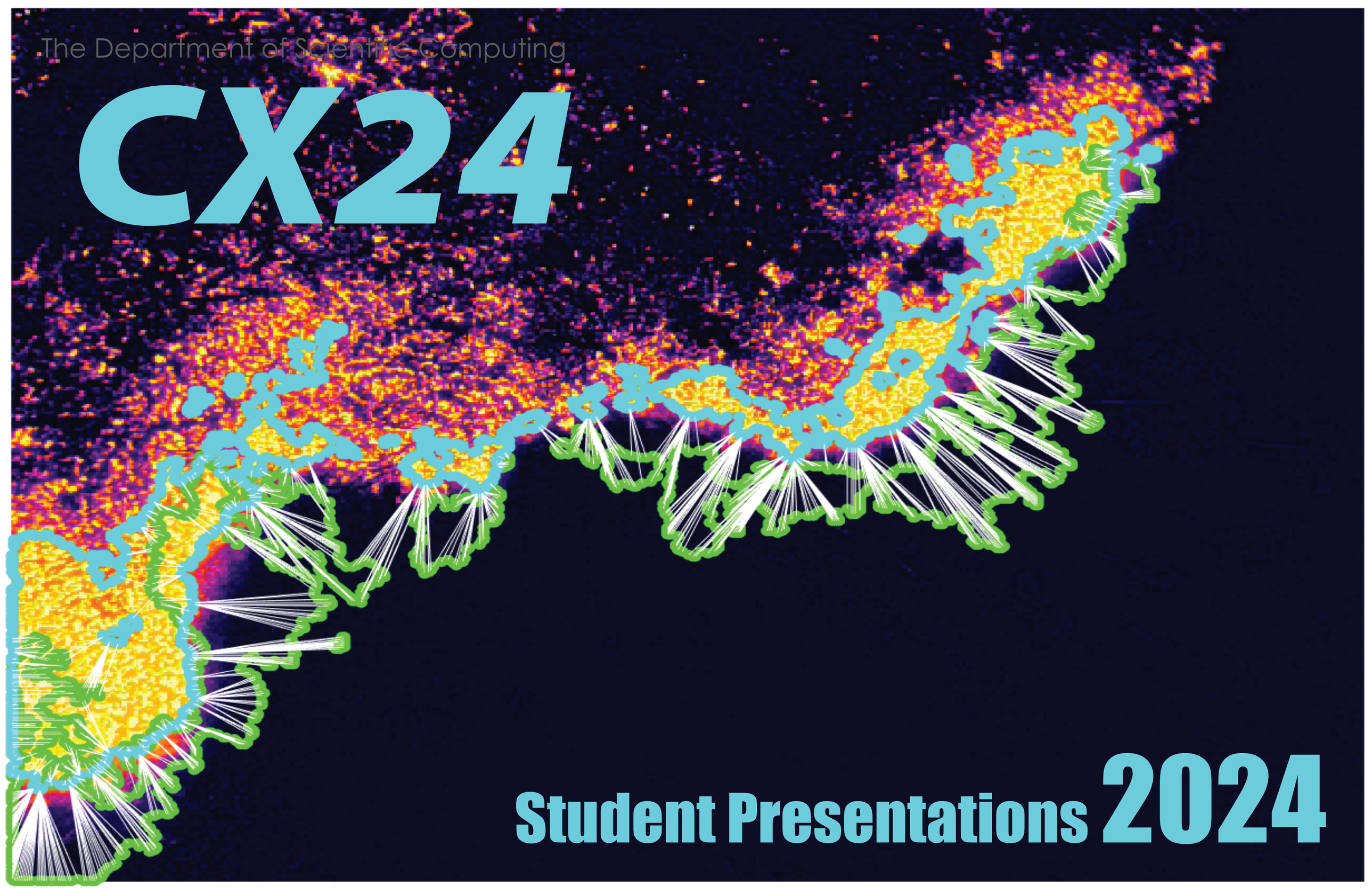
The Department of Scientific Computing

CX24



Student Presentations 2024

Welcome to Computational Exposition 2024, a yearly event where students in the Department of Scientific Computing (DSC) showcase the results of their research in the past year. This research covers a broad spectrum of disciplines, but shares a common thread: they concentrate on algorithm development and blend the computational and the mathematical to solve problems in the applied sciences. The innovation displayed is broad and remarkable. Our students make us proud!

The student posters reflect the breadth and depth of the research carried out in the DSC. They are the direct result of a fulfillment of our two most important missions: providing world-class interdisciplinary research and training in computational science.

Our graduate degree programs and the success of our students bolster our confidence that we are becoming a premier institution for the training of the next generation of computational scientists. Indeed, looking at what our current students have achieved over the past several years serves as evidence that we are already there!

So, enjoy the presentation, interact with the students, challenge them, learn from them, and reflect on the fruits of their intelligence, skills, dedication, and labor, and join us in thanking them for their contributions to the DSC, to FSU, and to science.

