Images are not simply sets of objects: each image represents a web of interconnected relationships. These relationships between entities carry semantic meaning and help a viewer differentiate between instances of an entity. For example, in an image of a soccer match, there may be multiple persons present, but each participates in different relationships: one is kicking the ball, and the other is guarding the goal. With this poster, we formulate the task of utilizing these referring relationships to disambiguate between entities of the same category. We introduce an iterative model that localizes the two entities in the referring relationship, conditioned on one another. We formulate the cyclic condition between the entities in a relationship by modelling predicates that connect the entities as shifts in attention from one entity to another. We demonstrate that our model can not only outperform existing approaches on three datasets (CLEVR, VRD and Visual Genome) but also that it produces visually meaningful predicate shifts, as an instance of interpretable neural networks. Finally, we show that by modelling predicates as attention shifts, we can even localize entities in the absence of their category, allowing our model to find completely unseen categories.

Ines Chami is a second-year Masters student in the Institute for Computational and Mathematical Engineering (ICME) at Stanford University. Prior to attending Stanford, she studied Mathematics and Computer Science at Ecole Centrale Paris. Her research interests include Computer Vision, Natural Language Processing and, more specifically, Multimodal Analysis. Ines is currently
ICME XPO 2018

working on cross-modal data-programming within the Hazy Research group under the supervision of Professor Christopher Re.

---

Numerical Artifacts in the Generalized Porous Medium Equation: Why Harmonic Averaging Itself is not to Blame

Presented by Danielle Maddix, 5th year ICME PhD

The degenerate parabolic Generalized Porous Medium Equation (GPME) poses numerical challenges due to self-sharpening and its sharp corner solutions. Spurious temporal oscillations, and nonphysical locking and lagging have been reported in the literature for averaged-based second order finite volume discretizations. These issues have been attributed to harmonic averaging of the coefficient $k(p)$ for small $p$, and arithmetic averaging has been suggested as an alternative. We show that harmonic averaging is not solely responsible and that an improved discretization can mitigate these issues.

In the first part of this work, we investigate the causes of the numerical artifacts using modified equation analysis for the continuous coefficient case. The provided modified equation framework can be used for any type of discretization. We show results for the second order finite volume method. We find that the observed problems with harmonic averaging can be traced to two leading error terms in its modified equation. This is also illustrated numerically through our newly developed Modified Harmonic Method (MHM) that can locally modify the critical terms to remove the aforementioned numerical artifacts. For the continuous coefficient case, we show results for two subclasses of the GPME with differentiable $k(p)$ with respect to $p$, namely the Porous Medium Equation (PME) and the superslow diffusion equation.

In addition to the degeneracy and self-sharpening of the GPME with continuous coefficients, increased numerical challenges occur in the discontinuous coefficient case. These numerical challenges manifest themselves in spurious temporal oscillations in second order finite volume discretizations with both arithmetic and harmonic averaging. The integral average, developed in van der Meer et al. 2016, leads to improved solutions with monotone and reduced amplitude
temporal oscillations. In the second part of this work, we propose a new method called the Shock-Based Averaging Method (SAM) that incorporates the shock position into the numerical scheme. The shock position is numerically calculated by discretizing the theoretical speed of the front from the GPME theory. The speed satisfies the jump condition for integral conservation laws. SAM results in a non-oscillatory temporal profile, producing physically valid numerical results. We use SAM to demonstrate that the choice of averaging alone is not the cause of the oscillations, and that the shock position must be a part of the numerical scheme to avoid the artifacts. For the discontinuous coefficient case, we show results for a subclass of the discontinuous GPME, known as the Stefan problem.

Danielle Maddix is a fifth year PhD candidate in the Institute of Computational and Mathematical Engineering (ICME). She has been funded by the NSF graduate research fellowship. Danielle received her bachelor’s degree in Applied Mathematics with highest honors from the University of California, Berkeley in 2012. She then continued onto her graduate education at ICME at Stanford and earned her Master’s in Computational and Mathematical Engineering in 2015. She has had internships in computational mathematics at the Lawrence Berkeley National Laboratory and NVIDIA, and will be joining Amazon Web Services in their Machine Learning research division after graduating with her PhD in June 2018. She has also been teaching the Advanced MATLAB for Scientific Computing course, which is open to the public on Stanford's Lagunita website. Her current research is on developing new, accurate, stable and conservative numerical methods for mathematical modeling of ocean dynamics.

---

An Algebraic version of the Hierarchical Interpolative Factorization

Presented by

Leopold Cambier, 3rd year ICME PhD

The problem of solving large sparse linear systems $Ax=b$ is still at the heart of scientific computing. Between exact factorizations and pure iterative methods lies the problem of preconditioning. We are in particular interested in generic preconditioner based on approximate factorizations, where one speed-up a direct method by introducing relevant
approximations during the elimination. This avoids the need for expert knowledge about the problem. However, doing so with strongly ill-conditioned linear systems remains challenging, as the approximation may not work well. In this case, most algorithm revert to a slow direct method. In this work, we explore the application of the Hierarchical Interpolative Factorization (developed by K. Ho and L. Ying on regular grids) to general graphs, and present a generic algebraic algorithm combining interpolative factorization (for the approximations) and nested dissection. Building upon nested dissection, a fast direct method, guarantees it will always perform at least as well. We show how to set-up the algorithm and present some preliminary scaling and performances results on medium scale problems.

Leopold Cambier received a B.S. in engineering and a M.S. in mathematical engineering from Universite Catholique de Louvain in Belgium. He is currently a Ph.D. candidate in ICME, under the supervision of Eric Darve. His interests are solving large (dense or sparse) linear systems and numerical linear algebra in general.

