BRIN MATHEMATICS RESEARCH CENTER

Scientific Machine Learning: Theory and Algorithms

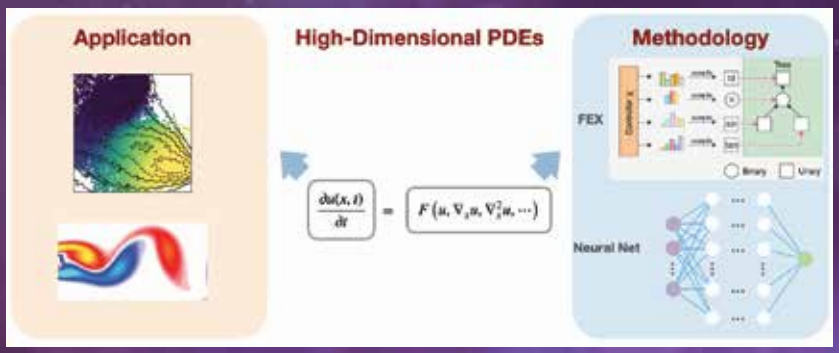
February 21-23, 2024

About the Workshop

Scientific machine learning combines computational science and machine learning to create a unified set of high-performance algorithms and implementations for solving complex tasks across science and engineering. Designing scientific machine learning with a provable capacity of going well beyond the available data is an active reserach field and an emerging educational task.

Organizers

Maria Cameron, University of Maryland
Chunmei Wang, University of Florida
Haizhao Yang, University of Maryland



Speakers

Ke Chen, University of Maryland
Ramani Duraiswami, University of Maryland
Jiequn Han, Flatiron Institute
Boris Hanin, Princeton University
Tom Hickling, University of Oxford
Yuehaw Khoo, University of Chicago
Samuel Lanthaler, Caltech
Holden Lee, Johns Hopkins University
Ling Liang, University of Maryland
Jianfeng Lu, Duke University
Mauro Maggioni, Johns Hopkins University
Reza Malek-Madani, Office of Naval Research
James Murphy, Tufts University
Deep Ray, Universiy of Maryland
Zezheng Song, University of Maryland
Shashank Sule, University of Maryland
Alex Townsend, Cornell University
Rebecca Willett, University of Chicago
Jinchao Xu, Penn State University
Yunan Yang, Cornell University
Margot Yuan, University of Maryland

CSIS Building 4th Floor
8169 Paint Branch Drive
University of Maryland
College Park, MD 20742

