Caltech
Student-Faculty Programs

Poster Session Abstracts

Frequency of Small Magellanic Cloud analogs in simulated Milky Way systems hosting a Large Magellanic Cloud analog
Selassie A. Ackah
Mentors: Andrew J. Benson and Lior S. Pachter

The Milky Way anchors studies of galaxy formation and dark matter, yet its satellite system may be unusual because it contains both Magellanic Clouds. We investigate how often a Small Magellanic Cloud analog accompanies a Large Magellanic Cloud analog in simulated Milky Way systems. From 10,000 Galacticus merger-tree realizations, we selected candidates with subhalo masses of 5-8 billion solar masses, three-dimensional separations of 20-40 kiloparsecs, and evidence of prior association with the Large Cloud analog. In the current analysis, this definition identifies 618 candidates in 606 systems, a preliminary candidate-hosting fraction of 6.06%. Removing the association requirement raises the fraction to 27.65%, indicating that mass and distance alone select many objects lacking a shared accretion history. These estimates remain subject to validation of the mass definition, association criterion, and selection pipeline. Nevertheless, inferred Small Magellanic Cloud frequency is highly sensitive to how physical association is defined. Robust criteria are necessary for judging whether the Milky Way is representative and for testing galaxy-formation and dark-matter models.

Composition optimization of an FeCoNiMo high-entropy alloy for hydrogen evolution
Jaehwan Ahn
Mentors: William A. Goddard III and Sejun Kim

High-entropy alloys of earth-abundant transition metals could replace platinum in hydrogen production by water splitting, but their compositional space is too large to search by experiment alone. This project optimized the molybdenum content of an iron-cobalt-nickel-molybdenum alloy on its close-packed (111) surface. Activity was judged by the Volmer step, in which water dissociates and leaves a hydrogen atom bound to the metal. DFT nudged elastic band calculations located the transition state along nine reaction paths. Grouped by the metal atom that receives the hydrogen, the barriers rise from nickel to cobalt to iron. They also follow the local arrangement around each site rather than the overall composition. A MACE interatomic potential fine-tuned on these calculations produced 100 equilibrated surfaces across 10 molybdenum concentrations, and a score built from the barriers ranked them. Activity peaks at 25% molybdenum. Below that value too little molybdenum reaches the surface to form sites, and above it molybdenum atoms sit next to one another and the sites are lost. Work continues on the Heyrovsky step, in which the bound hydrogen leaves as hyrdogen gas.

Phylogenetic comparative methods for the 8-Cube mouse founder transcriptome atlas
Akshay Anand
Mentors: Lior S. Pachter and Nikhila Swarna

Phylogenetic comparative methods, including Felsenstein's independent contrasts, require that samples be related by a hierarchical tree. We show that the five laboratory strains among the eight Collaborative Cross founders have extensively admixed genomes incompatible with tree-like descent and inference. The three wild-derived founder strains, representing the domesticus, musculus, and castaneus subspecies of Mus musculus, together with the reference strain C57BL/6J form a resolved four-taxon phylogeny suitable for expression-based tree inference. Using single-nucleus RNA-seq data from the 8-cube founder dataset, we find that the correct tree topology is recovered from gene expression across all cell types and gene sets. However, not all genes carry equal phylogenetic signal; entropy-based specificity metrics from ember effectively isolate genes with the strongest. Without such selection, we find that male and female transcriptomes diverge between strains at comparable rates while a substantial fraction of divergence is cell-type-contextual, varying by cell type rather than reflecting a uniform transcriptome-wide shift. For instance, we see divergence in the kidney can be isolated to proximal tubule and thin ascending limb epithelial cells. We extend this decomposition across all eight tissues in the 8-cube dataset to ask how cell-type divergence profiles compare across organs.

Neural circuit mapping and behavioral characterization in activity-based anorexia and restraint stress models
Ava Arabshahi
Mentors: Yuki Oka and Tongtong Wang

Stress can suppress appetite through multiple neural pathways, but whether distinct stressors recruit shared or separate brain circuits remains unclear. This project compares restraint stress (both acute and chronic immobilization) with activity-based anorexia (ABA), an established model of stress-induced hypophagia used here as a positive control. In ABA, time-restricted food access paired with a running wheel drives paradoxical hyperactivity and rapid weight loss. Whole-brain Fos immunofluorescence is used to map neural activation across the two models and their respective controls. Mice were assigned to five groups (ABA, food restriction, wheel access, restraint, and control). Wheel-running revolutions were tracked using a self-built Arduino logger. Additionally, body weight, food intake, and water intake were recorded daily for 14 days. ABA mice lost approximately 19% of baseline body weight at their lowest point, compared with about 6% in restraint-stressed mice and negligible change in controls. Restraint-stressed mice maintained near-normal, slightly increased food intake despite this weight loss, indicating that their weight loss may arise through a mechanism separate from ABA's intake-driven loss. Brain sections are being processed to quantify Fos-positive cell density across hypothalamic, amygdalar, and brainstem regions to determine whether the two models engage dissociable pathways.