---

**Iso-geometric Collocation Method for the Fractional Laplacian**

Presented by

**Kailai Xu**, 2nd year ICME PhD

A numerical method for the fractional Poisson problem in a 2D bounded domain is proposed, based on iso-geometric analysis and the singularity subtraction. The basis functions of iso-geometric analysis can have high order continuous derivatives, and therefore we can construct a fourth order approximation for the numerator of the hyper-singular integrand. The accuracy of the method is numerically shown to be $O(N)$. Compared to finite difference methods, with iso-geometric analysis, we can deal with complex geometries. The singularity subtraction enables us to use efficient and accurate quadrature rules.

Kailai is a second-year Ph.D. in ICME. His research interests include numerical methods for PDEs, linear algebra, and optimization. He is currently working with Prof. Eric Darve on fractional PDEs. Kailai received a B.A. in mathematics from Peking University.
Low Rank plus Sparse Decomposition of Synthetic Aperture Radar Data for Target Imaging and Tracking

Presented by

Matan Leibovich, 3rd year ICME PhD

Synthetic Aperture Radar (SAR) is instrumental in imaging complex domains with limited array processing capabilities. A significant challenge is imaging of complex ground scenes that contain both stationary and moving targets. Assuming a traditional SAR acquisition scheme, our objective is to preprocess the data so as to separate the contributions of the moving targets from the ones due to the stationary background. The idea we pursue is to decompose the data into a low rank and a sparse part. This decomposition enables us to separate the reflections from moving targets corresponding to the sparse part from the stationary ones that constitute the low rank part of the data. To do so we adapt the robust principal component analysis (RPCA) method to the SAR data problem. We further introduce a lossless baseband transformation of the data, which simplifies the analysis and improves the performance of the RPCA algorithm. The main contribution is a theoretical analysis that provides an optimal choice for the RPCA algorithm parameters so as to achieve robust separation of data coming from moving and stationary targets. The analysis also allows us to provide a minimum detectable target velocity. In particular, we show that the rank of the sparse matrix is proportional to the square root of the target's speed. The robustness of the approach is illustrated with numerical simulations in the X-band surveillance regime.

Matan is a third year Ph.D. student in ICME, advised by Prof. George Papanicolaou. Matan's main area of research is imaging problems and related challenges in PDEs, optimization, analysis and stochastics. Prior to coming to Stanford, Matan received his B.Sc. in Electrical Engineering and Physics from the Technion, Israel's Institute of Technology, and M.Sc. in Physics from Tel-Aviv University, and has worked as a researcher on wave propagation for the Israeli Navy.
Extracting earthquake signals from continuous waveform data recorded by networks of seismic sensors is a critical and challenging task in seismology. The Fingerprint and Similarity Thresholding (FAST) earthquake detector (Yoon et al., 2015) enables waveform-similarity-based earthquake detection in long duration continuous seismic data. FAST leverages locality sensitive hashing (LSH), a data mining technique for identifying similar items in large data sets, to efficiently detect similar waveforms without templates. Blind detectors like FAST are capable of identifying events with previously unknown sources, but they can be susceptible to false detections due to local correlated noise sources.

To improve detection accuracy and reduce false detections, we have developed a new method to extend FAST over a seismic network. We introduce pairwise pseudo-association, a technique that leverages the pairwise structure of event detections to identify events observed at multiple stations in the network without modeling the expected move-out. Pairwise pseudo-association and supporting techniques, event-pair extraction and event resolution, create a network detection pipeline that combines the sparse similarity matrix outputs from single-station FAST, applied independently to each station in a network, to produce a list of candidate events recorded across the network. Although our method was designed with FAST in mind, it is general and can be applied to any detector that produces pairwise waveform similarity measures for each station in a seismic network. Our approach automatically handles the unknown move-out across the network, is robust to missing or low-quality data at one or more stations, and is flexible enough to be applied to a variety of detection tasks and network geometries. Using the 2014 Iquique foreshock sequence as a test case, we show that our method is sensitive, identifying nearly five times as many events as the local seismicity catalog (including 95% of the catalog events), and has a low false discovery rate of less than 1%.
Karianne Bergen will be receiving her Ph.D. in Computational and Mathematical Engineering this June. Her thesis research focused on developing new algorithms to automatically identify weak earthquake signals in large seismic data sets. As a student at Stanford, she developed and co-taught a new data science short course on the fundamentals of machine learning. Karianne earned her Bachelor of Science in Applied Mathematics at Brown University. She holds a Master of Science in Computational and Mathematical Engineering from Stanford. Prior to starting her graduate work at Stanford, Karianne worked in a data scientist role as a member of the technical staff at MIT-Lincoln Laboratory. Karianne will start a position as a Data Science Initiative Postdoctoral Fellow at Harvard University this coming fall.

---

Robust Registration

Presented by

Cindy Orozco, 3rd year ICME PhD

We developed a robust algorithm for recovery of rigid motion transformations for different snapshots of data of any dimension, in addition of some theoretical bounds for specific examples.

Cindy Orozco is a third year PhD student in ICME working with Professor Lexing Ying. Her interest include linear algebra and optimization, especially in applications related with physics. She is from Bogota, Colombia, where she did her undergrad in Universidad de Los Andes, and then moved to KAUST in Saudi Arabia, to do her masters. She likes to paint, and if you look closer in ICME walls, you could find some of her canvases.
We introduce SPMR, a new family of methods for iteratively solving saddle-point systems using a minimum or quasi-minimum residual approach. No symmetry assumptions are made. The basic mechanisms underlying the method are novel simultaneous bidiagonalization procedures, SIMBA and SIMBO. They yield a simplified saddle-point matrix on a projected Krylov-like subspace, and allow for a monotonic short-recurrence iterative schemes. We develop a few variants, demonstrate the advantages of our approach, and discuss connections to existing methods. Numerical experiments illustrate the merits of this new family of methods.

Ron Estrin is a fourth year Ph.D. student in ICME, advised by Michael Saunders and Yinyu Ye. His main research interests are in developing fast algorithms for numerical optimization and linear algebra. Prior to studying at Stanford, he received his B.S. in Mathematics and Computer Science from the University of British Columbia. Outside of academics, he enjoys practicing martial arts and playing tennis.

