaster*, yet the underlying logic of its structure remains largely unknown. The recent advancement in high-throughput neuron tracing and computing power has enabled a connectomics-based investigation of whether the neural connectivity of the olfactory system follows a random or an ordered, odor-dependent structure. Specifically, this paper seeks to understand the structure of early olfactory processing between glomeruli and multi-glomerular projection neurons in the antennal lobe, as well as the connectivity in the lateral horn. Projection neurons gather input from olfactory receptor neurons in structures called glomeruli and transfer the information to higher order processing centers such as the lateral horn. As such, this project investigates whether a projection neuron with multiple glomerular inputs gathers its glomerular input randomly or clusters glomeruli associated with similar behaviors. Statistical analysis via k-means clustering yielded preliminary data suggesting a non-random clustering of glomeruli among projection neurons.

Improving Lower Bounds for Linear Size Additive Spanners and Shortcutting Sets
Lin Lin Lee
Mentor: Virginia Vassilevska Williams and Thomas Vidick

To improve efficiency in graph algorithms, graph compression schemes such as graph spanners have been used to approximate the true graph within some specified error. Spanners are sparser subgraphs of the original graph (i.e., having fewer edges), designed to preserve information about distances in the original graph. Understanding the trade-off between the sparsity of the spanner and its accuracy is an area of interest particularly for additive spanners, which preserve distances within an additive amount of error. A related graph theory problem involves shortcutting sets, sets of edges added to a graph to reduce the distance between pairs of vertices, which have practical applications similar to spanners (e.g. communication networks). We examine the trade-off in sparsity and accuracy in both spanners and shortcutting sets in the case of a linear number of edges in the number of vertices, by attempting to improve on the existing lower bounds of $\Omega(n^{1/11})$ for O(n)-size spanners and $\Omega(n^{1/6})$ for O(n)-size shortcutting sets, respectively. We modify graph constructions used to show previous lower bounds in order to close the gap to the current known upper bounds for both problems.

The Development of Test Mechanisms for Earth-Analog Mars Helicopter Mid-Air Deployment
Regina Lee
Mentors: Joel Burdick and Jacob Israelevitz

This project will investigate the feasibility of a mid-air drone deployment as part of future Mars helicopter missions. This is an initial proof-of-concept experiment and demonstrates an earth-analog version of deployment. Using self-stabilization concepts from a tangential project, Streamlined Quick Unfolding Investigation Drone (SQUID), the goal is for the drone to stabilize itself after being deployed mid-air. If successful, this removes the need for landing gear and cuts the overall cost and weight needed to send drones to mars. This project focuses on creating a tether assembly to help test the drone's self stabilization aspect. This tether will deploy the drone when directed and allow the option to deploy the drone at an angle, simulating a more realistic descent. The prototype assembly was designed in Solidworks and manufactured using either 3D printed material or off-the-shelf parts. The assembly's height can be controlled remotely during testing, and drone release angle can be manually adjusted prior to testing. The drone is oriented in the same angle as the tether assembly, and is released via a pin mechanism controlled by a servo. The backshell prototype to simulate its aeroshell was made of foam as a quick and accurate housing for the tether and drone assembly.

Isolating the Role of Volatiles in the Erosion of Bedrock Chutes on Fresh Martian Crater Rims: 3D Comparisons Between Mars and the Moon
Janette Levin
Mentors: Michael Lamb and James Dickson

Bedrock chutes are channel-like features on the walls of martian craters which could be indicators of volatile activity. However, the presence of volatiles such as H_2O or CO_2 in the martian environment is not well constrained, and it is unknown whether bedrock chutes can be carved by dry rockfall. To investigate the origin of bedrock chutes and whether their morphology reflects volatile activity or not, we measured attributes of martian bedrock chutes on the global scale to determine their spatial distribution, as volatile activity has a strong latitude dependence on Mars. We further categorized the cooccurrence of bedrock chutes with channelized fans – features previously argued to indicate volatile activity. We also analyzed bedrock chutes on the Moon, which is presumably devoid of significant volatile activity, as a control environment. We produced high-resolution Digital Elevation Models from HiRISE, CTX, and LRO NAC data using stereo photogrammetry to measure depth, width, and slope of the bedrock chutes. Results show that dry rockfall can form bedrock chutes, and that deeper chutes correspond to areas with possible volatile activity. We conclude that bedrock chutes can form through dry processes, but systematic trends in their incision depths can be used as an indicator of volatile activity.

Implementing Remote-State Preparation on a Noisy Intermediate-Size Quantum Device
Laura Lewis
Mentors: Thomas Vidick and Alexandru Gheorghiu

In this research project, we investigate the implementation of a protocol for remotely preparing quantum states, with applications to blind and verifiable delegated quantum computation. The remote-state preparation protocol relies on the use of cryptographic primitives known as *Noisy Trapdoor Claw-Free* (NTCF) functions. These functions are based on conjectured post-quantum secure problems, such as *Learning with Errors* (LWE). The goal of the project is to implement the evaluation of such functions in superposition on a *Noisy-Intermediate Size Quantum* (NISQ) Device, allowing us to assess the potential of NISQ devices for cryptographic applications. We designed and optimized quantum circuits for the implementation of the NTCF functions and analyzed their performance using the quantum programming language Q#. We also considered different post-quantum secure problems as the basis for these functions (such as *Learning Parity with Noise*, *Learning with Rounding* and *Ring Learning with Errors*) to find circuits that are optimal in terms of depth and required number of qubits.

Understanding Photoassembly and Oxygen Evolution in Photosystem II Through Molecular Dynamics
Anna Li
Mentors: Woodward Fischer and John Magyar

Photosystem II performs the role of water oxidation in photosynthesis, with the overall reaction $2H_2O \rightarrow O_2 + 4H^+ + 4e^-$. The formation of O_2 occurs at the oxygen evolving complex (OEC), a tetra-manganese cluster, which is ligated by residues from the D1 and CP43 sub-proteins of the PSII complex. In this process, the OEC is oxidized four times by the redox-active Tyrosine$_z$ progressing through 5 intermediate S states (S_0 - S_4). Experimental data from EPR spectroscopy has elucidated many details on the intermediate structures and mechanisms of the OEC in this process, as well as the assembly of the complex, but due to the short-lived nature of some states and the difficulty of getting EPR data for Mn in different spin states, computational chemistry modeling can be used to supplement experimental results. We have done several optimization steps and molecular dynamics simulations on the X-ray structures of the D1 and CP43 sub-proteins to achieve an apo structure of the OEC active site. We continue to perform molecular dynamics on various theoretical states of the OEC to elucidate more information on the assembly and mechanisms of the OEC.

Monte Carlo Simulations of Electron and Phonon Thermalization in Graphene at the Diffusive-Ballistic Crossover Point
Sophie Li
Mentors: Michael Roukes, Stevan Nadj-Perge, and Raj Katti

The ability to detect single photons has applications in a wide range of disparate fields. One approach for pushing the limits of single-photon detector (SPD) energy resolution is to exploit the temperature rise in a low-heat capacity material upon absorption of a photon. A collaboration between the research groups of Prof. Michael Roukes and Prof. Stevan Nadj-Perge has developed a novel thermal detector based on graphene, due to its attractive electrical, thermal and material properties. In contrast to most existing graphene thermal detectors, this SPD accesses the novel regime in which the mean free path of charge carriers exceeds the dimensions of the flake. In this regime, boundary scattering dominates the behavior of phonons and electrons and determines the temperature distribution across the graphene flake. By way of Monte Carlo simulations, we have quantified the effect of diffuse scattering and specular reflections for circular and rectangular flake boundaries. The results strongly support existing literature regarding phonon heat transfer and surface scattering.

Blackhole Video Reconstruction With Particle Filtering
Erich Liang
Mentors: Katherine Bouman and Aviad Levis

Although radio interferometry is often directly applied in the imaging of celestial bodies, imaging blackholes is much more challenging. This is due to the rapid evolution of black holes over time and the noisy measurements of the Event Horizon Telescope (EHT) due to atmospheric conditions. As a result, blackhole imaging much resembles solving a Hidden Markov Model (HMM) problem, where the sequences of true image of the blackhole serve as the hidden states and the sparse measurements made of the blackhole serve as the emissions. One approach to the HMM problem is through particle filtering, a Monte Carlo type method. Though particle filtering is more computationally heavy and only approximates the true solutions to HMMs, the method reduces the reliance on assumptions of image priors that more traditional methods such as expectation-maximization (EM) rely on. The challenging part of applying particle filtering to blackhole video reconstruction is deciding how to assign weights to and resample particles; several current methods of doing so and their efficacy will be discussed.

Modelling the Vibrational Relaxation of O_2 Behind a Shock Wave
Yuying Lin
Mentor: Guillaume Blanquart

The heat shield is an essential component of any spacecraft, protecting the astronauts and payload within as the vehicle enters the atmosphere at hypersonic speeds. To design a spacecraft, we must understand the amount of heat it receives. Hypersonic atmospheric entry creates a shockwave, and the temperature behind this wave does not reach equilibrium immediately due to energy transfer from translational and rotational energy to vibrational energy. Therefore, the heat experienced by the heat shield depends heavily on the rate at which gas molecules behind the shock wave reach vibrational equilibrium. We modelled the vibrational relaxation of oxygen and the associated time constants using the FlameMaster program. Our model is based on the Landau-Teller (LT) equation and Schwartz-Slawsky-Herzfeld (SSH) theory, as well as empirical data from Milikan and White. Modifying the constants of the LT equation, found in work by Doraiswamy et al., has led to good agreement with Milikan and White.

Shedding Light Upon a Gigantic Cell: How Patterned Illumination Affects Caulerpa's Growth
Jianbang Liu
Mentors: Elliot Meyerowitz and Eldad Afik

Caulerpa, a single-celled alga, can grow to the length of two feet. We have previously observed that crowds of chloroplasts pulsate through Caulerpa's blades (leaf-like organs), and the frequencies of these pulses are related to external illumination conditions. Our goal is to investigate how illumination patterns impact the morphology of Caulerpa blades, and their chloroplasts' movements. We grew Caulerpa samples under controlled temperature and illumination, and documented their growth using a camera mounted above. We found that during alternating day-like illuminations, the chloroplasts pulsed exactly once a day. However, once we switched to continuous lighting, the pulsation reduced in magnitude and increased in frequency. We have also observed a phase shift in the blades' growth – they initially grew in all directions at equal rates, but after roughly two days, stopped expanding in width and only extended in length. We suspect the morphological shift could be related to a critical blade size, or illumination patterns. To verify these hypotheses, we will repeat the experiments with different illumination patterns, and attempt to induce different frequencies of chloroplast pulses. We also plan to control the samples' growth rates by raising them in different liquid mediums.

Constructing Deviant Women: The Intersections of Victorian Medicine and Literature
Victoria Liu
Mentor: Kevin Gilmartin

During the Victorian era, scientific theories began backing the socially constructed concept of "deviant women." Social scientists and doctors started treating their charges as objects to be objectively examined, and this medical gaze gave white, male scientists the authority and language to solidify a natural standard for women that emphasized marriage, motherhood, and morality. Although Victorians championed this bourgeois feminine norm, eugenic sentiments delineated who should and should not participate in the cult of domesticity. Deviant women, consisting of those considered "fallen", "unsexed", independent and ambitious, highly educated, disabled, non-white, non-heteronormative, and otherwise unconventional, were often advised against marriage or motherhood. Victorian novels both subscribed to and rebelled against these standards, and the research discusses how the trajectories of various female characters reflected a fear of reproduction in deviant women. The research attempts to elucidate how Victorian novels borrowed from contemporary medical discourse to construct the literary trope of deviant women. It provides an overview on relevant factual literature, followed by discussions of specific novels, including *Hard Times*, *Oliver Twist*, and *She*.

Deep Complex Networks Using Polar Representation and Multiplicative Update Method
Xiaoqi Long
Mentors: Anima Anandkumar and Yujia Huang

Real-valued neurons have been employed in most deep learning networks. In recent advancements of deep learning, complex-valued neural networks have been introduced as an alternative. Most state-of-the-art complex deep learning models use the cartesian representation of complex numbers $z = x + iy$ and update the real and imaginary parts of a neuron in separate channels. Given the current advancements in the field of complex-valued neural networks, we decide to focus on the problem of implementing such models with the polar representation of complex weights in the product manifold of scalar multiplication (magnitude) and angular rotation (phase) where a complex number z is mapped to $(|z|, arg\ z), z = |z|e^{iargz}$.

Meanwhile, multiplicative update methods in machine learning is proven to be particularly effective in learning compositional functions via gradient descent, training neural network architecture without learning rate tuning. Since our complex-valued network is highly compositional with magnitude and phase of a single weight viewed as two independent channels, such method should be useful. We compare the performance of three optimizers based on the multiplicative update method on our complex network: ComplexMadam, ComplexMadagrad, and ComplexMadamax -- complex multiplicative versions of the Adam, Adagrad, and Adamax optimizers. We also compare our polar-represented networks with the common deep complex networks in the cartesian form.

Approximating Enzyme Kinetic Systems
Nathan Lopez
Mentors: Dinakar Ramakrishnan and Robert Tanner

Kinetic systems are ubiquitous in chemistry and biological applications. Enzymatic systems can be mathematically modeled with the simple kinetic model: $E + S \rightleftharpoons_{k2}^{k1} ES \rightarrow^{k3} E + P$. The behavior of the rate of formation of product, dP/dt, can only be studied through non-linear differential equations, solutions that are impossible to obtain analytically. Hence, numerical simulations were created to study the behavior of such systems. Rough analytic solutions could then be created to approximately portray the behavior of the kinetic systems at three points of interest: $[S] = 0$, $[S] = S^*$, and $[S] = S_{peak}$. Furthermore, once rough approximations were created to explore the dependence of dP/dt on $[S]$, the dependence of dP/dt on the rate constants could be further explored. All the rate constants were discovered to have a unique effect on dP/dt and three-dimensional simulations were created to acquire a handle on dP/dt's dependence on both $[S]$ and a rate constant. Several observations were made regarding the numerical simulations and three-dimensional plots.

Automated Data Extraction of Chemical Literature
Julen Lujambio
Mentors: Sarah Reisman and Michael Maser

Currently, there is an increasing amount of research exploring how to leverage machine learning in the Chemistry field; however, many of these applications have run into bottlenecks when it comes to data acquisition. While there is a plethora of chemistry reaction data available, it is not formatted in a way that is easy for machines to interpret. To that end, the goal of this research was to develop a program that could accurately extract reaction data from total synthesis papers and format it in a machine-readable format through a combination of Natural Language Processing (NLP) and Optical Structure Recognition (OSR) techniques. The project specifically focused on total synthesis literature to serve as a template for future more-generalized extraction programs and how to set up their extraction pipeline. One major goal for this program was for it to be as seamless and easy to use as possible, so that no advanced technical computation skills were required. Finally, I would like to thank John Stauffer SURF Fellow for his support that gave me the opportunity to undertake in this research as well as Professor Sarah Reisman and my mentor Mr. Mike Maser.

Developing and Publishing Statistical Models of the COVID-19 Pandemic
Samuel Lushtak
Mentor: Yaser Abu-Mostafa

The ongoing COVID-19 pandemic is one of the most serious and disruptive public health crises of the modern era, and health officials require the most accurate models possible to shape their decision-making. Building on the models created by this year's CS 156b students, Professor Abu-Mostafa's research group is seeking to build the most accurate model of daily COVID-19 deaths at the county level in the United States under current policies, along with counterfactual predictions. My contributions to the project consist of helping to develop statistical models of the pandemic; processing, scraping, and formatting data for use in other models; and helping to develop the website on which the group's results will be displayed to the public.

Simulating Chemical Cross-Talk in High Density Enzymatic Biosensors
Sonali Madisetti
Mentors: Michael Roukes and Jessica Arlett

Irregularities in neurotransmitter release often indicate the presence of a neurological disorder. To study the brain's chemical environment, compact biochemical sensors are being developed to minimize tissue damage and maximize the amount of chemicals tested for. In high-density, multielectrode arrays of sensors, the hydrogen peroxide which comprises each sensor's output can diffuse to adjacent sensors, causing chemical cross-talk and skewing results. To study the causes of chemical cross-talk and possible methods to reduce it, a random walk program reflecting sensor conditions for hydrogen peroxide molecules was created. With 1.5 μm Parylene walls, a 1.5 μm enzymatic layer, and a 0.15 μm m-phenylenediamine layer, cross-talk was approximated at 18-19% per adjacent sensor. By thinning the enzymatic layer, where the hydrogen peroxide is generated, or introducing catalase onto the walls to neutralize diffusing molecules, cross-talk was shown to drastically reduce to as low as 1% per adjacent sensor, which is promising for experimental research into cross-talk reduction.

Extending Ring Polymer Molecular Dynamics Rate Theory to the Marcus Inverted Regime of Electron Transfer
Ananth Venkatesh Malladi
Mentors: Thomas F. Miller III and Xuecheng Tao

Ring Polymer Molecular Dynamics (RPMD) is a computationally efficient alternative to full-scale quantum simulations to obtain dynamical information about a system. However, previous studies have found that RPMD does not capture dynamics in the Marcus inverted regime in electron transfer reactions. By adding ring polymer beads in the normal and inverted regime such that their contributions to the mean field potential cancel out, we can include effects from the inverted regime in the mean field ring polymer instanton. This allows us to extend RPMD to describe the inverted regime.

Classifying Transient Sources With Novel Filters
Scott Martin
Mentors: Andy Connolly, Stuart Vogel, Bryce Kalmbach and Sara Frederick

Classifying and measuring redshifts to type Ia supernovae will be one of the key challenges in using supernovae as high precision cosmological tools. With hundreds of thousands of SN detected by telescopes such as the Rubin Observatory we can't do this spectroscopically. Our goal is to design a novel set of photometric filters to accurately measure supernovae redshifts. Using the code package SIGgi (Spectral Information Gain Optimization code), we set out to design a set of filters that provides maximum information gain when used with the existing set of LSST bandpass filters for photometric observations. We began by generating a model of a type Ia supernova spectrum and used SIGgi to optimize a single trapezoidal broadband filter to the spectrum as it's redshifted from 0 to 1. We then determined that a set of narrowband filters that captures the features in a spectrum that provide the highest information gain allows for more precise redshift estimations than a single broadband filter. We will show this improvement is real by comparing actual photometric redshift performances. After experimenting with different shapes and widths for the narrowband filters, we intend to design a comb filter that includes these narrowband filters, which we hope to be installed at the Apache Point Observatory to be used with the LSST bandpass filters to reduce uncertainty in photometric redshift estimations for more precise supernova classification.

The Atomistic Level Structure for the Activated Glucagon-Like Peptide 1 (GLP1) Receptor Bound to the Full Gs Protein and the GLP1 Ligand
Krystyna Maruszko
Mentors: William Goddard and Soo-Kyung Kim

The glucagon-like peptide 1 (GLP1) receptor is a class B GPCR which couples to a Gs protein. GLP1R is known to play a role in type 2 diabetes and obesity through glucose homeostasis and regulation of gastric motility. Using the experimentally determined CryoEM structure as a base, we computationally added in the missing residues and missing amino acid side chains. We then optimized the structure in order to search for the formation of new hydrogen bonds and salt bridges. We found that our new and improved model had a 147.62% increase in the number of interchain hydrogen bonds with the most notable being the formation of bonds between the Gs-alpha subunit and the IC2, IC3, TM2, TM6, and H8 regions of the receptor and between the GLP1 ligand and EC0 domain of the receptor. Additionally, we found an 80% increase in the number of interchain salt bridges with the most notable including bonds formed between the ligand and regions TM5 and EC3 of the receptor: the the bonds in the EC3 region are especially of interest since they were not resolved in the original CryoEM. Through the implementation of Molecular Dynamics, we expect to see further formation of salt bonds within the molecule.

Low-Power Bluetooth Core Body Temperature Sensor for Constant Temperature Monitoring
Kevin Marx
Mentors: Azita Emami and Arian Hashemi Talkhooncheh

Core body temperature is an essential vital sign for detecting fevers, a sign of many different diseases. Constant temperature monitoring opens the door for tracking changes in the core body temperatures of members of the population and can show how COVID-19 and other diseases spread.

My project focused on evaluating whether the nRF 52832 microcontroller can be used as a long-term temperature monitor, periodically measuring a user’s core body temperature for a month using a single CR2032 battery. The hardware architecture necessary to pull off this low-power feat was constrained by power and size requirements, as well as restrictions imposed by the new temperature sensor developed by the Daraio lab.

After verifying that the nRF 52832 can measure the small phase delays properly, the hardware components necessary to create a working prototype were chosen. This demonstrated the nRF 52832 indeed can operate as a temperature sensor, albeit with greater power consumption than initially intended. This project can be continued by developing a prototype and refining it until the device functions properly and can be tested on people.

Quantum Proofs of Space
Jack Maxfield
Mentor: Thomas Vidick

Proofs of space have been developed as a more energy-efficient alternative to proofs of work protocols. They have applications in blockchain technology and deterrence of denial-of-service attacks. A proof of space protocol allows one to demonstrate they have dedicated a significant amount of disk space to a problem. We hope to leverage the properties of quantum information to create a quantum proof of space, a quantum analogue to proofs of space in which use of quantum memory is demonstrated. After developing a definition for quantum proofs of space, we developed a quantum proof of space protocol and proved its correctness and security.

Dispersed Pressure Sensing for Flow Field Estimation
Amritavarshini Mayavaram
Mentors: Morteza Gharib and Peter Renn

Fan array wind tunnels offer an ideal test environment for small unmanned aerial vehicles (UAV) due to their high configurability and ability to simulate a multitude of different flows. However, flows currently must be manually measured and adjusted to achieve desired conditions. Measurements have to be taken by sweeping across which can be inaccurate and does not provide a temporally synchronized reading of the entire field. The development of an array of pitot-static tubes can provide accurate pressure and velocity data across an entire flow field at a single instance, giving a better picture of the entire field. This measurement technique has significant potential for machine learning applications on fan arrays. This pitot-static tube array was implemented on a small fan array in Caltech’s Center for Autonomous Systems and Technology (CAST) with the primary design conditions of minimizing interference with the flow produced by maintaining a small profile and structural integrity and stability.

Machine Learning Based Morphological Classification of Type Ia Supernova Host Galaxies
Gavin McCabe
Mentors: Syed Uddin and Paul Sternberg

Morphological classification of galaxies yields important information in the understanding of the evolution of the universe. To-date, morphological classification is primarily done visually by volunteer scientists through programs such as Galaxy Zoo. To reduce the reliance on volunteer scientists, increase accuracy, and efficiently classify the millions of sources detected by sky surveys, machine learning was used in conjunction with the paremetrization package CyMorph to automatically classify galaxies. With CyMorph, parameters such as smoothness (S), asymmetry (A), gradient pattern analysis (GPA), and Shannon Entropy (H) were calculated for a training set of pre-classified galaxies and used as a training set for a machine learning algorithm created with scikit-learn. After completing the training and testing of the machine learning algorithm, the overall accuracy was calculated as [+– INSERT ACCURACY –+]. The algorithm will be deployed to automatically classify supernova host galaxies detected by the Carnegie Supernova Project and various sky surveys and aid in the understanding of the evolution of the universe.

Bayesian Inferencing PNT Estimation System
Austin McCoy
Mentors: Soon-Jo Chung and Anthony Fragoso

It is difficult to navigate UAVs without a reliable ground-truth GPS measurement. Estimates of PNT can be refined using on-board aerial imagery which recognize learned features through color-invariant convolutional neural networks, and attempt to use the ground-truth location of such learned features to localize PNT. While forming a trivial PNT estimate based on immediate features is a viable strategy, integrating Bayesian inferencing into such an existing system allows for PNT estimates to be more precise and reliable. As a prerequisite to such a system, a versatile testing pipeline is needed to generate aerial imagery observed over arbitrary flight paths. Such a pipeline would integrate existing DSM and orthophotography databases to construct an artificial terrain, over which a camera model could be flown, rendering an accurate depiction of what an undermounted nav-cam would observe over such a flight path. Once such imagery is generated, it can be used to train bayesian filters to produce a series of alternative PNT maximum-likelihood estimators. The effectiveness of such a system is then evaluated by comparing the accuracy, precision, and robustness of the PNT estimates produced by the two systems. Future work could involve investigating using such inferencing techniques to also produce confidence estimates for measurements, indicating the level of trust that can be placed in such measurements.

Multi-Scale Analysis of Mechanical Properties of *Arabidopsis thaliana* VASCULAR-RELATED NAC-DOMAIN7
Robin McDonald
Mentors: Eleftheria Roumeli and Chiara Daraio

Biological-matrix materials (BMMs) offer a competitive replacement for fossil fuel-based materials with greater sustainability. BMMs incorporate biological bulk matter and biopolymers, yet the biomechanical properties of the components are not well understood. Using a multiscale approach, we aim to characterize the structural and mechanical properties of the plant bulk matter, in this case a suspension cell culture of transgenic VASCULAR-RELATED NAC-DOMAIN7 (VND7) *Arabidopsis Thaliana* cells. We used optical microscopy to characterize plant cell morphology, microindentation to obtain the Young's Modulus of combined turgor pressure and plant cell walls, and atomic force microscopy to isolate the Young's Modulus of the cell walls. Thus far we have shown that cell developmental stage, cell culturing media, and cell shape play a role in the final mechanical performance of plant cells. Expanding the fundamental knowledge of plant cells will enable the development of better performing BMMs.

Development of Template Matching Algorithm for Simultaneous Estimation of Pose and Shape Algorithm Lidar Initialization and Scale Solving
Krish Mehta
Mentors: Soon Jo Chung, Vincenzo Capuano, and Alexei Harvard

The Simultaneous Estimation of Pose and Shape Algorithm (SEPS) proposed by Aerospace Robotics and Controls Lab uses a monocular camera to estimate the pose and shape of uncooperative objects in space. Since a monocular camera is used, the pose can be estimated only to a scale. A single beam lidar is used to solve for the scale. The proposed module for the SEPS algorithm uses a novel template matching algorithm that is better suited for space applications that can match to subpixel accuracy. The scale solving algorithm template matches the lidar point on sequential images and then uses triangulation to solve for the scale. This is then fused with the SEPS algorithm to improve the SEPS pose estimation. The template matching algorithm proposed is made to be computationally efficient and has high accuracy.

Analysis and Design of Novel Integrated 3D Magnetometers in CMOS for Position Sensing
Hayward Melton
Mentor: Azita Emami

Fully integrated three-dimensional vector magnetometers are a crucial piece of many magnetic systems. Several new topologies for integrated CMOS magnetometers are analyzed, including novel metallic Hall effect sensors and coil-based structures that take advantage of modern fabrication processes. These magnetometers are designed with the goal of biomedical contactless 3D positioning; by creating a precise magnetic field around the sensor, both the position and orientation of the sensor can be determined. This necessitates a fully-integrated system, limiting the size, power consumption, and material composition of the magnetometer element. It was determined that integrated metallic Hall effect sensors are nonviable due to in-circuit limitations, and that coil-based sensors hold promise for future research due to their simplicity and natural elimination of the effect of constant fields.

Complex Rumen Microbiome Assembly Driven By Strain-Level Dynamics
Liana Merk
Mentors: Benjamin Good and Richard Murray

The principles that underlie microbial community assembly are of great interest to microbiome research. Complex population structures arise throughout life, often impacted by a range of internal and external forces. One such important community is the cow rumen microbiome, which supports more than two-thirds of the host's energetic requirements. Previous rumen microbiome work showed that deterministic factors, such as diet and age, can impact the taxonomic composition in a manner constrained by early stochastic colonization. However, little is currently known about whether strain-level processes follow these species-wide ecological dynamics. To address this gap, we studied genetic dynamics in the rumen microbiome of ten cows throughout two years of life. We tracked the longitudinal dynamics of thousands of single nucleotide variants using deep metagenomic sequencing across hundreds of species per host. The effects of diet, age, and delivery method on within-host evolution were investigated, providing new insights to the population dynamics of microbiome assembly. By comparing genome-resolved fluctuations with species-level ecology, our findings allow a greater understanding of complex community assembly.

Sensor Fusion for Object Detection on Autonomous Vehicles
Esmir Mesic
Mentors: Soon-Jo Cheung and Anthony Fragoso

Autonomous vehicles use either radar-based sensors or image-based sensors to detect objects on the road. Visible imagery is well-suited for seeing objects from afar and making out details in objects like signs but is severely hampered in extreme weather conditions like heavy snow or fog. Thermal imagery cannot clearly see details or objects from afar but excels in weather conditions. LiDAR is good at giving basic distance and direction metrics for objects and is computationally inexpensive but lacks in detail. Due to the benefits and drawbacks of the sensors, they act as complements to each other. By fusing their inputs using convolutional neural network models, we improve object detection performance in extreme weather environments for autonomous vehicles.

Automated Data Extraction From Chemistry Literature Using Natural Language Processing and Optical Structure Recognition
Arya Mevada
Mentors: Sarah Reisman and Michael Maser

Due to the nature of Organic Chemistry research, much of the data generated during experiments is placed in text documents in bulk paragraphs, while actual chemical procedure and such are kept in reaction diagrams. In order for someone to access this information, they would need to manually read the data and input it wherever is necessary. However, with the advancements in Natural Language Processing and Optical Structure Recognition, it is now possible to automate that reading. Using these processes, data can be read directly from the papers and interpreted, allowing a computer to read through a paper and extract information about the reactions that took place, from reactants and products, all the way to procedures, catalysts, reagents, and conditions. Creating a pipeline that can extract this data from a paper and store it in a machine readable and easy to access way can allow chemists to quickly read and store thousands of reactions in a data base without having to manually input it. This clears the way for large reaction databases that are continually populated with new reactions, while also giving chemists and computer scientists large sets of reaction data to create new predictive and analytical models with.

CT Airway Segmentation Using Cascaded 2D U-Net Neural Networks
Prashanth Mohan
Mentors: Albert Hsiao, Lewis Hahn, and Judith Campbell

Chest CT is a valuable imaging modality for diagnosing and assessing severity of chronic obstructive pulmonary disease (COPD). However, characterization of the airways as visualized on chest CT of COPD patients is currently a manual, qualitative task. We sought to automate assessment of airway caliber and wall thickening using machine learning. The first step towards this goal is segmentation of airways. Our proposed strategy first segments the trachea and then recursively segments successive generations of airways. This study was IRB approved and HIPAA compliant. Ground truth segmentation of airways taken from 100 chest CTs was generated by a board-certified radiologist. A U-Net pretrained for purposes of liver CT segmentation was retrained for segmentation of the trachea and the mainstem bronchi. The Dice similarity coefficient (DSC) was used for evaluation**.** A preliminary algorithm which only segments the trachea demonstrated a DSC of 0.96 (95% CI: 0.71-1.0) on test set images from 20 CTs. A second algorithm which segments both the trachea and mainstem bronchi had an average DSC of 0.93 (95% CI: 0.48-0.99) on test set images. Future work will build upon these algorithms to segment subsequent generations of airways.

Enhancements of Scanning Magnetic Microscopy (SQUID) Data Inversion Using Nanotechnology and Real-Time Noise Cancellation Techniques
Noah Moran
Mentors: Joseph Kirschvink and Isaac Hilburn

Scanning magnetic microscopy using superconducting quantum interference devices (SQUID) is extremely sensitive to weak dipole sources, making measurements extremely vulnerable to noise. This noise can negatively affect Matlab routines used to fit 3D magnetic dipoles to the 1D, vertical-axis magnetic field intensity measurements. Often times, such noise will result a large vertical height difference between the raw and unfiltered data. The SQUID microscope systems at Caltech and the Tokyo Institute of Technology (TiTech) both use two 50 micron SQUID sensors, arranged in a gradiometer configuration, to filter out background noise from the vertical magnetic fields that they measure. On the TiTech system, the filter and the algorithm used to fit a 3D dipole magnetic moment to the filtered data, result in anomalous dipole vertical positions, locating the dipole inside the actual microscope system instead of in the sample (~300 – 500 microns below the microscope). To determine the cause of this anomaly, numerous tests have been conducted, such as wire-scans to better constrain the actual physical height between the sample handler stage and the SQUID sensors, dipole fitting of various samples, and cross analysis between data measured on the SQUID microscopes at Caltech and at TiTech.