BRINMRC.UMD.EDU

DEPARTMENT OF MATHEMATICS

Contents

Workshop Overview **5**

Workshop Schedule **6**

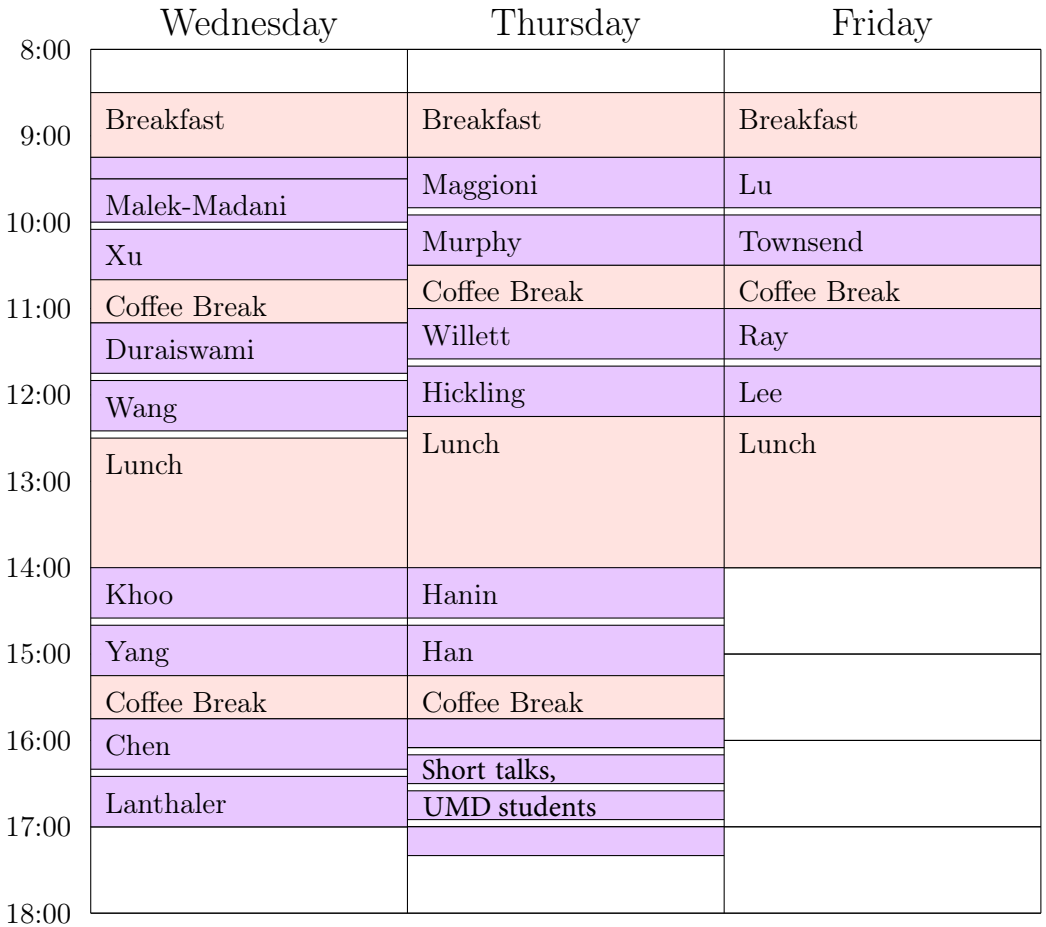
Abstracts of talks **11**

Reza Malek-Madani . . . 11
Jinchao Xu . . . 11
Ramani Duraiswami . . . 12
Chunmei Wang . . . 13
Yuehaw Khoo . . . 13
Yunan Yang . . . 14
Ke Chen . . . 14
Samuel Lanthaler . . . 15
Mauro Maggioni . . . 15
James Murphy . . . 16
Rebecca Willett . . . 16
Tom Hickling . . . 17
Boris Hanin . . . 17
Jiequn Han . . . 18
Ling Liang . . . 18
Zezheng Song . . . 19
Shashank Sule . . . 20
Margot Yuan . . . 21
Jianfeng Lu . . . 21
Alex Townsend . . . 22
Deep Ray . . . 22
Holden Lee . . . 23

The Brin Mathematics Research Center **24**

List of Participants **25**

Schedule at a Glance



Workshop Overview

Scientific machine learning combines computational science and machine learning to create a unified set of high-performance algorithms and implementations for solving complex tasks across science and engineering. Empirical successes have been made in various application domains with notable breakthroughs over traditional computational tools. In these applications, dynamics are complex and multiscale; function domains have high dimensions and complex geometry; data are heterogeneous, noisy, and expensive to acquire; models are nonlinear and decisions have high uncertainty. Designing scientific machine learning with a provable capacity of going well beyond the available data is an active research field and an emerging educational task. This workshop responds to the needs above with a schedule for invited talks, panel discussions, and poster presentations.

Organizing committee

Maria Cameron, University of Maryland

Chunmei Wang, University of Florida

Haizhao Yang, University of Maryland

Workshop Schedule

Wednesday, February 21, 2024

8:30 - 9:15	Breakfast
9:15 - 9:30	Doron Levy, Maria Cameron & Haizhao Yang *Opening*
9:30 - 10:00	Reza Malek-Madani (Office Naval Research) *Research Opportunities at ONR*
10:05 - 10:40	Jinchao Xu (Pennsylvania State University) *Deep Neural Networks and Finite Elements*
10:40 - 11:10	Coffee Break
11:10 - 11:45	Ramani Duraiswami (University of Maryland) *Making Scientific Computing Models Differentiable for Deep Learning*
11:50 - 12:25	Chunmei Wang (University of Florida) *Pseudo-differential Integral Autoencoder Network for Inverse PDE Operators*
12:30 - 2:00	Lunch
2:00 - 2:35	Yuehaw Khoo (University of Chicago) *High-dimensional PDEs, tensor-network, and convex optimization*
2:40 - 3:15	Yunan Yang (Cornell University) *Neural Inverse Operators for Solving PDE Inverse Problems*
3:15 - 3:45	Coffee Break
3:45 - 4:20	Ke Chen (University of Maryland) *Towards efficient deep operator learning for forward and inverse PDEs*

