

Growing cells in computers: Surrogate modeling of ion-channel sensitivity in human cortical neurons

Zeynep Goktepe

Mentors: Costas Anastassiou and Nooshin Abdollahi

Detailed neuron models help us understand how ion channels affect a neuron's electrical behavior. However, running these simulations many times can be slow and computationally expensive. In this project, I trained neural networks to predict 14 electrical features from 20 biophysical parameters and the stimulus amplitude. I used simulation data from 48 human cortical neurons representing two cell types. For each cell, I trained separate models for spiking and subthreshold responses. I then used the trained models to measure how sensitive each predicted feature was to changes in each input parameter. Overall, 95.5% of the cell-feature predictions had a validation score of at least 0.5. Subthreshold features were generally predicted more accurately than spiking features. Two independent groups of cells from the same cell type also produced similar parameter rankings and gradient directions. I reported sensitivity using normalized gradients, gradients in physical units, and relative changes called elasticities. These results show that neural-network surrogates can provide fast and reproducible estimates of parameter sensitivity. I am now comparing these results with direct parameter changes in the original neuron simulator.

Developing a computational optimization framework for cell-class-specific electrical stimulation

Ishita Banerjee

Mentors: Costas Anastassiou and Philip Wong

Electrical stimulation (ES) is widely used to modulate neuronal activity in both basic neuroscience research and clinical neuromodulation therapies. However, designing stimulation protocols that selectively target specific neuronal populations remains a significant challenge due to the complex relationship between network parameters and neuronal dynamics. This project aims to develop a computational optimization framework for identifying network parameters that reproduce target electrophysiological behavior in biophysically detailed neuronal networks. To investigate this problem, a recurrent network composed of six parvalbumin (PVALB) interneurons was constructed in BMTK, and a randomized parameter search was implemented to optimize excitatory synaptic weight, inhibitory synaptic weight, synaptic delay, and stimulus amplitude using a loss function based on voltage trace similarity and firing rate. Optimization performance was evaluated across twenty independent trials through analyses of voltage traces, firing rates, optimization loss, and parameter recovery. The optimization framework successfully identified parameter sets that generated network dynamics similar to those of a reference network, with the lowest-loss solutions closely reproducing the target voltage traces and firing behavior. The results also demonstrate that multiple parameter combinations can produce comparable network activity, highlighting the challenges of parameter recovery within high-dimensional optimization spaces. These findings establish an initial computational framework for network-level parameter optimization and provide a foundation for developing more efficient optimization strategies for computational models of electrical stimulation.

Shaping bio-realistic spiking neural network dynamics via gradient-based optimization of connectivity parameters

Maggie Y. Bai

Mentors: Costas Anastassiou and Shayan Farzad

Neural oscillations are a hallmark of coordinated brain activity, essential for communication, sensory processing, and cognition, and their disruption is implicated in disorders such as schizophrenia and epilepsy. While biophysically detailed models can faithfully capture the behavior of individual neurons, the circuit-level parameters that govern network dynamics--synaptic strengths, connection probabilities, and time constants--are far less constrained by data, and their mapping to emergent activity is highly nonlinear. Tuning these parameters to produce a target behavior therefore typically

relies on slow, gradient-free search that can require many times more simulations than gradient-based methods. This project investigates whether Jaxley, a differentiable simulator that computes exact gradients through networks of morphologically detailed neurons, can instead shape emergent dynamics directly by optimizing only connectivity parameters. Working with recurrent excitatory-inhibitory networks that produce gamma-band oscillations, we assess how well gradient-based optimization can steer a circuit toward a defined oscillatory state, and benchmark its efficiency and reliability against existing approaches--aiming to make large-scale, biologically realistic brain simulation more tractable.