Precise Simulation of the High-Contrast Coronagraphic Testbed for Segmented Telescopes Using PROPER
Grady Morrissey
Mentors: Dimitri Mawet and Jorge Llop

The direct imaging of exoplanets is the best suited technique to reveal exoplanet population characteristics and atmospheric compositions. Direct imaging poses a unique problem for scientists; most exoplanets have incredibly small angular separations and high planet-to-star contrast ratios, e.g. approximately a 10^{-10} contrast ratio for an Earth-like planet around a Solar-type star. Along with new techniques in coronagraphy, which aims to suppress bright starlight to reveal accessory objects nearby, future telescopes (TMT, LUVOIR, etc.) will also be segmented to allow for large apertures necessary to reach small angular separations. The High-Contrast Coronagraphic Testbed for Segmented Telescopes (HCST) at the Exoplanet Technology Laboratory in Caltech uses a vector vortex coronagraph design along with adaptive optics to demonstrate the techniques necessary for the direct imaging of circumstellar objects. I updated the optical propagation modeling code (based on the PROPER library) for precise, end-to-end simulations using HCST's current layout. These simulations determine the necessary specifications for a theoretical Lyot coronagraph in the current system, the implementation of which is being considered for future experiments. The current model developed from this work will allow for a better understanding of HCST's capabilities and limitations, and also for more coronagraph technology demonstrations.

SAIR Epidemic Dynamics on Group-Structured Networks
Siqiao Mu
Mentors: Franca Hoffmann, Mason Porter and Heather Zinn-Brooks

Network structure is critical to understanding the spread of infectious diseases through a community. In particular, if a community is partitioned into groups by some characteristic such as age or ethnicity, the "mixing" patterns within groups and between different groups can drive or deter disease transmission. We develop mean-field approximations for a susceptible–asymptomatic–infected–recovered/removed (SAIR) compartmental model that reflects some key traits of the COVID-19 pandemic. To investigate the effect of mixing patterns and heterogeneous model parameters, we simulate SAIR epidemics on networks that consist of two groups and different intergroup/intragroup edge densities and recorded measures of disease prevalence. We determine that when the groups have different sizes, the larger group and the community are generally more protected from infections when reducing intragroup connections, whereas the smaller group suffers. Additionally, when we adjust group-specific parameters that govern disease transmissibility and recovery time—these encode the effects of a disease on a heterogeneous population more realistically—we discover that adjusting the mixing parameters can help protect vulnerable groups or, by contrast, disproportionately favor healthier groups.

Homomorphisms of Braid Groups and Mapping Class Groups of Surfaces
Aru Mukherjea
Mentor: Lei Chen

The mapping class group $MCG(S)$ of a compact orientable surface, or 2-manifold, of genus g, can be described as the group of path-components of the group of orientation-preserving homeomorphisms of S to itself that fix its boundary pointwise. A related object is the braid group on n strands B_n, which can also be viewed as the mapping class group of a disk with n punctures. We aim to classify homomorphisms $B_n \rightarrow MCG(S)$, exploiting the totally symmetric set formed by the odd-indexed generators $\sigma_1, \sigma_{3,}, \ldots, \sigma_n$ or σ_{n-1}, and obtain a characterisation for $k \geq 8, g \leq 2k-3$, where $k = \lfloor\frac{n}{2}\rfloor$. Time permitting, we aim to further extend this result to the case $g = n-2$. In doing this, Castel's result for $n \geq 6, g \leq \frac{n}{2}$ is extended, a step towards a possible rigidity result in the vein of Mirzakhani's conjecture that nontrivial homomorphisms between mapping class groups arise only from maps between the underlying surfaces.

Improved Thompson Sampling for Linear Quadratic Problems
Maya Mutic
Mentors: Anima Anandkumar and Sahin Lale

Thompson sampling (TS) is a method of reinforcement learning (RL) in which an estimated distribution is updated according to the reward value of randomly drawn actions. This is used to balance what is known as the "exploration-exploitation" tradeoff in RL. We consider TS in a linear quadratic control setting, in which we attempt to estimate the state parameter matrices *A* and *B* by randomly sampling values that converge to a solution. In order to keep the system stable, we reject values that do not fall within a certain set through a process known as rejection sampling. This project extends upon previous work in rejection sampling algorithms by expanding the set of possible values, which increases the difficulty in converging to a viable solution. To compare to previous results we calculated the probability of our algorithm being optimistic, meaning it favors actions that have a higher reward. We also examine the regret bound for this problem, which was previously theoretically determined to be $O(\sqrt{t})$ but, to the best of our knowledge, never before implemented experimentally. Finally, we examine the applicability of our algorithm in the linear quadratic Gaussian setting.

Fast and Robust Trajectory Optimization for the Rocket Soft-Landing Problem
Daniel Neamati
Mentors: Zachary Manchester Melany Hunt

The rocket soft-landing problem describes the control challenge of safely landing a vehicle on the surface of a solar system body. Previous works losslessly reformulate this problem as a convex second-order cone program. Current trajectory optimizers pair model predictive control with off-the-shelf interior-point solvers. However, augmented lagrangian solvers are better suited to model predictive control applications since the predicted trajectory can warm-start the optimizer. Such an augmented lagrangian trajectory optimizer for second-order cone programs is not currently available. The augmented lagrangian solver presented in this paper is specialized for fast and robust trajectory optimization to tackle the rocket soft-landing problem. This new solver handles second-order cone constraints and is implemented in the Julia programming language to achieve both the speed and robustness needed for real-time rocket landing applications. This software package was tested via Monte Carlo rocket landing simulations and will add to the ALTRO trajectory optimization Julia framework to expand ALTRO's capabilities. The method is also general enough to handle second-order cone constraints in other trajectory optimization applications, such as the friction problem in legged walking-robots.

The Genetic Basis of Chemical Biosynthesis in Aleocharinae Rove Beetles
Jack Nguyen
Mentor: Joseph Parker

Current methods of cell analysis of 3d tissues are tedious and produce inaccurate results with poorly individually segmented cells requiring additional manual curation. Furthermore, in the context of most experiments, this is not practical as data is often on the order of terabytes making the task essentially impossible. Resolving this issue would be of great importance due to the widespread use of cell imaging in most biological experiments to show various phenotypes. Thus, by making automation of the process viable we can expect a decreased need for manual curation which ultimately will lead to increased throughput. For our purposes we will use manually curated 3d images of organoids to serve as training data for a deep learning model to eliminate the inaccuracies most methods face. By continuously training the models with data and curating the results we ultimately should resolve a viable method of imaging. Ultimately, the model should be able to successfully segment cells from images of tissues in 3d and qualify data with increased accuracy and speed when compared to present methods of image analysis.

Development of a Novel Robotic Fish Fin Propulsor for Autonomous Underwater Vehicles
Tyler Nguyen
Mentors: Morteza Gharib and Cecilia Huertas-Cerdeira

The development of Autonomous Underwater Vehicles (AUV's) for use in ocean surveying, military applications, and for exploration of ocean worlds such as Europa and Enceladus, has been spurred on by advancements in miniaturized electronics and computational autonomous decision making. Though most AUV's are propelled via screw-propellers or thrusters, biologically inspired fin-based propulsion has been proposed as a more energy efficient and effective method of propulsion through water. A novel fin propulsor has been constructed that can fit in a streamlined fish body and perform complex 3-dimensional trajectories in water. Based on a custom spherical parallel manipulator, the fin can be aimed in all directions within a 25° cone and can be spun continuously about a fixed center of rotation. During SURF, electronics and sensors were installed and control software was developed to enable functionality of the propulsor for experimentation. Software was also developed to numerically solve the inverse kinematics problem. The fin propulsor will be used for experimentation to optimize fin movement trajectories for thrust, efficiency, and mobility; study the effects of varying fin geometries; and to investigate the fluid mechanics behind fish propulsion. Results from experimentation may demonstrate fin-based propulsion to be a beneficial alternative to conventional AUV propellers.

A Search for Infant Radio Galaxies
Sandra O'Neill
Mentors: Anthony C. Readhead and Sebastian Kiehlmann

At the center of some galaxies there is an active galactic nucleus (AGN), a central engine consisting of a supermassive black hole, accretion disk, dust torus and sometimes, jets. Along these jets, knots, hotspots, or lobes can form. This activity often persists for ~10^7-10^8 years, resulting in a zoology of AGN objects. In this project, we focus on two: compact-symmetric objects (CSOs) and double-double radio galaxies (DDRGs). CSOs are tiny (< 1 kpc), young (~10-10,000 years), and symmetrical—representing the infant stage in the life of the fascinating symmetric objects. Relativistic beaming makes interpreting radio emission from AGN jets difficult. CSOs are unbeamed, providing a laboratory for testing theories of the birth and evolution of AGN jets. However, an unfortunate result of the label floating around in the literature has been its misuse, which has led to much confusion. The aim of this study is to fix this by refining the definition of this class and providing an authoritative catalog of CSOs. This paper also analyzes several mechanisms for DDRG formation. DDRGs exhibit two sets of lobes, suggesting intermittent periods of activity in AGNs. Studying DDRGs and CSOs simultaneously can yield new insight into the early evolution of AGN jets.

Influences of Sand Grain Kinematics During Shearing
Xin Hui Ooi
Mentors: José Andrade and Estéfan Garcia

There are numerous natural hazards involving shear deformation that are relevant to our well-being, such as landslides and earthquakes. In order to better protect civil society from natural hazards, it is important to first understand the fundamental aspects of how these natural disasters occur. In addition, upcoming challenges such as building habitable shelters on other planets, which is one of NASA's goals for its missions to Mars, necessitate a fundamental understanding of naturally occurring granular materials. By understanding the behaviour of granular materials, we can better prepare for the challenges both on our own planet and on others that involve granular materials. The shear rupture process, in which a distinct shear band develops, is a critical component of how granular materials deform. Through analysis of both simulations and data from experimentation of granular materials, we can gain insight into the kinematics (behaviour) of the grains during shear rupture and thus mitigate the effects caused by natural disasters.

SYK & Holography: Study of the Sachdev–Ye–Kitaev Model on Sparse Graphs and Small-World Networks
Patrick Orman
Mentors: John Preskill and Hrant Gharibyan

The Sachdev–Ye–Kitaev (SYK) model, consisting of N Majorana fermions all interacting in groups of 4, is a widely studied toy model in many realms including condensed matter physics, quantum chaos, quantum gravity, and holography. A variation of this model was recently proposed by Swingle in which interactions are randomly deleted, called the Sparse SYK model. Alongside significant computational advantages, one overarching question this development introduces is: What is needed for gravity/chaos to emerge, if all interactions are not? Though Swingle, Susskind, and Witten are currently studying its holographic properties, our aim is to study numerically its chaotic properties, particularly seeing the critical degree of sparseness at which chaotic (and perhaps by proxy holographic) behavior ceases as indicated by changes in its spectral form factor. Computations for several values of N were carried out in order to also determine the scaling of this critical value with the number of Majoranas. Other quantities of interest include nearest-neighbor distributions and out-of-time-order correlators.

Finite Element Analysis of Arthropod and Scorpion Exoskeleton Inspired Composite Microstructures
Lucas L.B. Pabarcius
Mentors: Julia Greer, Haolu (Jane) Zhang, and Matthew Hunt

Arthropod exoskeletons achieve their remarkable stiffness, toughness and versatility through a range of multi-scale composite structures composed of α-chitin embedded in variably cross-linked and mineralized protein matrices, notably the helicoidal Bouligand microstructure and nested sub-helicoids recently discovered in Scorpio Palmatus. To assess the structure-function mechanisms enabling the unique mechanical characteristics observed in exoskeletons across Arthropoda, we work to conduct high fidelity simulations of elastic moduli, stress concentrations and fracture toughness in computer modelled cuticles. We develop and validate an ABAQUS finite element method workflow for evaluating the material properties of direct-modelled bio-composites and produce an automated parametric 3D-modelling script to generate mesh-discretizable cuticles with arbitrary structure and test specimen geometry. We generate 10-500μm three-point bending and tensile test specimens with various structural and mechanical parameters and assess each parameter's effect on bulk mechanical robustness. We aim to further incorporate failure criteria into our testing process and generate additional models to assess more complex natural structures such as pore canals, as well as non-natural structural augmentations. These models and resultant mechanical parametrizations can be validated in future testing of biological specimens, and hopefully inform future bio-inspired engineered and 3D-printed materials with amplified mechanical performance.

Using Machine and Deep Learning Approaches to Improve the Estimation of Social Media Geographic Locations
Chase Pagon
Mentor: R. Michael Alvarez

The legitimacy of democracy depends on the efficacy of any given election. Thus, the legitimacy of any election must always be in question. There is no better group to make sure elections are fair than the people themselves. In the era of modern politics, social media plays a large role in the discussion of politics, essentially becoming the voice for the people. For my project I set out to geolocate twitter data using supervised machine learning in order to understand where problems regarding voter fraud may occur. For training data I used twitter data from a previous research paper and test data provided by professor Alvarez for my proposed machine learning algorithm. I then formatted the labeled training data and put it in the form of chunks of tweets per user pointing to a singular location. My continuation of the project will focus on completing the supervised machine learning algorithm and testing it against the twitter data set on voter fraud that professor Alvarez has.

Automating a Search for Young Planets
Elsa Palumbo
Mentors: Benjamin Montet and Lynne Hillenbrand

Studying planets around young stars is one important step in trying to better understand planetary formation, evolution, and habitability. However, only of a handful of such planets have been discovered so far, partly because young stars are typically quick rotators and very active, making it challenging to separate out planetary transits from the effects of star spots and stellar flares. Our project aims to get around this problem by automating the detection process, from the selection of flux type/aperture and the modeling of each target star's variability with Gaussian process regression to the identification of potential transits and the preliminary vetting of planet candidates. After writing the code and testing it on previously discovered young planetary systems, we begin our own search for planetary systems, with a focus on Young Moving Groups in the 30-minute cadence data of NASA's Transiting Exoplanet Survey Satellite (TESS). And finally, after identifying potential transits in the data, we do more in-depth vetting of each of our planet candidates, with the goal of either identifying the source of the false positive or of confirming our detection.

Decode-Verify: A Tree-to-Tree Neural Network With Model Interpretability
Alexander Pan
Mentors: Anima Anandkumar and Forough Arabshahi

Tree-to-tree neural models have achieved state-of-the-art performance on program synthesis and mathematical reasoning tasks by using an encoder-decoder architecture to learn a distribution between input and output trees. Although these models exploit semantic information expressed in the tree-structured data, they have not fully leveraged its potential. With the goal of greater supervision and model interpretability, we propose adding a pretrained verifier architecture to the encoder-decoder. Given an output tree, the verifier determines the probability a decoding error occurred for every node in the tree. Based on these probabilities, the decoder modifies its distribution and decodes a new tree. This iterative decode-verify procedure runs until the verifier determines the output tree to be correct. The verifier improves model interpretability by highlighting uncertain nodes. We train our model end-to-end on the program completion and differential equation solving domains and compare our model to beam-search baselines.

System Identification of Bacterial Gene Expression as a Function of Population Dynamics
Katherine Pan
Mentors: Richard Murray, Chelsea Hu, and Ayush Pandey

In living organisms, adjustments in gene expression are necessary in response to changes in the environment and affect the cell's physiological state. These changes in expression are often an undesirable complication for synthetic gene circuits. Recent research has shown that genetic circuits are not strictly insulated from their "host" cell, but rather are coupled to the physiological state of the cells as a whole. In this study, *E. coli* were transformed with a yellow fluorescent protein plasmid, grown in various mediums and temperatures, and fluorescence values and population densities were measured over 12 hours. The system was modeled deterministically as a set of ordinary differential equations, and stochastically using Markov Chain Monte Carlo. Parameters for mRNA transcription, protein translation, and mRNA degradation were identified across samples grown in each condition, and the deterministic and stochastic models were combined to allow more precise predictions of a cell's physiological state.

On an Analogue of the Titchmarsh Divisor Problem
Mayank Pandey
Mentor: Maksym Radziwill

Let (lambda(n))$_{n>1}$ be the Fourier coefficients of a holomorphic cusp form. We seek to improve bounds on the shifted convolution sum

$$\sum_{p < X} \text{lambda}(p-1)$$

that were obtained by Pitt. Specifically we seek to improve the saving of X^(1/392) obtained by Pitt. We improve this by improving Pitt's treatment of so-called "Type d_3" sums, in which one is led essentially to bounding sums of the form

$$\sum_{n_j \sim N_j \ (j<=3)} \text{lambda}(n_1 n_2 n_3 - 1)$$

We improve this by using the Kuznetsov trace formula to bound sums of Kloost-erman sums. This is distinct from Pitt's treatment of the same sum, in which he reduces the problem to showing cancellation in a certain exponential sum, which is achieved via appeals to deep results in algebraic geometry. Our technique, which is closely based on one of Bykovski and Vinogradov as implemented by Topacogullari, turns out to yield substantially better Type d_3 bounds, which leads to a significantly better saving of at least X^(1/501)

Detection of Metastable Helium Reveals Ongoing Mass Loss for the Hot Jupiter HAT-P-18b
Kimberly Paragas
Mentors: Heather Knutson and Shreyas Vissapragada

We use the transit method to study gas giant planets that are in very close proximity to their host stars, which are known as "hot Jupiters." Atmospheric mass loss is likely to be significant in sculpting this population of close in planets. By observing the wavelength-dependent decrease in light when the planet passes in front of its host star, we can detect excess absorption from the escaping atmosphere. We use an ultra-narrow band filter centered on the helium (He) 1083 nm transition line to observe transits. In this line, planets' upper atmospheres are opaque, causing an increased transit depth relative to broadband transit observations at nearby wavelengths. The transit depth in this line can be used to map the size of the region containing metastable helium and quantify the corresponding atmospheric escape rate. We observed two transits of the hot Jupiter HAT-P-18b in this line using the 200" Hale Telescope at Palomar Observatory and report the first-ever detection of outflowing gas from its upper atmosphere. This is only the seventh exoplanet with detected helium absorption, and the faintest system for which such a measurement has been made.

Discovering Genetic Signatures of Human Blood Cells
Eunice Park
Mentors: Wei Wei, Xiaowei Yan, and Jared Leadbetter

Discovering detectable markers for blood cells is the first step to isolating cells with abnormal characteristics, which may be circulating tumor cells. Several methods were used to analyze datasets and confirm signature genes of different cell types in human blood samples. The first method used to calculate signature genes includes clustering blood cells based on their RNA-seq expression profiles and running a UMAP (Uniform Manifold Approximation and Projection) algorithm on the data, then overlaying expressions of marker genes calculated for each cluster onto the overall UMAP layout and manually confirming what type of blood cell each cluster may be. The second method is to identify marker genes for each cell type based on Monaco's data set (Monaco et al 2019), using both t-test and fold change as metrics. Comparisons of these marker candidates were also made to confirm their consistency with pre-existing datasets and to check their reliability. A set of characteristic genes for various different cell types in human blood were generated for further analysis of these cells; one possibility would be filtering out negative tumor cells from real circulating tumor cells.

Computational Modeling of Two-Stage and Fluidized Bed Reactors for Carbon Sequestration in Cargo Ships
James Park
Mentors: Jess Adkins and Sijia Dong

In an effort to reduce carbon emissions, a carbon dioxide sequestration reactor is being computationally modeled for implementation in cargo ships. The fundamental design of the reactor involves two steps, dissolving carbon dioxide in seawater and reacting solid calcium carbonate in this water to convert carbon dioxide into carbonate ions. Currently, two different types of reactors are being modeled: two-stage reactors, which have the two steps occur in separate chambers, and fluidized bed reactors, which combine the two steps together in a single chamber. The model for the fluidized bed reactor presents several challenges, due to the complexity of fluid dynamics behavior and phase interactions. Once an initial implementation of the reactor has been completed, additional factors can be accounted for as well, such as physical limitations and engineering constraints.

Applications of Magnetic Co/C Hybrid ROMP-Derived Boronic Acids as Diol Protecting Groups in Polyol Synthesis
Jolly Patro
Mentors: Paul Hanson and Gihan Dissanayake

The development of technologies that provide safe, sustainable, green, and economically viable synthesis has gained attention during the past few decades. Amongst several approaches, the use of immobilized reagents aids synthetic challenges by incorporating facile purification methods into the synthetic route and allowing the potential of recycling the reagent. Co/C surfaces show potential in simplifying the purification process through magnetic decantation. Boronic acids have been widely utilized as diol protecting groups in synthesis. This project demonstrates that employment of Co/C Hybrid ROMP-derived boronic acids (Co/C-OBAn) as a versatile 1,3 and 1,2 diol protecting group. This effective capture-release technique reduces further purifications steps and does not require chromatographic separations. Co/C-OBA_n can be utilized for an efficient and stream-lined synthesis of advanced polyol subunits that relies on an iterative multifunctional aspect of a phosphate tether.

A Fundamentally New Approach to Designing Optical Filters With Random Nanoparticle Films
Elijah Paul
Mentors: Harry Atwater and Parker Wray

Nanoparticles which satisfy the Kerker condition with first-order modes can have strong forward-to-backward scattering ratios in particular wavelength regimes. We study the coupling behavior of these particles for varying radial distributions in 2-D films of randomly-placed nanoparticles through simulations based on generalized Mie theory. Since different-size particles will have these ratios in different wavelength ranges, by optimizing the proportions of particles of particular sizes and materials we demonstrate the possibility of a completely new approach to designing the basic types of optical filters.

Predicting Interannual Sea Surface Temperature Anomaly Using Historical Tide Gauge Records
Joshua Pawlak
Mentors: Mark Merrifield and Ben Hamlington

While sea level has risen steadily over the past several decades at a global scale, considerable deviation from this trend occurs at regional scales. Understanding the mechanisms behind this variation is important to predicting how rising sea levels will affect coastal communities around the world, both directly and through linked changes to other coastal processes. In this project, we attempt to establish and explain a connection between coastal sea level trends and regional sea surface temperature within the Southern California Bight. Using over a century of historical data, we show that tide gauge-measured sea surface height combined with regional wind stress serves as a strong proxy for the dominant mode of sea surface temperature anomaly in the region. This allows us to infer the magnitude and frequency of historical regional marine heatwave events and describe how these events have evolved with the changing climate over the last century. These findings suggest that existing predictive models of sea surface height within Southern California may be easily extended to sea surface temperature, allowing for more accurate forecasts of future marine heatwaves.

Development of Powered Ankle Exoskeleton and Cane-Assisting Device
Toussaint Pegues
Mentor: Aaron Ames

Millions of people worldwide are affected by injuries or disabilities that impede their ability to walk. Devices that restore mobility take many forms, from the simplest canes to exoskeletons. Even with the use of a cane, many people still fall, further injuring themselves. We are developing a device that clips onto a cane and vibrates to warn the user when the cane approaches unsafe sideways angles. The device will also be able to be attached to a user's torso to give immediate posture feedback during physical therapy. Exoskeletons can also provide assistance to individuals with conditions ranging from muscle weaknesses to paraplegia. The goal of our exoskeleton is to provide additional torque at the ankle to reduce the amount of effort users have to expend by supplementing the torque exerted around the ankle during walking or running. This additional torque leads to a reduction in the metabolic cost, allowing the user to walk farther or carry more weight with less effort. The augmentation is achieved using cables attached to a motor mounted on the body to pull the heel during the push-off phase of walking.

Atomistic Identification of Next-Generation Hh Cancer Therapeutics
Daniel S. Peng
Mentor: Alison E. Ondrus

The Hedgehog (HH) signaling pathway is a key regulator of animal embryonic development. Pancreatic, colon, and prostate cancers have been known to utilize aberrant HH signaling to initiate tumorigenesis. Full activity of the Hedgehog (Hh) protein, a key morphogen of the HH pathway, requires an unique post-translational modification in which the protein is cleaved by a cholesterol molecule. As such, small molecules that interfere with this cholesterolysis reaction have the potential to block the activity of Hh oncoproteins.In this project, protein-ligand interactions between full-length Hh protein and cholesterol were analyzed through in silico docking experiments. The resulting docking platform was refined through extensive docking of known cholesterolysis effectors, revealing key properties for cholesterolysis inhibition. Using this platform, a large library of drug and drug-like molecules was screened for viability as a competitive inhibitor of Hh protein cholesterolysis. Top hits were optimized via molecular design to yield a small list of molecules with potential for use as chemotherapeutics against HH signaling cancers.

Mapping Water Ice on the Lunar Surface: Combined Results From Existing Remote Sensing Datasets
Marcos A. Perez
Mentor: Bethany L. Ehlmann

Two decades of remote sensing data from several satellites show that lunar water ice and other volatiles exist within the 30,000 km^2 of the Moon's surface that is permanently shadowed. However, the precise spatial distribution of water ice on the Moon is not yet well understood. Initial radar observations from Arecibo and the Clementine Bistatic Radar Experiment did not reproducibly detect macroscopic water ice deposits at or near the Lunar surface. Later, infrared spectra directly detected surface water ice in select craters using spectroscopy in UV, near-infrared albedo, and temperature also consistent with ice. Additionally, the depth-diameter ratios of simple craters on the Moon, shallow significantly with latitude, suggesting that macroscopic deposits of volatiles in-fill craters on the Lunar surface.

To create a comprehensive map of water ice indicators on the Lunar surface, datasets collected by the following instruments were utilized: M^3 (infrared spectra), LAMP (far-ultraviolet spectra), LEND and NS (neutron spectra), LOLA (topography and albedo), both Mini-RFs (albedo and polarization), LROC (surface visible albedo), and Diviner (surface temperature). Although the raw data for each of these instruments are publicly accessible from NASA's Planetary Data System Geoscience Node, we utilize previously published derived products when possible to efficiently build upon the data processing already conducted.

Measurements and observed surface conditions consistent with the presence of water ice were spatially coregistered in ArcGIS and their correlations tested for statistical significance using Python. Since these instruments possess different spatial resolutions and different depths below the surface at which they can detect water ice, we aim to differentiate areas with water ice at the surface versus buried below the Lunar surface.
For future robotic missions to study the Moon's current and past geology and potential in situ resource utilization (ISRU), ascertaining with certainty the composition, location, and abundance of the water ice deposits on the Lunar surface is crucial.

In Silico Circuit Evolution
Joseph Peterson
Mentors: Richard Murray and Rory Williams

In vivo, it is observed that subjects of genetic engineering often experience selective pressure to remove or inactivate the synthetic circuit they have been programmed to express. This is because the new genetic circuit requires usage of resources and machinery that would otherwise have been contributing to cell growth. In order to provide a platform for researchers to easily explore this important aspect of engineered circuits, we have designed an extension of the biological circuit simulation software Bioscrape that allows the growth rate of cells to be modulated by circuit expression through the sequestration of cellular machinery, and random mutations to be modeled through adjusting rate constants. Taken together, this enables the simulation and testing of any genetic circuit against this selective pressure such that genes/architectures which are especially unstable may be identified, and variants aimed to improve the evolutionary half-life tested. Future projects should focus on improving the fidelity of the growth-oriented metabolism in the simulated cells.

Simulation of Fuel Transfer for Applications in In-Orbit Spacecraft Servicing
Martin Peticco
Mentors: Soon-Jo Chung and Yashwanth Kumar Nakka

Satellites developed over the past few decades have had their functionality greatly restricted by their notably short lifespans of around five to fifteen years, often attributed to their limited fuel capacity. As a result, satellites that are otherwise functional get retired and thrown into graveyard orbits to then be replaced by brand new satellites. This is incredibly time consuming, costly, and less than ideal. The aim of this project is to develop a proof-of-concept system that can demonstrate the extension of a satellite's lifespan through in-orbit refueling. This system will be designed around the spacecraft simulators developed at the Center for Autonomous Systems and Technologies lab at Caltech. A compatible fuel arm is designed in SOLIDWORKS, then an expression for a time-dependent moment of inertia tensor is derived based off the flow of fuel from one spacecraft to the other and simulated in MATLAB. This, and similar simulations, will allow for the spacecraft controller boards to take the changing moment of inertia into account for precise movement control. Further work would consist of building proof-of-concept systems for other areas of autonomous servicing, such as replacing parts on a satellite or adding components to it to expand its capabilities.

Investigation of ETF Pairs and Implications in ETF Pairs Trading
Jack Pierson
Mentors: Vladimir Cherkassky and R. Michael Alvarez

Exchange traded funds (ETFs) are a type of investment fund that are traded on stock exchanges. The objective of this project is to understand underlying relationships between pairs of ETFs based on mathematical phenomena. The first of these is cointegration, which occurs when a stationary relationship exists between two non stationary time series. In order to determine whether two time series are cointegrated, a regression is performed on a pair of ETFs over a given time frame and an augmented Dickey-Fuller test is used to test for stationarity in the residuals obtained by the regression. It was found that though the tested ETFs do exhibit cointegration over certain time periods and especially long time frames, there is no consistency that would indicate any short-term trading strategies. The second phenomena is correlation. Certain ETF pairs exhibit particularly strong correlation. Because of this correlation, one can expect some sort of correction mechanism to days which stray far from the correlation fit line. The rest of the project will revolve around exploring how a correction mechanism could be taken advantage of and if there is also a lagged correction between US and international markets.

EUV-Net: Predicting Extreme Ultraviolet Solar Emission From He I Absorption Lines Using Deep Learning
Anthony Pineci
Mentors: Peter Sadowski and Christopher Umans

Planet formation and evolution is impacted by radiation in the extreme ultraviolet (EUV) and far ultraviolet (FUV) from the host star. This high-energy radiation is only observable from space, yet remains partially obscured by the interstellar medium. This motivates a calibration between a ground-observable proxy to EUV and FUV emissions. In this work we predict EUV emission from contemporaneous observations of HE I absorption, using deep learning to learn this non-linear relationship from historical data acquired by SOLIS instruments operated by NISP/NSO/AURA/NSF and data provided by SDO and processed by JSOC. The model is a physics-informed fully convolutional neural network with skip connections that uses spatial information and accounts for the curvature of the Sun. Using normalized target values, results indicate an average pixel wise median squared error of 1.53% and a mean disk-integrated flux error of 6.43%. Qualitatively, the model learns the correlation between He I absorption and EUV emission, and is able identify high-absorption filaments that do not result in EUV emission.

Linking Neural Activity and Synaptic Clusters of Oscillatory Behaviors in the Leech
Amanda Piyapanee
Mentors: Daniel Wagenaar and Pegah Kassraian-Fard

To understand the computations of a neural circuit, it is necessary to reconstruct the connectivity between the neurons while simultaneously recording their activity. To this end, utilizing modern techniques, we combined scanning electron microscopy (SBEM) and voltage-sensitive dyes (VSDs) imaging in one ganglion of the medicinal leech. The VSD data was collected during three distinct behaviors: swimming, crawling, and local bending. Specifically, we focused on one output motor neuron, the dorsal excitor motor neuron (DE-3R) and all of its presynaptic partners. DE-3R is of particular interest due to its involvement in a multitude of different behaviors. In this project, VSD and anatomical data were linked by assessing the functional relevance of anatomical properties such as synaptic location and clustering. In addition, neurite morphology was analyzed to understand the growth process underlying the neurites' development. After an automated cleaning of the manually traced connectivity data, we aim to apply a machine learning model to predict DE-3R output activity based on the combination of presynaptic partners and the synaptic clustering.

Developing a Gerdien Condenser for Atmospheric Ion Measurement

Geoffrey Pomraning
Mentor: Paul Bellan

One problem in plasma physics is the measurement of atmospheric-pressure plasmas. Although conventional instruments are unable to characterize these plasmas, recent research has shown that Gerdien condenser-inspired devices can be effective for determining the ion mobility and density of an atmospheric pressure plasma. This presents opportunities for applications of Gerdien condensers to measure characteristics of atmospheric plasmas in general. For this project, we developed a Gerdien condenser to determine air ion density, with potential future applications to plasma diagnostics and measurements. We developed a ventilated cylindrical capacitor which induced an electrical potential and a continuous airflow for the condenser. Preliminary results indicate that a small current, likely induced by air ions being repelled from the outer electrode and contacting the central electrode, was detected. Further analysis is being undertaken to verify the accuracy of this signal as a measurement of air ion concentration.

Creating New Features for COVID-19 Case, Death, and Hospitalisation Predictions

Max Popken
Mentor: Yaser Abu-Mostafa

A key issue modelling the spread of the COVID-19 pandemic is the lack of reliable and informative data. The purpose of this project is to create new proxies to be used in these models. The main area of focus is the total number of new infections per day; this differs significantly from the total reported cases because of testing shortages, asymptomatic cases, and time delay between infection and positive test. At this point, we have found the probability of a transmission, hospitalisation, and death occurring at any time after initial infection. This makes it possible to find new infections as a function of previous infections once some inherent parameters are defined. These parameters are the effect of mitigation strategies on case growth; the amount of spread between vs among susceptible and non-susceptible population; and the baseline rate of spread of the virus. Determining these parameters is the current area of research; their discovery would allow for a reliable estimate of the total number of infections – and by extension the total number of hospitalisation and deaths – over time.