Diffusion guided posterior sampling for Bayesian inverse problems
Nicholas Z. Bai
Mentors: Anima Anandkumar and Jiachen Yao

Bayesian inverse problems seek to infer unknown physical fields or parameters from indirect, often sparse observations through a forward model—for example, reconstructing a permeability field from pressure measurements governed by a PDE. Because these problems are typically ill-posed, the solution is characterized by a posterior distribution rather than a single maximum-a-posteriori estimate. Existing sampling methods can recover locally plausible solutions but often fail to capture the full posterior when it is strongly multimodal. We address this challenge by separating posterior sampling into two complementary objectives: correctly allocating probability mass across modes and accurately resolving the distribution within each mode. Through theoretical analysis and controlled experiments, we find that gradient-based particle transport is effective for local refinement but poorly suited to reallocating mass between well-separated modes. Motivated by this observation, we develop a hybrid framework that combines mode-level resampling for global mass allocation with particle transport for local refinement. We further investigate how learned generative priors can accelerate inference while maintaining reliable posterior sampling.

Towards cryo-EM structural determination of FAM136A using a nanobody-based fiducial strategy
Sandra E. Baltazar
Mentors: Rebecca M. Voorhees and Jichen Zhang

The mitochondrial membrane consists of two types of membranes: the outer membrane (OM) and the inner membrane (IM). Between these membranes exists an inner membrane space (IMS) in which various proteins, such as FAM136A, conduct their functions. Recently, it has been found that FAM136A is an IMS chaperone that helps VDACs insertion into the OM. Without this interaction, the mitochondria would not be able to maintain proteostasis, leading to stress signal responses and human disease. This project aims to better understand this mechanism through the lens of structural biology. Cryo-EM is a tool that typically allows determination of protein structures that are the size of 100kDa or greater. Still, protein structures < 30 kDa, like FAM136A, are usually hard to analyze by cryo-EM. A developed solution is to bind a universal nanobody, NabFab, along with a binder that is specific to the target protein in order to increase its size without interfering with its functionality. We have been able to use a computational method to predict binders that are likely to bind to FAM136A. If successful, our efforts will display the accuracy of generated binders and their capabilities in solving the structure of FAM136A.

Investigating the molecular and neural mechanisms of social behavior in *Drosophila*
Kamarya A. Buckley
Mentors: David J. Anderson, Yousef Ouadah, and Ziqing Zhong

Social behaviors such as aggression and courtship are regulated by neural circuits and pheromone signaling, yet the molecular mechanisms underlying their evolution remain poorly understood. *Drosophila santomea* provides a valuable model because males exhibit increased same sex courtship and produce reduced levels of cis-vaccenyl (cVA), a pheromone that promotes aggression and suppresses male-male courtship in most Drosophila species. This study investigates the cellular basis of cVA production by characterizing the ejaculatory bulb, the organ responsible for cVA biosynthesis, using single-cell RNA sequencing (scRNA-seq). Current work focuses on optimizing tissue dissociation and validating cell viability to generate high-quality single-cell suspensions for sequencing. These data will ultimately identify candidate genes involved in cVA biosynthesis, with particular interest in epithelial cell populations.

In parallel, in *Drosophila melanogaster*, we examined the neural circuitry underlying aggression through an epistasis-based optogenetic approach by activating common aggression-promoting (CAP) neurons while silencing the downstream neuron SMP054; we tested whether SMP054 is required for aggression- related behaviors. Preliminary behavioral observations suggest SMP054 may not be required for CAP-mediated aggression and approach behaviors. Together, these complementary approaches will advance our understanding of the molecular and neural mechanisms that shape the evolution of social behavior in *Drosophila*

PTNO: Training neural operators with noisy Monte Carlo estimates for particle transport problems
Yubo Cao
Mentors: Anima Anandkumar and Xi Deng
Particle transport under multiple scattering is central to radiative transfer and plasma physics, yet high-fidelity Monte Carlo (MC) simulations demand tracing prohibitively large numbers of particles. Learning-based surrogates offer a path to amortize this cost, but their training typically requires expensive, well-converged MC solutions. We propose Particle Transport Neural Operator(PTNO), a neural operator framework for learning particle transport surrogates directly from noisy, low-cost MC labels. Training on such labels presents two coupled challenges: (1) the MC labels have high variance, which destabilizes standard supervised learning, and (2) particle transport solutions exhibit high dynamic range (HDR), spanning many orders of magnitude and making the solution difficult for a surrogate to resolve. To address the first issue, we train on noisy labels of many different configurations to learn the solution operator of the PDE, amortizing the cost of MC runs and enabling zero-shot generalization. For the second issue, a standard approach is to apply a non-linear transformation, such as log, to compress the HDR. However, this makes the resulting supervision biased. Instead, PTNO learns in the original space to keep the labels unbiased by adding an output softplus layer to PTNO, enforcing positivity. This allows small positive values to be represented smoothly and avoids nonlinear transformations of noisy MC labels. We demonstrate PTNO's effectiveness on representative particle transport problems, including neutron transport in fusion-reactor settings and radiative transfer in spatially varying participating media, showing that neural operators can learn accurate cross-domain transport surrogates while delivering 105--106× faster inference compared to converged MC simulations.