4:25 - 5:00 Samuel Lanthaler (California Institute of Technology)
Data-Complexity Bounds for Operator Learning

Thursday, February 22, 2024

8:30 - 9:15 Breakfast

9:15 - 9:50 Mauro Maggioni (John Hopkins University)
Learning Interaction laws in particle- and agent-based systems

9:55 - 10:30 James Murphy (Tufts University)
Intrinsic Models in Wasserstein Space with Applications to Molecular Dynamics

10:30 - 11:00 Coffee Break

11:00 - 11:35 Rebecca Willett (University of Chicago)
Deep Stochastic Mechanics

11:40 - 12:15 Tom Hickling (University of Oxford)
Adjoint Optimization of Deep-Learning Sub-Grid Scale Models for Large Eddy Simulation of Compressible Flows

12:15 - 2:00 Lunch

2:00 - 2:35 Boris Hanin (Princeton University)
Principled Hyperparameter Transfer Across Depth and Width in Neural Networks

2:40 - 3:15 Jiequn Han (Flatiron Institute)
A Neural Network Warm-Start Approach for Inverse Scattering Problems

3:15 - 3:45 Coffee Break

3:45 - 4:05 Ling Liang (University of Maryland)
On the Stochastic (Variance-Reduced) Proximal Gradient Method for Regularized Expected Reward Optimization

4:10 - 4:30 Zezheng Song (University of Maryland)
A Finite Expression Method for Solving High-Dimensional Committor Problems

4:35 - 4:55 Shashank Sule (University of Maryland)
Sharp error estimates for target measure diffusion maps with application to the committor problem

5:00 - 5:20 Margot Yuan (University of Maryland)

Optimal control for sampling the transition path process and estimating rates

7:00 - 9:00 Conference Dinner

Friday, February 23, 2024

8:30 - 9:15 Breakfast

9:15 - 9:50 Jianfeng Lu (Duke University)
Representation of symmetric and anti-symmetric functions

9:55 - 10:30 Alex Townsend (Cornell University)
Elliptic PDE learning is provably data-efficient

10:30 - 11:00 Coffee Break

11:00 - 11:35 Deep Ray (University of Maryland)
Learning WENO for entropy stable schemes to solve conservation laws

11:40 - 12:15 Holden Lee (Johns Hopkins University)
Theoretical foundations for diffusion models

12:15 - 2:00 Lunch

Abstracts of talks

Research Opportunities at ONR

Reza Malek-Madani

Office Naval Research

Wednesday, February 21, 2024 @ 9:30 AM

Deep Neural Networks and Finite Elements

Jinchao Xu

Pennsylvania State University

Wednesday, February 21, 2024 @ 10:05 AM

In this talk, I will report a new joint work with Juncai He on the connection between finite element and deep neural network (DNN) functions. In our earlier works, we reported that any linear finite element function in any dimension can be expressed in terms of a DNN using ReLU activation function. How to generalize this result to finite element function of arbitrary order has been an open problem. In this talk, we will report a solution to this open problem. Namely, we will show that any finite element function of any order on very general grids in any dimension can be expressed in terms of a type of DNN using some appropriately chosen activation functions. Furthermore, we will show that our new DNN can generate finite element functions of arbitrary order and can also generate any global polynomials of arbitrary degrees.