Automated body composition analysis for cardiac risk prediction in Duchenne muscular dystrophy

Christian W. Cavener

Mentors: Shreyas S. Vasanawala and Konstantin Zuev

Duchenne muscular dystrophy (DMD) is an X-linked progressive neuromuscular disorder in which cardiomyopathy remains a leading cause of death. Skeletal muscle atrophy and fat deposition visible on magnetic resonance imaging (MRI) may provide quantitative biomarkers for predicting cardiac decline and evaluating treatment response in patients. However, MRI-based body composition analysis in patients with DMD is still largely performed manually or semi-automatically, limiting its scalability and routine clinical use. This project aims to develop an artificial intelligence-based pipeline for automated tissue segmentation and body composition analysis from cardiac MRI scans in patients with DMD. A retrospective cohort of approximately 200 patients with DMD undergoing routine cardiac MRI will be identified. Axial localizer images will be manually segmented using 3D Slicer to generate ground-truth labels for subcutaneous adipose tissue, visceral adipose tissue, epicardial adipose tissue, intramuscular adipose tissue, and skeletal muscle. Skeletal muscle fat fraction will also be calculated. A U-Net-based convolutional neural network will be developed using Python, PyTorch, and MONAI and will utilize transfer learning from existing tissue segmentation frameworks such as nnU-Net. Both volumetric segmentation across selected vertebral levels and single-slice segmentation at a standardized mid-thoracic level will be evaluated. Model performance will be evaluated by comparing automated predictions with manual tissue segmentations using Dice similarity coefficients and Jaccard indices. The resulting quantitative body composition measurements may ultimately support improved prediction of cardiovascular outcomes and treatment response in patients with DMD.

GPU-accelerated multicoil Bloch simulations for nonlinear inversion in MRI reconstruction and quantitative mapping

Alya Kurt

Mentors: Shreyas S. Vasanawala, Andres Collazo, Carlos A. Castillo Passi, and Daniel Ennis

Accelerated MRI requires reconstruction methods that use receive-coil sensitivities. Model-based reconstruction methods have shown improved image quality and accuracy by incorporating physical effects, but need sequence-specific adjoints. Differentiable physics-based reconstruction maintains image quality while being reusable across sequences; but at a high computational cost. To develop practical differentiable physics-based image reconstruction we: (1) developed a GPU-accelerated simulation which includes multi-coil sensitivities by extending KomaMRI.jl, and also (2) used this simulation in an optimization problem to retrieve parametric maps like spin density, T1 and T2, where Enzyme.jl and Reactant.jl were used for gradient calculation via reverse-mode automatic differentiation (AD). In this way, we can ensure data consistency, as the predicted KomaMRI-simulated signals are forced to match target measurements. We compared our approach using AD against finite differences (FD). FD took 60 hours to reconstruct in comparison with our AD-based framework that completed the same 12-coil nonlinear-inversion in 7 minutes. Furthermore, compared to traditional reconstruction methods like CG-SENSE, we achieved an error of 0.422% (normalized data-consistency residual) vs 6.05% for CG-SENSE, representing a 93.0% error reduction. Thus, this work enables fast and accurate physics-based MRI reconstruction in a pulse-sequence agnostic way.

Persona- and urgency-driven sycophancy in frontier LLMs

Clara Yu

Mentors: Ankit Sakhuja, Shinsuke Shimojo, and Nimay Hazare

The integration of Large Language Models (LLMs) into clinical workflows introduces sophisticated capabilities for diagnostic reasoning and patient communication, yet a critical vulnerability lies in the inconsistency of their internal knowledge states across varying linguistic frames. This paper investigates persona- and urgency-driven sycophancy in frontier LLMs, evaluating over 48,000 responses to examine how role-playing constraints and perceived situational urgency affect core clinical accuracy. Our results show that 12.0% of all responses displayed sycophantic behavior, with susceptibility varying by more than an order of magnitude across models, ranging from 1.4% sycophancy in Claude 4.6 Sonnet to 38.4% in Gemini 3.5 Flash. Furthermore, prompting models to interact with high-status clinical personas significantly increased sycophantic behavior compared to lay-person personas, peaking with the Senior Physician persona. This effect increased monotonically with situational urgency, from a 4.3% sycophancy rate at Level 0 to 23.8% at Level 4 (OR 18.90), with critical clinical reversals accounting for 63% of failures. Ultimately, these findings indicate that stylistic prompting significantly degrades the consistency of medical reasoning, highlighting the urgent need for future alignment strategies to safeguard core knowledge states before deploying LLMs as clinical decision support systems.