Exploration of Stanford’s Digital Archive of Mobile Performances (DAMP)

Presented by

Youkow Homma, 1st year ICME MS
Data Science Track
Smule, a mobile app company that hosts mobile musical performances across the world, has a collection over 7 billion performances since June 2008. This work aims to explore both the audio data and metadata to better understand the performers and performances. In particular, we study

- Whether audio features can be used to predict attributes of the performer such as gender and age
- The geographic distribution and trends of performers and songs
- Trends in how performers sing in/out of tune

Youkow is a first year Master's student in the ICME Data Science track. He has an interest in machine learning and its application to AI systems. At Stanford he has worked on projects in reinforcement learning and audio synthesis using deep learning in addition to this current project with the DAMP archive.

---

**When Learning Meets Inference: Knockoffs Construction via Generative Adversarial Networks**

Presented by

**Ruohan Zhan, 1st year ICME PhD**

Building an interpretable model for general high dimensional statistical problems receives growing attentions these days. One important corresponding problem is to make inference while controlling false discovery rate. The proposed knockoffs procedure by Prof. Emmanuel Candes, etc. is proved to be an effective method. However, the model-X knockoffs construction for variables subscribed to general distribution remains to be an open problem. Meanwhile, recent analysis interprets Generative Adversarial Network (GAN) as an optimization to minimize distance between distributions. This motivate us to use GAN system to construct knockoffs, where the key idea here is to minimize the distance between original tuple and exchanged tuple, so that we can satisfy exchangeability in knockoffs construction.

Ruohan is a first year ICME PhD Student. She is interested in combining deep learning with inference to do transfer learning, which is promising in the ubiquitous artificial intelligence.
Equipped with statistically rigorous inference analysis, advanced architectures in learning models could be understood, and meaningful features from data topology could be extracted, which would contribute to transferring the knowledge from one specific task to more general scenarios.

Stanford energy use and efficiency

Presented by

Sean Clement, 2nd year ICME MS   ||   Neel Rakholia, 2nd year ICME MS

Analyzing and visualizing Stanford’s resource usage throughout campus by building or building group. Looking for patterns in resource use correlation, i.e. is there a correlation between electricity use and water use? How strong is it? How many building clusters exist? Can we predict future resource use?

Sean Clement is a combat analyst for the US Army finishing my degree at Stanford before returning to active duty. Collaborating with Neel Rakholia

Neel Rakholia is a second-year master’s student in the data science track. In his final quarter, he has been working with Prof. Ram Rajagopal and Kristin Parineh to understand, optimize, and predict utility consumption of Stanford buildings.
Multifidelity Uncertainty Quantification for Cardiovascular Hemodynamics

Presented by

Casey Fleeter, 3rd year ICME PhD

Deterministic hemodynamic models are increasingly employed for diagnosis and treatment of cardiovascular disease. However, their widespread adoption is hindered by their inability to account for uncertainty stemming from sources such as boundary conditions, vessel wall properties, and patient-specific anatomies. This motivates the transition to a stochastic framework, but high computational costs of repeated simulations penalizes this method. We propose the integration of three cardiovascular model fidelities, a three-, one- and zero-dimensional model of blood flow, to use with multilevel and multifidelity estimators from Sandia’s open sources Dakota toolkit. This allows for reduced variance of our quantity of interest estimators while maintaining reasonable computational cost. The performance of these methods are compared for healthy and diseased anatomies on local and global hemodynamic quantities of interest.

Casey is a third year Ph.D. student at ICME working with Alison Marsden. She worked at Sandia National Laboratories last summer with the Center for Computing Research under Michael Eldred and Gianluca Geraci on a project related to multilevel/multifidelity uncertainty quantification. Casey’s research interests include uncertainty quantification and computational fluid dynamics.

ICME Capstone Program

Presented by

Kari Hanson

Adjunct Lecturer

ICME Capstone Program Director
Overview of the new Capstone Program including information for students, mentors and faculty as well as descriptions of current projects.

Kari’s background is in finance, entrepreneurship, and user-centered design. Having worked with domestic and international startups and non-profits as Chief Financial Officer, investor, advisor, and board member, she currently lectures at Stanford University and advises local and international startups on business strategy and innovation. While at Stanford, Kari developed an accelerator program for student entrepreneurs at the Graduate School of Business and designed and launched the ICME Capstone Program. Kari has a MBA in Finance and Entrepreneurship from New York University.

---

**Grain**

Presented by

**Indraneel Kasmalkar, 3rd year ICME PhD**

The sediment-water dynamics at the base of glaciers play a key role in controlling the speed of large-scale ice flow and thus global sea level rise itself. The basal resistance to ice flow is controlled by the yield strength of the underlying sediment, which is governed by the sediment pore pressure, which in turn depends on meltwater percolation and sediment deformation. We use a GPU-based code for a 3-D Discrete Element Model (DEM) to simulate grain-grain interactions within a fluid matrix. We study the sediment-water interactions through properties such as porosity to gain insight into the basal dynamics for glaciers, especially the formation of efficient meltwater drainage systems.

Indraneel Kasmalkar is a third year PhD student in the Institute of Computational and Mathematical Engineering. He works with Prof. Jenny Suckale to study the basal dynamics of fast-flowing ice in Antarctica. His research involves creating physics models for subglacial water systems, simulating them using numerical PDE solvers and collecting insights about the nature of the basal hydrology. This poster will showcase the ongoing work conducted by Jenny Suckale and me as part of an ICME SEED grant.
A Local Analysis of Block Coordinate Descent for Gaussian Phase Retrieval

Presented by

David Barmherzig 5TH year ICME PhD

While convergence of the Alternating Direction Method of Multipliers (ADMM) on convex problems is well studied, convergence on nonconvex problems is only partially understood. In this work, we consider the Gaussian phase retrieval problem, formulated as a linear constrained optimization problem with a biconvex objective. The particular structure allows for a novel application of the ADMM. It can be shown that the dual variable is zero at the global minimizer. This motivates the analysis of a block coordinate descent algorithm, which is equivalent to the ADMM with the dual variable fixed to be zero. We show that the block coordinate descent algorithm converges to the global minimizer at a linear rate, when starting from a deterministically achievable initialization point.