Modeling experimental asset bubbles through Large Language Models with persona optimization
Yi-Jhe Chen
Mentors: Colin F. Camerer and Cady H. Ngo

Experimental asset markets are an effective way to study human trading behavior. With recent advancements in Large Language Models (LLMs), simulating human experiments with these models has become a new paradigm in social science. This project aims to build a robust environment with LLM agents in the experimental asset market, providing a proxy for studying human behavior in stock markets. By parametrizing and clustering experimental human data, we identified four types of traders in the experimental market. An optimizer LLM was then adopted to optimize natural-language personas, prompting LLMs to behave in a way that resembles humans. Our empirical results show that LLM agents are capable of reproducing human behavior and market paths with our optimized personas, demonstrating the promise of explaining human behavior with LLM agents.

Hardware-native quantum machine learning methods for anomaly detection in active galactic nuclei light curves
Joaquin Gabriel Malachi A. de Castro
Mentor: Matthew J. Graham

Active galactic nuclei (AGN) are bright galactic centers powered by accretion onto supermassive black holes. While AGN naturally exhibits stochastic brightness variations, unusual deviations from this variability are potentially indicative of rare or physically meaningful changes in the accretion system. Distinguishing expected damped random walk behavior from flares, repeating outbursts, or sustained state changes has broad implications for understanding the physics of black-hole accretion. This challenge is compounded by the millions of AGN light curves produced by modern time-domain surveys. This project investigates potential advantages quantum machine-learning methods provide to this problem, and circumscribes the cases and conditions under which they can compete with classical approaches. We adopted quantum variational rewinding with lagged-innovation and found comparable performance across anomaly strengths to one-class SVM and isolation forests. To develop an approach native to modern quantum hardware, we developed a novel quantum reservoir that maps two-band AGN light curves onto 128 data-encoding qubits and 28 additional interaction qubits using the native connectivity of IBM Fez. Across repeated noisy runs, the reservoir preserved 100% anomaly-ranking consistency, with quantum/classical score correlations in both bands. Rather than reproducing a classical classifier, this framework investigates quantum advantage through how quantum hardware represents and distinguishes light curves.

Investigating the origin of seismic clustering in the Salton Trough
Claire D. Fagan
Mentors: Joann M. Stock and Allen Husker

California's Salton Trough (ST) region experiences oblique extension between the Pacific and North American Plates, geothermal and magmatic activity, changes in hydrological loading, and high seismicity. We aimed to identify causes of seismicity in this geologically complex region. We first used Southern California Earthquake Data Center (SCEDC) catalog data from 2023-2026 to identify seismicity and separate events into (1) swarms, (2) mainshock-aftershock sequences, (3) mixtures containing attributes of both (1) and (2), and (4) background seismicity. We correlated borehole well data from ST geothermal fields with the seismic data, finding that swarms and mixtures occur within a 3 km radius of geothermal injection wells twice as often as mainshock-aftershock sequences and background seismicity. Swarms and mixtures within the injection radius occurred at shallower depths and lower magnitudes. Many clusters outside the injection radius were planar. To investigate these clusters, we mapped new 3D faults in the ST using relocated seismicity from the SCEDC catalog and compared the clusters to all updated regional faults. Overall, we interpret that seismic clusters within the injection radius of geothermal wells and fields occur due to increased pore fluid pressure generated by geothermal activity, while clusters outside of injection radius occur via strike-slip fault friction.

SafePPO: Anytime constrained proximal policy optimization via safe gradient flow
Varun N. Gabbita
Mentors: Aaron D. Ames and Pol Mestres

Reinforcement learning (RL) policies for legged locomotion are typically trained without formal safety guarantees, allowing falls and instability to persist throughout training. This project presents SafePPO, an extension of Proximal Policy Optimization (PPO) using Reinforcement Learning-based Safe Gradient Flow (RL-SGF), an anytime-safe constrained optimization algorithm. At each training iteration, SafePPO solves a closed-form quadratically constrained update that balances task reward against a safety constraint, projecting the policy back toward safety whenever the constraint becomes active. We validate this approach first on a reduced-order linear inverted pendulum model, then on a full-order G1 humanoid walking task using center of pressure (CoP) as the safety constraint. Across both settings, SafePPO matches PPO's task performance while incurring significantly fewer safety violations, in simulation and on hardware. Additionally, we explore encoding reward terms as hard safety constraints, reducing the need for excessive reward tuning while preserving safety guarantees. Together, these results point toward a path for deploying learned locomotion policies with the rigor and predictability expected of classical safety-critical control.