Peter Beerli

Chair, Department of Scientific Computing

Cover: Infrared image of fire from UAV, courtesy of Daryn Sagel.

Presenting Researchers

RESEARCH

Abstracts & Graphics

Joshua Briley
Undergraduate in Computational Science
Advisor: *Bryan Quaife*

Extending Four Bugs on a Square: The Infinite Bug Chase

The classical problem of 'Four Bugs on a Square' is often used to motivate systems of differential equations. Initially positioned at the vertices of a square, the four bugs pursue one another in a cyclic fashion and with a constant speed. The bugs converge to one another, and the asymptotic behavior of this convergence is well-understood. An extension of this problem is to initialize and constrain the bugs to a surface. With this simple modification, interesting dynamics are revealed for geometries as simple as a circle.

When confined to a circle, the bugs eventually assume one of two states: convergence to a single point or an infinite chase loop. Using Monte Carlo methods, I will estimate the probability that bugs initialized randomly will enter each of these two states. Notably, I will show that the likelihood of the bugs converging to a point decreases as the number of bugs increases. I will also show how probabilities can be calculated analytically, and I will carry out these calculations for the three-bug and four-bug scenarios.

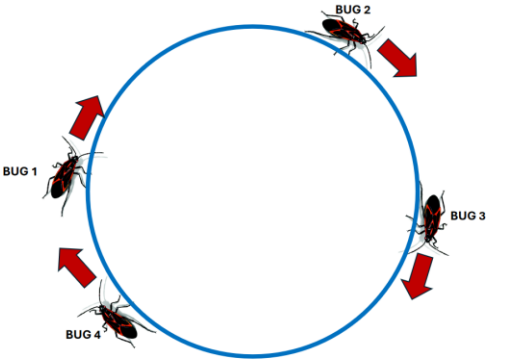
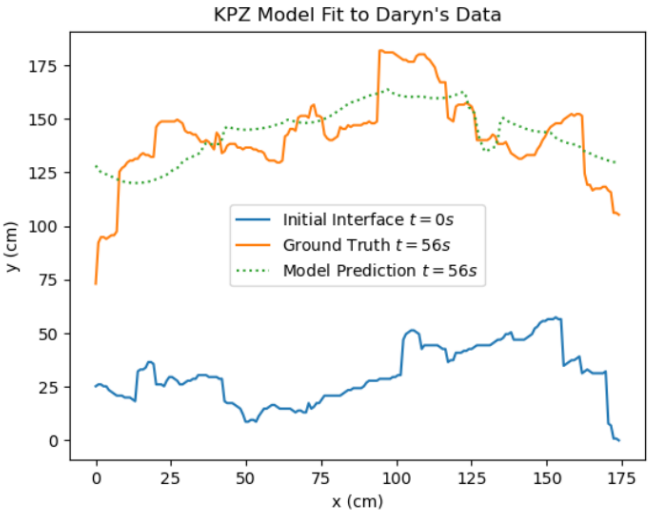


Figure 1, Above: Bugs on a Square.

William Callender
Undergraduate in Computational Science
Advisors: *Bryan Quaife & Kevin Speer*

Efficient Wildland Fire Spread Modeling Using the Kardar-Parisi-Zhang Equation

The project explores the applicability of the Kardar-Parisi-Zhang (KPZ) equation to fire spread dynamics. Building on the work of Campos et al., I estimate the coefficients of the KPZ equation by applying an inverse problem to a dataset collected by Daryn Sagel. Starting with reasonable initial parameter values, the inverse problem uses a forward solver for the KPZ equation to refine the parameter values so that the error between the simulations and data are minimized. The inverse method works reliably well for synthetic data, and demonstrates a reasonable fit to real-world data, with room for improvement. The figure shows an example of the initial condition (blue), final condition (orange), and predicted numerical solution (green). Future work should focus on applying this method to other moving interface problems arising in fire science, such as the plume and larger-scale fire lines.



Pankaj Chouhan
Ph.D. in Computational Science
Advisor: *Sachin Shanbhag*

Designing Gaussian Processes based Surrogate models for the Molecular Models of Polymer Viscoelasticity

Since their introduction in the early 1900s, synthetic polymers have become ubiquitous in almost every sector of the modern economy. Understanding their physical and chemical properties accurately has become crucial due to their wide use. Molecular models have played a significant role in expediting this understanding, particularly in fields like rheology. Here, these models help study the stress-strain relationship of polymers. Molecular simulation models for polymer rheology use the molecular structure of a polymer mixture to predict its linear or nonlinear viscoelasticity. However, the computational cost of molecular simulations can be high for some applications. Surrogate models offer a solution by approximating the behavior of molecular simulations at a lower computational cost. This study focuses on building surrogate models based on Gaussian processes for molecular models of polymer rheology. A key aspect is that each training data point is a vector-valued observation, and the training dataset is limited. The modeling approach involves using spectral decomposition techniques like KL-expansion and B-splines in conjunction with Gaussian processes. These techniques project the vector-valued observations into a lower-dimensional space, where the coefficients of the projection are used to build the surrogate. When given a test point, the surrogate infers the coefficients and then transforms them back to the original space. The study explores how the accuracy of the surrogate changes with increasing complexity in the data. It also compares the proposed surrogate with previous work that used separable kernel-based Gaussian processes and a recently proposed neural network-based surrogate for vector-valued data.