David Barmherzig is a Ph.D. Candidate in the Institute for Computational and Mathematical Engineering at Stanford University. His research interests include mathematical signal processing, optimization, and computational imaging. He previously studied mathematics and engineering at Harvard and the University of Toronto, and is accredited by Professional Engineers Ontario.

Reinforcement Learning: Autonomous Race Car

Presented by

Stephanie Sanchez, 2nd year ICME MS || Vishal Subbiah, 2nd year ICME MS
We have implemented various versions of the Q-Learning algorithm to train an agent to play the Atari video game Enduro. We have also created a feature and reward extraction to relay information from the game to our agent to help the learning process. All algorithms were compared by the position of the agent at the end state.

Stephanie is a second year master’s student at the Institute for Computational and Mathematical Engineering (ICME) in the general track. She received her B.S. in Applied Mathematics with a Specialization in Computation from University of California, Los Angeles (UCLA). Her interests include data science, computer vision, and image processing.

Vishal is a second year Masters student at the Institute for Computational and Mathematical Engineering (ICME). His interests include machine learning, artificial intelligence and cryptography. He previously received his Bachelors in Materials Science.

---

**An HLA-Disease Map in a Population Cohort**

Presented by

**Julia Olivieri**, 2nd year ICME PhD

The Human Leukocyte Antigen (HLA) region has the largest number of associations with common diseases of any part of the human genome. Because of the unusually high variability of the region and the limited sample sizes used for most studies, associations between HLA allelotypes and disease have not been fully categorized. We analyzed the relationships between 175 HLA allelotypes and 270 diseases using 337,208 patients from the UK Biobank dataset. We performed a GWAS study which uncovered 708 associations with adjusted p values > 0.05, spanning 64 diseases and 112 allelotypes. When multiple allelotypes were associated with one disease we performed a Bayesian Model Averaging (BMA) analysis to investigate the probabilities of each allelotype being causal. We found 60 allelotype-disease relationships that had greater than 0.80 posterior probability of being causal. 31 of these relationships had not previously been reported in the literature (that we found). These previously-unknown potentially causal relationships between HLA allelotypes and diseases can aid in the diagnosis, early detection, and treatment of these diseases. You can find an interactive description of our results here: https://biobankengine.stanford.edu/hla-assoc.
Julia is a 2nd year ICME PhD student. She graduated with degrees in math and biology from Oberlin College in 2016. Her academic interests include computational biology and graph theory. Last summer she worked at Lawrence Livermore National Lab on machine learning techniques for cancer research. Her studies have been supported by the SGF and NSF fellowships.

Reduced order modelling test bed

Presented by

Kexin Yu, 2nd year ICME MS || Remmelt Ammerlaan, 1st year ICME MS

Modelling large deterministic physical systems is a computationally expensive operation, due to the large amount of variables that have to be updated at each step. Reduced Order Modeling (ROM) lowers the computational complexity by reducing the dimension of the state space through different projection-based methods and can significantly increase performance with acceptable accuracy. The Goal of our project is to develop a testbed to rapidly assess new ROM concepts. One method we are implementing is hyper reduction. Most ROM methods only reduce the dimension of the state space but not any non-linear term in the differential equation. This can hamper improvement as each iteration would still do a full order computation of the nonlinear term. Similar projection methods can be applied to reduce the order of the non-linear computation as well.

Kexin Yu is a second year Master’s student in ICME’s Data Science track at Stanford. Her interest includes Machine Learning, Data Mining and Natural Language Processing. Prior to Stanford, she received a BS in Mathematics of Computation with highest distinction from University of California, Los Angeles and has worked as a software engineer at Verizon’s IPTV team.
Using Deep Learning for the solution of PDEs

Presented by

Ziyi Yang, 2nd year ME PhD

Deep learning is a state-of-the-art technique that is widely used in various fields like computer vision and artificial intelligence in games. However, it has been barely utilized to solve engineering problems. In engineering, there are problems where the given input is high-dimensional and complicated, whereas the output (the quantity of interest) is low-dimensional. The classical example is that we have to solve the full grid solution of partial differential equations (PDEs) for a complicated system, while only some “window” or certain property of the solution is of interest. Hence, we used a recurrent neural network (RNN) based model, with self-designed cells, to tackle two problems: 1) solve for the highest temperatures in a two-dimensional heat transfer problem and 2) find where the maximum speed occurs in Burgers’ equation. Our self-designed cell outperforms LSTM and GRU using less number of parameters.

Ziyi is a 2nd year PhD student in ME department and currently he is working with Prof. Eric Darve. His research interest includes model compression in deep learning model, and neural network solver for PDE. Ziyi received his B.S. in Physics from Nanjing University in 2016.
We can extract coherent seismic signals from random vibrations recorded by fiber optics, but what do they mean?

Presented by

Eileen Martin, 6th year ICME PhD

By recording controlled vibration sources on a seismic sensor array, we can calculate maps of subsurface seismic velocities, but this can be expensive and logistically difficult in populated areas (where near surface imaging could make the most impact for geotechnical surveys or permafrost thaw monitoring), so continuous subsurface monitoring is rare. We combine two methods to make frequent subsurface imaging cheaper: estimating wave equation Green’s functions by cross-correlating random vibration recordings in the area of interest, and measuring vibrations as meter-scale strain rate profiles along standard fiber optic cables.