Employing novel spectroscopic techniques in tracking biomarkers through coalification
Gabriella G. Giampietro
Mentors: John M. Eiler and Annie Edwards

Coal is a heterogeneous material made up of carbonaceous components derived from ancient flora. Quantifying biomolecule abundance in coals opens new avenues into understanding terrestrial paleoclimate at the time of deposition. However, current methods of quantifying these biomolecules are largely destructive, posing problems when analyzing sample-limited materials. Spectroscopy offers a non-destructive alternative for analyzing the geochemistry of fossil leaves and coals. We present new insights in the application of a range of traditional and novel spectroscopic methods, including bond selective fluorescence-detected infrared-excited (BonFIRE) spectroscopy, to the study of coal geochemistry. ATR-FTIR spectroscopy and Raman spectroscopy were conducted on coal rank samples and chemical standards including chlorophyll, lignin, and cellulose. Resulting spectra were analyzed and compared to chemical spectra of known compounds in order to predict visible signals when using novel spectroscopy techniques. Following literature review, BonFIRE spectroscopy was selected due to its potential to examine chemical specificity with greater sensitivity than traditional vibrational spectroscopy, making it well-suited to fossil applications. BonFIRE was then performed on the sample set and analyzed compared with established spectroscopic libraries. Preliminary results identify challenges of using BonFIRE in coal geochemistry, including selecting appropriate dyes to stain target molecules, navigating around machine blind spots, and sample preparation methodology. There is potential in applying BonFIRE to distinguish between molecules contributing to broad composite bands, such as aromatic carbon-carbon skeletal vibrations at 1593cm-1. Future directions include conducting additional novel spectroscopy methods on the sample set.

Single-shot universal quantum computation with 4D lifted product codes
Ian C. Gill
Mentors: John P. Preskill, Madelyn G. Cain, and Muzhou Ma

Quantum low-density parity check (qLDPC) codes provide a low spatial overhead path to fault-tolerant quantum computation by taking advantage of nonlocal connectivity to produce high encoding rates. However, present schemes for universal quantum computation at finite size using high-rate qLDPC codes require a time overhead which scales with the code distance in order to protect against time-like measurement errors. In this work, we present a framework for performing universal quantum computation with low time overhead using univariate quadracycle (UQ) codes, which are a subclass of 4D lifted product codes. We identify high rate, high distance UQ codes with check weight 12, including examples with parameters [[306,66,15]], [[306,54,18]], and [[510,54,28]]. For these codes, we construct surgery gadgets which use the underlying metacheck structure of the UQ codes to require a number of logical cycles per Pauli measurement which is significantly below the code distance. Furthermore, we develop a single-shot CCZ state preparation and injection procedures for the high rate codes presented using lower rate UQ codes equipped with transversal logical CCZ gates, enabling universal quantum computation in conjunction with the native fold transversal gates and surgery gadgets. Together, these results establish UQ codes as a concrete finite-size route to universal fault-tolerant quantum computation that retains the low spatial overhead of high-rate qLDPC codes while substantially reducing the temporal overhead of their essential logical operations.

Evaluating causal machine learning algorithms
David Gvadzabia
Mentors: Leonard J. Schulman and Erik Jahn

Causal discovery is the problem of inferring causal relationships between variables from observational data, offering insight beyond correlational or standard machine-learning methods. The problem is typically framed as recovering the directed acyclic graph (DAG) that represents the system, where a directed edge from X to Y indicates that X has a direct causal effect on Y. In recent years, a variety of algorithms have been proposed for this task, but they typically rely on strong assumptions about the data and data-generating process, such as independent and identically distributed (i.i.d.) samples and an acyclic ground-truth graph. These algorithms are also mostly validated on synthetic data, which lacks the complexity and noise of real-world measurements. To better understand how well they actually perform, we introduce computationally efficient baselines that are simplified versions of causal algorithms. We evaluate both the causal algorithms and these baselines on real-world datasets whose data-generating processes are independently known, rather than simulated. Our findings suggest new insights on the challenge of causal discovery on real-world datasets, and that new approaches may be required to achieve better performance.

Modelling the effects of tungsten alloying and surface oxidation on coverage-dependent hydrogen evolution in $Ni_{17}W_3$
Aditya Jain
Mentors: William A. Goddard, III and Sejun Kim

Electrolytic hydrogen offers a promising route to storing renewable energy, but current practice relies on platinum for the cathode reaction, whose cost and scarcity create a bottleneck to deployment at scale. Nickel-tungsten alloys are among the most promising non-precious replacements: nickel is abundant and already used in electrolysers, but binds hydrogen too strongly, and alloying with tungsten weakens that binding toward the Sabatier optimum. $Ni_{17}W_3$ is reported to reach 10 mA cm^{-2} at an overpotential of only 14 mV. Under operating conditions, however, the surface carries a substantial hydrogen coverage, and X-ray photoelectron spectra of the same material show tungsten in partially oxidised states alongside the metal. Here we use spin-polarised periodic DFT to build hydrogen coverage up site by site on pure Ni, on $Ni_{17}W_3$, and on $Ni_{17}W_3$ with two surface tungsten atoms each bearing an oxo oxygen, enumerating the symmetry-inequivalent sites at every step rather than assuming a filling order. We find that alloying and oxidation act principally by removing adsorption sites: of the sixteen face-centred-cubic hollows on pure Ni, ten remain accessible on the alloy, and oxidation displaces the additional hydrogen out of the hollow sites entirely. Because each surface reaches thermoneutrality at its own coverage, the descriptor that distinguishes these catalysts is where saturation falls. Density of states, charge analysis, and constant-potential calculations characterise the electronic origin of this behaviour, and nudged elastic band calculations probe the kinetics of hydrogen recombination.