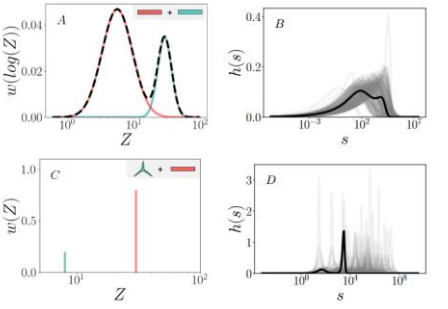


Figure 1: Subplot (A) illustrates the molecular weight distribution (dashed black line) of a linear-linear polymer. The contribution of the two linear components is shown by the red and teal curves. Subplot (B) shows the relaxation spectrum (dark black line) corresponding to the blend shown in (A). The gray curves represent the relaxation spectrum for other input values. Subplot (c) represents the molecular weights of the monodisperse star (teal) and linear (red) fractions in a star-linear blend. Subplot (D) shows the relaxation spectrum of this particular star-linear blend (dark black line) and the other 159 datasets (thin gray lines) considered in this study.

Jhamieka Greenwood
Ph.D. in Computational Science
Advisors: *Bryan Quaife & Kevin Speer*

Adverse Effects on Erosion and Biogeochemical Fluxes Post-Fire

Understanding the complex dynamics of wildfire spread is critical for predicting fire behavior and improving management strategies. Traditional models often rely on simplified assumptions or extensive computational resources, which may not accurately capture the nonlinear and turbulent nature of fire spread. By focusing on a controlled head fire spreading across a 2m x 2m plot of pine straw, we seek to unveil the intricate mechanisms driving fine scale fire behavior. The high-resolution spatial and temporal data of the firespread were collected by Daryn Sagel, using advanced imaging and sensing technologies.

This study aims to leverage the Sparse Identification of Nonlinear Dynamical systems (SINDy) algorithm, specifically its weak formulation (weak SINDy), to identify the governing equations of fire spread.

The weak SINDy method, an innovative data-driven approach, is applied to this dataset to extract the underlying dynamical systems governing the observed fire spread. This method's advantage lies in its ability to identify sparse representations of these systems, providing insight into the critical processes involved. The application of weak SINDy to fire spread data represents a promising step forward in the predictive modeling of wildfires. By uncovering the fundamental equations governing fire dynamics, this method opens new avenues for the development of accurate, efficient, and interpretable models. Future work will focus on validating these equations in various environmental conditions and incorporating them into broader wildfire simulation frameworks.

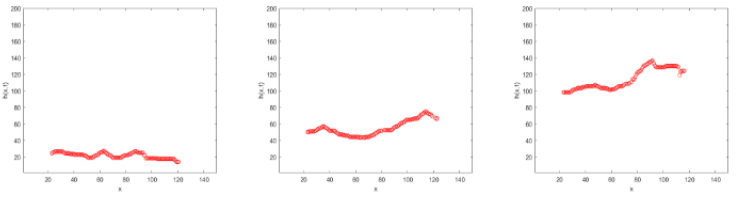


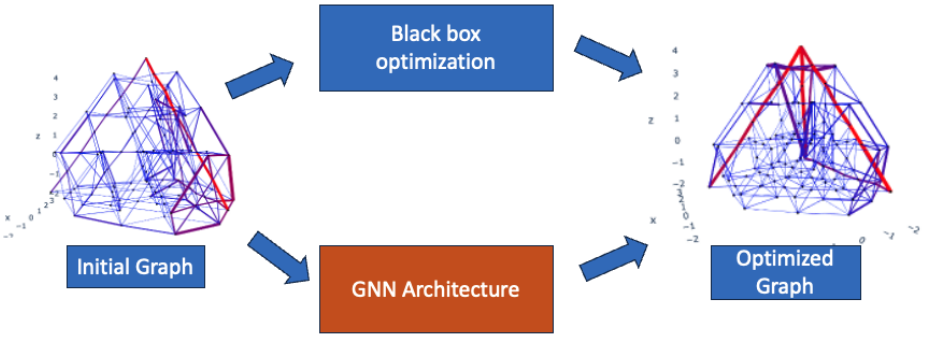
Figure 1: Head fire front propagating through fuel load.

Kunal Kanawade
Ph.D. in Computational Science
Advisor: *Nick Dexter*

Approximation of Structural Topology Optimization through Graph Neural Network Techniques

In the field of additive manufacturing engineering, the design and optimization of micro-structures subjected to various loading conditions has been increasingly represented using graph and constraint-based models. These representations facilitate the manipulation and exploration of structural topologies, enabling the optimization of micro-structures for desired performance characteristics. This research work investigates the application of Graph Neural Networks (GNNs) to learn the mapping from initial undirected graph representations, augmented with edge attributes, to their optimized topological solutions. By employing the GNNs, this work aims to capture the intrinsic relationships and constraints inherent in structural designs. By learning the mapping from data generated by a black-box solver, we aim to predict the optimized topologies that satisfy the design requirements and loading conditions for a given structure represented with the undirected graph model. This is ongoing research work, and more results are in progress.