Making Scientific Computing Models Differentiable for Deep Learning

Ramani Duraiswami

University of Maryland

Wednesday, February 21, 2024 @ 11:10 AM

Deep learning via deep neural networks has achieved great success in learning complex functions from data in areas like computer vision and natural language processing. A key aspect is the use of automatic differentiation and backpropagation on differentiable computational graphs to obtain gradients for optimization. However, many scientific computing problems have complex non-differentiable operations like meshing, matrix decompositions, and boolean decisions. Enabling differentiability in such scientific computing pipelines could allow the incorporation of deep learning for tasks like parameter optimization, cost function minimization, inverse problem solution, and learning predictive models. We propose techniques to make formulations in areas like computational electromagnetics, acoustics, and signal processing amenable to differentiation. Broader adoption would enable efficient physics-constrained deep learning, and end-to-end differentiable scientific computing. Potential problems that could benefit include electromagnetic design, seismic imaging, computational fluid dynamics, and multiscale materials modeling. This interdisciplinary approach combining scientific computing and deep learning could enable new applications at the intersection of physical modeling and data-driven methods.

Pseudo-differential Integral Autoencoder Network for Inverse PDE Operators

Chunmei Wang

University of Florida

Wednesday, February 21, 2024 @ 11:50 AM

Partial differential equations (PDEs) play a foundational role in modeling physical phenomena. This study addresses the challenging taskof determining variable coefficients within PDEs from measurement data. We introduce a novel neural network, "pseudo-differential IAEnet" (pd-IAEnet), which draws inspiration from pseudo-differential operators. pd-IAEnet achieves significantly enhanced computational speedand accuracy with fewer parameters compared to conventional models. Extensivebenchmark evaluations are conducted across a range of inverse problems, including Electrical Impedance Tomography (EIT), optical tomography, and seismic imaging, consistently demonstrating pd-IAEnet's superior accuracy. Notably, pd-IAEnet exhibits robustness in the presence of measurement noise, a critical characteristic for real-world applications. An exceptional feature is its discretization invariance, enabling effective training on data from diverse discretization schemes while maintaining accuracy on different meshes. In summary, pd-IAEnet offers a potent and efficient solution for addressing inverse PDE problems, contributing to improved computational efficiency, robustness, and adaptability to a wide array of data sources.

High-dimensional PDEs, tensor-network, and convex optimization

Yuehaw Khoo

University of Chicago

Wednesday, February 21, 2024 @ 2:00 PM

This talk presents new computational approaches for high-dimensional partial differential equations (PDEs), employing tensor networks and convex relaxations. Specifically, based on these approaches, we demonstrate the construction of inner and outer approximations to PDE solutions using low-order statistics. These in turn effectively address the curse of dimensionality.

Neural Inverse Operators for Solving PDE Inverse Problems

Yunan Yang

Cornell University

Wednesday, February 21, 2024 @ 2:40 PM

A large class of inverse problems for PDEs are only well-defined as mappings from operators to functions. Existing operator learning frameworks map functions to functions and need to be modified to learn inverse maps from data. We propose a novel architecture termed Neural Inverse Operators (NIOs) to solve these PDE inverse problems. Motivated by the underlying mathematical structure, NIO is based on a suitable composition of DeepONets and FNOs to approximate mappings from operators to functions. A variety of experiments are presented to demonstrate that NIOs significantly outperform baselines and solve PDE inverse problems robustly, accurately and are several orders of magnitude faster than existing direct and PDE-constrained optimization methods.

Towards efficient deep operator learning for forward and inverse PDEs

Ke Chen

University of Maryland

Wednesday, February 21, 2024 @ 3:45 PM

Deep neural networks (DNNs) have been a successful model across diverse machine learning tasks, increasingly capturing the interest for their potential in scientific computing. This talk delves into efficient training for PDE operator learning in both the forward and the inverse problems setting. In the forward setting, we address the curse of dimensionality and demonstrate that certain PDE structures require fewer training samples through an analysis of learning error estimates. In the inverse setting, we propose a novel algorithm for solving inverse problems through data regularization. We analyze and provide explicit learning error estimates for the LASSO problem for sparse linear regression and Bayesian inversion for a nonlinear operator.

Data-Complexity Bounds for Operator Learning

Samuel Lanthaler

California Institute of Technology

Wednesday, February 21, 2024 @ 4:25 PM

Operator learning frameworks leverage neural networks and define a methodology for the data-driven approximation of operators. How much data is necessary to learn operators in such a purely data-driven manner? In this presentation, I will report on recent work that provides first answers on this question. Upper and lower bounds on the data-complexity of operator learning will be discussed.