Mapping the inner geometry of active galactic nuclei using Wiener Deconvolution of X-ray reverberation
Adam M. Janik
Mentors: Fiona A. Harrison and Demet Kirmizibayrak

Observations of X-ray reverberation time-lags provide a powerful probe of the geometry of space surrounding supermassive black holes in the center of active galactic nuclei (AGN). Current techniques for measuring these lags primarily rely on Fourier-domain cross-correlation analyses. While effective, these methods recover only aggregate time-lag properties of the system rather than directly resolving the impulse response function that encodes its underlying geometry.

The novel astronomical Wiener Deconvolution (WD) technique not only improves the accuracy and precision of time-lag measurements by suppressing instrumental noise, but also directly estimates the response function, which holds physical information about the system's geometry and light-travel delay processes. In this work, we apply the WD method to AGN with previously detected X-ray reverberation lags. By reconstructing the response function and searching for lags, we investigate whether Wiener deconvolution can place tighter constraints on the geometry of the innermost regions surrounding the central black holes.

Dual-space orbital learning for electronic band structure prediction of periodic materials
Minhyuk Kang
Mentors: Anima Anandkumar and Beom Seok Kang

Despite the importance of electronic band structures in computational chemistry, direct learning of full band dispersions remains underexplored, particularly at high levels of theory such as HSE06 or G0W0 quasiparticle methods. Here, we introduce OrbNet-Crystal, an orbital-informed framework predicting higher-fidelity electronic structures of crystals. Our approach employs dual-space orbital learning, combining low-cost orbital features in real and reciprocal space to capture both local chemical interactions and global band dispersion. The model predicts band structures at arbitrary levels of theory by learning corrections to a relatively computationally efficient PBE baseline.

We demonstrate the effectiveness of OrbNet-Crystal on more than 4,000 spin-resolved HSE06 band structure samples from three-dimensional crystalline materials, achieving a state-of-the-art band-energy mean absolute error of 0.18 eV. We additionally trained the model with beyond-DFT quality G0W0 quasiparticle band structures on more than 300 two-dimensional materials, achieving a strong band-energy mean absolute error of 0.09 eV. Finally, we extrapolate the model on previously unseen experimentally characterized perovskite materials testing on reported carrier effective masses and band gaps, demonstrating strong generalization beyond the training distribution. Overall, OrbNet-Crystal provides a scalable and physics-informed framework for predicting high-fidelity electronic band structures across multiple levels of electronic structure theory.

Characterizing astrocyte-like glia morphology in the *Drosophila* larval antennal lobe
Brianna N. Newton
Mentors: Carlos Lois and Aubrie L. De La Cruz

The brain consists of two primary cell types, neurons and glia, that work together to maintain normal neural activity. Similar to humans and other animal models, the *Drosophila melanogastor* nervous system is composed of neurons and glia, with the largest glial subtype being astrocytes.. Although neurons are well characterized for their structure and function, few studies have investigated the physiology of local glial cells within a subset of neurons. In particular, the larval *Drosophila* olfactory system consists of 21 olfactory receptor neurons (ORNs), each of which projects to exactly one glomerulus in the antennal lobe (AL). To investigate the morphology of astrocyte-like glia, we utilized the TRanscellular ACtivation of Transcription (TRACT) system to visualize neuron-astrocyte contact. This genetic tool uses Gal4-UAS and lexA-lexAop systems to drive expression of a ligand and receptor in neurons and astrocytes, respectively, in third instar *Drosophila* larvae. Our next step is to induce apoptosis in astrocytes interacting with the ORNs to understand the local cellular response to astrocyte death. Following dissection, larvae brains were immunostained and scanned using confocal microscopy. While our study characterized astrocyte morphology in the larval AL, future directions include investigating astrocytes in other conserved larval brain areas or studying other local glial cells.

Reconstruction of muon detector showers using machine learning
Diya Pai
Mentors: Maria Spiropulu and Soham Bhattacharya

In the CMS experiment, muon reconstruction is central to many physics analyses, and the problem grows harder with the HL-LHC upgrade, where higher luminosity will increase pileup, background rates, and detector occupancy. This project is motivated by a reconstruction bottleneck in converting low-level detector readouts into accurate particle-hit positions. The classical RecHit algorithm reaches matching efficiencies of approximately 8-25% within 1 cm and 40-65% within 5 cm depending on shower density, motivating learned methods capable of greater accuracy. We train a detector-aware transformer to predict hit positions and existence directly from detector signals. Work over the course of the project spanned diagnosing the model's weaknesses, tuning against the metrics of interest, and revising the model accordingly. Learned predictions are benchmarked against RecHit on matching efficiency, fake rate, residuals, and median and mean distance to truth, with the longer-term goal of improving reconstruction accuracy beyond the RecHit baseline.