Figure 1: Schematic Diagram for the GNN based approximation for black-box optimization.



Marzieh (Tara) Khodaei
Postdoctoral Associate in Computational Science
Advisor: *Peter Beerli*

Multilocus Phylogenetic Tree Estimation Using Topic Modeling

Inferring the evolutionary history of species or populations with genome-wide data is gaining ground, but computational constraints still limit our abilities in this area. We developed an alignment-free method to infer the genome-wide species tree and implemented it in the Python package TopicContml. The method uses probabilistic topic modeling (specifically, Latent Dirichlet Allocation or LDA) to extract 'topic' frequencies from k-mers, which are derived from multilocus DNA sequences. These extracted frequencies then serve as an input for the program Contml in the PHYLIP package, which is used to generate a species tree.

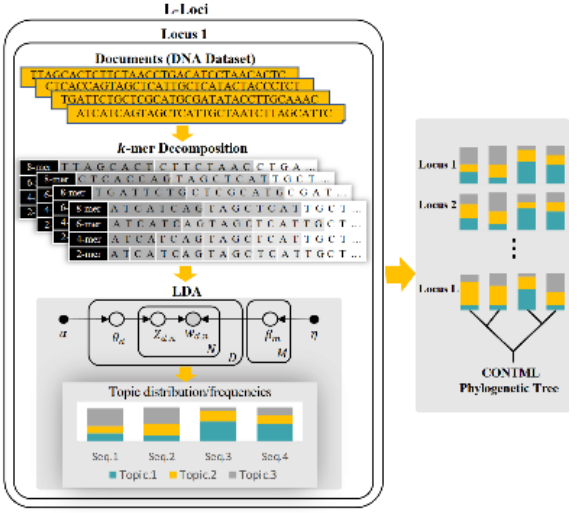


Figure 1. TopicContml workflow.

Kathryn Merritt & Hannah Squier
Undergraduate in Computational Science
Advisor: *Alan Lemmon*

New Genome Assembly Method

Current methods of genome assembly do not differentiate between the different regions of the genome (i.e., satellite regions, multi-copy regions, gene-rich regions, etc.). Instead, all regions are treated as equally helpful and informative. While this approach is sufficient for smaller genomes, assemblies for larger genomes face difficulties as it is difficult to distinguish the highly repetitive regions of the genome from one another. Our approach focuses on using single copy regions of the genome as anchor points, which allow for relative position information to be obtained and the other regions of the genome to be overlaid on the single copy regions towards the end of the assembly.

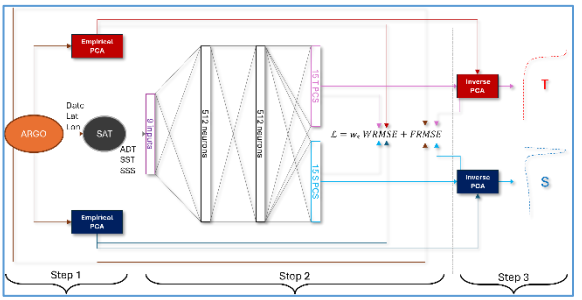
Jose Miranda
Ph.D. in Computational Science
Advisor: *Olmo Zavala-Romero*

Neural Synthetic Profiles from Remote Sensing and Observations (NeSPReSO) - Reconstructing temperature and salinity fields in the Gulf of Mexico

Accurate circulation modeling in the Gulf of Mexico (GoM) is hampered by the limited availability of in-situ subsurface data, leading to inaccuracies in subsurface representations. These inaccuracies diminish the reliability of ocean models and constrain the duration of dependable forecasts. This study introduces NeSPReSO (Neural System for Predictive Rendering of Subsurface Oceanography), a data-driven method to efficiently and accurately estimate subsurface temperature and salinity profiles using satellite-derived surface data. This provides an alternative to conventional synthetic data generation techniques.

Principal Component Analysis (PCA) is applied to extract the main features of temperature and salinity profiles of an Argo dataset. Then, a neural network is trained to predict these principal features using inputs such as time, location, and satellite-derived absolute dynamic topography alongside sea surface temperature and salinity. The model, evaluated using additional Argo profiles and glider data collected in the Gulf of Mexico, over-performs other traditional synthetic data generation methods, such as the Gravest Empirical Modes (GEM) and Improved Synthetic Ocean Profile (ISOP), in terms of root mean square error and bias. Our findings indicate that our method effectively captures the main variations of subsurface fields, and that synthetic profiles generated by the model align well with actual observations, accurately capturing key features such as thermoclines, haloclines, and temperature-salinity structure of the region. This new method will be implemented in GoM data assimilative models and is expected to improve the accuracy of modeled subsurface current.