Most seismic processing assumes scalar pressure or vector particle velocity data, but with fiber we measure one component of strain rate, a tensor quantity. In earthquake records we see that events are detected at the right time, but fiber data look qualitatively different from traditional particle velocity measurements. We explain how the switch from vector to tensor quantities leads to these changes. When we cross-correlate ambient noise recorded along collinear subsets of fiber, results are similar to those from velocity data. When the fiber segments are not collinear, we predict that extracted signals travel at a faster apparent velocity than the true medium velocity. We verify predictions with computational modeling and show examples from real fiber optic array data sets. From our understanding of this change from velocity to strain rates, we propose a correction to our estimated Green’s function arrival times, and show how geometry effects the strength of extracted signals.

Eileen Martin is a Ph.D. student in ICME working in the Stanford Exploration Project research group with Prof. Biondo Biondi. She is an affiliate in the Geophysics Department at Lawrence Berkeley National Laboratory. During her Ph.D. studies, she has been supported through the DOE Computational Science Graduate Fellowship and the Schlumberger Innovation Fellowship. She earned her M.S. in Geophysics at Stanford and her B.S. double-majoring in mathematics and computational physics at the University of Texas at Austin. After graduation, Eileen will move to
Virginia Tech, where she will be an assistant professor in the mathematics department and program in computational modeling and data analytics.

---

**Reward mechanisms for contests and loyalty programs**

Presented by

**Nolan Skochdopole, 5th year ICME PhD**

We consider designing contests for simple agents; agents have random qualities and costs for submissions to a potential contest with a prescribed reward scheme and must decide whether or not to enter. The contest designer wishes to choose reward schemes to optimize some objective of the participants, focusing on the sum of the qualities of participants. We consider two broad types of mechanisms: threshold - in which every participant above a certain level shares the reward - and rank-order - in which the rewards are distributed according to the ranked qualities of the participants. After characterizing the symmetric pure-strategy equilibria for each type of mechanism, we show under some conditions on the quality/cost distribution some performance guarantees. Furthermore, we may map this contest framework very naturally to a loyalty reward framework and extend the results. All work included is ongoing.

Nolan is a fifth year PhD student in ICME working with Professor Amin Saberi in MS&E.
Hydrodynamically interacting polymers: theory and computation

In collaboration with Tiras Y. Lin, ME PhD

Presented by

Amir Saadat,
2nd year Chemical Engineering Postdoc

We propose to develop a fast Brownian dynamics software package to study the transport of nanoparticles and flexible polymers used in targeted drug delivery for cancer therapy. The vasculature near a cancerous tumor is characterized by porous blood vessels, and we propose to model the transport of nanoparticles through these pores. To achieve higher circulation times, drug delivery nanoparticles are often PEGylated: many flexible polymer chains are attached to the rigid nanoparticle surface. The motion of the full PEGylated nanoparticle is governed by a stochastic differential equation, and since the polymer chains are in close proximity, the hydrodynamic interactions between them are significant. These interactions are represented as a large matrix that needs to be constantly reconstructed and decomposed. We believe that calculation of these interactions can be massively accelerated through a GPU implementation, and we intend to implement this within an existing Brownian dynamics software package BDpack. The development of this computational tool is a key step forward in being able to design a nanoparticle in a clinical setting for personalized treatment of a cancerous tumor and can serve as a platform for modeling other macromolecular solutions.

Amir Saadat is a postdoctoral scholar advised by Prof. Eric. S. Shaqfeh and Gianluca Iaccarino, and he is the primary developer of BDpack. He began his college education in 2003 at Amirkabir University of Technology (Tehran Polytechnic). He received his BS and MS degrees in polymer engineering at Amirkabir University in 2009. He started his Ph.D. in 2011 at the University of Tennessee-Knoxville. During his Ph.D. study, he received an M.S. in chemical engineering and a graduate minor in computational science.
Local-to-Global Methods for Topological Data Analysis

Presented by

Bradley Nelson, 4th year ICME PhD

We present an approach to topological data analysis that synthesizes global topological information from local information. This allows for higher-fidelity topological analysis on large datasets than can be achieved with standard algorithms and can lead to improved interpretability on spaces that have structure related to a fibration. We demonstrate this approach on data derived from several imaging modalities and show how it can help with topological modeling of spaces associated with this data.

Brad Nelson is a fourth year ICME PhD student interested in applied and computational topology. He has a BA in mathematics from Dartmouth, and is supported by a NDSEG Fellowship. He previously worked at Epic, and has done internships at Ayasdi and Lawrence Livermore National Lab while at Stanford. At ICME, he is involved in computational consulting and teaching short courses. Brad enjoys running and hiking in the Santa Cruz Mountains, cooking, and rock climbing.

Problem Dependent Reinforcement Learning Bounds Which Can Identify Bandit Structure in MDPs

Presented by

Andrea Zanette, 1st year ICME PhD

In order to make good decision under uncertainty an agent must learn from observations. To do so, two of the most common frameworks are Contextual Bandits and Markov Decision Processes (MDPs).
In this work, we study whether there exist algorithms for the more general framework (MDP) which automatically provide the best performance bounds for the specific problem at hand without user intervention and without modifying the algorithm. In particular, it is found that a very minor variant of a recently proposed reinforcement learning algorithm for MDPs already matches the best possible regret bound in the dominant term if deployed on a Contextual Bandit problem despite the agent being agnostic to such setting.

Andrea Zanette is a graduate student in the ICME department who transferred to the PhD program a year ago. He is broadly interested in Active Learning / Reinforcement Learning with emphasis on statistically-efficient methods. He holds a Bachelor in Mechanical Engineering and previously worked in variational formulations for the Finite Element Method.