Investigating the role of insulin signaling in zebrafish meningeal and neural development
Delia N. Parco
Mentors: Marianne Bronner and Ayyappa Raja

Insulin-like growth factor (IGF) signaling is an evolutionarily conserved pathway crucial for animal growth and development. Previous research has demonstrated that IGF binding protein 2 (IGFBP2) plays a major role in brain development and tumorogenesis; however, the specific role it plays in cell type development is not yet known. Using zebrafish (*Danio rerio*) as a vertebrate model, the lab generated a genetic knockout of *igfbp2a*, a zebrafish homolog of human *IGFBP2,* using CRISPR-Cas9 genome editing to investigate its role in embryonic development. 7-base-pair deletion mutants exhibit optic tectum closure defects by 48 hours post-fertilization, with phenotypic severity distributed in a pattern consistent with expected inheritance. In situ hybridization chain reaction experiments reveal a dramatic decrease in *igfbp2a* expression in the meninges of homozygous mutants, accompanied by a decrease in differentiated neuron markers and an increase in neural stem cell (NSC) markers, indicating self-cycling of NSCs. These findings suggest that *igfbp2a* may regulate brain development by modulating the brain meninges axis. Future work will be done in the lab to better understand the role of *igfbp2a* in insulin signaling during brain development. Together, these results establish zebrafish *igfbp2a* mutants as a viable model for investigating IGF signaling during development and neurogenesis.

Sharpening the search: Optimizing missed NEO detections for NEO Surveyor
Nikhita Penugonda
Mentors: Joseph Masiero and Ishan Mishra

NASA’s upcoming mission NEO Surveyor aims to significantly expand the catalog of Near-Earth Objects (NEOs) by detecting fainter bodies using observations in two bandpasses, NC1 and NC2. A key challenge in the completion of the software is optimizing the NEO Surveyor Survey Data System (NSDS), with a specific focus on its Moving Object Detection Pipeline (MODP). MODP links individual detections into triplets that encompass the object’s movement throughout the observation period, known as a quint. This project investigates why simulated NEOs are consistently missed by NMODE, the NEOS Moving Object Detection Engine, and identifies opportunities to improve efficiency of the pipeline. Using the simulated survey data, diagnostic scripts were developed to analyze the characteristics of missed detections, including average velocity, trajectory curvature, source crowding, moving point spread function accuracy, and more. Preliminary results indicate that the missed objects tend to cluster around an average velocity of 1deg/day and exhibit larger trajectory RMS values, suggesting that the current velocity grids and triplet allowance thresholds may need to be tweaked. These findings will guide refinements to pipeline parameters, with the goal of improving MODP completeness and supporting the mission objective of reliably detecting at least 90% of simulated NEOs.

Engineering fluorescent proteins for robust magnetic field sensing in mammalian cells
Francesco Plastina
Mentors: Mikhail G. Shapiro and Hengyu Li

Abstract withheld from publication at mentor's request.

Dual-frequency SWV measurements of aptamer sensor drift during extended continuous monitoring in artificial and real sweat
Conor K. Sargent
Mentors: Wei Gao and Hyunah Ahn

Continuous, non-invasive biomarker monitoring could transform cancer detection and management by replacing invasive blood draws with wearable diagnostics. A wearable sensor could let survivors continuously monitor for recurrence, or catch a new cancer at Stage 1 before it progresses. Electrochemical aptamer-based (EAB) sensors suit this well, since aptamers can be engineered to bind nearly any target antigen. This project validates an EAB platform using cortisol as a proof-of-concept biomarker before extending the approach to cancer markers such as PSA and HER2. The gold electrode sensors attached with a cortisol aptamer were tested using a square wave voltammetry frequency sweep (10–60 Hz) to identify the strongest signal pair, then tested continuously (3–24+ hours) in artificial sweat (.25x PBS) and real sweat. Our code corrected or eliminated faulty channels and calculated a dual-frequency ratiometric signal (I_R) to compensate for drift. Over a 24-hour PBS run, I_R remained within approximately 5% of baseline, closely matching published benchmarks. One key limitation was identified, being the fact that the ratiometric signal becomes unstable when either frequency's current crosses zero. With the remaining project time, focus has narrowed to fully validating long-term cortisol stability, with the goal to eventually implement this in cancer detection.