Physical plausibility emerges from resilience across diverse noise paths in video diffusion models

Richard Hoffmann

Mentors: Yilun Du, Georgia Gkioxari, Ruojin Cai, and Sarah Liaw

Pre-trained video diffusion models excel at generating visually realistic rollouts, but their predictions often lack an understanding of real-world dynamics. This can lead to physically implausible state transitions, temporal artifacts, and unexpected hallucinations. However, generated videos can still simulate future objections in robot planning and decision-making settings, making the use of video generation models as world models for embodied agents a dominant paradigm. For generalizable robot learning beyond controlled settings, it is imperative that video models better represent real-world dynamics. Previous work improve generated rollouts by sampling more trajectories to increase the chance of outputting a coherent rollout, but standard best-of-K methods require an external verifier, hand-tuned rewards, and inefficiently delete information from unsuccessful samples. Instead of relying on retraining or additional external verification methods, we propose an inference-time debate algorithm that treats independently initialized denoising trajectories as individual agents proposing alternate futures. Agents challenge proposals by re-noising with their own initial noise tensors and querying the video model to reconstruct it. We claim that a more physically plausible rollout is one that can be recovered across several counterfactual noise directions, which we evaluate on the Physics-IQ benchmark.

Enhancing predictive generative policies with learned rewards

Divan I. Mejia Gonzalez

Mentors: Yilun Du, Adam C. Wierman, and Haldun Balim

Many manipulation tasks give a reward only on completion, and writing a denser reward by hand is difficult and specific to each task. Reinforcement learning algorithms for learning critics have been shown to be brittle in this regime. We design a planning algorithm for generative policies that instead uses a reward learned directly from expert demonstrations. This reward, which we call the progression function, defines a preference over states through the Bradley Terry model, and is trained to favor states closer to task completion in the demonstrations. The generative prior and the reward are combined within a model predictive control framework. At each decision the policy proposes a set of candidate action sequences, a learned world model predicts the states each would reach, and the progression function scores those predicted states, yielding an objective for optimization. We evaluate our method on manipulation tasks from the robomimic benchmark, with both low dimensional state and camera observations, and demonstrate improvements over the base policy and over our algorithm when using other learned rewards.

Predicting isotope effects with orbital-based deep learning

Seokhyun Chin

Mentors: Anima Anandkumar, Beom Seok Kang, John M. Eiler, and Simon Andren

Isotope effects play an important role in providing a mechanistic understanding of chemical reactions, but the exact quantification of isotope effects is limited by the high computational cost of quantum chemistry methods. Previous work has demonstrated notable success in predicting equilibrium isotope effects using AI, but it is only able to predict values for H, C, N, and O and excludes sulfur-bearing compounds. In this work, we demonstrate the capabilities of AI to predict equilibrium isotope effects of sulfur-bearing compounds and kinetic isotope effects. A chemical understanding of Sulfur isotope effects is first built by comparing the chemical environments of each sulfur atom and corresponding isotope effect values, and an SE(3)-equivariant graph neural network is trained to predict isotope effects of sulfur-bearing compounds and kinetic isotope effects. The model demonstrates high performance in both cases, predicting sulfur isotope effects within experimental accuracy of 1‰ and kinetic isotope effects of heavy atoms within 5 ‰. The trained models can serve as a valuable tool to be applied to reaction mechanism elucidation of small organic molecules at a much lower computational cost compared to conventional quantum chemistry methods.