Algorithms to Identify Oxysterol Binding Pockets in Proteins

Sarida Pratuangtham
Mentors: Alison Ondrus and Yu-Shiuan Cheng

Oxysterols are relevant to a number of biological processes, including physiological processes, such as cholesterol synthesis; diseases and pathologies, such as cancer, and development, such as in the Hedgehog pathway. However, the functional protein targets of oxysterol signaling *in vivo* are largely uncharacterized. In order to predict novel targets of 20(S)-hydroxycholesterol, a hydroxylated cholesterol metabolite with potent activities *in vivo*, we explored computational methods for predicting protein-ligand interactions. Machine learning algorithms were examined for their favorable properties of high speed of computation and large set of proteins tested during development. However, an example of a machine learning algorithm designed for this analysis, SSNet, did not reliably predict protein-ligand binding for a small set of test proteins with known ligands. We therefore tested a number of web servers designed to predict protein-ligand binding, including SwissTargetPrediction, Pose&Rank, and AutoDock Vina, due to their favorable property of high speed and volume of computation. Each of these algorithms calculates likelihood of ligand interaction by different metrics; they are based on similarity to known ligands, atomic distance from ligand to protein, and steric/hydrophobic interactions, respectively. By comparing the results of these servers against one another, we were able to compare their reliability in predicting interactions between known protein-ligand pairs. Future work will focus on curating a list of likely 20(S)-hydroxycholesterol targets based on these algorithms, and refining it against a list of protein targets detected from experimental crosslinking/mass spectrometry data.

Searching for Magnetic Waves in Historical Astronomical Images

Anna Preuss
Mentors: Anthony Readhead and Georgia Panopoulou

Organized, linear features (striations), have been observed in nearby regions of the Interstellar Medium (ISM). The striations appear to be aligned with the magnetic field within interstellar clouds and are most likely caused by the excitation of magnetosonic waves. In theory, the striations should move over time as the waves propagate; however, their motion has not yet been detected. We will determine if shifts in the striations are measurable over a timespan of 100 years and, if possible, measure the changes to gather data on the waveform. A visualization of the striations was created in Python using estimated values and later two images of the Pleiades, one old and one recent, were overlaid to measure the changes. Limits on the shift in the striations provide the first direct test of the theory of these magnetosonic waves.

Flat Bands in Honeycomb Superstructure
Zihao Qi
Mentor: Yi Li and Jason Alicea

Flat bands on a honeycomb lattice have provided useful platforms to study strong correlation and topological physics. Honeycomb superlattices of s-orbitals have been discovered to host multiple flat bands. We study the origin of the robustness of such flat bands in the presence of local perturbations respecting lattice symmetry. Analytically, we constructed compactly supported localized states (LS) as the eigenstates of flat bands due to interference hopping in real space. We also pointed out the topological loop states responsible for flat bands. The numerical result exactly agreed with our construction. Our results could be used to engineer flat bands in real materials, a task that has been difficult so far, and to construct interaction-dominant phases of matter.

Using Toehold-Mediated Strand Displacement to Construct Oscillations in a Magnetic Panel System
Kavya Rajagopalan
Mentors: Michael Brenner, Erik Winfree, Chrisy Xiyu Du, Ofer Kimchi, and Agnese Curatolo

Due to the vast applicability of nanoscale devices, DNA nanotechnology has become a field of growing interest among various scientific disciplines. However, it has been recently discovered that DNA Origami, one such mechanism often used to construct nanoscale devices is limited. In order to develop a robust alternative to DNA Origami, we will be simulating the behavior of DNA oscillators in a magnetic panel system. To meet this end, we will examine the key mechanism in DNA oscillators: "Toehold Mediated Strand Displacement." Using ample knowledge of this process from Srinivas *et al 2013*, we have successfully constructed a model of strand displacement rates as a function of "toehold" length and binding energies. Likewise, we were able to construct plots of the Rock Paper Scissor Oscillator, a fundamental oscillator used to construct more complex oscillators, at different reactions rates. In analyzing these two models, we are and will be able to further translate and predict the oscillatory behavior of a magnetic system by simply selecting for specific and desired binding energies.

A Comparison of Accumulator-Based and Register-Based RISC Architectures
Malia Rebollo
Mentor: Glen George

Accumulator-based and register-based CPU architectures and instructions sets were investigated to compare execution speed and implementation and programming efficiency. In order to make a fair comparison, both architectures were written in Verilog with similar instruction sets. The architectures were then tested for functionality in simulation. In addition, a sample program was written using each instruction set to compare typical program sizes. Finally, a printed circuit board with an FPGA was designed to allow for comparisons in real-world applications.

A Small Primes p-Converse to a Theorem of Gross-Zagier, Kolyvagin, and Rubin
Jacob Craig
Mentor: Ashay Burungale

Let E be a CM elliptic curve over the rationals and p a good ordinary prime for E. In the case $r = 0,1$, we show that
$$\crk_{\ZZ_{p}}\sel_{p^\infty}(E_{/\QQ} = r \implies \ord_{s=1}L(s,E_{/\QQ}) = r$$
for the p^{∞}-Selmer group $\sel_{p^{\infty}}(E_{/\QQ})$ and the complex L-function $L(s,E_{/\QQ})$. In paticular, the Tate-Shafarevich group $\Sha(E_{/\QQ})$ is finite whenever $\crk_{\ZZ_p}\sel_{p^\infty}(E_{/\QQ}) = 1$. This completes proofs of the p-converse theorem by Rubin and Burungale, Tian for the small primes $p = 2,3$.

NTopology Software and Mid-Sole Project
Darren Rhodes
Mentors: Hongbing Lu and Guruswami Ravichandran

The objective of this research is to find an optimal design for the Mid-sole of a shoe. Taking an existing design for a mid-sole we used the same dimensions and created a CAD model for the Mid-Sole. Instead of making the mid-sole as one part, the mid-sole was split into three parts from top to bottom. The mid-sole was designed using SOLIDWORKS, then the middle part was imported to nTopology and made into a lattice structure. Based on material simulations from nTopology, the lattice structure will be edited to strengthen the weak spots. The project is split up so that everyone has unique designs. Once each of us complete the optimal structures for each of our designs, the designs will be compared. We currently do not have results yet, but I would suggest that after the different designs are compared, we should take the best aspects of each design and create a new model for the mid-sole.

Tracking Meandering Streaks in Wall-Bounded Turbulence
Sydney Richardson
Mentors: Beverley McKeon and Jane Bae

Turbulence is a natural phenomenon where fluid moves in irregular and chaotic motions, which can be observed in everyday life as well as in practical engineering applications. Turbulent flow close to a solid wall tends to show some signs of a pattern, in specific, the formation and breakdown of quasi-periodic three-step cyclic streaks. The main focus of this project is to analyze a time-resolved three-dimensional data set of a low Reynolds number turbulent channel flow, identifying those areas of high turbulence intensity. We then quantify defining characteristics of the movement of these streaks such as the time-evolution of streak size and meander ratio and study how they are generated and broken up.

Determining the Effect of COVID-19-Asssociated Cytokines on the Drug Responses of Immune Cells
Philippa Richter
Mentor: Matthew Thomson

In patients suffering from severe cases of COVID-19, an immunological phenomenon called a 'cytokine storm' has been observed. A cytokine storm is an overreaction of the immune system to a pathogen; an overproduction of cytokines triggers excess inflammation, which, if severe enough, causes cell death in the affected tissue. Currently, a variety of drugs are being tested in patients for their ability to control such an overactive immune response. However, it is not currently known how the transcriptional states induced in cells by COVID-19 may affect drug function. In this study, we systematically evaluate the effects of COVID-19-associated cytokines on human immune cells. We then expose activated immune cells to several immuno-suppressing drugs and assess their impact on gene expression. From these experiments, large quantities of high-dimensional single-cell mRNA expression data were collected and visualized, revealing several trends in the expression data; for example, IL-6 generally had little impact on transcriptional state, TNFA and IFNG had a large impact on myeloid cells, and IL-7 had a large impact on T-cells. From the drug trial, we found that CD3 seems to have a significant activating effect on both T-cells and myeloid cells, while a combination of other cytokines only had a significant impact on myeloid cells. The transcriptional response of myeloid cells to each pairwise drug-signal combination varies greatly. In T-cells, when CD3 was used as an activating signal, cyclosporine appears to be the most effective immunosuppressant tested; however, in samples exposed to other cytokines, cyclosporine has little effect. These results highlight the importance of testing drugs in the context of different transcriptional states, which will allow a more accurate assessment of a drug's efficacy in treating a patient afflicted by a disease such as COVID-19.

Comparing Reaction Mechanism for Nitrogen Reduction on Different Ruthenium Catalyst Surfaces
Michael Rose
Mentor: William A. Goddard, III

The Haber-Bosch process is the primary means of artificial ammonia production; however, its drawbacks include the intense temperature (400 - 500 ℃) and pressure (100 - 200 atm) under which the reaction occurs favorably, making it energy intensive and requiring large infrastructure. Electrochemical reduction, as opposed to this thermal reduction, would potentially solve these problems; however, currently the catalysts for an electrochemical alternative are not competitive. To this end, we investigated potentially improving the catalyst by using different crystal surfaces of the ruthenium catalyst: (0001), (1-100), and (2-1-10). The reaction mechanism of nitrogen reduction to ammonia was then modeled out on these surfaces and the energetics for these cases were calculated using Vienna Ab initio Simulation Package. While the initial calculations to outline these energetics have been completed, more work in the coming weeks is required before these mechanisms are fully understood and potential improvements to the catalyst are found.

An FPGA Architecture for a Kalman Filter-Based Neural Decoder
Joshua Rosenberg
Mentors: Azita Emami and Sahil Shah

Brain-machine interfaces (BMIs) require an efficient means to decode the user's desired instructions from neural signals. To perform this decoding, current studies often use neural networks, shown to yield high accuracy. However, these algorithms require significant computational complexity, a barrier to creating practical, small-scale, and low-power implementations. Requiring fewer parameters for computation than neural networks, the Kalman filter, a linear algorithm, offers a potential alternative. This project expands upon existing research on the Kalman filter as a neural decoder by investigating its effectiveness in hardware via a field-programmable gate array (FPGA) implementation. We offer a design that follows a modular approach, dividing the algorithm into computational blocks comprising distinct matrix operations. To continue this work, subsequent steps should optimize the design to minimize area and power consumption and test it in real time to evaluate the Kalman filter's effectiveness. Additionally, future work should investigate on-chip learning, which would eliminate the need to update model parameters externally. If the Kalman filter proves viable as an accurate and efficient neural decoder in hardware, this advance would provide an important step toward the greater goal of achieving high-performance, real-time neural decoding using an implantable application-specific integrated circuit (ASIC).

Techniques of Local Equivalence: Understanding Knot Concordance and Homology Cobordism
Daniel Rostovtsev
Mentors: Ian Zemke and Yi Ni

Understanding the structures of the homology cobordism group and knot concordance group are major open problems in topology. Up to local equivalence, elements of the homology cobordism group can be mapped to iota complexes, and elements of the knot concordance group can be mapped to knot complexes. Understanding the algebraic structure of these complexes up to local equivalence allows us to find infinite free summands in the homology cobordism group and in the knot concordance group. However, the classification of possible complexes up to local equivalence is incomplete. Here, we present new partial results in the classification of these complexes and describe a potential way to realize the classification problem in terms of Fukaya categories.

Modeling a Glucose Metabolic Pathway and an ATP Synthase Mechanism Shows ATP Life Extension in Synthetic Cells
Ankita Roychoudhury
Mentor: Richard Murray

In synthetic cell protein synthesis, a common limiting factor is the energy supply for transcription and translation. By studying computational and mathematical models of various ATP regeneration mechanisms in synthetic cells, we aim to propose experimental methods for ATP life extension. We use available software tools to study two models. These allow us to develop and study mass action models by implementing simple chemical reaction networks. Our simulations show that a glucose metabolic pathway is able to extend lifetime of ATP up to about 60 hours. Integrating ATP synthase can also independently lengthen the lifetime of ATP to various times depending on the implemented proton gradient mechanism. To ensure prolonged synthetic cell protein synthesis, either the glucose pathway or ATP synthase mechanism can be used. In the future, it will be useful to perform wet-lab experiments in order to compare our model to data.

Reinforcement Learning and Investor Behavior in Financial Markets
Rafael Santiago
Mentor: Lawrence Jin

Reinforcement learning (RL) - a type of machine learning where an agent learns from the environment - has been extensively studied in different fields such as computer science and neuroscience. The project aims to understand whether a subset of RL algorithms, those with strong neural foundations, indeed help explain investor behavior in financial markets. Convergence is tested in both model-free and model-based learning algorithms by comparing simulated valuations of various sample periods with their analytical counterparts. Additionally, regression analyses of final wealth allocation on past market returns, past portfolio returns, or past actions are performed. Although convergence of valuations usually happens in other fields, it only happens in financial settings under special circumstances. Furthermore, regression results show that the dependence of final wealth allocation on past returns and actions differs significantly between model-free and model-based learnings. Our findings provide a novel explanation for the quantitative differences documented empirically between the literature on extrapolative beliefs and the literature on experience effect.

Characteristics and Impacts of Far Side Lobe Features of the BICEP Array 1 Telescope
Eve Schoen
Mentors: Jamie Bock and Ritoban Basu Thakur

The physics of inflation is a central problem of modern cosmology. Measuring B-type polarization of the Cosmic Microwave Background (CMB) can provide direct evidence of inflation. The BICEP-Array (BA) telescopes will measure the CMB in 4 frequency bands to characterize foregrounds and signal (B-modes). Inflationary B-modes are expected to be faint (~1:10^8 in temperature variance units). Thus, causes of spurious polarization in BA must be measured and understood precisely. In this paper we focus on far side lobe (FSL) effects, resulting in temperature-to-polarization leakage. We analyze FSL data collected at the South Pole with the first BA telescope to make data-driven models of the large-angle FSL polarized beams. Using these, we simulate temperature to polarization leakage to understand systematic polarization contamination. To model temperature to polarization leakage, a simplified model of beam difference was convolved with a model of the ground shield and galaxy from PLANCK data. The first half of the paper details how the FSL maps were constructed using a polarized K-band source at South Pole. The second half outlines the temperature to polarization leakage calculations that are being pursued. We expect to quantify the level of spurious B-mode power that will be generated by the temperature differentials seen between the CMB, the telescope ground-shield and the galaxy due to the polarized FSL beams. Techniques developed here informs systematics control studies for the BA program.

Microcontroller Approach to Mossbauer Spectroscopy
Louise Schul
Mentors: Brent Fultz and Cullen Quine

Mossbauer spectroscopy has been used to characterize iron compounds in robotic space missions including the Mars Exploration Rovers, but upcoming lunar missions require smaller and lighter spectrometers. This project uses a Teensy 3.2 microcontroller along with analog circuitry to process and record pulses from a Mossbauer detector. Pulses are sorted both by amplitude and time of detection, for x-ray fluorescence and Mossbauer analysis, respectively. The software uses interrupt routines to efficiently handle detected pulses, to precisely control the velocity of a piezo involved in the detector, and to periodically output stored pulses to a computer. In place of a detector, a 555 timer circuit was used to generate 2 μs pulses, which can be processed by the microcontroller in around 10 μs. Microcontrollers are thus a realistic approach to Mossbauer spectrometry, although require additional testing with a Mossbauer detector.

Time-Series High-Resolution Spectroscopic Variations of Outbursting Pre-Main Sequence Stars
Jerome Seebeck
Mentor: Lynne Hillenbrand

We present a high-resolution spectral analysis of two variable stars: HBC 722 and V960 Mon. These objects were identified as FU Orionis stars: variable stars that have undergone extreme changes in magnitude caused by an increase in accretion rate. Observations were taken post-outburst at optical wavelengths (.3μm-1μm) with the Keck High-Resolution Echelle Spectrometer and Lick Automated Planet Finder between 2010 to 2017. In both stars, we have found evidence of decreasing wind-driven accretion through P Cygni profiles, and evolving disk features through double peaked absorption profiles. The strength of disk lines has been shown to increase in V960 Mon but not HBC 722. Measurements of half-width half-depth (HWHD) do not show any significant change with respect to wavelength (79km/s [HBC 722] and 45km/s [V960 Mon]), which is in conflict with the standard Keplerian rotation disk model. Several low excitation potential broad absorption lines with additional central dips blueshifted by 5 km/s and HWHDs of 76 km/s were identified in HBC 722 with no current explanation. Decreasing strength of high excitation potential lines, wind velocities, and source brightness have led to the interpretation of V960 Mon as a hotter and quicker cooling star than HBC 722.

3D Object Encoding in Deep Neural Networks and the Brain
Aditi Seetharaman
Mentors: Doris Tsao and Erin Koch

The two-stream hypothesis is a model of the visual processing system. The dorsal visual stream, which originates in the occipital cortex and leads to the temporal cortex, is primarily responsible for encoding spatial information, including three-dimensional structure. In previous research, the results from a feedforward deep neural network trained on object classification served as a framework for how neurons in the inferotemporal cortex are organized. In our research, we use a variational autoencoder generative adversarial network to develop a framework for spatial information processing along the dorsal stream. For this purpose, we focus on the encoder, which takes a two-dimensional image as an input and returns a latent representation vector. By analyzing the images fed to the encoder and their corresponding representation vectors, we propose a hypothesis for how the primate brain encodes three-dimensional structure. We expect future research to test the implications of our hypothesis through experimental methods.

Conditional Value-at-Risk Constrained Optimization
Anish Senapati
Mentor: Jose Blanchet

In this project, we consider chance constrained optimization problems with heavy-tailed distributions within the risk factors. The usual chance constrained optimization problem with value-at-risk constrains has limitations in modeling and tractability motivating a conditional-value-at-risk (CVaR) risk measure which prevails in many real-world applications. In this project, we aim to look at the generic CVaR constrained optimization problem. We transform the optimization problem into its corresponding Lagrangian relaxation problem and design an algorithm to solve this relaxation minimization using stochastic gradient descent. To find the optimal solution as the chance constraints become tighter, rare event simulation techniques were developed improve our algorithm efficiency in Monte Carlo evaluation required beyond that of a simple Monte Carlo approach.
Anish Senapati

Smart Error Messages for Data Structure Implementations via Codified ADTs
Archie Shahidullah
Mentor: Adam Blank

CS 2 is Caltech's course that teaches fundamental data structures, which are implemented in the Java programming language. It is absolutely crucial that students completely understand how these data structures function, and to this end, robust test modules are necessary. My project is two-pronged. The first objective is to identify current lackings in existing test modules, and the second is to write new modules to ensure students' implementations match with abstract data types (ADTs). In pursuit of the first, all class projects were re-implemented and errors encountered were carefully documented. This serves as both a reference for future course staff as to what issues to expect when assisting students, and as my own reference for which tests to improve/add. The latter involved using the Java Reflection library to access the internal state of classes and provide specific error messages detailing functional bugs. This allows the invocation of methods and analysis of attributes such as visibility, inheritance, adherence to specified constraints, etc. We expect these improved test modules will improve student learning, and to this end, student progress will be monitored and studied as part of my CS 81 project next academic year.

Development of a Powered Ankle Exoskeleton
Lorenzo Shaikewitz
Mentors: Aaron Ames and Maegan Tucker

Powered exoskeletons can enhance the human ability to carry load, improve metabolic efficiency, and help the injured to a quick recovery. However, they rely on control algorithms that are difficult to test on systems with many degrees of freedom. We designed an autonomous ankle exoskeleton to reduce the metabolic cost of walking and provide a one-degree-of-freedom platform for controls testing. The lightweight device actuates two moment arms attached to the foot to generate a mechanical torque during the transition from stance to swing in the walking gait. The flexible design allows electronics to be concentrated at the waist to maximize efficiency or at the shin to minimize complexity. Basic controls simulations show it can reduce biological ankle torque by as much as 40%, making use of a sleek angled strut design. Once constructed, this exoskeleton will provide a platform for complex controls to optimize its own function and aid in the development of larger exoskeletons.

Construction of Hecke Characters for Three-Dimensional CM Abelian Varieties
Zhengyuan Shang
Mentor: Matthias Flach

It is well-known for an elliptic curve with complex multiplication that the existence of a Q-rational model is equivalent to its field of moduli equal to Q, or its endomorphism ring equal to the ring of integers of 9 possible fields. Murabayashi and Umegaki established similar results for CM abelian surfaces. For 3-dimensional CM abelian varieties with rational fields of moduli, Chun narrowed down to a list of 37 possible choices of the CM fields and obtained a partial converse. In this project, we attempt to prove that these abelian varieties have Q-models by constructing certain Hecke characters and thus complete the analogy.

Determining Outcomes of Cooperative Games With and Without Externalities
Jason Shi
Mentors: Kevin Tang and Federico Echenique

The study of solution concepts for cooperative games has mainly focused on outcomes that are desirable by certain metrics, such as fairness and efficiency. However, since players in such games seek not to satisfy such abstract norms, but rather to maximize their own gain, alternative solution concepts are needed to predict the actual outcomes of games. We build a program to simulate games and compute their outcomes. Then we use the results generated by this program to form, test, and refute theoretical conjectures about game outcomes; for instance, we prove that all 3-player games do in fact converge on an outcome after a finite number of steps, while showing that this is not true for general games.

Structural Szémeredi-Trotter Theorem for Lattices
Olivine Silier
Mentors: Adam Sheffer and Nets Katz

Incidence problems provide a framework for characterizing an underlying geometry and find applications beyond discrete geometry, spanning combinatorics, number theory, computer science, and harmonic analysis. A point and a line form an *incidence* if the point is on the line. When |P| = |L| = n, the Szemerédi-Trotter theorem states that the number of incidences between points from the set P and lines from the set L is $O(n^{4/3})$. The theorem is tight since there exist configurations with $\Theta(n^{4/3})$ incidences. Only two such configurations were known, one from Erdős, the other from Elekes. In this work, we find a family of constructions (including these two) that spans all maximum-incidence constructions with a lattice of points. Moreover, while the Szemerédi-Trotter theorem has been known for nearly four decades, hardly anything is known about the *structural problem*: characterizing the configurations with $\Theta(n^{4/3})$ incidences. Here, we use an energy variant to derive a tight point-energy bound which depends on the geometry of the configuration. We also derive a variety of structural properties where the point set is a Cartesian product.

Guaranteeing Safe Behavior for Self-Driving Vehicles
Giovanna Silveira Amorim
Mentors: Richard Murray and Karena Cai

Previous works have formulated a mathematical framework for self-driving vehicles which guarantees safety through assume-guarantee profiles. This design assumes that each vehicle has complete information on other cars. This project migrates the original framework from a deterministic environment to a probabilistic one. Characterizing the original framework's robustness to errors allowed for the design of a new framework capable of dealing with uncertainties caused by lack of information on other vehicles' dynamics, profiles, and final destination. In this more realistic approach, the new system can provide safe operation with only partial information.

Quantum Error Correcting Codes for Large Spins
Aditya Sivakumar
Mentors: John Preskill and Victor Albert

The field of quantum information science uses quantum mechanical phenomena (such as entanglement) to perform computational tasks. Quantum information can be encoded using any quantum-mechanical physical system. Unfortunately, many of these systems tend to be easily damaged by external noise sources. Nevertheless, for many cases, there are general methods one can use to protect all information encoded in a quantum system against a given type of noise. This research project investigates the necessary and sufficient conditions for which such a general method of error correction can successfully be applied to a specific type of quantum system: a quantum rigid body (e.g. a molecule) suffering from noise which alters its angular momentum state. It turns out this situation is mathematically very similar to a different type of quantum system: a quantum simple harmonic oscillator (SHO) suffering from noise which alters the energy state of the system. For the SHO system, there is already a well-established method to protect any encoded information against the energy noise. By analyzing this method, this research project developed an analogous method for a quantum rigid body system, which protects any encoded information against the angular momentum noise.

Analyzing Straylight X-ray Binary Candidates With NuSTAR
Catherine Slaughter
Mentors: Fiona Harrison and Brian Grefenstette

The Nuclear Spectroscopic Telescope ARray (NuSTAR) is the first space telescope to observe in the hard X-ray bandpass (3-79 keV). NuSTAR was designed to have a deploy-able mast that is open to the space around it. This means that stray photons from objects that lie 1-4° from the line of sight can fall on the detectors of the scope without first going through the optics. Many of these serendipitous straylight observations have been taken over the past eight years, but have not yet been used for research purposes, despite containing a potential wealth of useful data. This project is two-fold. First, I loosely recreate the results of a Homan et al. 2018 paper, which looks at source GX5-1 with focused NuSTAR observations. This serves as a proof-of-concept for the efficacy of straylight research and gives us the opportunity to test the data reduction code written by Brian Grefenstette. Secondly, I perform new time series and spectral analysis on Z-source GX17+2. In particular, use the shape and spread of an iron-line spectral reflection feature (around 6.4-6.97 keV) to determine values for the rate of rotation and innermost stable circular orbit of the neutron star binary system's accretion disk.

Increasing Efficiency of Memetic Algorithms
Julia Sloan
Mentors: Pablo Moscato and S. George Djorgovski

Machine learning, a subset of artificial intelligence, is an unparalleled tool allowing computer scientists to process and analyze vast quantities of data. Memetic algorithms are an advanced type of machine learning methodology to find solutions of computational problems, employing optimization methods that work in a cooperative and competitive way. The Moscato group's current memetic algorithm is extremely accurate, but is very costly in terms of time. Testing the algorithm on a variety of datasets after any changes is essential to guarantee that it is applicable in a wide variety of implementations. Datasets involving the folded structure of a protein, correlation of body fat percentage with BMI, and the critical temperature of a superconductor are investigated here. With some data preprocessing, we find our algorithm to be largely effective in each of these domains, and these datasets help provide a range of data to study algorithm performance.

Identifying Novel dCas9 Guide RNAs to Target Repetitive Regions of DNA for Optical Barcodes
Whitney Sloneker
Mentors: David Van Valen, Edward Pao, Geneva Miller, Morgan Schwartz, and Emily Laubscher

While microscopy is useful to visualize the effects of knock-down genes, there are throughput challenges. The process is time consuming, as the number of samples that can be visualized at once are limited. This project aims to address the challenges through optical barcoding using dCas9 CRISPR display. dCas9 can bind to repeating regions of the genome, labelling them with fluorescent proteins. We aim to determine novel binding sites for dCas9 to add to our existing library of optical barcodes. These sequences must be repetitive to generate adequate fluorescence signal. They must also be adjacent to PAM sites – a GG or CC sequence. Finally, structural regions of the genome are preferable to ensure the labelling does not interfere with the genes required for cell viability. We used Biopython, a bioinformatics Python package, to determine the presence of different repeating regions proximal to PAM sites to find the optimal dCas9 target given the necessary features. However, while potential targets have been observed, wet lab testing is still necessary to confirm results.

Effects of Visual Saliency in Human Behavior
Nina Solovyeva
Mentors: Colin Camerer and Xiaomin Li

Visual saliency is a powerful tool that can be used in the prediction of human behavior. To further explore it, we designed an online behavioral experiment asking participants to make a choice between two piles of fruits presented in an image. Each fruit had an established value that was made known to the participants and the pile worth was the summation of the individual fruits' values in it. SAM (Saliency Attention Model) was used to predict which pile (right or left) was more salient based on the images we generated. Twenty images were presented to each subject, with 10 being congruent (more salient pile had a higher worth) and 10 being incongruent (less salient pile had a higher worth). The results showed that the pile with higher worth was chosen more often amongst congruent images rather than incongruent. Thus, people made more mistakes when saliency was conflicting with the reward attribute. Saliency and valuation were found significant through logistic regression in regards to the choice made.

Developing a Pipeline for the Measurement of Signaling Dynamics and Patterns of Gene Expression Upon Introduction of Pooled Barcoded Knock-Downs
Johnathon Soro
Mentors: David Van Valen, Emily Laubscher, Edward Pao, and Morgan Schwartz

In order to study the effects of a genetic knock-down and understand a gene's function, a single gene is typically knocked out in a culture of cells and the effects on the cell are measured. This can make investigating many perturbations hard since it requires each one to be done separately. However, identifying the effect of a single genetic perturbation in a pool of many perturbations is possible with spatial barcoding. Different barcodes are made by labeling various cell compartments with different colors, which allows for identification of the various perturbations. SeqFISH is used to observe changes in gene expression by measuring mRNA transcription. To determine how many different barcodes exist, a formula was developed to compute the number of barcodes for any number of fluorophores and compartments. To predict the appearance of these labeled cells, a program was made to visualize labeled chromosomes. To optimize the process of designing seqFISH probes, a program was developed to take any number of primary probes and target genes and design oligo probes with secondary probes for temporal barcoding. Future work includes identifying more cellular compartments that can be labeled, which would increase the number of available barcodes.

Interfacing Wavemeter for Stabilization of Multiple External Cavity Diode Lasers
Emily Springer
Mentors: Nicholas Hutzler and Yi Zeng

Lasers' frequencies naturally drift over time due to thermal and pressure changes in the environments. To correct the drift for precision measurement experiments, the laser must be constantly stabilized. A wavemeter is used to periodically read the current frequency. To lock one laser while still reading the state of others, there needs to be a way to automatically switch through different lasers. A virtual instrument (VI) was built in Labview to interface the fiber optic switch and wavemeter, which allows the switching between different channels. Then in another VI, a PID feedback loop was used to stabilize the desired laser on one channel on the switch. Since the drift is negligible in a switching period, the laser is well stabilized. These VIs can be used in the lab to automatically lock and read different lasers.

Simultaneous Estimation of Pose and Shape Rotation Filter
Shiva Sreeram
Mentors: Soon-Jo Chung and Alexei Harvard

A vital component of sending drones to space is their ability to detect objects around them. In order to do so, the drone's camera needs to be able to estimate the three-dimensional position from the image it captures in two-dimensions. The aim of the project is to improve an existing flow filter system that was created in python to increase accuracy in the detection of objects with the aforementioned cameras for space applications. In order to do so, the first objective was to profile measurement errors by running Monte Carlo simulations. Then, new techniques, such as a generalized total least squares algorithm, are implemented to better account for these errors. These new techniques are then tested and analyzed through simulations.

Automated Image Processing Pipeline for Fluorescent Microscopy Images
Aubrey Stevens
Mentors: Mitchell Guttman, Joanna Jachowicz, and Mackenzie Strehle

Advancements in microscopy produce an immense amount of image data, which is tedious to manually process, reducing data available for analysis. Many automated image processing systems struggle to distinguish between adjacent cells, as well as fluorescent markers from background noise, greatly reducing accuracy. Combining automated and manual processing techniques into a hybrid model could maximize the efficacy of image processing for segmentation and marker detection. Maximum intensity projections, filtering, thresholding, and detection of specific markers were applied with automated techniques. A hybrid method segmented cells, reducing manual inputs, while maximizing accuracy. These methods were integrated together with high flexibility allowing application to many datasets. The boundary segmentation method utilizing a hybrid approach obtained 93.5% accuracy in segmentation, with only 40.9% of cells requiring manual input from the user. High accuracy with little manual processing appears possible through primarily automated methods with minimal user input. Detection of fluorescent markers amidst background noise and the accuracy of current methods require further experimentation.

Characterization of Human Genome Regulatory Regions Using Convolutional Neural Networks and LSTMs
Kristina Stoyanova
Mentors: Ron Weiss, Sebastian Palacios, and Matt Thomson

The Human Genome Project has led to the sequencing of 94% percent of human DNA. However, there are still parts of the genome that have not been fully annotated. A recent multivariate Hidden Markov Model, chromHMM, has annotated 25 different chromatin states (each represents a genomic element such as a promoter or enhancer). This was done using epigenetic marks – such as ChIP-seq data of histone modifications. We used these labeled regions of chromatin states as input for a Convolutional Neural Network (CNN) and Long short-term memory (LSTM) model that we created for further categorizing regions of the human genome. We specifically chose to work with epigenome E003, which is the H1 cell line, due to it being a frequently used human Embryonic Stem Cell (hESC) line. By creating this new CNN-LSTM, we anticipate finding new regulatory regions that could aid in synthetic biology genetic circuit development which could then be experimentally verified.