Encoding and replicating hierarchical spatial information in reaction-diffusion systems
Eric F. Segrest
Mentors: Erik Winfree and Daichi Hayakawa

Living systems store information in geometry: when a cell divides, it copies not only its genetic sequence but also the spatial arrangement of its parts. Chemical mixtures that react and diffuse may spontaneously organize into spots, stripes, and other related patterns, some of which even reproduce themselves. However, no reaction-diffusion system has been shown to replicate a pattern with complex internal structure, in which distinct labeled regions are inherited by each copy. We ask whether such information can be copied by local chemical interactions alone, without templates or external control. Our target pattern, an American flag, is built hierarchically: a large chemical spot defines the boundary of each copy, smaller interior spots of two distinct types divide that boundary into labeled regions, and further reactions fill those regions with finer patterns. In simulation, asymmetric coupling between the two interior types favors one region of the first type and three of the second inside each dividing spot, breaking the symmetry that defeated earlier designs. Hierarchical coupling of simple chemical patterns may therefore preserve spatial organization across replication, suggesting a route toward chemically plausible self-replication of complex geometries.

Neural and behavioral correlates of social connection during human-AI interaction
Anika Shethia
Mentors: Dean Mobbs and Emily Towner

Social disconnection is a growing global concern, and people are increasingly turning to artificial intelligence (AI) for social support. Yet the neural and psychological mechanisms through which social connection forms during naturalistic social interactions, whether with humans or AI, remain poorly understood. This project applies switching linear dynamical systems (SLDS), a mathematical framework for modeling how internal brain states evolve over time, to test whether discrete states of connection and disconnection can be recovered empirically from neural data. We further ask how social connection alters the way we cognitively represent others, adaptively mentalize about them, and respond to their influence. Thirty participants hold text conversations with artificial partners designed to foster connection or disconnection while undergoing functional magnetic resonance imaging (fMRI). Participants also complete a face-viewing task before and after conversation, a Rock-Paper-Scissors task, and an artwork-rating task, indexing self-other neural overlap, adaptive mentalizing, and susceptibility to social influence respectively. Pilot behavioral data from a separate online sample (N = 40) confirm that the manipulation produces the intended difference in felt connection, validating the paradigm for the scanner phase. Scanning is ongoing; analysis will test whether latent connection states predict mentalizing depth and influence.

Lactadherin-modified biosensor for detection of inflammatory biomarker in perspiration
Jessica Slaughter
Mentors: Wei Gao and Jeonghee Yeom

The release of phosphatidylserine (PS)-exposed microparticles (MPs) during cell death and stress responses is a key biomarker of inflammation and has been associated with conditions such as severe decompression sickness (DCS), which disproportionately affects US Navy divers and individuals engaged in prolonged SCUBA diving. Current methods for measuring these biomarkers rely primarily on blood analysis, which is invasive and requires trained personnel, sterile sample handling, and laboratory processing, thereby limiting its suitability for continuous and in situ monitoring. Wearable electrochemical biosensors offer a promising alternative by enabling real-time, noninvasive detection of biomarkers in accessible biofluids such as sweat. Here, we present a wearable biosensor for the electrochemical detection of sweat MPs. The working electrode is functionalized with lactadherin, an affinity protein that selectively captures PS exposed on the outer leaflet of inflammatory MPs. To achieve sensitive detection, we employ a competitive binding assay in which methylene blue (MB)-loaded liposomes serve as signaling probes. As the concentration of PS-rich MPs increases, these MPs compete with MB-loaded liposomes for fixed binding sites on immobilized lactadherin, consequently decreasing the electrochemical signal. This platform has the potential to enable early identification of individuals at risk of severe DCS and continuous monitoring of inflammatory conditions in point-of-care and remote settings.

A pose-based pipeline for automated tennis shot detection and serve analysis for coaching applications
Theodore K. Stoltz
Mentors: Pietro Perona and Ilona Demler

Coaches currently rely on manual video review to evaluate player technique, a process that does not scale to the hours of practice and match footage a team generates and that offers no structured, quantitative record of a player's motion over time. We are developing the CalTennis Coaching Tool, a pipeline that automatically extracts and classifies tennis shots from raw video to support this kind of coaching feedback at scale. Each frame of an uploaded video is passed through RTMPose, a real-time multi-person pose estimation model, to recover 2D keypoints for every visible player, who are then tracked individually across frames. From each player's wrist, elbow, shoulder, and hip keypoints we compute a torso-normalized wrist-velocity signal, and use peak detection together with a set of motion-based rules to identify and classify shot events as serves, groundstrokes, or volleys, with a separate motion-boundary detector giving each serve a precise start and end timestamp. This pipeline has been used to build a large, structured dataset of labeled serve clips and paired pose sequences spanning professional, intermediate, and novice players, which is now being used to train a machine learning classifier to replace the current hand-tuned detection thresholds, with the further goal of testing whether pose-derived motion features can serve as a quantitative marker of player skill level. The pipeline has also been ported to a native iOS app for live, on-device shot tracking, a step toward giving coaches real-time feedback during practice itself.