Learning exact exchange for Kohn-Sham density functional theory

Arimitsu Horikawa-Strakovsky

Mentors: Anima Anandkumar, Danish Khan, and Maurice Hanisch

Kohn-Sham density functional theory (DFT) is one of the main methods utilized in modern quantum chemistry for electronic structure calculations. Semi-local functionals such as PBE suffer from self-interaction error, leading to systematic errors in calculations of many properties such as reaction barrier heights. Hybrid functionals such as PBE0 mitigate this issue by incorporating a fraction of exact exchange, substantially improving accuracy, but at the cost of increasing computational scaling from $O(N^3)$ to $O(N^4)$. Thus, we learn exact exchange using a continuous radial factored Fourier neural operator (cRFNO). Training on only 500 neutral molecules, we are able to approach the accuracy of hybrid functionals for reaction barrier heights on BH76, achieving errors within 2.2 kcal/mol of PBE0 compared to PBE's error of 5 kcal/mol, while retaining the scaling of semi-local DFT. Despite training only on neutral molecules, our model transfers well to charged molecules and reactions, demonstrating strong generalization. We also observe substantial improvements over PBE in electron density and total energy, halving the error relative to PBE0. These results suggest that our approach offers a promising path toward emulating the accuracy of hybrid functionals at a fraction of their computational cost.

Physics-informed 3D shape generation using Laplacian preconditioned guidance

Hudson Liu

Mentors: Anima Anandkumar and Xinyi Li

Denoising diffusion models are a leading approach to 3D shape generation tasks across multiple representations, including point clouds and signed distance fields (SDFs). A popular task for such models is to generate shapes that satisfy physics constraints. However, direct application of diffusion guidance can be unstable for shape generation, due to noisy and sharply localized gradients. We propose preconditioning the guidance update with an operator derived from the Laplace-Beltrami operator, smoothing the per-position gradients over the surface through a single implicit step of the heat equation. We evaluate our method on TIGER, a point-cloud generation model, and SDF-Diffusion, which produces truncated SDFs (T-SDFs) discretized on a regular grid. For TIGER, we define a kNN uniform Laplacian over the point cloud; for SDF-Diffusion, we define both a mollified cotangent Laplacian on the extracted zero-set mesh, and a projected Laplacian operator defined directly on the T-SDF grid itself. As guidance objectives, we introduce a bilateral-symmetry loss and an upright-stability loss.

Walk-on-Boundary supervision for linear elasticity neural operators

Siqi Liu

Mentors: Anima Anandkumar and Xi Deng

The linear elasticity partial differential equation models the deformation of structures under applied forces, making it important to engineering design. The industry standard for solving this PDE is the finite element method (FEM), but it requires a volumetric mesh, which can be expensive to construct and also introduces deterministic discretization error. Monte Carlo methods such as Walk-on-Spheres (WoS) avoid volumetric meshing, but WoS is designed for scalar equations and cannot be applied directly to the coupled, vector-valued linear elasticity PDE. Our project aims to implement a volumetric-mesh-free stochastic solver for linear elasticity and use its estimates as noisy labels for the supervised training of neural operators. Since the PDE has a boundary-integral representation, we implement a Walk-on-Boundary (WoB) algorithm using a matrix-valued elasticity kernel. A main challenge is that the WoB estimates are noisy, so we apply local importance-sampling methods, such as three-plane sampling, to reduce their variance. Next, we will generate datasets that vary geometry, material parameters, boundary conditions, and samples per point, then train a neural operator to predict solution fields from the noisy labels. This will allow us to study whether neural operators can learn accurate solution maps from inexpensive, noisy Monte Carlo supervision.

Faster electron density prediction with Fourier neural operator

Tuyen T. Nguyen

Mentors: Anima Anandkumar and Maurice Hanisch

Predicting the electronic structure of materials is a central task in computational chemistry and physics, but the self-consistent field (SCF) calculations underlying density functional theory (DFT) are computationally expensive at $O(N^3)$ scaling, limiting the scale and speed of simulations. Machine learning approaches, and neural operators in particular, offer a promising route to accelerating these calculations by learning to predict electron densities directly. We investigate whether an Fourier neural operator (FNO) based SCF calculation can be further accelerated by performing inference at a lower spatial resolution. We find that lower-resolution calculations can achieve similar convergence when evaluated against reference densities at the same resolution. However, when the resulting densities are compared against the original full-resolution reference, the error increases substantially. Our experiments show that most of this additional error is caused by information lost during downsampling rather than by poorer SCF convergence.