Figure 1, Below: General diagram of NeSPReSO. Step 1 computes the empirical PCA of the Argo database. Step 2 trains a dense neural network from interpolated satellite data (Absolute Dynamic Topography (ADT), Sea Surface Temperature (SST), and Sea Surface Salinity (SSS)), location and date to predict the Principal Component Scores (PCS). Step 3 reconstruct the profiles using the predicted PCS and inverse PCA.



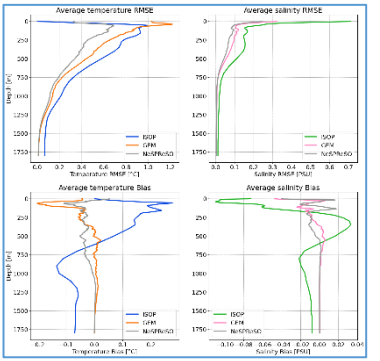
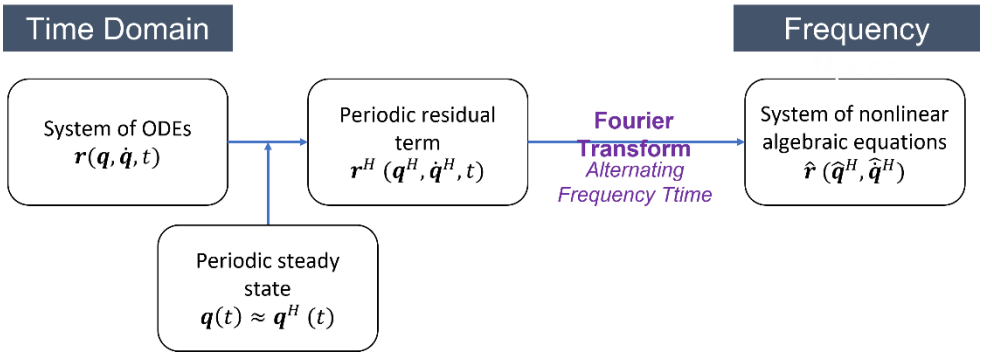


Figure 2, Above: Average RMSE for temperature and salinity predictions (top), and average bias (bottom) as a function of depth.

Shivangi Mittal
Postdoctoral Associate in Computational Science
Advisor: *Sachin Shanbhag*

The method of Harmonic Balance for Differential Constitutive equations in Oscillatory Shear flow

Rheological studies are aimed at understanding the flow characteristics of a soft material subjected to a deformation field in terms of experimentally measurable properties such as viscosity or modulus. To surpass the anomalies associated with conventional numerical integration approaches, we present the method of harmonic balance (HB) for solving differential constitutive equations subjected to LAOS. This technique is based on the spectral Galerkin method and uses Fourier basis functions to transform the initial value problem in time domain to an optimization problem the frequency space. Hence, by incorporating the principles of symmetry, the complicated set of nonlinear differential equations in terms of system parameters and oscillatory terms is reduced to a system of nonlinear algebraic equations in terms of Fourier coefficients. This approach reduces the computational expense by about 1-3 orders of magnitude and reduces error by several orders. It has been implemented on the Giesekus model, the Phan Thien Tanner model, and the Temporary network model. The formulation can be adapted for any differential CM with any form of nonlinearity using the Alternating Frequency Time (AFT) technique. The use of HB provides a computationally affordable alternative for model calibration and model selection for LAOS data It further opens new avenues for theoretical studies of nonlinear CMs under LAOS such as thermodynamic stability analysis and occurrence of mathematically unstable solutions.



Sanjeeb Poudel
Ph.D. in Computational Science
Advisor: *Xiaoqiang Wang*

Phase-field modeling of elastic bending energy with a deviation energy constraint

In vesicle membranes, the minimum of the elastic bending energy determines the equilibrium shape. We can reformulate the energy using a phase-field function and optimize it using the standard gradient flow approach. The method, in its typical formulation, allows topological changes in the membrane; however, in certain events, like in the simulation of blood cells, maintaining the initial topology of the membrane is crucial. In this study, we add a constraint on the phase-field method to preserve the topology. We note that even if the phase-field method is formulated using a hyperbolic tangent function, during a change in topology, the membrane's profile deviates from the tanh function. Owing to this idea, the constraint imposes an additional requirement on the profile to maintain the tanh shape, thus preserving the topology. We perform extensive experiments in two-dimensional and three-dimensional scenarios to demonstrate that our method preserves the topology during the simulation.