The Effect of Length to Height Ratio on Thrust Produced by a Caudal Fin
Christian Stromberger
Mentors: Mory Gharib and Cecilia Cerdeira

Autonomous underwater vehicles (AUVs) are growing in popularity as they eliminate the need to put humans in hazardous situations during military, scientific, and commercial missions. These AUVs are typically powered by classic rotary propellers, but driving mechanisms modeled after fish fins could provide more efficiency, speed, and stealth than the rotary propellers. One of the steps toward successfully implementing a bio-inspired propeller is understanding the morphology of a fish's caudal (tail) fin and how exactly each of its parameters effects swimming. This project studies the effect of different length to height ratios on the thrust produced by a tail. We do this by creating a mechanical fin that can alter this ratio while swimming but keeps other parameters (like area) constant. Then, with the aid of a previously created MATLAB trajectory optimization code, the forces produced by this fin in different aspect ratios can be measured and recorded. This data is then processed so that further conclusions can be drawn.

Designing a Magnetic Lens to Collimate a Molecular Beam for Larger Signals in the Search for Symmetry Violations
Victoria Su
Mentors: Nick Hutzler and Ashay Patel

Discovery of certain permanent electromagnetic dipole moments would violate charge-parity (CP) symmetry and help to explain the baryon asymmetry of the universe. In this search for new physics beyond the Standard Model, one of the leading approaches is to use a low energy, tabletop experimental setup for high precision measurements. For this, polyatomic molecules that are both laser coolable and sensitive to such measurements are used, in particular YbOH. A lot of resources are put into cooling these molecules to the lowest temperatures possible and running the experiments. However, many molecules are lost through divergence of the molecular beam off the beamline before they can be measured. Decreasing thi loss would increase the efficiency of data collection. YbOH has a magnetic dipole moment, allowing it to experience a force in a changing magnetic field. Particle tracing simulations have shown that the magnetic field of a cylindrical k=4 Halbach array would successfully act as a magnetic lens by collimating a spread of divergent particles, thus preventing molecule loss through wall collisions. The simulations will allow the lens design to be further optimized for maximum measurement yields. The final design will be built and mounted onto the experiment.

Modeling a Cancer Detector Circuit and Searching for Detector Targets
Maggie Sui
Mentors: Michael Elowitz and Michael Flynn

Abstract withheld from publication at mentor's request.

Evaluation of F1-ATPase Rotary Motion Using Automated Methods to Detect Hidden States
Nathan Suiter
Mentors: Sandor Volkán-Kacsó` and Rudolph A. Marcus

Within single-molecule imaging experiments of the F1 adenosinetriphosphatase (F1-ATPase) enzyme, experimental data is convoluted with a delay caused by a difference in size between the 4-nanometer enzyme and the 40-nanometer gold probe used to record the rotary motion. Due to this issue, we have developed a method capable of accounting for Brownian motion in order to distinguish whether *Paracoccus denitrificans* ATPase follows a three- or four-state model. This method utilizes multiple algorithms evaluating the enzyme’s angular jump and rotation, which allows the true behavior of the enzyme to be distinguished from the Brownian motion. The algorithms have been shown to be capable of countering Brownian noise, and are capable of identifying and analyzing when the enzyme is in a transition or dwell state. The effectiveness of these programs has been verified through an evaluation of the algorithm’s experimental values compared to theoretical values for the system’s spring constant, diffusion coefficient, and viscoelastic relaxation time. Within the comparison, the theoretical values are mathematically calculated, and the experimental values are determined through data returned by automated processes. After this process, we plan to institute a rotational correction to account for a potentially skewed angle in the initial recording of the probe. Through this evaluation, the possibility of a state distinguishing substeps of 80◦ and 40◦ has been demonstrated to be plausible.

Predicting Solubility Between Proteins and Small Molecules Through Transfer Learning
Jessica Sun
Mentors: Stephen Mayo and Aiden Aceves

Protein design is a promising tool to create proteins with desired structures and functions. Machine learning supplements classical computational techniques for rational protein design by more accurately making predictions about new sequences with desirable properties, and in many cases, removing the need for hand-engineering of input features. A specific type of machine learning, transfer learning, can be used to train a model on one domain and make predictions about a different domain. We hypothesize that a siamese neural network can be trained to compare protein solubilities between inputs and accurately predict solubility of small molecules. Understanding the conditions in which transfer learning works well for the network can provide insight into important residues and characteristics for molecule solubility. It will also elucidate the limitations of transfer learning as well more applications for it in protein design.

Self-Assembly Metamaterial With Aperiodic Microstructures
Rachel Sun
Mentor: Chiara Daraio and Ke Liu

Recent advancements of manufacturing technology enable new ways of constructing architected materials that possess desirable properties such as stiffness and strength by controlling their microstructures. Through numerical approaches, one can program aperiodic material microstructures towards desirable properties by a “growth”-like process that is encoded by “DNA”-like pairwise combination rules. To do this, a “growth” process similar to cellular automata is used after defining some fundamental building blocks and connectivity rules over a cellular space. This project achieves a similar "growth" process in the physical world, allowing for rapid production of strong programmable materials. Proofs of concept of 2D and 3D physical self-assembly methods are constructed via simulation. 2D self-assembly methods include "seed" and "shaker box" type scenarios in which notched tiles self-assemble according to local "growth" rules. A 3D self-assembly strategy is also created, in which notched cubes in a shaker box attract each other and connect according to prescribed rules. Varying tile type and quantity leads to different aperiodic microstructures that are random, yet ordered, allowing one to efficiently generate new metamaterials and effectively explore the infinite dimension design space, pushing boundaries of new material properties that can be applied to engineering applications such as shock absorption and acoustics.

The Chemical Logic of Input Integration in Third-Order Fly Olfactory Neurons
Sharne Sun
Mentor: Elizabeth J. Hong

Olfactory circuits have a particularly orderly anatomical structure, in which primary sensory neurons that express the same odorant receptor protein project to a common compartment in the brain, called a glomerulus. The glomerulus defines the fundamental unit of sensory input in the olfactory system. Higher-order olfactory neurons deeper in the brain pool input from multiple glomeruli, and the logic of this pooling defines the chemical feature selectivity of these neurons and thus how odors are represented in the brain. One such higher-order olfactory center is called the mushroom body; its principal neurons, called Kenyon cells, encode representations of odor objects that are used in olfactory associative learning and memory. In this project, we investigated the logic by which Kenyon cells integrate input across different glomeruli. We use a recently released dataset called the "hemibrain" connectome, which provides a high resolution, detailed, and complete map of neuronal connectivity of one side of a fly brain, which includes one complete mushroom body. Using a powerful web-based tool for the visualization and analysis of this data called Neuprint, we investigated the degree of structure in the glomerular connectivity matrix for the ~1754 Kenyon cells in the dataset. By comparing the observed connectivity matrix to the prediction of different types of simulated null models, which assume random connectivity statistics, we demonstrate that glomerular input to Kenyon cells is structured, in agreement with a recent analysis of an independently generated connectome of another fly mushroom body. We then use databases measuring the chemical sensitivity of each glomerulus, as well as their computationally predicted molecular feature preferences, to probe the logic of why certain glomerular inputs are more frequently co-integrated by Kenyon cells. We will test the hypothesis that Kenyon cells tend to receive co-input from glomeruli that tend to be concurrently active when the fly encounters odors from natural sources that are behaviorally significant, such as food or mates.

Search for Quirk Pair Production via a Standalone Reconstruction Algorithm
Nathan Suri
Mentors: Maria Spiropulu, Cristian Pena, Christina Wang, Michele Papucci, and Si Xie

Quirks are a class of heavy stable charged particles (vector-like fermions). Existing within a dark confinement scale significantly larger than Standard Model QCD, quirks are most easily identified by their oscillating trajectories resulting from the coupling of the quirks via a flux tube of dark gluons. This study seeks to identify key features that may allow for quirk signals to be most easily identifiable from Standard Model background. Since quirks are exhibited by non-helical tracks within the detector, one kinematic discriminant tested was the potential time delay of the quirks. The offline timing study showed high signal efficiencies for time delay cuts as large as 1 nanosecond when the timing resolution of the target MTD layer is 30 picoseconds. Further signal to background discrimination may be found by utilizing the correlation of the quirk positions and time delays since quirks are theorized to be significantly larger than Standard Model quarks.

Liposomes and the Communication of Genelet Circuits
Jeremiah Lyn Susas
Mentors: Richard Murray and Melissa Takahashi

A large goal in synthetic biology is creating a synthetic multicellular system, where synthetic cells communicate with one another to perform a larger function. These larger functions can range from synthesis of bio-molecular material to environmental monitoring. The objective of this project is to model DNA/RNA templates, genelet circuits, within liposomes and have two liposomes communicate via RNA transport. The chemical reactions and kinetics of the two different genelet circuits, a bistable switch and "Produce 2N", are individually simulated through bioCRNpyler and bioSCRAPE. These programs are used to optimize parameters to showcase different conditions of bistability for the bistable switch and repression of "Produce 2N". Through sub-SBML, a program that compartmentalizes chemical reactions into components, these genelet circuits are simulated within a liposome and communicating with one another. Moving forward, the next steps that can be beneficial for the synthetic cell community is to: find an optimal RNA transmembrane protein and, optimize genelet circuits under a cell-free extract environment.

Mining Extremely Deep Spectral Images of the Distant Universe
Isabel Swafford
Mentor: Charles Steidel

By studying galaxies at redshifts of z = 2 - 3, the peak of galaxy formation, we can learn about the processes by which they form and evolve. We want to study the baryon cycle, or the inflows, outflows, and accretion of gas that surround the galactic disk, and how this cycle affects star formation rates, metallicity, structure, and other characteristics. These flows of gas interact with the IGM and the CGM, and the dynamics of these interactions are not yet well understood. This redshift range also lends itself to the optical-band observation of rest wavelength UV-band light from the source. We can determine metallicity, gas dynamics, and make inferences as to many other characteristics by taking spectra of sources at this epoch. We have developed a tool that automates and streamlines the process of extracting 1D spectra from sources in KCWI data cubes, allowing the user to extract multiple spectra at once. By using these spectra to examine the mechanisms through which galaxies develop, we can work toward the standardization of the field and strive to understand how the very first galaxies formed.

Factors Influencing Urban Wildfire in the Los Angeles Basin
Madeleine Swint
Mentor: William Deverell

With the increasing frequency and severity of wildfire in Southern California, determining risk factors leading to the more aggressive fires and alleviating them have become of upmost importance. A catalog of wildfire in Griffith Park was created with the intent of isolating various fire-influencing factors while maintaining topography and fuel type. In doing so, fire intensity appeared to increase between 1933 and present day, the clearest cause being the intensification of drought in the region, possibly due to aerosols preventing atmospheric convection. Further comparison of Griffith Park fires to others under differing constants, such as precipitation and year, will lend further insight to determining contributing factors outside of drought. Accordingly, these factors could be addressed based off their respective natures to mitigate future fire risk.

Simulation of Battery Cycling of Different Electrode Architectures Using Porous Electrode Theory
Brandan Taing
Mentors: Julia R. Greer and Kai Narita

Lithium-ion batteries are attractive as a form of energy storage for various applications (e.g. power grid storage, electric vehicles, and portable electronics) because of its rechargeability, high energy density and high power density. The most adopted model for lithium-ion batteries is the porous electrode theory, which accounts for porous structures using mean structural properties such as tortuosity and porosity to estimate the transport and reactions. Here, we verify and utilize the LIONSIMBA suite, a pseudo 2-D, finite volume method-based MATLAB framework based on the porous electrode theory to simulate the battery cycling of lithium-ion batteries with various electrode architectures. Future research can be done to review the porous electrode theory itself, with its various assumptions.

Conductance of an Integer Quantum Hall Edge Proximitized by a Superconductor
Yuchen Tang
Mentors: Jason Alicea and Christina Knapp

An integer quantum Hall phase exhibits the remarkable behavior of being an insulator in the bulk and a conductor along the edge. In contrast, the electrical resistance of a superconductor disappears below a critical temperature and magnetic field, resulting in supercurrent flow. Combining these two exotic phases of matter has long generated interest in the condensed matter community, both on general grounds of understanding the implications for electronic transport, as well as for potential applications to realizing topological defects in related systems. Although it is experimentally difficult to combine the quantum Hall effect and superconductivity in a single device due to the presence of a large magnetic field, recent experiments have generated superconductivity in integer quantum Hall edges. In this work, we investigate the conductance of an integer quantum Hall edge proximitized by a superconductor, paying particular attention to the effect of superconducting vortices on the transport.

Using Geant4 to Simulate Propagation of Quirks in the CMS Detector
Kaden Taylor
Mentors: Maria Spiropulu, Michele Papucci, and Cristián Peña

Quirks are exotic vector-like fermions charged under both the Standard Model and a hidden confinement scale which arise in models beyond the Standard Model, such as folded supersymmetry or twin Higgs models. We assume the confinement scale Λ is in the range 100 eV – keV, and the quirk mass m is in the range 100 GeV – TeV, causing quirk production to result in macroscopic strings connecting the quirk pair. This gives rise to two separate quirks which curve towards each other, resulting in non-traditional detector signatures which break down under traditional tracking algorithms. Using the Geant4 framework, we model the propagation of quirk pairs in the CMS detector to find such signatures, including tracking four-momenta, positions, and detector hits over time. Further work must be done to validate the implementation and include additional detector energy loss signatures, but, when finished, the simulation would allow for experimentalists at the LHC to perform a first-of-its-kind dedicated quirk search.

Origins of Luster in Böttger Lusterware
Zane Taylor
Mentors: Katherine Faber and Celia Chari

Böttger lusterware was Europe’s first successful replication of Chinese porcelain with decorative over-glaze enamels, known for its rare “mother of pearl” and lustrous purple glazes. It gains its famous colorations from a layer of gold nanoparticles in its luster layer, but this phenomenon is not well understood. The gold nanoparticles of an authentic sample were found to occupy the outer ~500 nm of a ~1μm thick lead-rich layer with size distribution 34 ± 47 nm using number average and 292 ± 128 nm by volume averaging. Simple models considering surface plasmon polaritons or Mie scattering separately fail to adequately describe the color and luster of this authentic sample. This work implemented a more complex model preliminarily predicting that much of the specular iridescence of lusterware is due to thin-film interference caused by nanoparticle layers and stratification. Experimental results have demonstrated the importance of controlling lead content and reducing character of the atmosphere during processing. With further modelling and characterization, the luster and color of Böttger lusterware can be more fully understood, and then explored in the context of the chemistry and processing used by the original Meissen factory in the early 18^{th} century.

Modeling the Diffusion Within the GV-gel Drug Delivery Platform
Anna Tifrea
Mentor: Mikhail Shapiro

Abstract withheld from publication at mentor's request.

Non-Relativistic Geometry of the Gravitational Field
Marc Touraev
Mentor: Anton Kapustin

We formulate a non-relativistic spacetime as a fiber bundle equipped with an Ehresmann connection and a symmetric contravariant metric on the fibers. We write down the action for a massive particle in a gravitational potential and the action for a Schrödinger field. We define a family of transformations of the connection, gravitational potential, and Schrödinger field which preserve both actions. We show that any gravitational potential can be eliminated through these transformations, which as a consequence changes the Ehresmann connection. This exhibits the equivalence principle in this setting. We give an example in which we eliminate the gravitational potential of a point mass.

A Local Proof of the Jacquet–Langlands Correspondence
Justin Toyota
Mentor: Justin Campbell

The Langlands program is an important field of modern mathematics, conjecturing deep connections between number theory and representation theory. One of the earliest examples of this philosophy is the Jacquet–Langlands correspondence, which posits a unique correspondence between certain representations of $GL_2(F)$ (where F is a non-archimedean local field) and certain representations of $D\times$ (where D is the central division F-algebra of dimension 4). In this project we fill in an outline of this correspondence given by Bushnell and Henniart. Of note is that the resultant proof only uses local methods, and does not require global tools like automorphic forms or trace formulas. This fills a gap in the current literature, which does not contain a purely local proof of this result outside of sketches.

Creating Framework to Support Fluid Online Workspaces
An Tran
Mentor: Santiago Lombeyda

Online workspaces merit various benefits, especially under the current world and educational environment. In the past, the research group has produced classrooms and workspaces in virtual reality, complete with 2 different labs (a planetary simulation and MATLAB calculator) connected to central servers allowing direct interaction with 3D models and/or data. To address unification and platform problems, a new highly efficient central relay server was built from scratch, with the goal of supporting multiple backends and clients. The relay server was built on Rust and WebSockets with the goal to expose easy and performant interfaces for simulations, virtual labs, and academia usage. Due to the low adoption and accessibility of virtual reality headsets, a new web interface powered by BabylonJS and modern web technologies was built. A new basic physics-based game and an updated planetary simulation was written to demonstrate the capabilities of the relay server. We have thus created a highly efficient modular framework to enable collaboration and instruction across different services and front clients.

Utilizing Conflicting Domain Knowledge in Reinforcement Learning Settings
Albert Tseng
Mentors: Yisong Yue and Adith Swaminathan

We study the problem of how to best approach conflicting human domain knowledge in goal based reinforcement learning settings, where the aim is to utilize or otherwise resolve the existence of conflicts. In contrast to settings where domain knowledge is presented as what is essentially ground truth, we are concerned with settings where multiple domain experts exist and provide disagreeing and thus ambiguous information when labeling goals. As existing methods for utilizing labeled goals do not provide ways for addressing conflicting information, we propose a formalization of human disagreement in goal based settings on which we base our hierarchical curriculum algorithm. More specifically, we assume that human labelers are fully faithful to goal definitions, and show that goal disagreement is an indication of ambiguous goal definitions and ease of goal achievement. We demonstrate the efficacy of this approach in MineRL, an extremely hierarchical Minecraft environment, as well as simpler maze environments.

Identification of Neuronal Action Potential to Explore Electrophysiology in Head-Fixed Mice
Avedis Tufenkjian
Mentors: Michael Roukes and Alice Hsu

Intan recordings of electrical current from in-vivo craniotomies of wild-type mice are used to explore electrophysiology in the brain by identifying the exact soma where certain neuronal activity is generated. The raw cranial data is captured using a custom made 128-Channel probe-equipped PCB, which is inserted into a sedated mice's brain. The raw recording data is read and filtered through *Spiking Interface*, which identifies the action potential of specific firing neurons from the recording. The data is then visualized using *Matlab* and *Phy*. Further research will continue to contextualize the waveform electrical data and determine and locate the recorded action potential of specific neurons.

Spin-Symmetry Restored Many-Body Perturbation Theory
Shu Fay Ung
Mentors: Garnet Chan and Chong Sun

To accurately capture the behavior of electrons, an elusive goal of electronic structure theory is to achieve a balanced description of dynamical and static electron correlations efficiently. Unrestricted Møller-Plesset perturbation theory (MPn) can attain high accuracy in computed observables for non-degenerate systems, where the correlation is almost exclusively of dynamical character. On the other hand, an approach to account for static correlation in degenerate systems is to employ methods that use spin projection. We present a spin-projected unrestricted MP2 method capable of handling both dynamical and static correlations. This method utilizes an integral form of a simplified spin projector, leading to a computationally simpler scheme compared to previous spin projector formulae. We benchmarked our method against other quantum chemical methods and applied it to the study of selected molecular systems.

Analyzing Fukushima Blast Waves and Gas Flow Using Background Oriented Schlieren
Maxwell Vale
Mentors: Joseph E. Shepherd and Conor D. Martin

With recent developments in image processing software, the Background Oriented Schlieren technique has emerged as a way to visualize density gradients and other phenomena in transparent media. Unlike traditional schlieren image, which relies on a relatively expensive setup of cameras and mirrors to see the refraction of light in varying densities, the BOS technique requires a single camera, a point light source, and a speckled background to capture the desired phenomena, which is then analyzed by image processing software to produce a synthetic greyscale schlieren image. The objective of this study is to use the Background Oriented Schlieren technique on the raw TV images of the Fukushima nuclear disaster and receive more accurate and reliable information about the gas flow and shock waves from the explosion, providing insight about the internal conditions about the plant before the explosion occurred and the nature of explosion.

Influence of Substrate Curvature on Dynamic Cone Formation in Electrified Liquids by Finite Element Modeling
Lorenzo X. Van Muñoz
Mentors: Sandra M. Troian and Nicholas C. White

Above an initial critical field strength, the unstable, free surface of an electrified, viscous perfectly conducting liquid develops an accelerating protrusion resembling a cusp with a conic tip. Such conic tips are known to emit highly energetic ion beams from the conic apex. This process forms the basis for such modern day ablation and lithographic instruments as focused ion beam systems. At the conic apex, the stabilizing normal stresses from capillary forces lose out to the increasing pull of Maxwell stresses, which rapidly accelerate the liquid tip toward a counter-electrode in nanoseconds. Previous studies have focused on the formation of dynamic cones in liquids supported on flat substrates. Here we examine the influence of a curved substrate. Finite element simulations are used to model the flow of an incompressible, thin film of liquid metal on a curved substrate for a system in vacuo for various values of initial field strength and liquid film thickness. We investigate regions of the parameter space which appear to give rise to conic formations with trailing oscillations.

A TEve Based Event Display for Mu2e
Aditi Venkatesh
Mentor: David Hitlin

The Mu2e Experiment at the Fermilab seeks to find a Charged Lepton Flavor Violation through measuring the ratio of conversion of a neutrinoless muon to an electron relative to normal muon capture. In order to visualize data inside the experiment, an Event Display is created. Event Displays act as a visual interface between the user and the data and help physicists visualize the geometry of the experiment and particle interactions and events that occur in the detectors. An Event Display for the Mu2e experiment was created using TEve subsystem of the CERN ROOT Project.

Applications of Conductive Hybrid Hydrogels in Drug Delivery Systems: A Review
Polina A. Verkhovodova
Mentors: Xuanhe Zhao and Guruswami Ravichandran

During the past few years, the development of hydrogel-based controlled drug delivery systems have shown the potential to increase the efficacy of many therapies. Hydrogels are attractive materials for this application due to their high biocompatibility, adjustable physical properties, and controlled degradation. These systems provide a targeted release of drugs and other bioactive materials either temporally or in response to an external trigger. Many researchers are currently exploring the use of carbon nanotubes to create a “smart” hydrogel that will release drugs or other bioactive materials in response to changes in applied voltage and current. This Review will explore current developments in the use of carbon nanotube-hydrogel hybrid systems, specifically focusing on improving biocompatibility and how different systems provide better controlled release.

Exoplanets in Multi-Star Systems
Adrienne Vescio
Mentors: Calen Henderson and Julian van Eyken

Over half of Sun-like stars and nearly 10% of confirmed exoplanets in the NASA Exoplanet Archive (hereafter: Archive) exist in systems with one or more stellar companions. As a result, acknowledging the existence of stellar multiplicity in exoplanet research is fundamental in assessing the environments in which exoplanets form and evolve. In order to better understand the role that stellar multiplicity plays, we aim to construct the first catalog of physical, orbital, and observational properties of companion components in the Archive. In an effort to integrate these data into the Archive, we extract stellar parameters from over 200 refereed manuscripts and prepare them for ingestion through use of individually authored data files. Upon completion and validation, these files are used to incorporate information through the Archive’s bi-weekly data releases. The primary result of these efforts will be the first publicly available, collated dataset of binary, planet-hosting systems, providing an easily accessible, greater context for planetary systems than what presently exists. Following this effort, we will utilize data collected for this project to search for immediate patterns in companion star impact on orbital and physical parameters of exoplanets.

Using Deep Reinforcement Learning to Learn Active Vision for a Mobile Manipulator
Yasmin Veys
Mentors: Silvio Savarese, Roberto Martín-Martín, and Soon-Jo Chung

Many navigation tasks require humans to interact with their environment. Navigation through a building, for example, may involve opening a door or pressing an elevator button. In robotics, these types of tasks which combine navigation and manipulation require the use of a mobile manipulator consisting of an arm attached to a mobile base. Efficiently coordinating the movement of an arm with a mobile base, referred to as mobile manipulation, is challenging, especially in unstructured and dynamic environments. Recent work has demonstrated that deep reinforcement learning techniques can be used to learn mobile manipulation, but limitations arise due to partial observability of the environment. Our work aims to address these issues by using deep reinforcement learning to control not only the arm and base of a mobile manipulator, but also its head. We have shown in simulation that the mobile manipulator agent can successfully navigate towards and reach a randomized goal with its arm. We have also successfully trained the agent to move its head as it navigates through an environment with dynamic obstacles to increase its visibility. Our objective is to integrate these policies so that the robot can better perceive its environment to complete more challenging tasks successfully.

The Atomistic Level Structure for the Activated Smoothened Receptor Bound to the Full Gi Protein and Agonist
Amy Vo
Mentors: William A. Goddard III and Soo-Kyung Kim

Smoothened (SMO) is an oncoprotein and signal transducer involved in the Hedgehog signaling pathway, a pathway which regulates cellular differentiation and embryogenesis. Although the mechanism of signaling through SMO is not yet well-understood, it has been suggested that SMO may function as a G protein-coupled receptor (GPCR) to activate a G protein. To investigate this proposed mechanism, we start from experimental cryo-EM data and model SMO in complex with the full Gi protein, allowing us to find hydrogen bonds and salt bridges not resolved in the cryo-EM structure. Using nanoscale molecular dynamics simulations, we look for characteristic interactions between the GPCR and Gi protein and determine the stability of the binding conformation.

Predicting the Spread of COVID-19 Using Artificial Intelligence
Jagath Vytheeswaran
Mentor: Yaser Abu-Mostafa

The COVID-19 pandemic has rapidly become the most crippling and widespread crisis in recent history, and its devastation has disproportionately affected the United States, with the national death toll quickly approaching 200,000 victims. As the country tries to get the virus under control, understanding the spread of the disease and predicting future cases, hospitalizations, and deaths are of paramount importance. This project is part of a larger effort from a team of Caltech researchers to create a de-facto COVID-19 death prediction model for every county in the United States. More specifically, this project deals with the modelling of policy and mitigation effects of different demographic groups, using distributional assumptions on disease responses to model the changes in cases and deaths given non-pharmaceutical intervention strategies (NPI's). These models, in turn, inform counterfactual predictions of future cases and deaths given changes in policy, providing value to policymakers as they make decisions on how best to counter the disease. Finally, the policy-effect and mitigation-response models were used to generate features for the final COVID-19 deaths prediction model and bolster predictive power.

Integration of Distributed Multi-Agent Localization for Large-Scale Swarm Formation Flying
James Walker
Mentors: Soon-Jo Chung and Kai Matsuka

The implementation of formation flying techniques to spacecraft constellations requires an accurate representation of Earth's gravitational model as well as a relative controls algorithm architecture. This paper discusses the use of the most recent gravitational model for the Earth and the Harris-Priester model for atmospheric drag. These perturbations are used to create orbits for both a chief and a deputy satellite. Using relative orbital elements, the orbit of the deputy satellite is defined based on the position of the chief. Because the perturbations will affect each satellite differently, the two satellites will slowly drift apart. A control program must be designed to allow for the deputy satellite to perform maneuvers and correct its orbit. Because the deputy's orbit is defined by relative orbital elements, the maneuvers will adjust these values. This will allow satellites to have orbits defined by a chief orbit and conduct orbit reconfigurations based on that single orbit.

Measuring H_0 Using X-ray and SZ Observations of Galaxy Clusters
Jenny T. Wan
Mentors: Steven Allen, Adam Mantz, Jack Sayers, and Sunil Golwala

We use observations of 14 dynamically relaxed galaxy clusters in X-ray and radio/sub-millimeter wavelengths to determine the value of the Hubble constant through the Sunyaev-Zel'dovich (SZ) effect. The SZ effect involves the up-scattering of cosmic microwave background (CMB) photons by thermal electrons in the hot gas in clusters of galaxies. This inverse Compton scattering of CMB photons leads to a distortion of the CMB spectrum along the line of sight through a cluster. For an assumed spherical geometry, the combination of X-ray and SZ observations can be used to measure the distance to a cluster. The ratio of the predicted (based on X-ray observations) and observed SZ signals is proportional to the square root of the angular diameter distance, making this a powerful technique for probing the cosmic distance scale. Because of the near-spherical geometries of the clusters in our sample, most of the systematic error in our measurement of H_0 will likely be associated with absolute calibration of the X-ray data. With this, we will be able to obtain a value for H_0 that should have interesting statistical precision, and we will seek to identify steps to improve the calibration of future observations to minimize this systematic error.

Structural and Functional Correlates of Imaginative Suggestibility
Alexander Wang
Mentors: Michael Lifshitz and Tanya Luhrmann

Suggestibility is the susceptibility of an individual to suggestions, communicable ideas that can rapidly and either temporarily or permanently alter that individual's experience or behavior, and is believed to be a trait which varies among individuals. Hypnotic suggestibility refers to suggestibility in a state of hypnosis following hypnotic induction, in contrast with imaginative or waking suggestibility, for which there is no preceding hypnotic induction. Resting state functional magnetic resonance imaging (fMRI) has been used to investigate neurobiological profiles and structural and functional correlates associated with various psychological traits, including suggestibility. However, previous studies have been characterized by small sample sizes (fewer than 10 participants) or non-representative samples (e.g. entirely female), and there has been very limited investigation into imaginative suggestibility. In this study, we investigate the relationship between functional connectivity and imaginative suggestibility, correlating resting state functional magnetic resonance imaging data with imaginative suggestibility rating scores.

Investigating Giant-Celled Biology Using Serial Block-Face Electron Microscopy and Image Processing
Jackie Wang
Mentors: Elliot Meyerowitz and Tyler Gibson

The tropical alga *Caulerpa* is the world's largest single-celled organism. There are many unsolved questions with respect to the biology of *Caulerpa*, one of which is how the alga coordinates its growth with its metabolism. As a first step towards answering this question, we have collected both macroscopic and microscopic geometric data on the *Caulerpa* stolon and rhizoids. To generate the microscopic data, we used serial block face scanning electron microscopy to generate a stack of thinly sliced sections of the stolon. The stack was analyzed through human-guided image processing with an open-source software called MiB (Microscopy Image Browser). This allowed for the segmentation of the various organelles in a machine-assisted manner. We present a quantitative analysis of the various organelles of the stolon of the alga. We then used MiB to measure the growth rates of the stolon and the rhizoids. The linear growth rate of both structures reflects a controlled, systematic coordination of growth, both in terms of volume and in terms of surface area. The data can be applied to understanding the scaling of size and metabolic rate in *Caulerpa*.

Systematic Assessment of Genome Assemblies (SAGA) for Synthetic Genomes
Yinghan Megan Wang
Mentors: Kaihang Wang and Charles Sanfiorenzo

The reliability and accuracy of whole-genome assemblies stemming from long-read sequencing technologies are highly affected by the occurrence of repetitive regions and error-prone reads. Such a problem becomes more predominant in the case of meta-genome assemblies obtained from communities comprised of closely related species, or in the case of *de novo* assemblies obtained from the generation of synthetic chimeric genomes. Through the use of Cas9 and recombineering technologies, our group has developed the ability to generate genomic chimeras by interchanging or adding large genomic regions (>500-kb), a feat first carried out between *Escherichia coli* and *Shigella flexneri.* In order to effectively assemble, compare, and annotate synthetic genome assemblies in high throughput, our group has created a pipeline for the systematic assessment of genome assemblies (SAGA). SAGA performs assembly steps on size-filtered and unfiltered read sets utilizing several parameters and options for state-of-the-art *de novo* assemblers, Flye and Canu. Assembly is followed by downstream analysis of assembly quality through Quast metrics, graph entanglement, and genome topology configurations. Consensus genome assemblies are then fed into Prokka to obtain fully annotated genomes. By these means, SAGA achieves streamlined and effortless *de novo* assembly, comparison, and annotation of polished assemblies.

Synthetic Phase Contrast MRI for Deep Learning Reconstruction
Jack Warren
Mentors: Shreyas Vasanawala, Matthew Middione, and Brian Stoltz

Phase-contrast Magnetic Resonance Imaging (PC-MRI) measures blood flow velocity within blood vessels. However, in order to get high-quality measurements, patients must hold their breath for up to 25 seconds, which may be difficult due to various factors. Currently, deep learning (DL) neural networks are being developed to heavily undersample the data acquisition process in order to decrease acquisition times and enhance image reconstruction. Acquiring fully-sampled patient data to train these networks is not readily available. This work attempts to create applicable *synthetic data* to effectively train DL neural networks for PC-MRI. First, a computer program was written to generate anatomically correct 2D PC-MRI images with accurate tissue properties. Then, synthetic sensitivity maps were generated and the data augmented to better capture the range of images that would otherwise be acquired in clinical patients. Finally, the DL neural network was trained using the generated synthetic data and evaluated for its effectiveness. Successfully training the PC-MRI DL neural network with synthetic data has the capability of revolutionizing the entire MRI field. By substantially reducing scan times for all types of MRI images, there is potential to decrease cost, increase ease of use, and therefore boost MRI accessibility and diagnostic value.