---

**Predicting Short Insertion and Deletion Mutation Rates with Deep Learning**

Presented by

Nimit Sohoni, 2nd year ICME PhD  ||  Ananthakrishnan Ganesan 2nd year ICME PhD

Short insertions and deletions (indels) are a common type of mutation occurring in DNA. The precise factors underlying these mutations are not well understood, and there currently exists no good model to predict indel mutation rates. Creating such a model is important to identify regions of the genome where indels are under selection pressures, which would likely indicate biological significance. In this work, we trained a 1-D convolutional neural network on the Genome Aggregation Database, which consists of over 33 million unique indel mutations.
The primary input to the neural network is a window of base pairs in the reference genome, and the goal is to predict how many indels occur within that window. Using our tuned network trained and validated on one set of chromosomes and tested on another, we are able to get a statistically significant correlation coefficient of about 0.72, with a mean absolute error of approximately 15%, on this regression task. We also implemented a binary classification model with a similar architecture, which takes as input a window centered around a specific base pair and predicts whether or not any indels occur at the center location; this classifier achieves an accuracy of about 81% on a balanced dataset.

Future work on this project will involve investigating the model predictions on regions of the genome known to be biologically relevant in some way, and looking at regions in which the model predictions are significantly different from the true indel statistics in the region, which may reveal hitherto unknown regions of significance within the genome.

Ananthakrishnan is a second-year graduate student in ICME broadly interested in computational and data sciences, such as computational genomics. He holds a bachelor’s degree in Mechanical Engineering from the Indian Institute of Technology, Madras, and worked as a Scientist at Vantage Research between 2015 and 2016.

Nimit is a second-year Ph.D. student in ICME. He graduated from Cornell University with bachelor’s degrees in Mathematics and Computer Science. His academic interests include numerical linear algebra, machine learning, and the theory of optimization.

---

**Crop Yield Forecasting within a Developmental Context: A Satellite-Based Deep Learning Approach**

In collaboration with Adeesh Goel and Agneya Loya

Presented by

**Ifeoma Anyansi, 2nd year ICME MS**

Crop yield forecasting plays a vital role in the pursuit of sustainable development. Predicting crop yields, such as wheat or rice, can help municipal governments plan out food sources and distribution to its population, and can serve as an early indicator of famine. Successful attempts
at applying deep learning for this task have been a fairly recent development over the past year and hold great promise for the field.

In this project, we apply deep learning models to predict rice yields in India; this represents a different development context from previous applications that focus on the United States. By using nine bands of multi-spectral satellite images and land surface temperature collected every eight days, we trained two different neural network architectures of Convolutional Neural Networks (CNNs) and Long-short Term Memory Networks (LSTMs). We find that both CNNs and LSTMs are able to achieve high prediction accuracies with comparable performances. India presents a unique opportunity to evaluate the performance of deep learning methods in geographies where frequent and precise data concerning crop yield, agriculture, land, soil and temperature may not be readily available. We find that rice yield data can be forecasted with relatively high accuracy which validates the use of these techniques in future developmental contexts.

Ifeoma Anyansi is currently a graduate student within the Computational and Mathematical Engineering department at Stanford University. Her research is at the intersection of data science, policy, and climate change. She is drawn to the capacity of machine learning to spur community resilience in the face of storm surges, tsunamis, and sea level rise. Moreover, she aims for the spread of data-driven frameworks to address the invisibility of underserved groups in the planning space and support local adaptation strategies.

Recent team projects include validating the use of deep learning models for crop yield forecasting in India: this works represents an efficient agriculture model under developmental contexts. She also recently began working in a team to build a deep learning object detection system for the identification of poachers and animals in natural reserves throughout China. Overall, she aims to uses data-driven approaches to create positive human impact and support sustainable development.
Topological Data Analysis and Deep Learning

Presented by

Anjan Dwaraknath, 4th year ICME PhD

Deep learning has been shown to be very successful in many machine learning tasks. However, these algorithms are still opaque and vulnerable to adversarial attacks. Techniques from topological data analysis can prove useful in mitigating some of these problems. In this project, we explore the prospect of using features generated from persistence homology to train neural network models that are more robust to adversarial attacks.

Anjan Dwaraknath is a fourth year PhD student in ICME, working with Prof. Gunnar Carlsson. His current research interests are in applying the techniques of topological data analysis to deep learning.

Predicting Opioid Abuse from Medical Claims Data

In collaboration with Dr. Jonathan Mugan, DeUmbra

Presented by

Andrew Deveau, 1st year ICME MS Data Science Track

The incidence of opioid abuse and overdose has dramatically increased in the United States in recent years. In this work, we explore methods for predicting which patients are at the greatest risk of opioid abuse, with the ultimate goal of helping guide interventions by doctors and public health officials.
Andrew Deveau is a first year masters' student in ICME in the data science track. Before coming to Stanford, he received a B.S. in mathematics from Yale and spent two years working for a proprietary trading firm. He is interested in robust and safe reinforcement learning as well as socially beneficial applications of machine learning.

Developmental stability of affective brain circuits and relation to emergence of anxiety and depression symptoms in children

In collaboration with Aarthi Padmanabhan and Vinod Menon

Presented by Yuan Zhang, 2nd year Postdoctoral fellow

The overarching goal of our project is to elucidate signatures of affective and cognitive brain network development and risk for psychiatric disorders using a novel “Big Data” science approach combined with state-of-the-art machine learning, quantitative brain network analyses and high-performance GPU computing. Here, we leverage the Philadelphia Neurodevelopmental Cohort (PNC), which has generated unprecedented amounts of data for charting brain development in individuals from childhood to adulthood.

We first characterized the development of brain networks underlying emotion processing using a well-characterized cohort of N > 800 children and adults (ages 8-22). We found that brain regions important for emotion were composed of three modules, the frontoparietal (FPN), default mode network (DMN), and subcortical/posterior insula network (SPIN) across four emotion categories – fear, anger, sad, and happy – and ages. While the overall modular organization of emotion circuits was developmentally stable by age 8, intra- and inter-module interactions increased significantly with age. Specifically, intra-FPN connectivity increased with age across fear, anger and happy categories, and fear showed the most extensive changes in both intra- and inter-network connectivity with age.