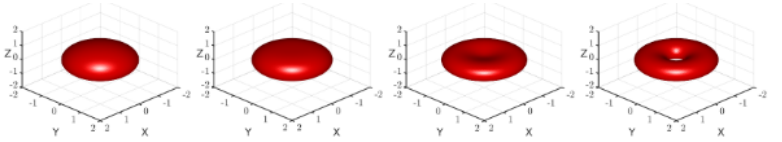


Fig: Dynamics of the membrane without the deviation energy constraint

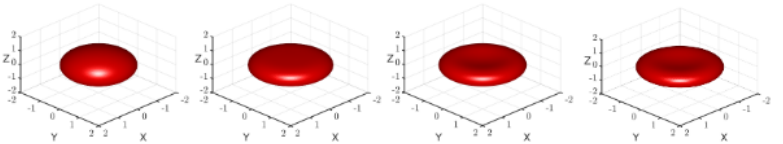
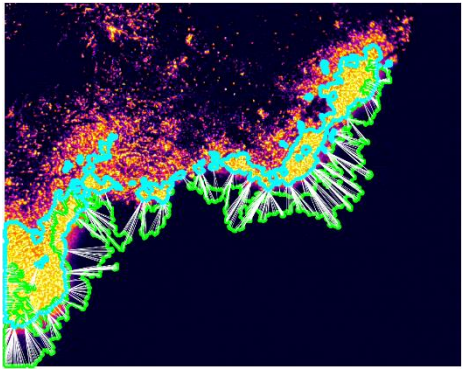


Fig: Dynamics of the membrane with the deviation energy constraint

Daryn Sagel
Ph.D. in Computational Science
Advisor: Bryan Quaife

Data Unraveling Fire and Plume Behaviors: A Multiscale Image Approach



Understanding the variety of structures and behaviors that occur during fire spread and plume evolution is crucial. We explore fire spread and plume evolution by analyzing data across diverse spatial and temporal scales in natural settings, capturing interactions among fire, plume, forest canopy, and atmosphere. Our approach integrates computer vision and graph theory to track these dynamics from visual and infrared data obtained with ground-based instruments or unmanned aerial vehicles (UAVs). Calculations that quantify heat transport, fire spread, turbulent statistics, and near-field plume structures contribute to our understanding of fire and plume behavior. A statistical analysis reveals patterns underlying fire and plume dynamics, energy transport, and fire-atmosphere coupling.

Figure 1, Left: Mid-wave infrared (MWIR) image of fire spread in a natural environment. An outline of the flaming region in the current timestep is denoted using blue points, with the following timestep denoted using green points. White lines show the connections, or rates of spread, between them. An analysis of the distribution formed by these velocities provides insights into the dynamics underlying fire behavior.

Farhana Taiyebah
Ph.D. in Computational Science
Advisor: Tomek Plewa

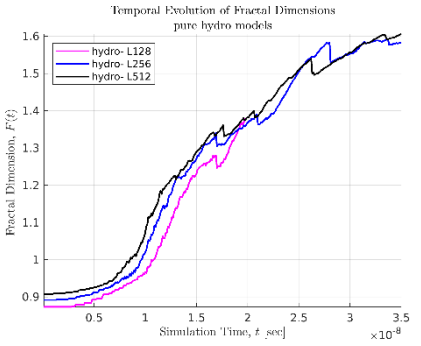
Fractal Characteristics of Rayleigh-Taylor Unstable Systems

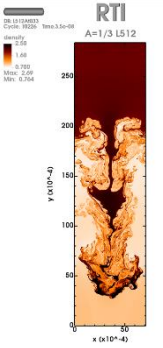
Understanding the intricate dynamics of stellar explosions is one of the fundamental problems in theoretical and computational astrophysics. In this work, we investigate the role of the Rayleigh Taylor Instability (RTI) in the acceleration of flames in thermonuclear supernovae, which is believed to be important in terms of comparing models with observations. Multi-dimensional simulations have shown that the burning rate associated with the laminar flame is insufficient to explain observations, leading to the consideration of turbulence, induced by RTI, as a necessary element of supernova combustion. It is believed that RTI, associated with the flame evolving in the gravitational field, plays a crucial role in this process.

In the context of supernovae and Rayleigh-Taylor instability, we investigate the presence of a unifying similarity across different physical systems exhibiting RTI evolution. Specifically, we compare three model setups: a classical hydrodynamic RTI case, a reactive flow with constant gravity and a flame model, and a fully integrated supernova simulation incorporating self-gravity and stratification. Our objective is to examine the continuity of behavior in terms of the fractal dimension characterizing the interface structure across varying scales, suggesting that the fractal dimension can serve as a scaling exponent for geometric complexity.

Our simulations across these three model setups reveal intriguing similarities and key differences in the evolving fractal dimension characterizing the interface morphologies. Initially, the fractal dimension increases across all systems as the interface becomes more convoluted, indicating a commonality in the development of complex structures driven by RTI. However, the attainable fractal dimension varies between the non-reactive and reactive flame systems, pointing to a potential analogy between the turbulent combustion's mixed regime and the disrupted interface structure in the hydrodynamic RTI case.

Future work should aim to expand our analysis into three dimensions, moving beyond the limitations of our two-dimensional simulations to better understand the inherently three-dimensional nature of turbulent flames. This expansion is anticipated to uncover differences in fractal dimension values, shedding light on the relationship between fractal dimension and fuel channel formation, potentially correlating higher dimensions with narrower channels.