An All-Sky Infrared Variables Catalog From Palomar Gattini-IR
Thomas K. Waters
Mentors: Kishalay De and Meredith Rawls

Palomar Gattini-IR is a robotic, near-infrared (NIR), time domain survey that can scan approximately 15000 square degrees with a median cadence of two days. This project gives us an unprecedented look into the infrared universe by detecting a vast number of transient sources obscured by dust and extinction and very cool, red variable objects, which remains a largely unexplored area of time domain astronomy. We use Source Extractor, a source detection software, to detect astronomical objects in Palomar Gattini-IR images and associate them into catalogs. However, as the Palomar Gattini-IR data reduction pipeline uses a PSF reconstruction algorithm, which introduces correlated pixel noise in the data, the traditional Source Extractor parameters are not optimal for source detection in Palomar Gattini-IR images. We use the Two Micron All-Sky Survey (2MASS) as a reference dataset to optimize critical Source Extractor parameters to maximize source detection in Palomar Gattini-IR stacked exposures and dithered images. Using this optimal parameter set, we implement Source Extractor on Palomar Gattini-IR nightly stacked science images to generate science quality light curves and associate them into the first-ever NIR variables catalog.

Level-Set Discrete Element Method (LS-DEM) Verification of Macroscopic and Inter-Particle Properties of Granular Materials Under Shear Stress
Mitchell Watson
Mentors: Guruswami (Ravi) Ravichandran and Zichen Gu

Granular material (such as sand) is the second most used material in Industry besides water, yet it lacks a universal model to give a complete theoretical understanding. Particle-by-particle discrete simulations have been used in the past, but their high computational cost and lack of accuracy with large systems reduce their efficiency and practicality. The Level-Set Discrete Element Method (LS-DEM) finds the inter-particle forces of the material by treating each interaction of particles as a spring system with 2 masses by calculating the inter-particle forces by a spring constant and the overlap of the particles. This, along with using a Level-Set function to describe the surface of the particles, allows large, complex systems to be simulated with lower computational costs and higher accuracy. Using a shear stress tensile test machine, we calculate the inter-particle properties as well as the macro-scale properties of the system and compare to the LS-DEM results of various shear stresses. Future work should include testing various 3D models as well as capturing wet materials, which adds capillary bridges between particles.

Error Analysis of the Positional Accuracy of the Zwicky Transient Factory Streaking Near-Earth Asteroid Fitting Method
Emily Whittaker
Mentors: George Helou and Gerbs Bauer

The Zwicky Transient Factory (ZTF) imaging camera is a mosaic of 16 CCDs mounted onto the 48-inch telescope at Palomar Observatory, and is a leading survey in the search for new near-Earth asteroids (NEAs). ZTF contains a pipeline called ZStreak which uses a stretched-PSF fitting model to detect small, close NEAs whose apparent angular velocity is high enough to create a streak within a single exposure. In this project, positional data produced by ZStreak regarding 125 well known NEAs are compared to the predicted positions of these NEAs as determined using the HORIZONS software. On average, datapoints tend to be offset from their associated HORIZONS position by 1.38 arcseconds, or 6% of the length of the streak. These errors tend to be in the opposite direction of the streak, assigning the streak direction as moving from the object's position at shutter open to its position at shutter close. Once all of the positional errors are analyzed and the systematic errors are identified, they will be used to improve the positional accuracy of the ZStreak model fits.

Generating Domain Representations for Functional Gas Vesicle Proteins Using Hidden Markov Models
Katie Wong
Mentors: Mikhail Shapiro and Rob Hurt

Abstract withheld from publication at mentor's request.

Many-Body Quantum Systems in Ratchet Potentials
David Wu
Mentors: Gil Refael

Ratchet potentials are periodic potentials that break spatial and temporal symmetries and can be used to drive the directed motion of Brownian particles in classical systems. Recent experimental results in solid state systems and atomic physics systems have shown that quantum transport in ratchet potentials give rise to interesting phenomena, such as tunnelling oscillations and quantum resonances that hint towards the construction of quantum motors. In this study, we investigate the directed motion of quantum particles in ratchet potentials in the context of many-body physics. We numerically simulate the quantum system on a one-dimensional lattice subject to quasi-periodically driven potentials that are spatially asymmetric. By examining the asymmetric responses from the evolution of the quantum states due to right and left directed potentials, we conclude the necessary conditions for realizing directed motion in many-body quantum systems subject to ratchet potentials.

Artificial Intelligence in Radiology: Deep Learning for Automated Localization and Segmentation of Abdominal Organs in Magnetic Resonance Imaging
Sulan Wu
Mentors: Albert Hsiao, Sophie You, and Rob Phillips

Magnetic resonance imaging (MRI) is a form of non-invasive medical imaging that produces three dimensional anatomical scans. Convolutional neural networks (CNNs) are a type of machine learning algorithm commonly used for image classification tasks. There are several stages in the deep learning pipeline: data collection, curation, annotation, preprocessing, training and validation, testing, and clinical application. First, 964 abdominal imaging series were collected and anonymized. Of those, 181 series containing 7,760 total images were annotated using software developed in-house, with labels corresponding to the liver, spleen, right kidney, and left kidney. Data and annotations were split into 70% training, 20% validation, and 10% testing. Transfer learning was conducted using weights and a previous U-Net architecture developed for liver segmentation on axial MRI images. Data augmentation (field of view and translation) and over- and under-sampling were also implemented to address class imbalance. Results were analyzed by computing the average Dice coefficient. Clinical applications of this algorithm include the automated prescription of imaging planes during the acquisition process, as well as automated measurements of the size and volume of organs to assist clinicians in the monitoring of abdominal organs and related diseases.

Computer Vision Analysis of Time-Dependent Shock-Wave/Boundary-Layer Interactions in Hypersonic Flow
Brit L. Wylie
Mentors: Joanna Austin and Joel Lawson

Shock-wave/boundary-layer interactions (SWBLIs) in high-speed flows have been studied for the past several decades in an effort to better understand effects of shock interaction on flow separation and thermal loading. Schlieren image techniques are commonly used to photograph these high-speed shock structures and can produce detailed videos comprising large numbers of still-frames. However, these images are difficult to analyze due the small scale of separation and reattachment shock features, noise from the light source, and imperfections in the flow field. Two widely used test geometries for this problem are the double wedge and double cone. At Caltech, Knisely collected schlieren images of these geometries at conditions simulating hypervelocity flight (~4 km/s) in the T5 shock tunnel. In this project, we implement a three-stage method of segmentation, edge detection, and feature tracking. We have modified existing techniques developed by Smith for our use on higher resolution datasets, and have studied several improvements to background subtraction which together will enable us to extract bow shock frequency and track triple-point location for several free-stream gas compositions. Further work will include research into refined methods to detect and track reattachment and separation shocks.

Myt1l as a Promising Transcription Factor for Beta Cell Transdifferentiation
George Wythes
Mentors: Marc Montiminy, Sam Van De Velde, and Kaihang Wang

Diabetes is one of the most prevalent diseases worldwide, caused by a lack of functioning beta cells. A promising type of therapy is to reprogram non-beta-cells into insulin producing cells through transdifferentiation Traditionally, studies have overexpressed the pancreatic transcription factors NeuroD1, Pdx1, and MafA, abbreviated as NPM, to promote transdifferentiation of various cell lines in beta cells. This study focuses on Myelin Transcription Factor 1-Like Protein (Myt1l), which is a neuronal transcription factor that has been shown to help differentiate fibroblasts into neurons and is also believed to play a role in islet cell development. To measure its capacity to induce beta cell transdifferentiation, mouse hepatocytes were transfected in vivo with combinations of Myt1l and other beta cell lineage determining transcription factors (LDTFs). Myt1l amplified the effects of the other transcription factors when expressed alongside them, more than doubling the number of significantly upregulated islet specific genes. Functional enrichment also showed that genes attributed to Myt1l were more focused on characteristic beta cell pathways and functions than that of the core transcription factors in NPM.

Construction of Chemical Database From Total Synthesis Articles Through Data Mining
Yuanzhe Xie
Mentors: Sarah E. Reisman and Michael Maser

Machine learning and deep learning, the two state-of-the-art data science tools, are attaining their rising influence on how researchers approach natural sciences, including chemistry. Their "magic power", however, originates from large databases on which models are trained to make predictions. Apparently, chemistry publications are not readily available as a source of training; as a result, data mining as well as data analysis are indispensable to convert publication languages into databases accessible to training. This project proposes a complete workflow for the construction of a database from organic chemistry publications by employing open-source computer vision (CV) and natural language processing (NLP) toolkits. We have performed reaction figure extraction, chemical entity recognition, and chemistry action identification on the supporting information documents of the total synthesis articles from the ACS publications, with an observation of promising preliminary results.

2D U-Net CNN Segmentation of the Blood Vessel Lumen
Elizabeth Yam
Mentors: Albert Hsiao and Adam Blank

This project seeks to implement a UNet CNN to automate segmentation of blood vessel (aorta, inferior vena cava (I.V.C.), and iliac vessels) from 2D slices of MRI images. This increases accessibility for radiologists to identity diseases (e.g., associated with abnormal narrowing/widening of vessel). With HIPAA-compliance and IRB waiver of informed consent, clinical MRIs performed at UCSD was used as the dataset. Six anatomic landmarks (channels) on around 150 series, each containing 20-200 axial slices, of MRIs were annotated. Data was split into 70% of cases for training, 20% for validation, and 10% for testing. Implemented in TensorFlow using Keras, the model uses the sum of DICE loss (1 - DICE score) for each channel to train on. The model uses under/oversampling to account for class imbalance (i.e., undersampling the iliac vessels and oversampling aorta and I.V.C.), staged learning (i.e., adding more slices to the training set in later model training), and transfer learning from model on cardiac segmentation. The model had an average DICE score of 0.814 ± 0.0102 over all the channels. Therefore, CNNs may have the potential to automate identification of blood vessels in MRIs and may complement other techniques used to identify blood-vessel-related diseases.

Designing Resonant Metasurface-Based Optical Modulators
Evan Yamaguchi
Mentors: Andrei Faraon, Hyounghan Kwon, and Tianzhe Zheng

Optical metasurfaces are thin optical devices that allow for efficient manipulation of light at the nanoscale. Previous work has demonstrated metasurfaces to be suitable compact analogs of bulky optical components such as lenses, polarizers, spectrometers, and other multifunctional and active optical elements. Recently, metasurface-based subwavelength dielectric scatterers have also been shown to be compact and efficient enhancers of light-matter interactions at the nanoscale. In this work, I design a resonant metasurface composed of subwavelength periodic scatterers to engineer a device with full phase and polarization control. This work contributes to novel methods of engineering resonances in metasurfaces for applications in nanophotonics and metaoptics.

Coupling Hexagonal Boron Nitride Quantum Emitters to Dielectric Metasurfaces
Frank Yang
Mentors: Harry Atwater and Pankaj Jha

Single-photon sources are emissive species capable of emitting photons on demand and are applicable for quantum cryptography, metrology, and simulation. Color defects in hexagonal boron nitride (*h*BN) exhibit room temperature single photon emission that is polarized, photostable, and has record brightness, making *h*BN a leading material platform for single photon sources. An important challenge in the integration of single-photon sources with photonic chips is the efficient extraction of photons.

In this work, we seek to address this challenge by coupling single-photon sources with dielectric metasurfaces to produce directional emission. Metasurfaces are two-dimensional subwavelength arrays of scatterers capable of controlling the polarization and phase of incident light. Using full-wave simulations, we design an efficient metasurface capable of controlling the far-field radiation pattern from a single *h*BN emitter, resulting in directional emission or lensing. This work demonstrates the potential of metasurfaces in enabling efficient photon extraction from quantum emitters.

Thermal Bioswitches For Localized Activation of Microbes
Michael S. Yao
Mentors: Mikhail Shapiro, Mohamad H. Abedi, David R. Mittelstein, and Avinoam Bar-Zio

Abstract withheld from publication at mentor's request.

Wearable "Team Flow" Real-Time Monitoring and Modulation System
Jessica Ye
Mentors: Shinsuke Shimojo and Mohammad Shehata

Working in a space environment with an extended duration of extreme conditions can lead to a decline in the cognitive functions, leading to poor teamwork. NASA's Human Research Program (HRP) identified this highly likely risk, which has consequences on both operations and health in the long-term. This project aimed to addresses this risk by providing an objective monitoring and modulation method to enhance team performance with high motivation and engagement. Team performance is a part of "team flow", or a state of high task engagement during team activities. The experience of being "in the zone", or flow state, is a pleasurable state indicative of task engagement for individuals performing a task. In a former study (Shehata et al, 2020), simultaneous EEG measurements were collected from teams performing in a music rhythm game. Team flow neural correlates using electroencephalograms (EEG) were identified (Shehata et al., 2020). Team flow was maximized by matching two participants (teammates) in gaming skill and music preference, and experimental manipulations were used to create three conditions: team flow, no-team flow and team no-flow conditions (Shehata et al., 2020). We have developed a method that uses parallel processing in Matlab to quickly (~5s) calculate flow indexes based on EEG data. Our algorithm is based on three components: auditory evoked potentials (AEP), which measure attention to task-irrelevant sounds; phase-locking values, which measure intra- and inter- brain neural synchrony; and the power of EEG signals from flow-associated brain regions. Our flow index can be calculated for both an individual and a team. The team flow index could be indicative of team flow, team no-flow, or no-team flow conditions. The rapid calculation of flow indexes enables us to provide the participants with an immediate real-time feedback on their individual- and team- flow states. Due to COIVD-19, we have been unable to test our real-time feedback with participants. We aim to confirm the plausibility of our objective flow indexes using Bayesian statistics. We are creating a predictive hierarchical Bayesian model based on the results of the flow calculations from the above study in order to predict individual flow and team flow.

Sequential Sampling for Accelerated Magnetic Resonance Imaging
Tianwei Yin
Mentors: Yisong Yue, Katie Bouman, and He Sun

Recent learning-based approaches have shown promises for accelerating magnetic resonance imaging. However, most previous works focus on designing reconstruction networks with fixed undersampling patterns, ignoring the subject-specific information acquired in the inference stage. In this project, we propose to learn a sequential sampling strategy to decide acquisition trajectories based on previous observations and reconstructions adaptively. We use a convolutional network-based sampler and train the whole differentiable pipeline end to end. We are evaluating our method on the large scale FastMRI dataset and plan to adapt our sequential sampling approach to the task of multi-modalities magnetic resonance image recovery.

Engineering Macrophages as Cellular Cancer Sensors With Mammalian Acoustic Reporter Gene Expression
Mei Yi You
Mentors: Mikhail Shapiro and Justin Lee

Abstract withheld from publication at mentor's request.

Discrete Fourier Transform on Schurrian Schemes
Hantao Yu
Mentor: Chris Umans

In 2019, Umans proposed an arithmetic algorithm to compute generalized Discrete Fourier Transforms with respect to a group G, using O(|G|^0.5ω+ε) operations, where ω is the exponent of matrix multiplication. We introduce the DFTs on a more general structure – association scheme. We narrow our research on schurrian scheme, which is of great importance in theoretical computer science and combinatorics. The fundamental methods used in previous work on groups include subgroup reduction. That is, apply the DFTs on a smaller subgroup, which requires less operations, and consider all translations such that the whole group will be covered. However, the concept of "coset" has not been defined on association schemes. Given a closed subset C of a scheme R, we try to find |R|/|C| many translations of the closed set to cover the whole scheme. We then apply similar approaches to optimize the operations needed for DFTs on Schurrian Schemes.

Power-Aware Scheduling of Precedence-Constrained Tasks on Multiple Machines
Jannie Yu
Mentors: Adam Wierman and Yu Su

The prominence of large-scale, general-purpose machine learning platforms on the cloud such as Microsoft Azure and Google Cloud Platform have accelerated the construction of increasingly computationally demanding models with considerable energy use. In the machine learning community, these model workflows are typically expressed via a directed acyclic graph, where vertices and directed edges represent the tasks in job batches and precedence relationships between tasks, respectively. The energy consumption of the platforms relies on how these precedence-constrained tasks are scheduled across machines, motivating the goal of developing speed-scaling scheduling designs that optimize a trade-off between energy consumption and performance. In this project, we focus on developing a new scheduling framework that can provide the first provable worst-case approximation guarantee to optimize for performance – via time metrics like makespan and mean response time — and power.

Our work provides performance bounds for the optimal task speeds and the objective function value of the scheduler with dynamic speeds, providing the first known results of power-aware scheduling of precedence-constrained tasks.

Using Computer Vision Analysis to Detect Hand-Stimming in Home Videos for Diagnosis of Pediatric Autism
Jennifer Yu
Mentors: Dennis P. Wall, Peter Washington, and Lior Pachter

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects an individual's ability to communicate and interact with their peers. Though symptoms may vary, the condition is normally characterized by repetitive behaviors along with deficits in social interaction and communication ability, such as difficulty making eye contact, persistent repetition of hand movements, and trouble engaging in social activities with peers. Recently the overall prevalence of ASD has increased. Recent advances in machine learning have led to new methods of rapid and scalable detection. Early intervention can alter the abnormal trajectory of children with ASD and help guide their development towards a normal pathway. Smartphones and other forms of technology are quickly becoming ubiquitous, so they can play a large role in providing socio-cognitive intervention that is accessible and effective. The Wall Lab has a suite of mobile games, notably a charades style game named *Guess What?*, which collects egocentric video data of children with ASD during gameplay, while simultaneously providing digital therapy to the affected child. My research will leverage existing computer vision libraries for hand detection, namely OpenCV and OpenPose, for hand detection in home video data obtained from the *GuessWhat?* application. This work will pave the way for future in-depth analysis of hand movements within unstructured video data using computer vision, namely for interpreting hand stimming in children as a possible risk factor for ASD.

The BSD Conjecture: *p*-converse Theorem for Special Primes
Yu Qiyao
Mentor: Ashay Burungale

The Birch-Swinnerton-Dyer (BSD) conjecture predicts a relation between arithmetic invariants of an elliptic curve E over a number field K and the behavior of its Hasse-Weil L-function $L(E, s)$ of E at $s=1$. Consider the case when K is the rationals and p a rational prime. One particular instance of the conjecture claims that the Z_p-corank of the p-infinity Selmer group of E being 1 (or 0, respectively) is equivalent to the order of vanishing of the Hasse-Weil L-function being 1 (or 0, respectively). We refer to the direction from the corank to the order of vanishing as the "*p*-converse." In this project, we study a *p*-converse theorem for an elliptic curve with complex multiplication, good ordinary reduction at a prime p, and Z_p-corank 1. Burungale and Tian have established this theorem in the case of $p>3$. We identify and resolve some of the difficulties encountered when generalizing this theorem to $p=2, 3$.

Recovering the Forgotten Sonorine Recordings
Shuyue Yu
Mentors: Adam Finkelstein and Adam Wierman

It is difficult to extract sounds from the Sonorine postcards, early 20th-century postcards with audio inscribed on the back, since the Sonorine cards are too rare and fragile to be played on a device. A non-tactile approach was used by Professor Finkelstein and his group, and they managed to scan the Sonorine postcards and successfully recovered some audio from the scans. In this project, we modify the existing architecture and build an improved recovery pipeline. The new architecture considers the warps on scans caused by uneven surface as offsets and is able to remove the offsets. We believe that our final pipeline will recover audio with high quality from scans of Sonorine postcards, facilitating researchers to study about the past.

Exploring Planet Formation by Simulating Dust-Gas Fluid Instabilities
Elizabeth Yunerman
Mentors: Philip Hopkins, Jonathan Squire, and Eve Lee

The processes behind forming planetary building blocks from micron-sized dust in gaseous planet-forming disks has been an ongoing debate for several decades. Once these dust grains reach millimeter-sizes, they tend to either inspiral into their star due to gas drag, or collisionally fragment before they can grow to larger sizes. Planet forming disks are dynamic systems driven by the interaction between dust and gas. When a resonance occurs between the two, a wide variety of feedbacks create an unstable environment with numerous fluid instabilities. The disk settling instability is one important feedback which occurs as dust vertically settles onto the midplane of the rotating disk. We examine whether the disk settling instability can play a significant role in clumping dust grains into high density regions, which can then either amalgamate grains or directly collapse into kilometer-sized boulders. We set up relevant initial conditions and analyze numerical simulations of dust-gas mixtures to better understand the effects of the disk settling instability. Exploring the various clumping and structures that emerge in our simulations can help us piece together the early staent.

Learning Visually-Guided Latent Actions for Controlling Assistive Robots
Albert Zhai
Mentors: Dorsa Sadigh, Dylan Losey, and Animashree Anandkumar

Assistive robotic arms enable people with disabilities to perform physical tasks on their own. However, teleoperation of such dexterous arms is challenging, as even basic manipulations require the human to coordinate numerous degrees of freedom. Latent action methods aim to solve this problem by leveraging autoencoder neural networks to learn a low-dimensional latent space of robot actions. Joystick inputs in the latent space can then be decoded to the corresponding high-dimensional actions for intuitive teleoperation. However, existing latent action methods focus on environments with deterministic initial states and cannot be applied when object properties are unknown beforehand. We extend the latent action approach by incorporating visual perception for sensing the environment state. We compare multiple vision network architectures and conclude that explicit object location combined with semantic category provides the most effective state representation for learning useful conditional latent actions. We evaluate our visually-guided latent action teleoperation scheme on manipulation tasks both in simulation and using a real robotic arm. We find that the network is able to learn actions that successfully adapt to different object arrangements.

Transfer Learning for Mouse Behavior Annotation
Isabella Zhang
Mentor: Pietro Perona

In behavior neural science, many videos are typically recorded during experiments to study neural basis of behavior. Biologists can annotate these behaviors manually, but the process is time-consuming and expensive. To help reduce the cost of these annotations, we aim to automatically generate behavior labels given tracking data from experiments. Since the data is labeled by different annotators with variability across annotation styles, it is difficult to train fully supervised models separately for each annotator. We investigate using transfer learning to transfer information from one annotator to another so that a fully trained model from a single annotator can be used to train models for a different annotator with less data. Using a model trained from one annotator with more data, we train a new model for a different annotator starting with the weights of the first model and explore the effect of different features. We anticipate that if there are promising results in this direction, classifiers can be shared more easily across annotators, which can improve efficiency in behavior annotation and reproducibility in science. Preliminary results show promising results that the classifiers have information that can be shared between annotators.

Theoretical Investigation of Oxidative Addition Mechanism in Decarboxylative Asymmetric Allylic Alkylation of β-Ketoester Substrate Class
Tianyi Zhang
Mentors: Brian Stoltz and Alexander Cusumano

Prior mechanistic investigations into the mechanism of decarboxylative asymmetric allylic alkylation of β-Ketoester have revealed a selective *trans*-N *anti*-displacement to incur the lowest activation barrier. In order to determine whether this observation arose from an artifact of the choice of DFT functionals, a *trans*-effect argument, or the inherent electronic and/or steric properties, DFT computations were performed on a simplified allyl acetate substrate with Pd(dmedmp) and Pd((*S*)-*t*-BuPHOX) as two contrasting catalyst systems – the geometries optimizations and frequency calculations were carried out at PBE0-D4//def2-SVP/def2-ma-SVP[O]/def2-TZVP-ECP[Pd] level of theory while single-point calculations on all stationary points were implemented at PBE0-D4//def2-TZVPP/def2-ma-TZVPP[O] level of theory – to calculate $\Delta G^{\ddagger}_{298}$ of *syn*- and *anti*-displacement for four configurations in gaseous and THF(CPCM) phases. Preliminary data revealed significantly lower activation barriers for *anti*-displacement in both Pd(dmedmp) and Pd((*S*)-*t*-BuPHOX) catalyst systems and in both solvent environments, in close agreement with prior investigations. Interestingly, for *anti*-displacement, *trans*-P was kinetically favored over *trans*-N pathway for Pd(dmedmp) in both gaseous and THF-solvated phases, while *trans*-N displacement had a moderately lower activation barrier for Pd((*S*)-*t*-BuPHOX), implying that *trans*-effect alone may be insufficient to explain the observed regioselectivity. Further investigations into the molecular orbital picture of the regioselectivity may provide insights into improving future catalyst designs.

Pseudoentropy and Log Space Computation
Tina Zhang
Mentor: Chris Umans

Answering the question 'Are deterministic algorithms as powerful as randomised algorithms?' would solve an important open problem in complexity theory (namely BPP ?= P). Conventional wisdom indicates that the answer o this question ought to be 'yes' (so that BPP = P). In this project, we study the *space-bounded* version of this problem, which asks, 'Are deterministic algorithms which use a very small (logarithmic in the problem instance size) amount of memory as powerful as randomised algorithms using a similar amount of memory?' This question is known as BPL ?= L, and it is believed to be easier to address than BPP ?= P, because the conclusion

BPL = L, unlike BPP = P, does not imply strong circuit lower bounds which the community currently believes to be far off. A popular technique for attempting to prove either BPP = P or BPL = L is that of constructing *pseudorandom generators*. In this project, we investigate a weaker form of these objects, namely *pseudoentropy generators*, and find that pseudoentropy behaves very differently in the setting of space-bounded computation (BPL ?= L) as compared with the time-bounded (BPP ?= P) setting.

Smooth Pathing and Dynamic Learning for Swarm Spacecrafts
David Zheng
Mentors: Soon-Jo Chung and Yashwanth Nakka

Swarm formation flying spacerafts consists of hundreds or thousands of small spacecrafts that orbit around a main/chief spacecrafts. A swarm's flexibility of equipment on spacecrafts, redundancy of spacecrafts, along inexpensive individual spacecrafts has made them far superior then their traditional monolithic counter parts. Despite the main benefits, swarm formation flying spacecrafts suffer from a complex guidance system. In our research, we implement a waypoint guidance system that takes in coarse motion plannings from the planning algorithm and uses smoothing and filtering to communicate desired states to the controller in the Ros hivemind package. This allows the movement of the spacecraft in simulations and in the spaceroom to be smoother and more defined for each time step (0.1 sec). In addition, learning algorithms are implemented to model unknown environment including determining waypoints for the controller based on vicon data. We analyzed and compared our model done using Gaussian Process (GP) and Robust Regression (RR). We analyzed the effects of dataset size on the efficiency and accuracy of the two different models in addition to the different parameters (ie learning rate, neural network layer, etc) on the RR that affects how it performs.

Interactive Robot Feedback System for Stroke Rehabilitation
Emily Zheng
Mentors: Laurel Riek and Aaron Ames

Serious games for stroke rehabilitation have been explored as a means of increasing the motivation and interest of a person engaging in stroke rehabilitation (PESR) by performing repetitive motions to regain motor control. These games are often simple and lack customization, and there are still open questions as to the best approaches and mechanisms for providing feedback. We have designed a novel interactive feedback system, embodied on a Double robot with an iPad display. As a PESR walks on a treadmill, the robot calculates the magnitude of the ankle force produced by the target leg as a measure of user performance and learning, and provides adaptive feedback in the form of a game. We built a framework for this application with customizable modalities of feedback so that they may be tested in future rehabilitation sessions, to develop a learning paradigm for walking. This system will ensure that PESR are more actively engaged during rehabilitation and promote long-term retention of motor learning.

Superconducting Qubits Control and Measurement Optimization
Joy Zheng
Mentors: Oskar Painter and Xueyue (Sherry) Zhang

Precise and high-fidelity control and measurement of superconducting qubits are crucial to realizing scalable and reliable quantum computers and algorithms. Control of multiple qubits can unleash unprecedented ability but also poses challenges to quantum control due to the presence of crosstalk from multiple sources. When operating multiple flux-tunable qubits, flux bias applied to one qubit may thread through an untargeted qubit and cause unwanted qubit frequency shift, known as flux-bias crosstalk. The first part of my project involves developing a software control package that characterizes and compensates the flux bias crosstalk between multiple qubits. Besides controlling qubits, measuring qubits' states at the end of an operation is also crucial to developing quantum algorithms. The second part of my project involves optimizing the readout pulse by changing its amplitude and length and simulating the resulting effects. We found that using an initial ring-up pulse with a larger amplitude than a constant square pulse leads to faster readout. We also characterized multi-qubit readout crosstalk by calculating the average dephasing rate. Lastly, we aimed to develop an automated qubit calibration process.

Irreducible Characters of GSp(4) Over Finite Rings
Daniel Zhou
Mentor: Dinakar Ramakrishnan

The analysis of matrix groups with coefficients over finite and p-adic fields has been developed. We know a lot about the complex representations of such groups. However, on the contrary, considering matrix groups over rings such as the p-adic integers Z_p is not well-developed. Since Z_p is the inverse limit of the finite rings $Z / p^k Z$, we will consider matrix groups over these finite rings as a first step. We will look into the case of the general symplectic group over finite rings $Z / p^k Z$ where $k > 1$. To simplify, we will consider only the case where p is odd since the case $p = 2$ is more nuanced. For future work, the case for $p = 2$ can be considered or the characters of the matrix group over Z_p can be described in terms of these results.

Molecular Sensing With Cascaded Half-Harmonic Optical Parametric Oscillator
Selina Zhou
Mentor: Alireza Marandi

Optical parametric oscillators (OPOs) are laser systems that convert an input laser wave with some frequency into two output waves of lower frequencies by means of second-order nonlinear optical interactions. The Marandi lab has previously built a mid-IR femtosecond frequency combs generator using parametric down-conversion of OPOs. Generally, the principle spectral region of interest for molecular spectroscopy is at wavelengths greater than 3μm, where molecules exhibit stronger absorption cross sections. However, many recently developed spectroscopy techniques still lag the more traditional Fourier Transform Infrared Spectroscopy due to limited spectral coverage into the mid-IR. Given that the output power of OPOs are cavity-length and wavelength-dependent, instead of generating and directly measuring an absorption spectrum, we propose that molecular sensing could be done through analyzing only the output power peaks as a function of cavity length detuning. We simulated the absorption and phase information of gas molecules using a generalized Sellmeier equation and placed the molecules in the cavity of a 4μm OPO. At the time of writing, a noticeable difference in the power peaks of OPOs with and without carbon dioxide molecules was observed.

Understanding Contrail Parameterizations in Linear Contrail Models
Fangyu Zou
Mentor: Yuk Ling Yung

Aircraft emissions have been proven to have a significant impact on the atmosphere and climate. The contrails that are emitted spread, contributing to the global fractional cloud coverage and increasing the radiative forcing at the top of the atmosphere. However, the extent of radiative forcing from linear contrails still remains uncertain. Using modeling is an effective way in estimating radiative forcing and determining its sensitivity to certain parameters. However, there exists uncertainties in models due to assumptions, remote sensing, and poor parameterization. In this study, 3 prominent linear contrail models (NASA Langley, Texas A&M, NCAR) and their calculations for net global linear contrail radiative forcing values are analyzed. The NASA Langley Model uses the Fu-Liou radiative transfer model with contrail parameter additions. The Texas A&M model runs a Global Climate Model (GMC) with a RRTMG Model containing contrail properties to simulate global contrail radiative forcing. The National Center for Atmospheric Research (NCAR) model runs an advanced GCM that incorporates aircraft emissions from datasets to simulate climate and chemical effects from linear contrails.

Determining the Precision of Prime Focus Spectrograph Measurements
Shion Andrew
Mentors: Evan Kirby and Mia de los Reyes

The PFS (Prime Focus Spectrograph) is a multiplexed, optical and near-infrared spectrograph designed to exploit the unique wide field capabilities of the Subaru Telescope. Starting in Fall 2022, approximately 100 nights will be dedicated toward measuring radial velocities and detailed chemical abundances of metal-poor red giants in the Milky Way and M31 to infer the dark matter mass profiles and evolution of these galaxies. In preparation for the survey, we quantified the precision to which radial velocities and chemical abundances can be measured with the PFS for different observational conditions. Using a grid of over 300,000 synthetic spectra spanning 4100-9100Å, we interpolated template spectra for ~60 red-giant stars of various temperatures, surface gravities, and chemical abundances. We used the PFS spectral simulator to calculate the expected signal-to-noise ratio of each pixel for a given template spectrum and set of observational conditions. We generated multiple realizations of the expected noise for each spectrum, and used a chi-squared fitting pipeline to measure the chemical abundances and radial velocities among the distribution of noisy spectra. We interpreted the standard deviation of the resulting measurements as the 1-sigma confidence interval that we can expect from random noise. In future work, we plan to additionally estimate the magnitude of systematic errors, such as template mismatch and imperfect sky subtraction.