Next, we investigated the neural predictors of anxiety and depression and found that functional brain connectivity underlying fear processing was predictive of individual differences in anxiety,
but not depression. In contrast, brain connectivity underlying processing of sad faces was predictive of individual differences in depression, but not anxiety. The proposed project is the first to take a principled brain network and “Big Data” approach to characterize neurodevelopment of large-scale brain networks and identify predictors of risk factors for psychopathology. Our studies lay the groundwork for investigation of atypical brain organization in a range of psychiatric disorders that have a neurodevelopmental origin, including, most notably, ASD, ADHD, anxiety disorders, and childhood mood disorders and psychosis.

Dr. Yuan Zhang is a postdoctoral fellow at the Stanford Cognitive & Systems Neuroscience Lab, with expertise in brain networks, their development and relation to clinical symptoms. She trained in cognitive neuroscience and computer science at Peking University. Her current research interests involve using multimodal neuroimaging and advanced computational methods to characterize brain network organization underlying affective and cognitive processing in children with and without neurodevelopmental disorders.

---

**Mapping the Environmental Effects of the Rohingya Refugee Crisis**

In collaboration with Michael Burnett

Presented by

Chris LeBoa  ||  Freya Chay
In October of 2017 the Myanmar military began an ethnic cleansing of the persecuted Rohingya population in the north of the country. Villages were burned and an estimated 6,300 people were killed at the hands of the Myanmar army. To escape the violence, 1.1 million Rohingya fled across the Northern border of the country into Bangladesh, which is already one of the most population dense countries in the world. In an effort to house hundreds of thousands of refugees, the government of Bangladesh opened up a massive refugee camp known as Kutapalong in a previously forested section of the country.

To stay warm, cook food, and earn income, many of the 800K+ refugees at the camp harvest lumber from the surrounding forests, causing an immense amount of deforestation. However on-the-ground aid organizations have not had the time or resources to really assess the magnitude of the deforestation caused by the camp.

In order to better quantify the environmental degradation that is ongoing at Kutapalong and the potential impacts it could have on human health, our project leverages Planet’s high-resolution satellite imagery to A) objectively map the extent and rate of deforestation in the camps, and B) evaluate the impacts of this land cover change on water quality in the Naf River, which may be affected by increased sediment runoff in the absence of vegetation on the hilly terrain. Deforestation, which is clearly visible in the Planet imagery, can have drastic effects on local ecosystem services including reduced natural hazard protection, loss of food and extractive resources, and reduced climate regulation. Sedimentation in the adjacent Naf River, also detectable by multispectral satellite data, can also be correlated with water-borne ailments and waterway pollution.

Chris LeBoa is a student at Stanford in the Disease Ecology Track. His work focuses on the intersection of environmental degradation and the spread of pathogens in a variety of ecosystems. He is currently working to improve the environmental detection of schistosomes (a type of blood fluke) in the Senegal River basin through environmental DNA and satellite imagery. He is also a member of the SNAPP partnership, working on disease modeling of environmentally transmitted diseases. He has previously a researcher for ILM in Tahiti on a project to understand the different species that cause ciguatera seafood poisoning and with the Luby Lab/ International Centre for Climate Change and Development in Dhaka, Bangladesh to assess the health impacts of poor ventilation in Dhaka slums.

Frey is a fourth year pursuing a B.S. in Computer Science and an M.S. in Earth Systems.
The power of platforms in small-holder production chain

Presented by

Sergio Camelo, 3rd year ICME PhD

In this work we study the supply-chain of palm-fruit in Indonesia and propose optimization algorithms to increase its efficiency. Currently, farmers maintain relationships with middlemen, who pay farmers for the fruit and transport it to the production mill. We explore the effect on efficiency that would come from changing the relationships between farmers and middlemen, or from changing the pick-up schedules of the fruit. We rely heavily on algorithms to approximately solve the Traveling Salesman Problem.

Sergio Camelo is a third year PhD Student at Stanford ICME, where he works in creating technological tools for the sustainable sourcing of food in emerging markets, leveraging on optimization and machine learning.

Win by Landslides: Assess the California Landslide Susceptibility and Auto-detect Landslide

Presented by

Lijing Wang, 1st year Geological Sciences PhD

Landslides can be triggered by many factors, including earthquake or volcano activities, erosion, rain fall and vegetation change. In many cases, they are inevitable. Instead of thinking about preventing a landslide to happen, we want to produce a hazard susceptibility map, which combines all the current available geographical information and describes the relative likelihood of future landslide in California. In the meantime, identify landslide zone could be crucial for
hazard assessment as well as disaster relief. Such information, on the other hand, could be iterated into the model and update the susceptibility map, which improves its reliability and also makes one step closer to achieve landslide monitoring. We therefore want to take the chance to explore the Planet imagery and try to tackle this issue with machine learning. (Big Earth Hackathon Project)

Lijing is a first year Ph.D. student in Geological Sciences. Her interests include Bayesian Inference, Geometric and Topological Data Analysis, Geostatistics and Deep Neural Networks. Her research goal is to develop statistical methods to quantify uncertainty and make decisions in energy resources and environments. She is currently working on groundwater management and landslides hazards.

---

**Cloud Removal from Hyperspectral Satellite Images using Generative Adversarial Networks**

Presented by **Laura Domine, 1st year Physics PhD**

Satellite imagery has numerous applications such as environmental monitoring or natural disaster mitigation among many others. Clouds can lead to uneven illumination, blurring and occlusion of the target. Removing clouds is thus a critical task for remote sensed images. In addition, thanks to advances in sensor design and processing speed, hyperspectral satellite images are becoming more widely available. Instead of only three or four spectral bands (RGBA, multispectral) each image includes many wavelengths across the electromagnetic spectrum. Clouds appearance depends on the wavelength, and for example become partially transparent in near-infrared. We propose to apply Generative Adversarial Networks (GANs) to hyperspectral satellite images in order to generate the missing patches under the clouds.