Fault-tolerant coordination for the Near-Earth Object Surveyor ground data system
Ahmad Wahab
Mentors: Joseph Masiero and Analetta Ehler

The Near-Earth Object Surveyor mission requires a ground data system that can continue operating despite computer failures, restarts, and network disruptions. Its cluster consists of a leader node that coordinates processing and worker nodes that execute data-processing tasks. Previously, worker membership depended on manual configuration or remote procedure calls, while coordination relied on a fixed leader, creating a single point of failure. This project developed a fault-tolerant coordination layer in Rust that replaces those mechanisms with automatic membership management and leader selection. Membership is maintained using the Scalable Weakly-consistent Infection-style Process Group Membership protocol, in which computers probe one another and spread status updates through randomized gossip. Direct and indirect probes, suspicion timers, and Lifeguard-inspired local health awareness reduce false failure reports during slowdowns or overload. Periodic full-state exchanges repair missed updates, while process-generation and incarnation numbers prevent outdated messages from reviving failed or replaced processes. Leadership is assigned through a PostgreSQL lease, and each new term receives a higher fencing epoch so workers reject delayed commands from former leaders. An elected leader must restore the processing state before accepting work. Tests demonstrate cluster formation, failure detection, restart recovery, view repair, graceful departure, and automatic leadership transfer after shutdown or failure.

The Type-Theory Virtual Machine: A logic-aware, compiler-like system for theorem-prover interoperability
Jiayi Wu
Mentors: Anima Anandkumar and Robert Joseph George

Formal mathematics remains fragmented across theorem-prover ecosystems, and the benchmark ITPEval shows that current language models struggle especially with translations involving ecosystem-specific dependencies. Therefore, we present the **Type-Theory Virtual Machine (TTVM)**, an end-to-end translation pipeline that preserves declaration dependencies and foundation-specific assumptions across provers: starting from exporter and checker evidence, TTVM builds a checked source unit containing a theorem and its required declarations; the Core Intermediate Representation (CoreIR) then normalizes supported logical structure while preserving provenance and keeping unknown symbols opaque; at the same time, AtlasTT evaluates transport between the source and target logical worlds, where effects record logical and trust obligations, morphisms declare transport paths, and residues identify unresolved incompatibilities. Together, the AtlasTT decision and dependency graph determine a target-neutral transport plan. A compatible backend then renders the ordered declarations and theorem in target syntax for checking by the target prover. This layered design replaces pairwise rules with reusable source decoders, foundation-aware transport, and target backends, establishing a scalable basis for cross-ITP statement translation. Overall, TTVM enables formal results to be reused across prover ecosystems without collapsing distinct logical foundations or hiding the assumptions on which they depend.

Optimizing sequential resin–latex extractions from guayule through analytical characterizations
Ruoyun Xu
Mentors: Julia A. Kornfield and Florence Mueller

Guayule is a promising natural rubber source alternative. Compared to Hevea brasiliensis, guayule is domestic to the US, less susceptible to diseases and its rubber has better mechanical properties and is hypoallergenic. The guayule plant produces two high value products: latex and resin; which both need to be extracted in the highest possible purity for usability and high performance. This work addresses the co-extraction of both products located in the bark through a sequential resin-then-latex extraction method. The aim is to reduce cross-contamination of resin in latex while preserving the resin.

Fresh guayule bark was first sliced, opening resin channels while keeping the latex encapsulated in the parenchymal cells. The bark slices were then submerged in an extraction buffer and immiscible solvent was added on top. The mixture was then agitated to extract the resin while keeping the bark hydrated to dissolve the resin into the solvent. Latex was then extracted through conventional methods and the resin was dried from the solution. This work gives a quantitative comparison of the variation of different process parameters, including the ratio of buffer to solvent and the agitation time. The purity of the latex and resin were analyzed using thermogravimetric analysis and thin layer chromatography, respectively.

By exploring the different process parameters of the sequential extraction process, this work informs important variables for a scale-up of guayule latex and resin extraction.

Exploring new nonreciprocal photonic platforms for controlling thermal radiation
Yiting Zhao
Mentors: Harry A. Atwater, Jr., and Simo Pajovic

Thermal radiation originates from the random motion of matter and thus pervades the world around us. Given its ubiquitousness, the control of thermal radiation has many important applications, such as radiative cooling and solar energy harvesting. Thermal emitters are constrained by Kirchhoff's law of thermal radiation, stating that the emissivity and absorptivity of a surface are equal for a given wavelength, direction, and polarization of light. Current approaches to lifting this constraint and achieving "nonreciprocal thermal radiation" emphasize materials such as InAs, which break the emissivity-absorptivity equality under a magnetic field. However, these approaches are limited, requiring a large magnetic field or only functioning for specific polarizations. Here, we explore two new platforms for nonreciprocity. The first achieves dual-polarized nonreciprocal thermal radiation using an array of silicon micropores on bulk InAs. We successfully fabricated this structure and experimentally demonstrated nonreciprocal absorptivity. The second is a theoretical thin film InAs structure which can be fabricated via spalling, a technique recently applied to III-V materials. We predict that this structure's nonreciprocal optical response differs from bulk substrates, suggesting a novel, experimentally accessible InAs platform. Our findings identify new ways to tailor thermal radiation, with potential applications in thermal management and energy conversion.