Covariant Approximate Quantum Error Correcting Codes for Qubits
Galit Anikeeva
Mentors: John Preskill and Sepehr Nezami

The development of quantum error correcting codes is necessary for building quantum computers.
In a quantum error correction protocol, we usually talk about encoding one logical space into several physical systems. For a protocol to be useful, we should be able to correct the error caused by noise in the physical systems. Additionally, we should be able to modify the logical space through simple operations on the physical spaces. Unfortunately, the Eastin-Knill theorem tells us that the we cannot have a protocol where these two things happen perfectly. In this project, we expand the results of arXiv:1902.07714. We consider approximate error correction instead of perfect correction, and we study how well can we correct under the assumption that any gate on the logical spaced can be transformed into local operations in the physical systems. In particular, we expand on the case of the qubit system, which was not previously covered. We apply these results considering erasure and angular momentum errors. Finally, we consider the problem of implementing the protocols in Schwinger boson systems.

Analysis of Fracture Orientation Trends on Earth as a Potential Analogue for Tessera Terrain on Venus
Johanna Baraga
Mentors: Joann Stock and Daniel Nunes

Due to the relative lack of recent missions, datasets imaging the Venusian surface are lacking either in extent or resolution. Therefore, in order to learn more about features on Venus, it is beneficial to study analogous features on Earth where more data are available. Three sites on Earth were chosen that display multiple intersecting fracture groups similar to heavily deformed areas of Venus known as tesserae. To determine the deformation history, the fracture patterns at each site were digitized and the orientations determined. These fractures were then put into a rose diagram to categorize sets of similar strikes. Observations of cross-cutting relationships were made to determine the relative ages of each fracture set. Indian Cove, an area of Joshua Tree National Park, was found to contain two main sets of fractures, one striking between 060 and 070 (or 240 and 250) and the other striking between 150 and 160 (or 330 and 340). A third but much smaller fracture set was found striking 110 and 120 (or 290 and 300). The first two sets of fractures are oriented 90° from each other, indicating they are likely the result of jointing. The cross-cutting relationships were inconclusive due to the lack of field data. Analyses of other sites are ongoing.

Developing a Sprayable Antiviral Coating in Response to the SARS-CoV-2 Pandemic
JC Daniel Calso
Mentors: Robert H. Grubbs and Christopher Marotta

The public health crisis of 2020 has dramatically altered daily life on a global scale, due to the deadliness of the novel 2019 coronavirus. The highly contagious virus spreads primarily through airborne droplets but infection is also possible via secondary “fomite” transmission; the virus remains viable and infectious on hard, nonporous surfaces for up to 72 hours. While surfaces can be sanitized with commercially available cleaning agents, it is unfeasible to constantly sanitize high-touch public surfaces such as doorknobs, buttons, and handrails. The Grubbs Research Group has thus proposed a long-lasting sprayable antiviral coating that can be applied to surfaces with high exposure to combat this secondary route of infection. It disrupts viral membranes using quaternary ammonium cations often found in consumer cleaning/disinfecting agents; these are coupled to a polypropylene backbone, imbuing its longevity on plastic surfaces. Multiple derivatives of varying carbon lengths off the quaternary ammonium group were synthesized and tested for their antiviral capabilities. The compound adheres strongly to common plastics, including those found in frontline medical worker PPE, but its ability to adhere to other substrates such as metal requires modification. Commercially available adhesion promoters were included and compared using industry standard adhesion test protocols. A digital image processing protocol was created to quantify adhesion testing results.

Potential Subduction in Western Astkhik Planum, Venus
Morgan Carrington
Mentor: Joann Stock

Venus is Earth's closest neighbor, yet the surface and geologic processes that have caused its deformation remain largely in question. A major part of this question centers around subduction areas: their formation, the structures that result from them, and how they contribute to the current deformation of the surface. The western portion of Astkhik Planum, where the Vaidilute Rupes trends north-northwest until being covered by lava flows, is a suspected subduction zone. My project utilizes synthetic aperture radar (SAR) image data from the 1989 Magellan mission to map and digitize structural features in this area that suggest compression or extension. Further understanding of this region and its features as an area of potential subduction presents the opportunity for other forms of research to support or disprove the hypothesis.

Efficient, Scalable Numerical Simulation of Formation Flying Spacecraft Swarm
Jacqueline Castellanos
Mentors: *Soon-Jo Chung and Kai Matsuka*

Multi-spacecraft mission architecture has various advantages such as robustness to individual spacecraft loss, high spatial-temporal coverage, and more science return. On-board guidance, navigation, and controls (GNC) algorithms for spacecraft swarms are the key technology that enables such missions. To validate the candidate spacecraft formation flying technology, we must validate the autonomous algorithms in realistic environments; however, validation of such GNC algorithms for large scale spacecraft swarms is not an easy task itself. Flight experiments in-space or even the ground-in-the-loop control of the individual spacecraft are technically challenging. A cost-effective alternative is high fidelity simulation model of the formation flying spacecraft. Such simulation architecture must account for high fidelity environmental models while ensuring that the simulation framework scales well for a large-scale swarm. The objective of this project is to develop and validate a python-based spacecraft swarm simulator and apply this simulator to ongoing formation flying missions, such as the Distributed Aperture Radar Tomographic Sensors (DARTS) mission at NASA JPL. The mission plans to use formation flying CubeSats to study the three-dimensional vegetation structure of rainforests. The data structures and speed available in Python were leveraged to improve the efficiency, flexibility, and scalability of the simulator.

Visualizing and Analyzing Pressure Fluctuations in Turbulent Channel Flows With Varying-Phase Opposition Control
Dory Castillo
Mentors: Beverley McKeon and Simon Toedtli

This study examines pressure fluctuations in a low Reynolds number turbulent channel flow with varying-phase opposition control. The goal is to analyze the pressure at the wall to understand the relationship between the wall-normal velocity and pressure. The data to be visualized and analyzed is collected from direct numerical simulations, where an active control scheme is implemented and another where there is a no-slip condition at the wall meaning no spanwise and streamwise velocities present at the wall and no normal velocity penetrating the wall; the former cases are associated with drag increase or drag reduction. We first validate the data with statistical analysis and methods by comparing it to written literature. Then, by switching from the physical domain to the Fourier domain, we calculate the amplitudes, phases and phase difference of the velocity and pressure profiles. The results demonstrate that in the drag increasing case the velocity and pressure demonstrate a specific correlation in the wall-normal direction. Understanding the physics between the phases of the wall-normal velocity and pressure can potentially lead to new methods of implementing the control scheme in an experimental set-up which eventually would be of great benefit to engineering applications.

Analysis of the Near-Wall Region in Turbulent Channel Flow With Wall Transpiration
Miya Coimbra
Mentors: Beverley McKeon and Yuting Huang

Surface roughness is known to significantly change the skin friction drag of bodies moving through a fluid, and as a result a significant area of research involves studying how roughness affects a flow. This work analyzes the effect that a transpiration boundary condition has on near-wall turbulent channel flow in an attempt to determine whether this condition can be used to simulate surface roughness. To study this, dominant wavenumbers from the Fourier series representation of velocity fields were compared to the wavenumbers implemented by the transpiration condition, and Reynolds stresses and variances of the streamwise and wall-normal velocities were observed as a function of channel height. Amongst significant results from the study, it was determined that immediate effects of the wall transpiration reach flow heights of approximately $y^+=20$. If wall transpiration effects can be accurately quantified in a turbulent flow, potential impacts include tuning the transpiration condition such that it can mimic the effects of different materials of various roughnesses. This would ultimate lead to lower experimental costs as new materials will not need to be manufactured and will increase the efficiency of experiments in roughness research.

A Low-Energy Saturable Absorber: Saturable Absorption via Nonlinear Mirror Modelocking in Nanophotonic Waveguides
Devin Dean
Mentor: Alireza Marandi

Existing saturable absorbers, key components in creating mode-locked lasers, typically require input energies in excess of 1 nJ in order for intensity-dependent absorption to shorten the pulse duration. We propose a saturable absorber in the form of a Lithium Niobate quasi-phase-matched nanophotonic ridge waveguide that has the potential to shorten pulses at input energies as low as 5 pJ. An adaptation of the nonlinear mirror mode-locking technique, the waveguide consists of a second-harmonic generation section, a small section that leads to a pi relative phase shift of the process and loss at the input wavelength, and a section identical to the first that due to the phase shift "undoes" the initial amplification of the second harmonic and transfers energy back into the input wavelength. Through numerical simulations we have demonstrated that this device can act as a low-energy saturable absorber. Such a device could enable a new class of compact low energy mode-locked lasers.

Towards a Deep Convolutional Neural Network to Predict Complex Contact Maps for Protein-Protein Docking
Serena Debesai
Mentors: Stephen Mayo and Aiden Aceves

Due to the inefficiency of experimental methods, protein-protein docking, a computational technique, has emerged as the method of choice for predicting protein-protein complex structures. However, the human proteome is vast, and the run time and accuracy of current protein-protein docking methods are not well-suited for proteome-scale screening for new pairs of protein-protein interactions. We propose using a deep convolutional neural network to improve the efficiency and precision of protein-protein docking. Previous work involving deep learning methods for protein-protein docking has only explored the use of sequence-based representations for network inputs. We use 3D structures for network inputs, allowing our model to learn spatial features that cannot be extracted from sequences; however, the sparsity of 3D structural representations leads to network collapse. To avoid this, our model predicts Complex-Contact Maps (Co-COMAPs), a 2D structure intermediate that maps interactions between residues of each protein in a complex. We trained several network variants of three popular high-level deep convolutional neural network architectures, the ResNet, UNet, and GoogLeNet, to optimize various hyperparameters such as the loss function, optimizer, and network depth. In the future, we will select the best network and continue to refine the network architecture to improve performance.

Calculating Reaction Barriers for Electrochemical CO_2RR on α-Brass Surfaces With Aryl-Pyridinium Molecular Additives
Cassandra Decosto
Mentors: Jonas C. Peters and Nicholas B. Watkins

Due to the increase of carbon dioxide (CO_2) in the atmosphere, the need for sustainable, economical, and long-lasting environmental change is crucial. One potential energy-efficient method is to sequester CO_2 and catalytically transform it into valuable feedstock chemicals such as C_2H_4, C_2H_5OH, and C_3H_7OH. Catalysts, such as Cu-Zn alloys, have shown promising results in electrochemically reducing CO_2 (CO_2RR) to C_{2+} products at low overpotentials. However, the competitive H_2 evolution reaction (HER) has limited achieving higher reaction selectivity. Electrodimerized aryl-pyridinium additives on copper electrode surfaces have been shown to suppress HER and increase faradaic efficiency for C_{2+} products. Here we investigated the effect of these molecular additives on α-brass surfaces using Vienna Ab-Initio Simulation Package (VASP). By computing the binding energies and transition state energies associated with the key C–C coupling step, we have worked to obtain insight into whether the addition of the molecular additive would afford higher faradaic efficiencies for C_{2+} products with this catalyst. Our computations so far have set up a framework for calculating the barrier towards C–C coupling at catalytically relevant potentials with and without molecular additives.

Developing Reliable Testing Methods for Autonomous Vehicles in an Automated Valet Parking Lot
Berlin Del Aguila
Mentors: Richard Murray and Apurva Badithela

Incidents in which autonomous systems have malfunctioned in real-world scenarios have shined the light on the importance of formal verification, testing and evaluation (T&E) for the development of reliable autonomous systems. Through the development of four testing scenarios, the T&E team at Caltech will apply these tests to an automated valet parking case study to gather data to finetune further tests, and compare with other formal T&E methods. We hope to show that our test generation scheme results in test runs that collectively cover a designated set of behaviors while bringing the AVP system as close as possible to violating its specifications. To do so, the T&E team developed an abstraction of the AVP case study that is consistent with the specifications of the actual AVP model and interfaced the AVP model's system code to the abstraction framework of the T&E algorithm and simulated the outcomes from the testing scenarios. Upon successful testing and simulation, the team developed a ROS(Robot Operating System) module for the AVP environment to be interfaced with the ROS module for the AVP system developed by the AVP team and construct simulations of the test cases in Gazebo.

The Stochasticity of r-Process Enrichment Events in Galaxy Formation
Yadira Gaibor
Mentors: Philip Hopkins and Andrew Emerick

A substantial fraction of heavy elements (heavier than iron) are a result of rapid neutron-capture (r-process). We know that these elements were produced somewhere along galaxy formation and were distributed as it evolved. However, the rates and production sites of such enrichment are still being studied. Spectroscopic observations of Reticulum II and detection of gravitational waves from a neutron star merger (NSM), seem to suggest that that NSM are the main production sites of r-process elements. Due to the rarity of these events, it is possible that they are stochastic in some regimes. However, there are questions about where they come from and what signatures correspond to said processes, which are the main motive of our project. We use a post-processing method (age tracers) on three simulated galaxies (m10q, m11i, m12q) of different sizes (halo masses: 8.3E+09, 7.2E+10, 1.8E+12 M⊙) from FIRE-2 to study the stochasticity of r-process events. We explore four r-process fields with different yields and delay times. Then, we compare the simulated and post-processed abundances to determine if there were any r-process events and how many of them could have occurred. We are currently analyzing some "critical cases" in which the simulation and age tracers disagree by more than 3 dex. Our goal is to better understand the yield of r-process elements and their rarity across the different types of galaxies.

A Three-Dimensional Convolutional Neural Network to Predict Fluorescent Protein Maximum Emission Wavelengths
Michelle Garcia
Mentors: Stephen Mayo and Aiden Aceves

The scarcity of far-red emitting fluorescent protein (FP) monomers is a major limitation for live-tissue imaging, where the near-infrared window is favorable for light penetration and necessary for deep imaging. The comparably smaller energy difference between the transition of HOMO and LUMO states in far-red FP fluorophores ensures a reduction of autofluorescence, light-scattering, and phototoxicity. We endeavor to contribute to the array of FP monomers with peak emission wavelengths (λ_{max}) between 650nm to 700nm by creating a three-dimensional convolutional neural network to predict FP λ_{max} given a molecular structure. We use TensorFlow, an open source machine learning platform, and its neural-network library to build Keras models consisting of 3D convolutional and pooling operations. Inspiration for network architecture was drawn from image classification state-of-the-art networks. While considering potential architectures, hyperparameters such as learning rate and dropout rate were automated using Keras Tuner's Hyperband algorithm. After training models on data from extant public databases, we will execute in-silico sequence perturbations and evaluate augmented molecular models. In addition to evaluating proteins with potential for subsequent synthesis and characterization, these experiments shed light on the use of three-dimensional convolutional models for the automated learning of features relevant to functional properties of biomolecules.

Flow Visualization of Vortices Transcribed by Artists
Victoria Ignacio
Mentors: Morteza Gharib and Chris Roh

Opposed to today's advanced, technological society, artists in the 18th and 19th centuries did not have the correct tools, besides their visual sense, to measure how accurately they were able to depict a natural phenomenon onto their canvas. Our current approach is to extract data points from the found artwork by artists, such as Leonardo da Vinci (15th century), Utagawa Hiroshige (Tokugawa Period, 1603-1868), and Ogata Korin (Tokugawa Period), who were able to transcribe visuals of visible vortex flow. The extracted data points are used to fit well-known spiral equations to determine the constancy in artists' visual measurements, which might reveal the innate bias in the way we perceive and depict the swirling motion.

A Search for Surviving Stars From SNe 1a in PTF and ZTF Surveys
Conor Larison
Mentors: Mansi Kasliwal and Danny Goldstein

Despite the importance of Type 1a Supernovae to the field of astronomy, the progenitor system that leads to SNe 1a is still a mystery. To determine which candidates for the progenitor system have merit, we measure high proper motion stars in Zwicky Transient Facility (ZTF) and Palomar Transient Factory (PTF) survey data to discover white dwarf candidates with proper motions above 100 mas/yr that are polluted with the products of SNe 1a, as these velocities and chemical compositions could only be attributed to the survivor of a SN 1a explosion. We have begun to measure the proper motions of thousands of stars and have calibrated our procedure with data from Gaia. We hope to achieve a depth of around 21.5 magnitude that would allow us to make accurate proper motion measurements for thousands of white dwarf candidates not visible to Gaia.

Separating Photon-Initiated and Neutron-Initiated Showers in the Light Dark Matter Experiment
Juan Lazaro
Mentors: David G. Hitlin and Bertrand Echenard

The Light Dark Matter eXperiment (LDMX) will use a fixed target approach to explore light dark matter and various scenarios of sub-GeV new physics. One class of new physics model consists of axion-like particles. I present an initial study of how machine learning could make use of the hadron calorimeter to separate the photon-initiated shower signature of axion-like particles from neutron-initiated shower background. Using an LDMX-specific variant of Geant4, both shower types have been simulated at various energies, directions, and initial positions. Distributions of values derived from these simulated events are analyzed, and those that show the greatest contrast between the showers of different particle origin are added to the feature list of a boosted decision tree. At this point, it has been possible to achieve a 10^{-4} rejection of neutron showers while preserving 55.2% of photon-initiated showers. Further development of the boosted decision tree is required, and a deep learning option may be explored to determine if it will be possible to improve sensitivity for testing axion-like particle models.

Implementation of Acoustically Targeted Chemogenetics (ATAC) in Mice and Nonhuman Primates
Christina Lee
Mentors: Mikhail Shapiro and Richard Li

Abstract withheld from publication at mentor's request.

Simulations of Electric Field in the Channel of a High Electron Mobility Transistor
Rohan Lopez
Mentors: Austin Minnich and Tomi Esho

A high electron mobility transistor is a field effect transistor that utilizes an interface between two materials of different band gaps as the channel. Due to their excellent transport properties, HEMTs find use in applications where low noise and high frequency is desired, such as radio astronomy, quantum computing and telecommunication. High electric fields within the channel of a HEMT is a common and undesirable effect. These high fields can lead to low breakdown voltage and degraded noise performance. This work will be carried out on the Sentaurus TCAD (Technology Computer Aided Design) program. This software is used for running simulations of thermal, electrical, and optical characteristics of semiconductor devices. Once a baseline model is made and several simulations have been run to ensure results are obtained as expected, appropriate modifications will be made to determine what techniques would potentially give the best results with respect to reducing peak electric fields in the channel of the HEMT.

Small Hot Spot Ignition Study
Alan Maida
Mentors: Joseph Shepherd and Silken Jones

Hot spots, which can be caused by electrical failure, can lead to ignition in the presence of vaporized fuel. This makes it an area of interest for aviation purposes due to the proximity of electrical wiring to the fuel tank. In this report, we'll be analyzing the critical conditions of ignition due to hot spots on a steel disk. Hot spots will be created utilizing resistive heating on a steel disk. To understand how heat is transferred away from the central hot spot, we worked in detail with the heat equation. The heat equation has shown that having a disk with varying thickness allows for a higher max temperature by trapping heat and allowing for a smaller hot spot. We have created a python script for a 1D heat equation to test the effects of varying certain parameters and to be able to track certain variables that are a function of temperature. Findings have shown that assuming certain variables such as electrical resistivity to be constant can lead to a $68\% \, error$ when calculating max temperature.

Developing a Simulation Platform for Flow Phenomena With Langranian/Eulerian Frames of Reference
Kelvin Martinez
Mentors: Mory Gharib and Cong Wang

Fluid mechanics has come a long way when studying its properties. To study the fluid motions, neutrally buoyant particles are often added to the fluid media as marks for tracking the fluid motion. Two common frame references, namely the Lagrangian and Eulerian frame of reference, offer two different perspectives on the fluid flow. These frames of references offer us the ability to analyze the flow motion, either following a specific fluid parcel over time (Lagrangian frame) or focusing on a selected region where the flow passes through (Eulerian frame). My work targets to build a virtual platform to study fluid motions with these two frame references. First, synthetic particle images that correspond to various flow phenomena, such as a linear translation or a rotational vortex, are

produced. Second, the flow motions are quantitatively extracted using cross-correlation. Lastly, the flow motions are visualized with the Eulerian/Lagrangian frame of reference. This platform is expected to be a handy tool, which produces insights in fluid dynamics research and education.

Investigating Returning Radiation in Black Hole X-ray Binaries Using NuSTAR Data
Jessie Miller
Mentors: Fiona Harrison and Javier Garcia

Through reflection spectroscopy, we can study the way light behaves in the inner regions of a black hole's accretion disk. In the average stellar-mass black hole we observe distant reflection from photons emitted at the outer regions of the disk, relativistic photons from the inner regions of the disk, and photons emitted by the corona. However, when the gravity of a black hole is particularly strong, photons emitted from the inner regions of the disk may be bent backwards, twisting around the black hole as it rapidly rotates. These photons are then reflected off of the accretion disk before finally reaching the observer. This process, known as "returning radiation," creates a self-illuminating disk first theorized in the 1970s and recently observed for the first time this past year. This project aims at identifying returning radiation in additional black hole binaries, thereby further developing our current understanding of black hole physics, accretion dynamics, and aiding in future measurements of black hole spin rates.

Comparison of Multiphase Flow Simulations of Acoustic Wave Propagation Through Bubbly Liquids
Franz O'Meally
Mentors: Tim Colonius and Spencer Bryngelson

Multiphase flows are found in engineering systems from turbomachinery to medical devices. To enhance our understanding of these flows and design future systems, it is necessary to develop accurate techniques for their numerical simulation. In this project, the Multi-component Flow Code (MFC) is used to model pressure wave propagation through bubbly liquids. The results of these simulations are compared to experiments from the literature. A newly implemented feature of the algorithm involves a statistical approach for the bubble state. Specifically, the initial bubble mass, radius, and radial velocity are taken as random variables whose distribution is tracked during the simulations, using a quadrature-based moment method (QBMM). This more general modeling of bubbles should allow for increasingly realistic simulations, since except under carefully controlled laboratory conditions, the occurrence of bubbles in real flows is a stochastic process. Initial findings using MFC show well established pressure oscillations during steady shock propagation in both monodisperse and polydisperse bubbly liquids. Further simulations will be conducted to determine the effects of the new QBMM bubble parameters on the solution of shock propagation and bubble-screen problems, with the goal of improving the understanding of multiphase flow physics.

Implementing the SINDy Algorithm to Identify a Model for Passive Vortex Induced Motion of Airfoil
Patrick J. Saade
Mentors: Beverley McKeon and Morgan Hooper

Nonlinear dynamical systems occur in a variety of locations in the environment. Though very common, identifying a model for this kind of system can prove difficult due to their erratic nature and indistinct relationship between the measured variables. Despite the difficulty, identification of a model is still achievable through the implementation of machine learning. Machine learning is the perfect tool for analyzing nonlinear dynamical systems because a computer can determine patterns and correlations that researchers typically cannot. Our implementation of SINDy has been tested on the Lorentz system and will be applied to a fluid-structure problem, namely the pitch/heave motion of an airfoil moving in response to vortex shedding from an upstream body. With the proper utilization of SINDy in this nonlinear dynamical system, enhanced modeling and understanding of the system's dynamics is expected. This will have implications for tasks such as passively harvesting energy.

Can LYSO Target Help on Identifying Fake Dark Matter Events?
Julian Sennette
Mentors: David Hitilin, James Oyang, and Dexu Lin

The Light Dark Matter Experiment (LDMX) seeks to detect and characterize dark matter through searches within the sub-GeV mass range. Current detection, however, is it's susceptibility to false positives due to photonuclear and electronuclear interactions as high energy electrons (4 GeV) interact with the target, depositing energy back into the target which are not read by the detector. These missing energy signatures falsely appear as potential dark matter events. However the use of active scintillation material, LYSO, rather than the currently used Tungsten is theorized to detect the them scintillation processes which are able to be read by the detection apparatus. This project uses Geant4 simulation software, in order to generate emission and energy signatures for tests of LYSO's ability to accurately detect fake dark matter events and remove background compared to Tungsten. Data collection is done through ROOT6 softawre implemented within Geant4. Statistical significance analysis will be done along a

95\% confidence interval in which releveant paramters will be compared between and LYSO and Tungsten targets. While no results have been recorded, significant false positive rejection by LYSO could result in the implementation of active LYSO into the LDMX detection apparatus.

Evaluating Potential High Albedo Jupiter Trojan Asteroids
Anna Simpson
Mentor: Michael E. Brown

Current models of Solar System evolution propose that the icy Kuiper Belt Objects and the Jupiter Trojan asteroids share an origin. Therefore, finding a compositional connection between these two sets of objects is relevant to understanding Solar System formation. In the absence of any apparent ice on the surface of Trojans, it may still be possible that recent collisions exposed previously subsurface ice, which would be evidenced by higher than expected albedos. In general, Trojans are similar in albedo; however, previous analysis of data from the space telescope WISE has suggested that some Trojans may have albedos that are higher than typical. Combining the WISE data with additional data taken with the radio telescope ALMA, we investigate twelve Trojans in order to search for further confirmation of their albedos. Fitting a thermal model to the data results in a set of parameters, including albedo, that best predict the observed thermal flux. By further refining the fitting procedure in pursuit of the best possible model fit to both the WISE and ALMA data, we aim to obtain albedo values that will enable evaluation of the possibility that recent collisions exposed ice within these objects.

Identifying and Characterizing Dusty Stellar Candidates in Spitzer Kepler Survey (SpiKeS)
Sydney Skorpen (Jet Propulsion Laboratory/California Institute of Technology and Moorpark College)
Mentors: Farisa Morales (Jet Propulsion Laboratory/California Institute of Technology) and Varoujan Gorjian (Jet Propulsion Laboratory/California Institute of Technology)

Characterizing debris disks within the *Kepler* field can further our understanding of the composition and evolution of planetary systems. The Spitzer Kepler Survey (SpiKeS) dataset, which captures the *Kepler* field in the 3.6 and 4.5 μm wavelength, contains 889 sources and presents the opportunity to identify and analyze dusty disk systems. Here, a subsample of 889 sources was isolated by its likelihood for infrared excess, a key characteristic of debris disks. Spectral Energy Distribution (SED) graphs were produced for each of the 889 sources and were organized by characteristics, of which 146 sources were selected for their clean SEDs containing infrared excess. A visual inspection of the 146 sources showed that 10 sources contain no obvious signs of infrared contamination and will require further analysis as a candidate for containing dust. Another 30 sources exhibited unusual "rising" SEDs, many of which were galaxies as confirmed by the SIMBAD archive. The results thus far show that sources within the SpiKeS potentially include circumstellar dust, expanding our knowledge of the *Kepler* field and the planetary systems within.

Thermometry Program Optimization for Cosmic Microwave Background Kinetic Inductance Detectors
Ysaris Sosa
Mentors: Sunil Golwala and Fabien Defrance

Cosmic microwave background (CMB) is a form of electromagnetic radiation that originates from the earliest stages of the universe, and thus provides crucial information for dark matter and astrophysics research.
B-mode polarization of CMB targets a signal of universal inflation, or spatial expansion of the early universe. This mode is difficult to detect because it is both expected to be very small and also because other sources of radiation can also yield B-mode signals. A new B-mode polarization experiment, called CMB-S4, has been developed for this detection purpose. One potential technology for the detectors for CMB-S4 is microstrip-coupled titanium nitride kinetic inductance detectors because of the expected sensitivity and amenability to mass fabrication and multiplexing. In order for these detectors to work under their best conditions, they must be cooled to about 250 mK. The detectors are cooled and monitored by code written in the LabVIEW program. The program that was previously in place calculated resistance values, and subsequently converted them into temperature, rather inefficiently. Not only has this since been optimized, but an additional thermometer was added into the mainframe to allow for additional measurements.

Highly Efficient, Two-Dimensional Transition Metal Dichalcogenide Emitters for Luminescent Solar Concentrator Photovoltaics
Rachel Tham
Mentors: Harry Atwater and David Needell

Recently, monolayer transition metal dichalcogenides (TMDCs) have been shown to achieve high radiative efficiency, with MoS_2 and WS_2 TMDCs achieving photoluminescence quantum yields (PLQY) greater than 95%. Given these high radiative efficiencies, TMDCs can be used as active luminophores in luminescent solar concentrators (LSCs) that absorb and emit light as photoluminescence (PL) at Stokes-shifted energies coupled to bandgap-matched solar cells. This study aims to leverage experimentally validated modeling methods to computationally test the first-ever TMDC-LSC device, optimizing the device design and performance through a validated Monte Carlo ray-trace model. We analyze the TMDC-LSC performance through key metrics (e.g., optical efficiency, concentration factor, power conversion efficiency) and vary waveguide parameters, such as thickness and photovoltaic cell coupling, in addition to testing TMDC homo- and heterojunction luminophore structures. With the optimized TMDC-LSC device, we identify initial targets for experimental fabrication and prototyping. To further improve TMDC-LSC devices, future research can quantify the effects of external factors (e.g., temperature dependence and 2D TMDC strain) as well as develop initial TMDC-LSC devices.

Designing Corroles to Target Cancerous Tissue
Hoang Tran
Mentors: Scott Virgil and Harry B.Gray

In previous research from Prof. H. B. Gray's group, corroles and their metal complexes have shown promise in cancer therapy. In this project, corroles and their metal complexes that could have better activities against cancerous tissues have been studied. With a goal to target cancer cells specifically, our research has led to the design of corroles that are expected to be more directed towards the acidic environment of cancer cells. We describe our research on the design of novel corroles that include basic groups into the ring with strong potential for activity against cancer cells.

The Barrel Timing Layer's Design and Performance of The MIP Timing Detector for CMS Phase-2 Upgrade
Matthew Valdez
Mentor: Maria Spiropulu

This report summarizes the design and performance of the Barrel Timing Layer (BTL) of the Minimum Ionizing Particle Timing detector for the CMS Phase-2 Upgrade. The performance of the BTL reconstruction is evaluated by using the Monte Carlo simulation. Caltech's BTL design maintains the current performance of the CMS detectors in time resolution, efficiency and background rejection. Simulations of the detected time of arrival for a given BTL cell includes different outcomes the time of arrival of the first energy deposit in the cell, the time delay for the scintillation light to reach the silicon photomultiplier at the crystal ends, and the time delay effect depending on the signal amplitude and the discrimination threshold of the readout units. The efficiency of the BTL from the Monte Carlo simulations of single muon events is at 90%, which is described as the probability from cluster energy deposits that are larger and 3 MeV and are associated to a single muon track of transverse momentum (p_t) that is greater than 0.9 GeV.

Using Gravity Measurements to Map Structure in the San Gabriel and San Bernardino Basins
Valeria Villa
Mentor: Robert Clayton

Sedimentary basins such as the San Gabriel and San Bernardino Basins in Southern California are thought to trap and amplify seismic waves. Recent studies suggest that a magnitude 7-8 earthquake on the southern San Andreas Fault will exceed model predictions by a factor of 4. Due to the low seismic velocities and concave shape of the basins, seismic energy can be channeled towards the densely populated downtown Los Angeles area. Therefore, obtaining an accurate model of the subsurface structure of the basins, in addition to seismic velocities within the basins, can help to provide better estimates of ground shaking in the event of large earthquake. As part of the Basin Amplification Seismic Investigation (BASIN) project, we set out to incorporate the gravity data in these basins to provide further detail on the structure, thereby obtaining a 3-dimensional model of the subsurface structure. Here, we combine the sediment-basement interface obtained from P-to-S receiver functions and Bouguer gravity data along the ten seismic lines to determine the 2D structure along the profiles. We then interpolate these results to a 3D structure using gravity as a constraint.

Designing a Magnetic Field Coil for the Nuclear MQM Experiment
Maya Watts
Mentors: Nicholas Hutzler and Chandler Conn

The asymmetry between baryons and antibaryons remains mysterious, and violations in CP symmetry may hold the answer to why we live in a universe composed of matter instead of antimatter. One method of detecting CP violations is through cooling molecules and using tabletop experiments to measure their magnetic quadrupole moments (MQM). Researchers at the Hutzler Lab use this method to shed light on this puzzling question about the composition of the universe. Measuring a molecule's MQM requires a uniform magnetic field with little variation in order to preserve the precision of the measurement. A magnetic field coil where the wires are placed in a $\cos\theta$ distribution will generate a constant magnetic field perpendicular to the direction of the current. The magnetic field expression for a $\cos\theta$ coil is derived and simulated and two figures of merit that measure uniformity are established. Preliminary results reveal that the magnetic field generated by the $\cos\theta$ coil has satisfactory uniformity in the middle of the coil and that increasing the number of wires improves uniformity. Magnetic field uniformity in the future can be improved by machine learning algorithms.