Mani Tyagi
Ph.D. in Computational Science
Advisor: *Chen Huang*

Embedded Atom Approximation for Strongly Correlated Systems

Density Functional Theory (DFT) has been instrumental in understanding the electronic structure of materials, but accurately modelling the exchange-correlation (XC) energy in strongly correlated systems remains challenging. In this study, we propose a novel method to treat the on-site XC energies in strongly correlated systems. Our method involves partitioning a system's electron density onto individual atoms to enable a more accurate treatment of on-site correlations, such as the correlation between d and f electrons in transition metals and rare earth elements. This allows us to better capture the complex on-site correlations essential for understanding strongly correlated materials, such as transition metals oxides. Our method presents a potential solution to the limitations of existing methods like the local density approximation (LDA) + U method and the exact exchange for correlated electrons (EECE) method, which are two major methods for treating on-site correlations. One limitation of these two methods is that they rely on the choice of local projectors for defining local correlated orbitals, while our density partitioning scheme is first-principle and unique. Another more prominent limitation with LDA+U and EECE is that they cannot treat on-site correlation by using advanced orbital-based correlation energy functionals, since they only have access to occupied orbitals from the projection. Our method is able to employ any advanced correlation functional, such as the random phase approximation correlation energy functional, since each embedded atom is a complete Kohn-Sham system. With our method, it is possible to fully calculate the on-site correlation energy without any empirical parameter. In this poster, we demonstrate some preliminary results of this new method by performing non-self-consistent calculations on CO and N2 molecules. To treat the on-site XC energies, a range-separated hybrid functional is used as the high-level XC functional. The binding energy curves and bond lengths agree with the benchmarks very well. The development of a fully self-consistent version of this new method is in progress in our group.

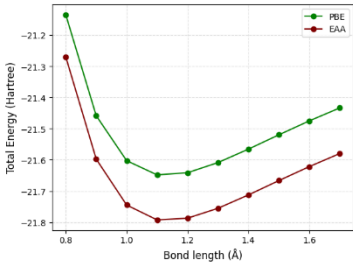


Figure 1, Above. Energy vs Bond length for CO.

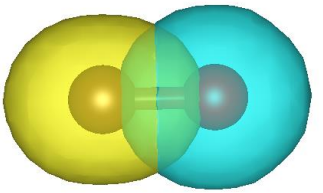


Figure 2, Above. Atomic election densities.

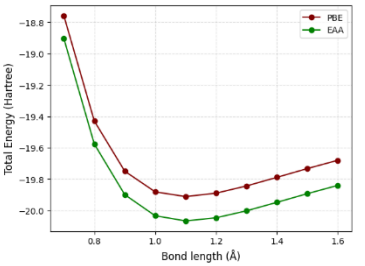


Figure 3. Energy vs Bond length for N_2.

Courtney Weintraub & Ivor Ho
Undergraduate in Computational Science
Advisor: *Alan Lemmon*

Modeling Neural Circuits to Understand Reproductive Isolation and Cryptic Evolution

As a certain species evolves over time, their neural circuits, or the structure of how neurons are arranged by synapses for a function, also evolve. In this project, we are exploring the characteristics of neural circuit evolution in chorus frogs by studying the relationship between evolution and neural circuits. We are measuring mating preferences of female frogs from the southeast region of the United States, and writing programs that compute likelihood scores of breeding compatibility and success between frogs from different or individual states, such as Alabama and Florida. Programming in R and MatLab allows us to visualize data that represent certain parameters and allows us to extract likelihood scores that we could use when evaluating chorus frogs' breeding behaviors. We plot these likelihood scores as a function of neural circuit parameter values in order to get a better picture of how behaviors map to neural circuit architecture. We have developed a method to evaluate the degree to which different peaks are connected by models that also fit the behavioral data. Neural circuits have redundant parameterization, and throughout the experiment, we have seen that several distinct parameter sets can produce neural circuits that respond in the same way to a suite of calls. The next part of the experiment will explore the way the behavior changes across parameter space to find if it is consistent from one peak to another. Recognizing the mapping between neural circuits and evolution allows us to predict and analyze evolution patterns in other species, and results from this project paves a way for future endeavors in the field of evolution and neuroscience.