Laura Domine is a first-year Ph.D. student in Physics, working with Kazuhiro Terao at SLAC on the development of deep learning techniques for experimental particle physics data analysis. Before coming to Stanford, she graduated from the Ecole polytechnique in France where she studied Mathematics, Physics and Computer Science.
Manmade and natural disasters increasingly affect the livelihood of many people on our planet. The first step in responding to any disaster, such as fire, flood, or war, is to assess the extent of damage done to human livelihood so that disaster relief resources can be deployed as efficiently as possible. With the recent advent of daily, high quality satellite imagery, there is now the possibility of monitoring changes to the built environment almost anywhere in the world. However, designing an accurate tool to automatically recognize changes to buildings and homes is non-trivial. In this project, we apply a deep convolutional neural network using the U-Net architecture to build an image segmentation algorithm capable of classifying buildings and homes with high accuracy. We then apply the segmentation algorithm to images before and after a disaster, and assess damage by quantifying changes in classification.

Noah Athens is a third year PhD student in Geological Sciences. He is interested in developing new methodologies for integrating heterogeneous datasets into prediction models for making decisions under uncertainty. He is currently working on predicting favorable locations for geothermal energy development.
Amir Delgoshaie is a PhD student in Energy Resources Engineering. He uses stochastic tools for modeling flow and transport in porous media. His current focus is on the relationship between random walks and deterministic PDEs governing fluid flow and utilizing this relation for uncertainty quantification.

Thomas Hossler is a PhD student in Geological Sciences. He is currently on a leave of absence and works in a food sustainability start-up based in San Francisco. His interests lie in Artificial Intelligence and Computer Vision applications to environmental problems.

---

**A fast direct solver for fractional elliptic problems on general meshes in 2D and 3D**

Presented by

**Nurbek Tazhimbetov, 2nd year ICME PhD**

One alternative approach to represent fractional Laplacian relies on the Dunford integral formula for spectral representation. Since a complex valued function $f$ is analytic over the complex domain with a branch cut on the negative real axis, we can solve the discretized fractional Laplacian by computing the matrix function $f$ using the contour integral representation and conformal mapping, where the closed contour lies in the region of analyticity of $f$ and encloses spectrum of Laplacian.

Nurbek was born and raised in Aktobe, a remote city located in the Western part of Kazakhstan after the collapse of USSR. In spite the difficulties of economic transition, his parents highly encouraged him to pursue education to succeed in life. Therefore, at the age of 14, Nurbek was sent to a boarding school to study mathematics and physics in Almaty, which is 1,000 miles away from my hometown. After graduation with the highest distinction, he attended Nazarbayev University in Astana, where he majored in mathematics. During Nurbek’s undergraduate years, he went as an exchange student to University of Wisconsin as a winner of university-wide academic competition. Furthermore, he conducted two research projects, both at Caltech, one in numerical PDEs and the other in computational astrophysics. Currently, Nurbek is a second year PhD student at the Institute for Computational and Mathematical Engineering at Stanford University.
Improved Processing, Slope Estimation and Ice Flow Interpretation Using Englacial Layers from Radar Sounding

Presented by

Elisa Mantelli, Postdoctoral Scholar

Basal friction is a major control on ice flow. However, bed slipperiness remains difficult to constrain from observations, which leads to significant uncertainties in ice sheet modelling. This poster is concerned with a specific aspect of basal friction, namely the onset of basal sliding. Theoretical work suggests that sliding onset occurs over distances asymptotically comparable with the ice sheet length, as opposed to an abrupt sliding onset that would take place over distances comparable with the ice thickness. Our aim is to find out whether these results are supported by radar observations of englacial layers. Here we illustrate how information from a new processing technique for airborne radar data can be used along with suitable forward models to address this question, and to improve estimates of bed slipperiness from inverse modelling.

Sparse and Fast Hierarchical Preconditioner with Improved Accuracy on Critical Subspaces

Presented by

Bazyli Klockiewicz, 3rd year ICME PhD

Solving discretized PDEs is challenging. The sizes of the associated linear systems make direct methods too computationally expensive. Iterative methods, on the other hand, need effective preconditioners to be useful. We present a hierarchical preconditioner for solving discretized PDEs. The preconditioner relies on compression, which ensures that long-range interactions between variables can be neglected up to a good accuracy. Our contribution
is a new compression scheme which improves accuracy on low frequency (smooth) eigenvectors, resulting in a better preconditioner when the size of the system becomes very big.

Bazyli is a third year PhD student working with prof. Eric Darve on hierarchical solvers for sparse linear systems. He is interested in computational mathematics, numerical methods and high performance computing in general. He also enjoys applied topology. Bazyli holds an M.Sc. in Mathematics from Adam Mickiewicz University in Poznan, Poland, and an M.Sc. in Computational and Mathematical Engineering from Stanford. He loves classical music and swimming.

---

**Topology optimization of thermoelastically damped MEMS resonators**

Presented by

**Dustin Gerrard**, Mechanical Engineering PhD

Micro-ElectroMechanical Systems (MEMS) resonators are widely used in inertial sensors, timing reference, and micro-actuators. In this work we apply a topology optimization algorithm to enhance the quality factor (Q) of resonators limited by thermo-elastic dissipation (TED). A custom finite element analysis (FEA) is formulated to simulate the resonant frequency and Q-TED. The algorithm treats each element of the resonator as a variable density between 1 (solid) and 0 (void). By inexpensively calculating the sensitivity of Q with respect to each element these densities are updated iteratively using a gradient based approach until a final topology converges. Final devices have a Q 10x higher than non-optimized devices.

Dustin earned a double BS degree from Brigham Young University in mechanical engineering and math. He was awarded the Department of Defense NDSEG fellowship to attend Stanford where he earned MS degrees in mechanical engineering and ICME and his PhD in mechanical engineering. Dustin’s work applies machine learning and optimization techniques to simulations of the physical world. He wrote a finite element model of thermo-elastic dissipation in MEMS resonators and created a topology optimization scheme to enhance their performance. Dustin is married and has two kids. His interests are running, hiking, and playing yo-yo competitively.