Neural Network Control for Seismometer Temperature Stabilization
Ella Buehner Gattis
Mentors: Rana Adhikari, Koji Arai, and Tega Edo

Seismic movement is a source of noise in LIGO detector readings. Seismometers stationed around the detector monitor seismic activity. Temperature fluctuations add noise to the seismometer readings. The aim of this project was to design a control system to maintain a stable temperature within the seismometer enclosure, thus eliminating noise from the seismic data so that seismic noise can be removed from LIGO data.
Control systems can be divided into two categories; linear and nonlinear control systems. Linear control systems obey the principle of superposition and are governed by linear differential equations. Nonlinear control systems do not obey the principle of superposition and are governed by nonlinear differential equations. The objective of this project was to design a nonlinear control system using a convolutional neural network controller to improve disturbance rejection and decrease overshoot/undershoot in comparison to a linear control system.
The neural network was trained by reinforcement learning, using a custom OpenAI gym environment to train the network for temperature stabilization. The network was built with RLzoo's framework, which uses the tensorflow machine learning library.
The end goal is then to evaluate the performance of the NN controller and compare it to PID control in terms of overall efficiency.

Machine Learning Applications in Gravitational Wave Detection to Reduce Noise
Nadezhda D. Dimitrova
Mentor: Hang Yu

The LIGO gravitational wave detector is contaminated in the sub-60 Hz range by control noises of auxiliary channels. The paper explores how machine learning techniques can be applied to reduce excess noise induced by the nonlinear coupling mechanisms between channels. One phenomenon of interest is the bilinear angular noise, which is caused by coupling the high-frequency angular motion of the mirror and the low-frequency beam spot motion to create a length signal that mimics a gravitational wave. We use machine-learning techniques to try to identify the low-frequency component and it serves as a critical first step of removing the bilinear noise from the real data. Minimizing the instrumental artifacts is crucial and paves the way toward a much deeper understanding of gravitational waves.

Detectability of Quantum Effects in Gravitational Waves Emitted by Binary Black Hole Mergers
Zoe Haggard
Mentor: Alan Weinstein

Gravitational wave detectors such as Advanced LIGO and Advanced Virgo provide a test of the theory of general relativity in the strong-field, highly dynamical regime, such as in compact binary coalescences. General relativity, a purely classical theory, does not incorporate quantum mechanics. It is thought, however, that quantum mechanics must modify gravity; quantum uncertainty must manifest itself during the merger of two black hole horizons. These quantum mechanical effects could be observable in gravitational waves detected by LIGO as small perturbations in the signal waveform and higher harmonics, not explainable by classical general relativity. We propose to study the phenomenology and detectability of such quantum mechanical effects from binary black hole mergers for future LIGO observations.

Effects of Different Data Quality Veto Methods in the PyCBC Search for Compact Binary Coalescences
Brina Martinez
Mentor: Derek Davis

The PyCBC search pipeline has been used since the first gravitational wave detection made by Advanced LIGO and continues to be used today in the search for gravitational waves. To identify gravitational waves from compact binary coalescences, PyCBC runs a matched filtering and chi-squared (c^2) consistency test to determine significant signal-to-noise ratios and compares triggers to previously modeled templates. To confidently detect gravitational waves, we need to mitigate noisy data, which in return improves the sensitivity of searches. Current veto methods use data quality flags to veto and remove triggers in LIGO data that are believed to have terrestrial origin, though these methods risk accidentally removing signals and must be finely tuned to prevent a decrease in the search sensitivity. In this investigation, we test different veto methods based on the current set of data quality flags and detector characterization tools. We analyze how simulated signals are recovered by the PyCBC pipeline and the overall change in the sensitivity of the pipeline. Our results show an improved veto method that increases the significance of signals and the overall number of detectable signals without removing data. The results of this investigation can be implemented in the PyCBC search pipeline in future observation runs held by LIGO as a data quality tool to improve the search for gravitational waves from compact binary coalescences.

Measuring Impacts of Glitch Removal on Gravitational Wave Parameter Estimation
Lilah Mercadante
Mentors: Jonah Kanner and Alan Weinstein

No scientific endeavor ever runs flawlessly. There are always malfunctions and interference that cause the data to be less than perfect. In the case of gravitational wave data, one of the defects often found in the signals are noise transients, called glitches. These glitches are often difficult to model due to their non-Gaussian nature. It is not currently routine practice to remove them, although sometimes glitch subtraction must be done when the glitch strongly interferes with the signal. Each glitch is unique, and there are certain glitch parameters that cannot be found using an algorithm, because the algorithm has not yet been constructed. If we were able to calculate these parameters in an automated way, it would greatly improve the process of glitch removal. Additionally, the process of glitch subtraction has not yet been tested and documented in a systematic way. We hope to add to the documentation on the effects of glitch removal on parameter estimation by running parameter estimation on a data set of simulated signals with glitches injected at varying distances from the signal. We will then remove the glitch from the data and run parameter estimation on the clean waveform. This will allow us to study how the distance between the glitch and the signal plays a role in the accuracy of the parameter estimation. We have yet to draw any conclusions, but we anticipate that there may be a distance at which the glitch subtraction has negligible effect on estimating parameters.

Detectability of Nonlinear Gravitational Wave Memory
Darin C. Mumma
Mentors: Alan J. Weinstein, Colm Talbot, and Alvin K. Y. Li

Gravitational waves passing through a region of spacetime leave behind a permanent distortion, with strain typically on the order of 10^{-23} , the so-called memory effect. Linear and nonlinear components exist in gravitational wave memory, the latter appearing as a non-oscillatory, cumulative signal. Current gravitational wave detectors have not yet been able to reliably detect and isolate this low-frequency, nonlinear component which skews the numerical inferences of gravitational wave source parameters. Because this effect is cumulative, it is non-negligible, and its non-oscillatory nature distinguishes it from the rest of the waveform, making it detectable, in theory. Though previous studies have quantified and suggested improvements for the detectability of nonlinear memory, more templates and new data are available than ever before. In this project, we apply Bayesian parameter estimation to simulated gravitational waves from compact binary coalescences with memory to determine nonlinear memory detectability.

A Bayesian Approach for Constraining the Population of Kilonovae Based on Zwicky Transient Facility Searches for 13 Neutron Star Mergers
Priyadarshini Rajkumar
Mentor: Shreya Anand

In their third observing run (O3), the Advanced LIGO and Virgo interferometers detected gravitational waves (GWs) from several neutron star-black hole (NSBH) and binary neutron star (BNS) mergers. 13 of these GW events were followed up by the Zwicky Transient Facility (ZTF), an optical time-domain survey telescope, in search of kilonovae (electromagnetic counterpart to GW event). However, no kilonovae were found. Previous studies have shown how to interpret the non-detection of kilonovae in ZTF searches to constrain the underlying population of kilonovae. However, these studies adopted an empirical model parameterized by certain kilonova observables (e.g. peak absolute magnitude and decay rate). In addition, important GW event properties, which describe the likelihood of certain GW events being terrestrial and having implausible amounts of ejecta mass to power a kilonova, were not folded into the analysis. We build on these existing frameworks by using Bayesian statistical methods and combine ZTF O3 observations with GW event properties and existing kilonova model-space to constrain the same observed properties of kilonova population. In the future, our formalism could serve as a model for placing kilonovae population constraints and could help develop an effective strategy for EM-counterpart follow-up.

Studying the Properties of Higher Order Modes in Gravitational Wave Emission From Binary Black Hole Merger Events
Jennifer Sanchez
Mentors: Alan Weinstein and Colm Talbot

Advanced LIGO and Advanced Virgo have confidently detected gravitational wave signals from ten binary black hole mergers and one merger from a binary neutron star. Each observation contains encoded information about the physical properties of the binary system. As the detectors continue to improve their sensitivity, these developments will allow us to detect rarer systems and make more confident statements regarding their source properties. In order to fully characterize the gravitational wave observations, we rely on numerical and analytical models that approximate the signal waveforms from the emitted source as specified by the source parameters (masses, spins, sky location, etc.). The dominant emission frequency of gravitational waves from compact binary coalescence is at twice the orbital frequency; however, recently published events have demonstrated subdominant higher order harmonic contributions. The primary focus of this study is to explore higher order modes in gravitational wave signals with newly improved signal models.

Optical Refrigeration for LIGO Instruments
Samuel Schulz
Mentors: Rana Adhikari, Yehonathan Drori, and Christopher Wipf

The objective of this project is to determine the viability of using optical refrigeration (OR) to cool various cryogenic elements for Voyager LIGO. We began by exploring the existing literature on optical refrigerants and implementations of refrigerators to determine which could potentially be useful for LIGO. Next, we designed a demonstration that could be used as a proof of concept for real implementations of optical refrigeration using but expanding on existing data and techniques. The final project examined the viability of OR for particular applications, namely cooling the mirrors of the proposed phase-sensitive optomechanical amplifier (PSOMA) as well as the test masses. We found that with some improvements in crystal quality from those currently available, OR could increase the amount of cooling power on the PSOMA mirrors by a factor of ~2-10 from what is possible only using radiative cooling. We additionally hope to examine the effects of radiation pressure to show that no significant displacement noise would be added by the fluorescence from the crystal hitting the reflective surface of the mirror. These combine to show that OR may be viable to cool PSOMA mirrors in Voyager. We show that although we do not anticipate optical cooling to be viable for cooling test masses in the near future, it may be if the technology continues to progress.

Measuring the Hubble Constant With Dynamical Tides in Inspiraling Neutron Star Binaries
James Sunseri
Mentor: Hang Yu

The so called "Hubble Tension" is a large, newfangled problem in astronomy which has even larger cosmological consequences in its eventual resolve. Currently all calibration methods of the constant rely on the use of the electromagnetic spectrum-a method that does not rely on EM light but instead solely on gravitational radiation could prove extremely useful as both a backup for the unreliability of multi-messenger astronomy at its current state and as a new lens into cosmology which could potentially expand our understanding of light, gravity, and the universe. To extract a cosmological redshift from a gravitational waveform, one can look at both the point particle approximation phase contribution and the tidal phase contribution to the total phase of a gravitational waveform which allows for a break in the redshift degeneracy found in the mass parameters, which we can exploit to extract a cosmological redshift and thus the Hubble constant. Our analysis incorporates both f-modes and r-modes into the tidal phase contribution that are found in binary neutron star inspirals. We use the Fisher Matrix Analysis to generate our relevant possible errors on each parameter of the waveform.

Using Adaptive Filtering to Track Noise Lines or Signals With Varying Frequency and Amplitude
Rebecca White
Mentor: Ling Sun

Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) and Virgo have observed gravitational wave (GW) signals from compact binary coalescences. Yet undetected GWs from other types of sources, including quasimonochromatic continuous waves (CW) from individual spinning neutron stars and long-transient signals from newly born neutron stars, are of interest. One of the issues with the LIGO data that plays a role in making it challenging to detect these narrow-band GW signals is the instrumental spectral lines, which could obscure astrophysical signals at the frequencies where they occur. We use an adaptive filter, named "iWave", to track, study, and remove these instrumental artifacts from the interferometric data. Further, we explore the capability of iWave to track and detect weak CW or long-transient GW signals and quantify the sensitivity. This new tracking method, operating on time-series data, provides an efficient alternative to existing frequency-domain matched filter search methods. We discuss its application to low-latency follow-ups of binary neutron star postmerger remnants.

Do Binary Black Hole Merger Events Observed by LIGO and Virgo in their Third Observing Run Agree with Waveforms from General Relativity? A Study of Residuals
Erin A. Wilson
Mentors: Alan J. Weinstein and Dicong Liang

We present the study of fitness of General Relativity-predicted waveform models to binary black hole signal data from LIGO and Virgo's third observing run (O3). The data series observed by LIGO composed of both a merger signal and instrumental noise. Our hypothesis is that, should a waveform template predicted by General Relativity be fitted to the merger signal and then subtracted from the data series, what should be left over is a residual made up of pure instrumental noise. Thus, our objective is to determine if GR waveform templates accurately model O3 binary black hole signal data by creating a digital signal processing script that will derive a best fit waveform template from event parameters, subtract the template from the event data series, and then run a variety of statistical tests to check the consistency between the residuals and instrumental noises.

Validation of Analysis Tools to Characterize Infrasonic Wind Turbine Signatures
Thea Adumitroaie, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Adrian Stoica and Didier Keymeulen, Jet Propulsion Laboratory, California Institute of Technology

Remote sensing via high-altitude balloons enables the detection of infrasonic signatures that propagate through the atmosphere. Wind turbines are one source of infrasound and present specific interest due to their harmonic tendencies, interference with seismic observations, and adverse effects on human health. Verification of the analysis tools' validity within Python libraries is essential in accurately characterizing wind turbine signatures in a dataset. Thus, utilizing a pressure transducer specialized for infrasound detection to record a ceiling fan played an important role in writing scripts containing specific algorithms to manipulate and visualize the targeted source. The objective of data processing is to characterize the infrasonic signatures, which can be accomplished through the generation of periodograms estimating power spectral densities of the data for identification of prominent peaks, as well as the computation of spectrograms with consecutive Fourier transforms to identify harmonic tendencies. Furthermore, a data analysis processing pipeline was constructed and then tested on a raw data set (obtained from a previous high-altitude balloon flight managed by Sandia Laboratory). Wind turbine harmonics had only been found in the higher frequency band of 10 to 30 Hz. However, detection of wind turbines from their narrow-band infrasonic signature was visualized between 1 and 2 Hz, hence verifying the tools' effectiveness regarding data set manipulation.

Analysis of High-Resolution Mid-Infrared Data of Jupiter's High Latitudes to Determine Evolution of Thermal Structure and Chemistry
Zade Akras, Jet Propulsion Laboratory/California Institute of Technology & Yale University
Mentors: Glenn Orton and James Sinclair, Jet Propulsion Laboratory/California Institute of Technology

Jupiter's polar atmosphere is influenced by the acceleration of high-energy ions and electrons from its magnetosphere and potentially the external solar-wind environment. These charged particles deposit their energy through particle precipitation, dissociation, ion drag, and joule heating. Previous work used the TEXES instrument on NASA's Infrared Telescope Facility to study the unique chemistry and thermal structure of Jupiter's auroral regions. These retrievals must be repeated with an updated and consistent atmospheric model and using a spatial grid that is consistent with those adopted by the Juno Science Team. We used the Non-linear Optimal Estimator for Multivariate Spectral Analysis (NEMESIS) model to invert mid-infrared emissions spectra and retrieve atmospheric properties. We visualized the temperature profiles at 5 mbar, 1 mbar, 0.1 mbar, and 0.01 mbar with polar projections and investigated temperature as a function of time and pressure. Overall, temperatures were elevated in the auroral zone at 1 mbar, 0.1 mbar, and 0.01 mbar, peaking at 0.01 mbar during the summer of 2017. Temperatures inside and outside the auroral region were similar at 5 mbar. We identified significant temperature variation at 1 and 0.01 mbar, compared to moderate variation at 5 mbar and minimal variation at 0.1 mbar. Future work will turn to investigating spatial and temporal variations of the hydrocarbon distributions.

Modeling Wide-Angle and Window Function Effects on Matter Power Spectrum Measurement in SPHEREx and Roman Surveys
Joshua Benabou (Ecole Polytechnique, NASA Jet Propulsion Laboratory, California Institute of Technology)
Mentors: Olivier Doré, Chen Heinrich, and Henry Gebhardt (NASA Jet Propulsion Laboratory, California Institute of Technology)

Modern redshift surveys are capable of measuring the redshift of millions of galaxies, whose positions can be deduced from recession velocities via Hubble's law. We can thus map the galaxy density field, and perform on it statistical measurements such as the matter power spectrum $P(k)$, which offers a powerful test of cosmological models. Accurate power spectrum measurement must take into account redshift space distortions, whereby peculiar velocities modify the measured distances to galaxies via Doppler shift. A recent area of research attempts to quantify such distortions on large-scale surveys, in which the small angle approximation fails, and corrections to the power spectrum multipoles arise due to both wide-angle effects and coupling with the survey window function. We first validate a recently developed formalism to compute wide angle corrections for the BOSS survey, and model these effects on the planned all-sky survey SPHEREx and the Roman survey. SPHEREx, set for 2024 launch, will measure over 450 million galaxy redshifts from near-infrared spectra over a large cosmological volume at low redshifts. The mission aims to place strong constraints on the f_{NL} parameter characterizing non-Gaussianity in the distribution of inflationary fluctuations, which will require measuring the power spectrum with sub-percent level accuracy.

Unsupervised Learning Approach to Predict Exoplanet Host Stars From Stellar Elemental Abundances
Ye Won Byun, Jet Propulsion Laboratory, California Institute of Technology, Brown University
Mentor: Yasuhiro Hasegawa, Jet Propulsion Laboratory, California Institute of Technology

It has been well recognized that the chemical composition of stars can provide important clues about the properties of planets that orbit around stars. A famous example is that the Fe abundance of stars is positively correlated with the presence of Jovian planets. However, systematic exploration about potential correlations between the presence of exoplanets and other stellar elemental abundances has not yet been conducted. Thus, it is still inconclusive how

stellar composition can be used to infer the presence of planets. To address this issue, supervised machine learning techniques have recently been employed in the literature. However, current data sets do not have data points of true non-host stars; it is questionable whether a supervised learning model can accurately learn the features of a binary classification task with access to data points of only one kind of label (i.e. true host stars). Thus, an unsupervised learning technique is a much more appropriate approach to develop a recommendation algorithm for this task. In this project, we utilize a variational autoencoder approach, which uses dimensionality reduction to learn the parameters of the probability distribution of the data. This project aims to solidify the correlation between the presence of exoplanets and stellar elemental abundances and produce a list of stars, which currently possess no known exoplanets, that have a large prediction probability of hosting an exoplanet. We will improve target selection methods for discovering more exoplanets, by developing this predictive tool with a more solid theoretical background than past works.

Time Dependence of Near-Infrared Zonal Waves in Jupiter's North Equatorial Belt
Joheen Chakraborty, Jet Propulsion Laboratory, California Institute of Technology; Columbia University
Mentor: Glenn S. Orton, Jet Propulsion Laboratory/California Institute of Technology

Images of Jupiter's North Equatorial Belt (NEB, 7–17° N) display recurring atmospheric wave patterns in the 2-2.28 micron near-infrared range. We investigate the time dependence of the wave patterns using the Lomb-Scargle periodogram and Fast Fourier Transform methods, paying special attention to how their behavior correlates with periodic NEB expansions (NEBEs), which exhibit a 3-5 year cycle. Lomb-Scargle wave powers of NEB cylindrical maps display an apparent periodicity (with a frequency of ~3) during the time span between sequential NEBE start dates. Applying the FFT to composite maps covering all longitudes yields similar results as for the longitudes covered by single images. Moreover, the wave powers exhibit a longitudinal dependence, with higher System-III longitudes favoring higher wave powers throughout the entire periodic cycle. Further observations of trends in wave power and wavenumber will improve our understanding of NEB wave behavior, e.g. differences between measurements that sample different altitudes in Jupiter.

A Python-Implemented Deep Earth Water Model to Aid in Determining the Composition of Icy Ocean World Interiors
Andrew Chan[1,2],
Mentors: Mohit Melwani Daswani[2] and Steven Vance[2]
[1]Div. of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA 91125
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109

Currently, the Helgeson-Kirkham-Flowers (HKF) model is used to calculate thermodynamic properties of aqueous complexes and ions at differing temperatures and pressures. However, this model has certain regions of density and pressure where, for certain temperatures, it is inaccurate at predicting the speciation of solutes (Miron, Leal, & Yapparova, 2019). The combination of a different model for the properties of water alongside the Deep Earth Water (DEW) model (Huang & Sverjensky, 2019) could prove robust in determining the composition and evolution of icy ocean worlds interiors. We specifically explore the combination of SeaFreeze (Journaux, 2019), a Local Basis Function evaluation-based package that allows a user to examine icy polymorphs at different temperatures and pressures, with the DEW model. We present an object-oriented implementation of the DEW model which allows for convenient array inputs of custom water properties, specifically the Gibbs free energy, dielectric constant, and water density. Our implementation behaves identically to the Excel spreadsheet-implemented DEW model while additionally simplifying and streamlining the process to input reactions and make changes to calculations quickly. Additionally, our model builds in the full set of minerals from the thermodynamic database SUPCRTBL.dat (Zimmer et al., 2016), to provide an easy way to access a full range of minerals thermodynamic properties without having to calculate them separately. We compare the calculated properties of water from the DEW model and to those calculated by SeaFreeze. We find discrepancy between the two models at high pressure and temperature that possibly indicates a need to reevaluate the parameterization of the DEW model. We also present a set of reactions relevant to icy ocean world interiors calculated with the SeaFreeze-predicted ΔG_r of water incorporated into the Python implementation of the DEW model and a cross-comparison between these calculations and those with the "original" DEW model implementation.

A part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Developing the Sphinx-Based Avionics Ecosystem for Space Exploration
Sandra Chea, California Institute of Technology
Mentor: Yutao He, Jet Propulsion Laboratory, California Institute of Technology

Sphinx is an advanced low SWaP (**S**mall **W**eight **a**nd **P**ower) single-board computer technology developed at JPL for deep space CubeSat/SmallSat exploration. To facilitate rapid prototyping and constructing complete avionics systems that meet various mission requirements at reduced cost, time, and risks, we have been developing a Sphinx-based avionics ecosystem that includes modular, reusable, and configurable reference designs composed of different peripherals. In this project, we have developed three reference designs: (1) motor control, (2) camera

capturing, and (3) CAN communication in the context of space missions. First, we have built and tested the Arduino-based reference designs; Second, we have identified and documented detailed differences between Arduino and Sphinx with respect to both hardware (pin-level signalmapping) and software (register-level memory-mapping); Third, we have been migrating these reference designs from Arduino to Sphinx. The methodology proves to be effective and the generated work contributes to further evolution of Sphinx-based avionics ecosystem for upcoming NASA/JPL space missions.

Evaluating Operational Implementation of the Deep Space Network's Delay Tolerant Networking (DTN)

Hannah Li Chen, Jet Propulsion Laboratory, California Institute of Technology
Mentors: J. Leigh Torgerson and Nathaniel J. Richard, Jet Propulsion Laboratory, California Institute of Technology

Future space communication systems will begin to use Delay Tolerant Networking (DTN) which allows for the extension of Internet-like automated protocols to constrained environments experiencing frequent disruptions in communication. This technology improves upon current point-to-point relay links by ensuring reliable internetworking between multiple nodes through Bundle Protocols, forming the foundation for a Solar System Internet. In preparation for DTN's first customer, the KPLO launching in July 2022, it is necessary to build a test capability for the DSN DTN implementation. The JPL Protocol Technology Lab is developing an end-to-end communications protocol testbed that will use DTN as a prototype for future missions. We configure Interplanetary Overlay Network nodes to analyze incoming data and noise loss in the deep space link. To ensure successful DTN implementation, we develop a spacecraft telemetry simulator that will be interfaced to the BP software. On the mission operations side, the telemetry and command processing will be displayed using the AMMOS Instrument Toolkit. Emulating spacecraft data flow in conjunction with the simulated DTN network facilitates the advancement of DTN and its infusion into the Deep Space Network.

Jupiter Data and Image Archival Through the Planetary Data System

Reeya Chenanda, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Glenn Orton, Jet Propulsion Laboratory, California Institute of Technology

Mid-infrared images were taken at the National Astronomical Observatory of Japan's (NAOJ) Subaru Telescope with the Cooled Mid-IR Camera and Spectrograph (COMICS). These data include thousands of infrared images and scientific observations of Jupiter that need to be archived in NASA's Planetary Data System (PDS). The archival process has entailed a long endeavor of several years for previous SURF students, which is now nearly complete. Raw, reduced (calibrated) and ancillary files were collected and prepared for the COMICS archive. Using Python and FORTRAN-based Interactive Data Language (IDL) programs, the data from this instrument were labeled with XML template files following PDS4, the PDS's latest archiving standard. The labels include information about how and when the data were taken. Once the final archive has been reviewed and accepted by the PDS, the observations from COMICS can be easily accessed by scientists and be used in future inquiries.

Analysis of 3D Point Cloud Segmentation Using Deep Learning for Improved Perception in Subterranean Environments

Alexander N. Dimopoulos, University of California San Diego, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Benjamin Morrell, Jet Propulsion Laboratory/California Institute of Technology

The rapid development of autonomous robotic systems and their demonstrated success has been aided by the integration of neural networks for computer vision tasks which require a perceptual understanding of a given environment. Object detection networks commonly utilize data from image sensors which are bounded by their reliance on outside sources of light. Addressing the limitation of visual cameras in dark areas would expand the range of applications for robotics in regions with poor light quality and improve the robustness of autonomous systems. A solution may be the use of lidar, active sensors that do not require a light source which are already in use in a majority of robotic navigation systems. Lidar provides rich 3D point cloud data but implementing object detection techniques has been difficult in the past. However, with advancements in processing techniques and hardware, several neural network architectures have demonstrated the speed and accuracy necessary for real time applications. Analyzing 3D networks requires massive labeled datasets which are costly to produce. This work focuses on automating the process of annotating lidar point clouds using a simulation, and the datasets are examined through the training of semantic segmentation models which perform object detection.

Delay Tolerant Network Implementation and Design for Intersatellite Optical Communicators

Wei Foo (Jet Propulsion Laboratory/California Institute of Technology)
Mentors: Jose E. Velazco and Mark Taylor (Jet Propulsion Laboratory/California Institute of Technology)

Intersatellite Optical Communicator (ISOC) is a multidirectional, optical communications device to be deployed into space on satellites, forming an interplanetary network (IPN). Up to 150 ISOCs will compose a local area network (LAN) swarm, 100 kilometers wide, that orbits planetary bodies in space. Intra-swarm ISOCs can share computational resources and combine optical signals to amplify their transmission range when communicating with other swarms over a wide area network (WAN), spanning hundreds of millions of kilometers. To design the most effective network layout, LAN and WAN topologies were simulated using CORE, quantified, and analyzed. Furthermore, Delay Tolerant Network (DTN) protocols will enable data transmission within and between swarms because conventional internet protocols (such as TCP/IP) suffer from long delays and high error rates when subject to frequent link disruptions and asymmetrical transmission in a space environment. As so, we seek to implement the Interplanetary Overlay Network DTN protocol suite into ISOC's Optical-Generation Operating System by integrating the Bundle Protocol and Convergence Layer Adapter into Linux kernels. Additionally, we plan on modifying optical small form-factor pluggable modules, which plug into a Network Interface Card, to complete the data link and physical layers. We believe implementing DTN with ISOC will establish the first low latency, terabit IPN.

Applying Machine-Learning to Predict High-Latitude Alaskan Ionospheric Irregularities
Annabel R. Gomez, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Xiaoqing Pi, Jet Propulsion Laboratory, California Institute of Technology

Plasma convection and associate dynamics in the high-latitude ionosphere can produce various instabilities depending on location, geomagnetic conditions, presence of field-aligned currents, temperatures, and velocities. These instabilities give rise to irregular structures in ionospheric density distribution, or ionospheric irregularities, which can cause scintillations of radio signals and degrade the integrity and accuracy of technology applications such as Global Positioning System (GPS) navigation. This research applies machine-learning techniques to develop a prediction system that uses historical GPS and magnetometer data to predict the location, time, and intensity of concerned irregularities, under various geophysical and space weather conditions. A correlation analysis was first conducted between ionospheric and magnetic field data sets. It was determined that globally distributed auroral electrojet data did not correlate well with the localized ionospheric data while magnetometer data from a nearby station did. Understanding how and why these different physical phenomena correlate is crucial in making accurate predictions. As the development of the machine-learning algorithm continues, its output will aid and contribute to the representation and understanding of the variations of the ionospheric state.

An Analytical Approach to Determining the Long-Term Behavior of Satellite Trajectories
Andrew Graven, Cornell University
Mentor: Martin W. Lo, Jet Propulsion Laboratory, California Institute of Technology

The determination of long-term statistical properties of satellite trajectories typically entails integrating the trajectory over long time scales, which is a computationally expensive and time-consuming process. In the course of SURF projects in summer 2018 and 2019, a fast analytical method of understanding the long-term statistical behavior of circular and elliptical satellite trajectories was developed. This summer, the method was applied to a variety of problems related to expected satellite-ground station view periods, passage times and gap times. In addition, MATLAB and C programs were developed which contain practical implementations of many of these techniques. For the vast majority of cases, these analytical methods exhibit strong agreement with the statistics generated by the method of direct trajectory propagation, while taking a fraction of the time. There is significant potential for continued work on this project. This may include investigating the efficacy of computing a wider range of statistics, such as mean magnetic field strength and atmospheric drag. In addition, there is interest in adding corrections which account for trajectory perturbations due to atmospheric drag and diurnal effects such as solar radiation pressure.

Acknowledgement: This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Summer Undergraduate Research Fellowship Program and the National Aeronautics and Space Administration (80NM0018D0004).

Autonomous VTOL Fixed-Wing sUAS for HAB Payloads and Long-Duration Flight Applications
Hunter Hall, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Adrian Stoica, Jet Propulsion Laboratory, California Institute of Technology

During the Summer of 2020 the Innovation to Flight (i2F) team designed a payload-carrying, VTOL-capable fixed-wing vehicle as a cheap and easily reusable solution for high altitude balloon (HAB) payloads to be safely recovered at predetermined landing locations. The vehicle ascends and descends with a HAB flight train, detaches after HAB flight data is collected, and autonomously flies to a destination near the launch site or predetermined safe landing location for efficient recovery and reuse. A solar-powered battery swapping station was also developed to be located on the ground and extend effective flight time. The vehicle's stability and efficiency was simulated in fixed wing flight, vertical take-off and landing, and the transition phase between the two modes, and results from each flight modality are discussed. Mission Planner and Ardupilot simulations along with custom and open-source HAB flight prediction softwares were implemented to determine optimal launch times during the day, ascent rates, max altitudes, etc. to analyze all potential VTOL flight scenarios. The vehicle also demonstrates precision landing capabilities, applicable for landing on the previously mentioned battery swapping station. Finally, the vehicle is to be tested both in a low altitude balloon scheme and in a full HAB flight test. This vehicle is the first step towards greater efficiency and reliability into future HAB flight experiments that entail: strict geofencing requirements, concerns of the payload landing in an unretrievable area, and expected long range travel of the balloon.

Development of Miniaturized Spatial Heterodyne Spectrometer for Water Detection in Comets
Kiran Hamkins, California Institute of Technology
Mentor: Seyedeh Sona Hosseini, Jet Propulsion Laboratory, California Institute of Technology

A miniature spatial heterodyne spectrometer is being developed to detect water and its byproducts in space with high sensitivity and high resolving power. Few measurements have been made regarding the ratio of D/H on comets, and we only have 11 data points with this measurement to date. Obtaining more of these types of measurements allows us to understand how the galaxy evolved and how water evolved in the galaxy. For the instrument to complete this goal, it must be prepared and tested for flight and mission readiness. This includes making measurements here on Earth by using a telescope to observe the Sun and making measurements in the Earth's atmosphere by flying on the Blue Origin New Shepard rocket. Requirements for each of these operations were compiled, and some of the necessary instrument tests were carried out through simulation, indicating the performance of the actual instrument under flight conditions.

An End-to-End Methodology for Simulating Particle Entrainment in Internal Rocket Fairing Environments
Kealan J. Hennessy, University of California, Berkeley
Mentors: William A. Hoey, John M. Alred, and Carlos E. Soares, Contamination Control Engineering (353D), Jet Propulsion Laboratory, California Institute of Technology

Effective contamination control is essential for the success of most aerospace programs because the presence of contamination, even in minuscule quantities, can degrade the performance of spacecraft hardware. We aim to highlight the requirements of effective contamination control analysis, and subsequently present a succinct methodology and proof of concept for carrying out such analysis within rocket fairing launch environments. This process is composed of initial CAD modeling and internal fluid volume construction, through set-up of a "decoupled" fluid simulation and visualization of results prior to the iteration of different particle regimes – all restricted within the first two minutes of flight. A series of experimental results and (in due course) fine-tuning of these components has produced a framework that meets contamination control criteria for typical mission success most effectively. As well as providing a general guide for future work, this also forms a basis for possible extension of the analytical process (e.g. defining additional particle-particle interaction) and facilitates the exploration of potential experimental validation methods.

Investigation of Autonomous Balloon Technology for Automatic Launching of Payload Carrying VTOL
Nicholas M. Hennigan, Milwaukee School of Engineering
Mentors: Adrian Stoica and Thomas Lu, Jet Propulsion Laboratory/California Institute of Technology

High-Altitude Balloon platforms (HAB's) can provide a realistic and cost-effective testing analog for a diverse range of technology. However, with the uncertainty of HAB payload recoverability and deployment, research teams may deem the exercise risky, especially near final hardware development cycles. Researchers were tasked to develop a new HAB Auto-Launching architecture for a payload caring VTOL. This research focused on developing & validating various designs with a specific focus on parachute size and latex balloon packing/deployment characteristics. Using non-linear drag analysis, traditional parachute design was determined ineffective and a proposed buoyancy-driven descent mechanism was designed. Additionally, two pilot studies were designed & executed: *Deployment Characteristics of Thin-Walled Latex Balloons Given Initial Packing Geometry & Balloon Membrane Deflection*. Study results revealed a novel deployment mechanism for latex balloons deserving further development. These technology innovations provide a step forward in redefining high-altitude exploration and experimentation.