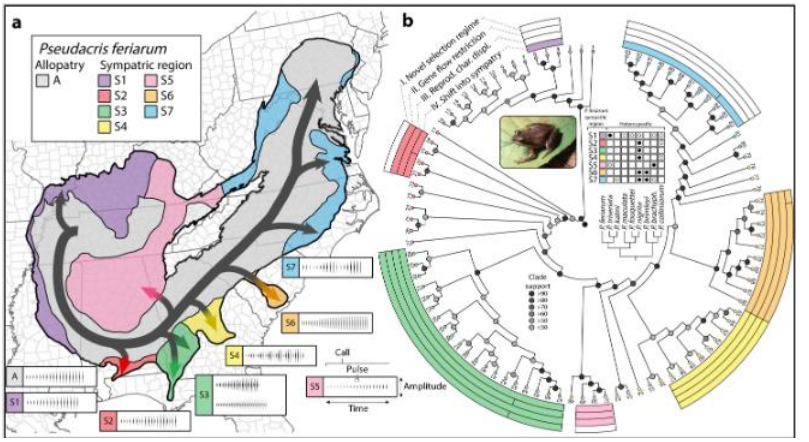


Figure 1. Diversification of mating signals in the upland chorus frog (P. feriarum). (a) The upland chorus frog has expanded from an ancestral region (gray) into the ranges of heterospecifc species (sympatry, colored ranges) multiple independent times. In many of these cases, the male mating call has diverged (see oscillograms representing the calls) in response to selection on females to avoid hybridization. (b) Phylogenetic relationships among P. feriarum sampled across the range, showing the independent expansion into sympatric regions. The inset shows phylogenetic relationships among the chorus frogs and the interactions of P. feriarum with those species.

Liam White
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

Level Set Grids for Hybrid Manufacturing

Hybrid manufacturing merges the design flexibility of additive manufacturing with the precision finishing of subtractive processes. Despite its innovative potential, the sequential nature of current hybrid manufacturing workflows presents significant efficiency challenges. This work introduces the Level Set Grid, a novel representation aimed at enhancing hybrid manufacturing by enabling the parallel execution of additive and subtractive processes.

The Level Set Grid innovatively combines explicit and implicit geometric representations to facilitate the precise modeling of objects and efficient collision detection. By leveraging a level set function, typically a signed distance function, the model offers a robust framework for representing complex geometries. This function's evaluations populate a sparse volumetric grid, blending the accuracy of implicit models with the practicality of explicit representations.

Preliminary findings underscore the model's capability to integrate seamlessly with existing manufacturing protocols, thereby facilitating the generation of additive slicing and subtractive toolpathing. This work outlines the foundational steps toward the model's application in hybrid manufacturing, highlighting its potential to revolutionize the field by improving process efficiency, flexibility, and innovation.

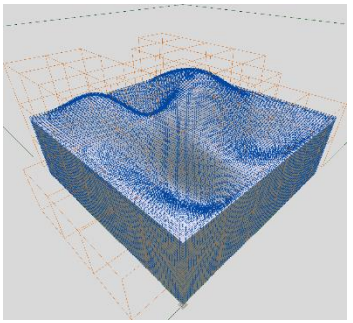


Figure above: A visual depiction of the Level Set Grid representation of an object characterized by intricate high-curvature features.

Kevin Ziegler
Ph.D. in Computational Science
Advisor: *Alan Lemmon*

Correcting the Draft Genome Assembly of Pseudacris *feriarum* with HiC Data

In this poster, we improve the draft de novo genome assembly of Pseudacris *feriarum* using HiC contact maps (High-throughput Chromosome Conformation Capture). The previous version of the assembly included large scaffolds, but not chromosome scale structures hundreds of megabases in size. Here we show the first step of a scaffolding algorithm, 3d-dna, which arranges contigs into pseudo chromosomes; these chromosomes are roughly the correct size and number for Pseudacris *feriarum* but have many mis-joins. After careful visualization of these mis-joins, we diagnose the root cause of mis-joins being multi-mapping HiC reads. Multi-mapping reads most commonly produce the distinct signature of heterozygosity, and less frequently cause mis-joins by mapping to "repeats" dispersed throughout the genome. We show steps used to filter out contigs containing the key signature of heterozygosity and remove multi-mapping repeat reads causing mis-joins. Once filtering is complete, future work will involve re-scaffolding using our cleaned dataset to produce an assembly containing less mis-joins.

Filtering out this data is challenging because our dataset is noisy and stochastic. HiC reads contain two fragments, each fragment mapping to different places in the genome. The distance between these fragments follows a negative binomial distribution. As a result, regions within the genome which are close in proximity will have many mapped HiC connections while regions far apart will have less. Along with stochastic read mapping, HiC data is known to have interaction biases for different genomic regions.

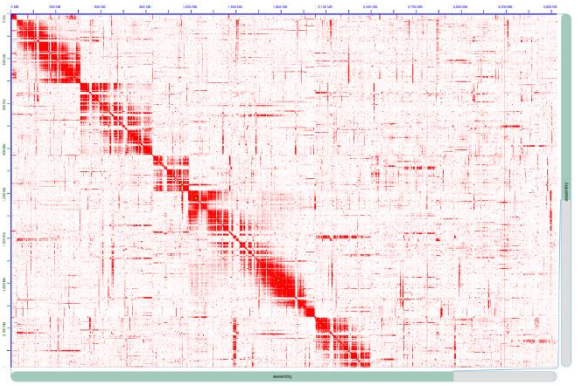


Figure Left: Pseudo Chromosome Genome Assembly of Pseudacris feriarum

The x and y axis give the location of the assembly. The red color indicates the interaction frequency of HiC reads. Notice the diagonal has the highest interaction frequency because most HiC reads map to locations which are very close together. Chromosome "squares" are visible in the heatmap because chromosomes interact with themselves more than other chromosomes.