Visualization and Analysis of Ocean Floor Core Samples for Microbiological Studies
Lukas Hermann, Carnegie Mellon University
Mentors: Victoria Orphan, California Institute of Technology, Hillary Mushkin, California Institute of Technology, Scott Davidoff, Jet Propulsion Laboratory, California Institute of Technology, Santiago Lombeyda, California Institute of Technology, Maggie Hendrie, ArtCenter College of Design

Deep sea vents are home to extreme conditions that allow for a variety of microbial life forms. The study of these life forms and their ecology occurs at a multitude of scales that range from methane clouds covering kilometers down to the DNA level of the taxa that inhabit the vent. Scientists from a variety of specializations have analyzed these communities using methods such as bathymetry scans and core samples, but their work does not exist in a unified environment that allows one to compare geochemical parameter changes associated with the undersea topology, as well as the taxa that exist in this spatialized view. Through the collaboration of the DataVis team and Orphan Lab, I have been implementing prototypes of a unified environment in which researchers can explore and analyze data collected during their expeditions through a holistic lens. These prototypes exist as functional representations of data using D3.js as well as using wireframe workflows designed by my colleagues. These prototypes allow a spatial understanding of geochemical parameters, along with a direct comparison of core data as distance increases from a vent. Future research will explore long term manageability of multiple expeditions and scale based understandings of statistical data.

Characterization of an Instrument Model for Exoplanet Spectrum Estimation Through Wide Scale Analysis on HST

Noah Huber-Feely, Columbia University, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Mark R. Swain, Gael Roudier, and Raissa Estrela, Jet Propulsion Laboratory, California Institute of Technology

Context. Instrument models enable the removal of systematic error in transit spectroscopy light curve data, but, since the model formulation can influence the estimation of science model parameters, characterization of the instrument model effects is important. Aims. Analyze a simple instrument model and assess its validity and performance across Hubble WFC3 and STIS instruments. Methods. Over a large, n = 65, sample of observed targets, the parent distribution of each instrument model parameter is computed using MCMC. Possible parent distribution functions are then fit and tested against the empirical IM distribution. Correlation and other analysis are then performed to find useful IM relationships. Finally, a light curve residual metric is determined that explains where the proposed IM works well and where it does not. Results. The model is shown to work well across the 2 instruments and 3 filters analyzed and, further, the Lorentzian distribution is shown to closely fit the empirical parent distribution of IM parameters. This parent distribution can be used in the MCMC prior fitting and adds confidence to wide scale atmospheric analysis using this model. Finally, we propose a simple metric to determine model performance and demonstrates its ability to determine whether a derived spectrum under this IM will agree well with the literature.

Simultaneous Retrieval of Telluric Parameters and Astronomical Science Observables

Emmy Li
Mentors: Mark Swain, David Thompson, and Robert Zellem, Jet Propulsion Laboratory/California Institute of Technology

The New Mexico Exoplanet Spectroscopic Survey Instrument (NESSI) is one of the first spectrographs designed for ground-based, high-precision spectroscopy of transiting exoplanets. For ground-based astronomical telescopes, distortions due to atmospheric absorption and scattering significantly muddle the science signal. Previously, transit spectroscopy has mitigated these effects in post-processing by removing airmass changes and other telluric variability. Most correction strategies are sequential. This means they first estimate the atmospheric state (via calibration stars or by fitting spectrally-resolved absorption lines to determines abundance of atmospheric gases), and then seek to calibrate the observation from the atmospheric estimation to retrieve the target.

We seek an alternative method which simultaneously estimates the earth's atmosphere and science observables. Using a mathematical model of the atmosphere, target, and instrument, we can disentangle the data by performing a Bayesian model inversion to find the Maximum A Posteriori solution for all free parameters. This is made possible by adopting advancements in downward-facing spectroscopy of the Earth's surface for astronomical measurements, generalizing them for exoplanet transit time series. A successful implementation of a simultaneous retrieval of the telluric state together with science observables can improve accuracy of ground-case astronomical measurements, eliminating the need for calibrator stars.

Searching for Gravitational Wave Progenitors: Supermassive Black Hole Binaries

Emma McGinness, University of California, Berkeley, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Joseph Lazio, Jet Propulsion Laboratory, California Institute of Technology

Most large galaxies contain supermassive black holes at their core, and galaxies, as they evolve, merge. It is predicted that, as a result of such a merger, the two galaxies' central supermassive black holes (SMBH) converge into a new central region, forming a binary that emits gravitational waves as they spiral together. Thus, further investigation of SMBH binaries may reveal the relationship between galaxy merger astrophysics and gravitational wave physics. Analyzing follow-up observation data collected by the Very Long Baseline Array (VLBA), which consists of 28 radio galaxy sources at the parsec level, whose original images suggested the existence of SMBHs, potentially allows further identification and observation of likely SMBH binaries. In this project, VLBA data for those galaxy sources, measured at frequency ranges of 4~5 GHz, 6~7 GHz, and 8~9 GHz, were collected, reduced (i.e. calibrated), and processed via an established pipeline and CASA (the Common Astronomy Software Applications package) functions to eliminate radio frequency interference (RFI) to prepare it so that high angular resolution images can be created and reviewed. Those later produced images that exhibit characteristics consistent with two proximate point sources can then be identified as SMBH binary candidates for additional observation with the goal of further evidencing the theory that galaxy evolution and gravitational wave emissions are linked.

Software Development for Radio Astronomy
Tyrone McNichols, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Kristen Virkler and Melissa Soriano, Jet Propulsion Laboratory, California Institute of Technology

NASA's Deep Space Network (DSN) is an international array of radio antennas that support interplanetary missions and provide radar and radio astronomy observations. Recently, a new high-resolution spectrometer was deployed at Deep Space Station 43 (DSS-43), the 70-meter antenna at the Canberra Deep Space Communications Complex (CDSCC). This new spectrometer consists of four Reconfigurable Open Architecture Computing Hardware (ROACH)-2 boards created by CASPER (Collaboration for Astronomy Signal Processing and Electronics Research). In this project, software improvements have been implemented to enhance the functionality of this spectrometer. Specifically, measures were taken to resolve existing issues with the software and firmware and increase the consistency of the data acquired from the spectrometer. Additionally, changes were made to better facilitate automation and improve the user experience.

Intelligent Mission and Scientific Instrument Classification
Pavle Medvidovic, University of Southern California, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Chris A. Mattmann and Valentino Constantinou, Jet Propulsion Laboratory, California Institute of Technology

The Foundry A-Team, along with subject matter experts, guides clients through the concept formulation process of JPL missions in studies. These studies result in curated products including small, but rich data sets in the form of Confluence Wiki pages that clients can use to further iterate upon mission or instrument concepts investigated during the study. The AI and Analytics group (174B) of the Artificial Intelligence, Analytics and Innovative Development Division (1740) is working with the A-Team to develop the A-Team Analytics Tool (AAnT) to allow for greater utilization of relevant annotations from previous studies. Our research aims to enhance the AAnT's search capabilities by automating named entity recognition (NER) on study data sets. We trained a statistical model from the natural language processing library spaCy to identify instruments and spacecraft from studies using spaCy's built-in capabilities. To provide the model's training data, we wrote Python scripts to clean the text and then manually annotated the data using Prodigy, an annotation tool from the creators of spaCy. With 393 accept/reject samples annotated, the model has a 77.3% accuracy with 100 iterations in training, strongly suggesting that we can expect to reach at least 80% accuracy, among other metrics, when using the model to annotate new data.

Software Development for Spatially Resolved Calibration of Near-Infrared Spectra of Jupiter's Atmosphere in Prism Mode
Mohini Misra, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Glenn Orton (Jet Propulsion Laboratory/California Institute of Technology), and Tom Momary (Jet Propulsion Laboratory/California Institute of Technology)

SpeX is a medium-resolution 0.8- to 5.5-micron spectrograph at NASA's Infrared Telescope Facility (IRTF) which is used to capture near-infrared spectra of Jupiter. The spectral images resulting from scanning the slit across the planet can be used to understand Jupiter's atmospheric properties. Although the IRTF has previously developed SpeXTool, a set of sophisticated software to reduce and calibrate spectral images, the software does not treat spectra with regard to each pixel. Of the several supported spectroscopic modes of SpeX, this project focuses on spatially resolving data taken in Prism mode. A relative calibration is achieved by dividing individual pixel flux values by a star spectrum taken on the same observing run. This is then transformed into an absolute calibration by multiplying the pixel values by the standard spectrum of the star. A wavelength correction is calculated using pre-existing methods in SpeXTool. Then, dividing by the Sun's radiance provides calibrated reflectivity values for the Jupiter spectra. Using these methods, all spectra taken in SpeX Prism mode can now be spatially resolved, allowing others to perform more accurate analysis of the planet's atmospheric properties, such as with NEMESIS, a retrieval software tool used by my advisor and his team.

Impact of Intrinsic Alignment on Self-Organizing Maps of Modified Gravity
Brigitte Montminy, Jet Propulsion Laboratory, California Institute of Technology, and Tufts University
Mentors: Agnès Ferté and Jason Rhodes, Jet Propulsion Laboratory, California Institute of Technology

We know the universe is expanding at an accelerating rate, which is explained by dark energy in our common ΛCDM model. An alternative to this elusive dark energy is to modify the most generally accepted theory of gravity, General Relativity. However, there are many ways to modify this theory, many 'modified gravity' theories. This project aims to create 2D self-organizing maps (SOMs) of various theories of modified gravity to determine how similar or different they are to each other. To study this behavior of the universe, we observe the shapes of galaxies. Due to the gravity of matter between us and the target galaxies, the light of the target galaxies gets distorted, a process known as weak gravitational lensing. One main source of error when modeling weak lensing is intrinsic alignment (IA), the initial shape and orientation of the galaxies previous to being lensed. The SOMs will be created both with and without taking IA into account in order to determine the impact of IA on these models of modified gravity.

VTOL Vehicle Range Extension System

Spencer Morgenfeld, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Virgil Adumitroaie and Adrian Stoica, Jet Propulsion Laboratory, California Institute of Technology

During the summer of 2020 a proof of concept VTOL vehicle range extension system was developed as part of the Innovation to Flight (I2F) team's work towards automating high altitude balloon launches. The system is composed of two parts: a drone-portable, solar powered battery swapping station, and a compatible vehicle fuselage. The swapping station will be flown into a mission area by the vehicle. Once in place, the station will extend the range of vehicle missions by swapping the vehicle's depleted battery for a freshly solar-charged battery. First, the VTOL vehicle configuration, airframe, motors and electronics were selected; the vehicle was later successfully flown by members of the I2F team. Second, the design work for the swapping station was completed in CAD, enabling reliable estimates for weight (2.25 kg) and cost ($500). Additionally, the station electronics were assembled and tested, and displayed adequate charging performance. Third, the vehicle fuselage was roughly modeled in CAD, showing that the swapping method is feasible to implement with our chosen vehicle airframe. Together, this work showed that the range extension system is feasible and merits further development.

Classifying Terrains, Rock Densities, and Concerns for the Mars 2020 Rover and Mars Sample Return

Madison Morris, Rice University, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Matthew Golombek (Jet Propulsion Laboratory, California Institute of Technology) and Nathan Williams (Jet Propulsion Laboratory, California Institute of Technology)

NASA's Mars 2020 rover, launched successfully on July 30, 2020, is planning to land within Jezero Crater (18.47°N, 77.438°E). Known as Perseverance, the rover will look for signs of habitable environments and indications of past microbial life on Mars while analyzing the chemistry and lithology of the surface. During its mission, Perseverance will collect samples of rock and regolith inside of tubes as it travels from Jezero to a second site known as Midway. These tubes will be dropped in predetermined locations where a Mars Sample Return rover mission could pick them up toward returning them to Earth. In order to safely and efficiently traverse the Martian surface, the prospective navigation paths must be closely surveyed for hazards such as ripples, steep slopes, and large rocks which would hinder the rover's ability to traverse. Through utilizing 25 cm/pixel images taken by the Mars Reconnaissance Orbiter High-Resolution Imaging Science Experiment (HiRISE), various terrain types, rock densities, and concern areas can be coded onto a map mosaic containing the possible rover paths. The characterization of the paths will further establish an understanding of where Perseverance and a future Mars Sample Return rover could go, and how they can get there.

Modeling Plume Redeposition and Photodissociation of Organic Molecules on Enceladus' Surface

Izzy Muise, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Edith Fayolle, Jet Propulsion Laboratory, California Institute of Technology

Enceladus, one of Saturn's moons, is an icy body with a subsurface ocean and rocky core, making it a potentially habitable environment. Enceladus has icy plumes which likely vent ocean material into the atmosphere where it eventually gets redeposited on the surface. Analyzing its surface would make it possible to study the composition of the ocean. Detecting amino acids such as glycine would show that complex organic molecules essential to life are being formed in the oceans, but glycine and other organic molecules near the surface are constantly being broken down by energetic particles, including photons. By modeling these processes using values from analogous laboratory studies, it is possible to constrain the depth below the surface and the amount at which organic molecules could be found. These models provide a more accurate picture of the required depth a rover would require to get a full picture of the chemistry occurring on Enceladus.

Calculating Stresses on Europa and Other Solar System Moons: Improving SatStressGUI

Caroline Nagib (Jet Propulsion Laboratory, California Institute of Technology)
Mentors: Robert Pappalardo and D. Alex Patthoff (Jet Propulsion Laboratory, California Institute of Technology)

Icy satellites undergo stresses that result in geological deformations that are evident on their surfaces. Some stresses include changes in the satellite's orbit, volume changes due to the melting or freezing of a subsurface layer, change in obliquity, or different rotations of the outer layer of the satellite relative to the interior, which is known as non-synchronous rotation. These deformations are evidence to the history of the stresses the satellites have experienced. SatStressGUI V5.0 is an open-source Python modeling software that offers a user-friendly interface that models stresses on satellites. The user can input the satellite's parameters and the types of surface stresses. For a specific location, SatStressGUI can compute the three components of the stress tensor, maximum and minimum tension, orientation of maximum tension, calculate surface Love numbers, generate cycloids, and create stress plots and lineaments. Using Python, additions have been made to automate some processes that were initially manual to enhance user experience. To better understand tension and compression, double ridge locations in Conamara Chaos on Europa are used to calculate where the maximum tension (α), obtained from SatStressGUI, is perpendicular to the double ridge, or the bearing.

Spectroscopy of Active Gas-Rich Radio Galaxies
Elena Romashkova, Massachusetts Institute of Technology
Mentors: Matt Bradford and Reinier Janssen, Jet Propulsion Laboratory, California Institute of Technology

A galaxy is said to have an active galactic nucleus (AGN) when the supermassive black hole at its center is actively accreting matter. There are two primary types of AGN: radiative-mode AGN and jet-mode AGN, which generally occur in distinct types of galaxies. Radiative-mode AGN are typically found in smaller, bluer galaxies with a high star formation rate, while jet-mode AGN are found in older, redder galaxies. Jet-mode AGN contain radio jets, causing their host galaxies to emit at radio wavelengths. In this project, we study a type of AGN that does not fit into these categories: jet-mode AGN found in blue host galaxies. These galaxies have a high star formation rate, which casts doubt on our understanding of the mechanisms that power their AGNs. To investigate this phenomenon, this project analyzes optical spectroscopy data taken for a sample of 7 such galaxies using the Isaac Newton Telescope. Using these spectra, properties of the host galaxies and their AGNs can be determined, such as the star formation rate as a function of galaxy position and the presence of emission lines from highly ionized atoms. A Python data reduction pipeline is developed to calibrate and extract results from this data set. The results of one galaxy in the sample are analyzed, yielding the possible explanation that activity in the edges of the galaxy masks high-excitation activity in its nucleus. By further analyzing the calibrated data set in future work, we will be able to better characterize the activity that causes the unusual behavior of these galaxies.

Patterns of Altitude and Ultraviolet Exposure in Oval BA's Color Change
Sahana Saikumar, Jet Propulsion Laboratory, California Institute of Technology
Mentors: Glenn Orton, Tom Momary, James Sinclair, Kevin Baines; Jet Propulsion Laboratory, California Institute of Technology

Oval BA, more popularly known as the Little Red Spot, is the second largest anticyclonic vortex on Jupiter's surface. In December of 2005, Oval BA's initial white pigmentation was observed to have become red and remained this color for more than a decade. In July 2018, it was observed that Oval BA's red color was fading and in December 2018, Oval BA was observed to have become completely white. To determine a correlation between Oval BA's color change and the altitude and ultraviolet exposure of the storm's particles, an image set containing Oval BA at various wavelengths that include epochs when it is red, transitioning to white, and completely white is compiled. These images are then sorted based on their respective wavelengths as a function of time and calibrated in order to account for the distortion that any surrounding light sources may create on the image. Emission-angle values are found at the center of Oval BA and are compared with the storm's reflectivity using the Minnaert function. Results suggest a change in altitude of particles between the red and white epochs.

Probing Cosmology With Large-Scale Galaxy Redshift Surveys
Isabel Sands[1,2]
Mentors: Olivier Doré[2], Chen Heinrich[2], and Henry Gebhardt[2]
[1] Yale University
[2] NASA Jet Propulsion Laboratory, California Institute of Technology

Galaxy clustering measurement are anisotropic with respect to the line-of-sight; this phenomenon, known as redshift space distortion (RSD) is important to understand in order to determine the cosmological redshift of observables in spectroscopic surveys. While fixed line-of-sight approximations have been sufficient for past galaxy surveys, the next generation of surveys will cover larger fractions of the sky that will require wide-angle corrections to the galaxy power spectrum. In this project, we utilize the Yamamoto estimator to compute the wide-angle galaxy power spectrum from mock data with window functions for the BOSS DR12 survey, as well as the upcoming Nancy Grace Roman Space Telescope and SPHEREx. Additionally, we use similar methods to estimate the window function multipoles, which are required to calculate the analytic prediction for the wide-angle power spectrum. Preliminary results for the wide-angle power spectrum agree with theoretical models at large scales.

Creation of a Ground Truth Navigation Filter for Mars Science Helicopter
Mohit Singh, Georgia Institute of Technology
Mentor: Jeff Delaune, Jet Propulsion Laboratory, California Institute of Technology

In robot navigation, a central problem is that of obtaining ground truth for the position, attitude, and velocity of the platform. For the Mars Science Helicopter (MSH), ground truth is required to evaluate the MSH state estimation framework, xVIO. There currently does not exist a ground truth filter in the robotics community which can evaluate ground truth over long distances. Using 2 RTK GPS sensors and an IMU fused together in an Extended Kalman Filter(EKF), we present a ground truth navigation filter that allows for sub-centimeter position accuracy and sub-degree attitude accuracy. The system can work for long distances and in the absence of a horizontal acceleration, which is essential for most UAV applications. Evaluation of the filter was done on real data and simulation data via a Monte Carlo analysis, where hundreds of trials were done to evaluate the worst case error. The ground truth filter was designed for Mars Science Helicopter, but is generally applicable to any autonomous robot.

Improving Upon the Multispecies Model for Organic Contaminants of Space Crafts
Nishita Sinha (Harvard University and JPL)
Mentors: Dr. John Alred (JPL), Dr. William Hoey (JPL Contamination Control Department, JPL/California Institute of Technology)

This project models molecular contamination of spacecraft technologies for various times and temperatures. Molecular contamination commonly arises as outgassing - the trapping of gaseous molecules within metal surfaces. These molecules can diffuse through metal surfaces and desorb off of faces. We developed python scripts to determine diffusion coefficients and initial concentrations of contaminant species from data on outgassing rates (OGR) and total mass loss (VCM) of contaminants. Our functions employed the Crank-Nicolsan numerical method to solve the diffusion equation and the Nelder-Mead minimization scheme to minimize error between predicted and actual data. Our models found overall diffusion coefficients and initial concentrations of best fit, and used these to predict OGR and VCM curves for long periods of time.

Contaminant samples often consist of many contaminants. We next studied these individual contaminants. Heating the sample at a constant rate and measuring evaporation provides information on individual species since peaks in evaporation rate graphs correspond to different contaminants. We updated Python scripts to better fit these evaporation peaks by fitting water evaporation as a zero-order desorption process and organic contaminants as first-order. We then used fits of these evaporation peaks, VCM data, and our Crank-Nicolsan functions to find diffusion coefficients and initial concentrations for individual species.This allows us to predict OGR and VCM of contaminants under varied conditions. We hope to use these outputs in 3D-modeling of spacecraft contamination.

Identifying and Characterizing Dusty Stellar Candidates in Spitzer Kepler Survey (SpiKeS)
Sydney Skorpen (Jet Propulsion Laboratory/California Institute of Technology and Moorpark College)
Mentors: Farisa Morales (Jet Propulsion Laboratory/California Institute of Technology) and Varoujan Gorjian (Jet Propulsion Laboratory/California Institute of Technology)

Characterizing debris disks within the *Kepler* field can further our understanding of the composition and evolution of planetary systems. The Spitzer Kepler Survey (SpiKeS) dataset, which captures the *Kepler* field in the 3.6 and 4.5 μm wavelength, contains 889 sources and presents the opportunity to identify and analyze dusty disk systems. Here, a subsample of 889 sources was isolated by its likelihood for infrared excess, a key characteristic of debris disks. Spectral Energy Distribution (SED) graphs were produced for each of the 889 sources and were organized by characteristics, of which 146 sources were selected for their clean SEDs containing infrared excess. A visual inspection of the 146 sources showed that 10 sources contain no obvious signs of infrared contamination and will require further analysis as a candidate for containing dust. Another 30 sources exhibited unusual "rising" SEDs, many of which were galaxies as confirmed by the SIMBAD archive. The results thus far show that sources within the SpiKeS potentially include circumstellar dust, expanding our knowledge of the *Kepler* field and the planetary systems within.

Autonomous Return to Home VTOL Fixed Wing UAV for HAB Payloads and Long Duration Flight Applications
Andrew Song, Columbia University
Mentors: Virgil Adumitroaie and Adrian Stoica, Jet Propulsion Laboratory, California Institute of Technology

The Innovation to Flight (i2F) team designed a payload-carrying, VTOL-capable fixed-wing vehicle as a cheap, easily reusable solution for high altitude balloon (HAB) payloads to be safely recovered at predetermined landing sites. The vehicle ascends and descends with a HAB flight train, detaches after HAB flight data is collected, and autonomously flies to a destination near the launch site or predetermined safe landing site for efficient recovery and reuse. A solar-powered battery swapping station was also developed to be located on the ground and extend effective flight time. The vehicle's stability and efficiency was simulated in fixed wing flight, VTOL, and the transition phase between the two modes using Ardupilot's Software in the Loop (SITL) simulation software, and results from each flight modality are discussed. The vehicle should demonstrate precision landing capabilities, applicable for landing on former mentioned battery swapping stations. Finally, the vehicle is to be tested both in a low altitude balloon scheme and in a full HAB flight test. This vehicle is the first step towards greater efficiency and reliability into future HAB flight experiments that entail: strict geofencing requirements, concerns of the payload landing in an unretrievable area, and expected long range travel of the balloon.

A Lobatto Interpolation Over Unstructured Grid
Felianne Teng (Jet Propulsion Laboratory, California Institute of Technology)
Mentor: Dragan Nikolic (Jet Propulsion Laboratory, California Institute of Technology)

Solar Wind Particles (SWP), Solar Energetic Particle (SEP) events, and high-energy Galactic Cosmic Rays (GCR) in outer space present numerous dangers to space exploration. The lack of Earth's protective atmosphere and geomagnetic field causes astronauts and equipment to be at high risk of being harmed by charged ions. Passive shielding materials are mainly effective against SWP but are largely ineffective against SEP and GCR. For protection against the latter two events, it is better to alter the particle trajectories, preventing collisions altogether through the generation of an electric field. Thus, we wish to develop an active shielding platform to simulate and determine the optimal electric field to minimize ion flux. To improve user interaction with the system, we implemented the ability of the project to import and run existing systems from .stl and COMSOL .txt files. We also wished to increase accuracy of our system. For this purpose, we utilized the relativistic Verlet algorithm to more accurately predict the trajectory of ionized particles. Finally, we selected points around a given location and constructed a Lobatto interpolation grid, using both linear and quadratic interpolation surfaces to accurately interpolate values.

Scientific Bias and Trends in the Development of Spectrometers for Space Exploration
Vincent Tieu, Jet Propulsion Laboratory, California Institute of Technology
Mentor: Seyedeh Sona Hosseini, Jet Propulsion Laboratory/California Institute of Technology

Spectroscopy has always been an integral part of space exploration and research. The scientific targets we are able to research are directly correlated with the capabilities of our current and past instruments, and as such, our science is significantly biased by our instrumentation. Using spectrometer specification data obtained from the NSSDC (NASA Space Science Data Coordinated Archive) and NICM (NASA Instrument Cost Model), we analyze the spectrometers developed by NASA/JPL and other space agencies to determine if there are any trends in the scientific space observed, examine the effects of these biases, predict the trajectory of future spectrometer development, and identify any potentially untapped or overlooked regions. To aid our analysis, we have created a web application (free to use by the scientific community and the public) to generate scatterplot matrices to compare instruments over various combinations of specifications. We will also perform PCA (Principal Component Analysis) to analyze the progression of spectrometers over time. By performing this analysis, we hope to encourage the development of spectrometers to address scientific biases and any unserved but feasible scientific spaces.

Improving Accuracy and Robustness of Detecting Exoplanet Transits Using Consumer-Grade Telescopes
Aaron Tran (Jet Propulsion Laboratory/California Institute of Technology)
Mentors: Rob Zellem (Jet Propulsion Laboratory/California Institute of Technology), and Kyle Pearson (Jet Propulsion Laboratory/California Institute of Technology)

Discovering exoplanets, a current hot topic in astronomy, using conventional telescopes is a time and cost-intensive process. Exoplanet Watch, a new initiative, aims to dramatically lower the amount of resources needed to observe exoplanet transits through the use of small (less than one meter) telescopes. Over the summer, Exoplanet Watch's data reduction software, EXOTIC, was overhauled to be able to directly query SIMBAD, a star database, in order to readily determine optimal comparison stars to use when generating light curves from observed data. In addition, the older analysis model was replaced by a nested sampling method that decreased runtime from a scale of hours to minutes. The culmination of these enhancements over the research period has resulted in a robust application that is capable of generating live curves in real time and using previously observed images. Widespread usage of this program could exponentially increase the rate at which exoplanets are discovered while helping to lower the entry barriers to astronomy.

Markov Chain Relaxation for Data Clustering in AIRS Radiances
Sophie Uluatam, Cornell University
Mentors: Amy Braverman (Jet Propulsion Laboratory, California Institute of Technology), and Hai Nguyen (Jet Propulsion Laboratory, California Institute of Technology)

The Atmospheric Infrared Sounder (AIRS) instrument, aboard NASA's Aqua satellite, collects vertical profiles of temperature and water vapor. The data collected by AIRS is partitioned into *granules*, each consisting of 1,350 *footprints* and 2,378 *channels*. Because the granules are so large, handling them demands a more parsimonious representation of the data. One method of concisely representing large datasets is clustering. In this project, we implement a clustering technique on the AIRS footprints that relies on Markov relaxation. To do this, we generate a Markov chain whose state space consists of all of the footprints in a given granule, and whose transition probabilities are exponentially proportional to the distances between footprints. We then perform a random walk on this Markov chain, at each time step computing the *mutual information* between the initial state and the state at that step. Finally, we identify locations where the rate of information loss slows, which correspond to meaningful clusters in the data. The results of this project can be used in other clustering algorithms— for example, the information bottleneck method— to determine the elements of each cluster. Ultimately, this project will serve as a tool for generating condensed representations of large datasets.

Weak Gravitational Lensing Cosmology to Constrain Modified f(R) Gravity
Leah Vazsonyi, California Institute of Technology, Jet Propulsion Laboratory
Mentors: Peter Taylor, Jet Propulsion Laboratory, and Jason Rhodes, Jet Propulsion Laboratory, California Institute of Technology

In this work, I constrain the free parameter of f(R) gravity, a modification of General Relativity, using data from weak gravitational lensing observed as the slight distortion in the apparent ellipticities of galaxies. I use the Cl statistic from the Subaru Hyper-Suprime Camera (HSC) year 1 data. To perform these calculations, I use

CosmoSIS, an open-source software package designed to compute predictions for cosmological functions based on input parameters of the large-scale structure of the Universe. I integrate a modified gravity emulator for CosmoSIS to generate a theoretical prediction for the Cl, as well as a module to generate the likelihood of a particular theoretical sample given Cl data to perform samplings of the parameter space. Since we are unable to model smaller scales, I apply the k-cut cosmic shear method to remove these. I have computed the likelihood of the corresponding k-cut Cls to constrain a 2 parameter ΛCDM model. In the future, this will be extended to constrain the free parameter in f(R) gravity.

Applying Machine Learning to Classify Land Cover in Panama
Angelina Ye (California Institute of Technology)
Mentor: Erika Podest, Jet Propulsion Laboratory, California Institute of Technology

Sustainable forest management is often tracked by mapping forest cover using optical satellite data; however, this presents many challenges in tropical areas due to cloud cover. Synthetic Aperture Radar (SAR) remote sensing can overcome this challenge because the signal can penetrate through clouds and map the surface. This project is focused on exploring different machine learning approaches to classify land cover in Panama using SAR images from the Phased Array L-band SAR (PALSAR) sensor onboard the Japanese ALOS satellite, which operated from 2006-2011. The PALSAR data has a 50 meter per pixel resolution and HH and HV polarizations. The land cover classes identified were open water, non-forest, agriculture, and forest. Training and validation data were labelled for each class using a validated reference classification map for Panama. Three machine learning classification approaches were used: 1) Random Forest, an ensemble learning model that uses a majority vote on parallel decision trees to classify data; 2) Gradient Tree Boosting, an ensemble method similar to Random Forest, but with sequentially connected decision trees; and 3) Convolutional Neural Network (CNN), a deep learning algorithm that convolves the image data to learn specific features of the image. Results indicated that Random Forest worked well but had difficulty with edge cases while Gradient Tree Boosting had similar results to Random Forest. CNN was more complex and needed much more training data, time and memory to train, validate and predict but had the potential to identify edge cases more effectively.

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Machine Learning for Mars Rover Pancam Images to Support Content-Based Search
Brandon Zhao, Duke University
Mentors: Steven Lu and Kiri Wagstaff, Jet Propulsion Laboratory/California Institute of Technology

NASA has acquired more than 4.3 million detailed panoramic images of Mars surface through the Pancam instrument mounted on the Spirit and Opportunity rovers. These images are currently labeled with metadata parameters such as image wavelength, position code and date, but lack content-based labels, such as which geological features or rover parts specific images may contain. Thus, we developed a convolutional neural network based classification algorithm to create these labels, supporting a content-based search on NASA's Planetary Data Science Image Atlas database. To do so, we consulted with planetary scientists to create a significant yet concise list of classes of interest, then created and labeled a dataset for deep learning. Next, we employed transfer learning techniques to adapt an existing neural network architecture pretrained on Earth images to a multi-label task on Mars images. The resulting classifier can, given an input panoramic image, accurately output confidence levels for matches with 25 different classes of interest, leveraging image metadata and class dependencies to make an informed classification.

Multi-Phase, Planetary-Chemistry Modeling for Biosignatures
Catherine Zheng, Yale University
Mentor: Pin Chen, Jet Propulsion Laboratory/California Institute of Technology

There are many factors of the Earth's atmosphere and geochemical processes that affect the oxidation state of the atmosphere, such as the abundance of volcanic gases and changes in sources and sinks of O_2. Looking at all these factors is necessary to get an idea of how the atmospheric system of Earth is affected by things like oxygen and reducing agent sources and sinks. Understanding the oxidation of Earth and the factors that affect its atmospheric chemistry is important in creating an accurate representation of the planetary system, and could be utilized in studying exoplanets. In modeling the atmospheric chemistry of early Earth, we used a detailed atmospheric kinetic model that incorporates solid and aqueous equilibrium chemistry, atmospheric chemistry, and atmospheric escape to accurately simulate atmospheric compositions and observable spectra. This utilizes the AROC (atmospheric-rock-ocean-chemistry) model that couples an aqueous geochemistry model with an atmospheric photochemistry model (PHREEQC). By coupling both of these models, we get a more accurate representation of the atmosphere and how it evolves when more complex parameters and conditions are taken into account.