Learning Interaction laws in particle- and agent-based systems

Mauro Maggioni

John Hopkins University

Thursday, February 22, 2024 @ 9:15 AM

We consider systems of interacting agents or particles, which are commonly used for modeling across the sciences. Oftentimes the laws of interaction between the agents are quite simple, for example they depend only on pairwise interactions, and only on pairwise distance in each interaction. We consider the following inference problem for a system of interacting particles or agents: given only observed trajectories of the agents in the system, can we learn what the laws of interactions are? We would like to do this without assuming any particular form for the interaction laws, i.e. they might be "any" function of pairwise distances, or other variables, on Euclidean spaces, manifolds, or networks. We consider this problem in the case of a finite number of agents, with observations along an increasing number of paths. We cast this as an inverse problem, discuss when this problem is well-posed, construct estimators for the interaction kernels with provably good statistically and computational properties. We discuss the role of the geometry of the underlying space, in the cases of Euclidean space, manifolds, and networks, even in the case when the network is unknown. We also consider extensions to second-order systems, more general interaction kernels, stochastic systems, and to the setting where the variables (e.g. pairwise distance) on which the interaction kernel depends are not known a priori.

Intrinsic Models in Wasserstein Space with Applications to Molecular Dynamics

James Murphy

Tufts University

Thursday, February 22, 2024 @ 9:55 AM

We study the problems of efficient modeling and representation learning for probability distributions in Wasserstein space. We consider a general barycentric coding model in which data are represented as Wasserstein-2 (W2) barycenters of a set of fixed reference measures. Leveraging the geometry of W2-space, we develop a tractable optimization program to learn the barycentric coordinates and provide a consistent statistical procedure for learning these coordinates when the measures are accessed only by i.i.d. samples. Our consistency results and algorithms exploit entropic regularization of the optimal transport problem, and the statistical convergence of entropic optimal transport maps will be discussed. We also consider the problem of learning reference measures given observed data. Our regularized approach to dictionary learning in W2-space addresses core problems of ill-posedness and in practice learns interpretable dictionary elements and coefficients useful for downstream tasks. Applications of optimal transport to compression of molecular dynamics simulations will be considered.

Deep Stochastic Mechanics

Rebecca Willett

University of Chicago

Thursday, February 22, 2024 @ 11:00 AM

In this talk, I will describe a novel deep-learning-based approach for numerical simulation of a time-evolving Schrödinger equation inspired by stochastic mechanics and generative diffusion models. Unlike existing approaches, which exhibit computational complexity that scales exponentially in the problem dimension, our method allows us to adapt to the latent low-dimensional structure of the wave function by sampling from the Markovian diffusion. Depending on the latent dimension, our method may have far lower computational complexity in higher dimensions. Moreover, we propose novel equations for stochastic quantum mechanics, resulting in linear computational complexity with respect to the number of dimensions. Numerical simulations verify our theoretical findings and show a significant advantage of our method compared to other deep-learning-based approaches used for quantum mechanics. This is joint work with Elena Orlova, Aleksei Ustimenko, Ruoxi Jiang, and Peter Y. Lu.

Adjoint Optimization of Deep-Learning Sub-Grid Scale Models for Large Eddy Simulation of Compressible Flows

Tom Hickling

University of Oxford

Thursday, February 22, 2024 @ 11:40 AM

Direct numerical simulation (DNS) of the Navier-Stokes equations for real-world engineering conditions and geometries (e.g., a wing of an aircraft) is typically computationally intractable. Large-eddy simulation (LES), which resolves only the largest turbulence scales, is computationally tractable but introduces unclosed, sub-grid scale (SGS) terms which must be modeled. SGS models usually introduce errors which reduce the accuracy of the LES prediction. There is a growing interest in leveraging deep-learning sub-grid scale (DL-SGS) models to improve the predictive accuracy of LES at coarse grid resolutions. We develop and implement an adjoint optimization method for training DL-SGS models for compressible flows. To address computational challenges, the adjoint equations for the LES compressible flow are parallelized via domain decomposition across multiple GPUs. The deep learning LES (DL-LES) model is trained on DNS data for a NACA 0012 airfoil and then evaluated out-of-sample for higher angles-of-attack, new airfoil geometries, and different Reynolds numbers. The DL-LES model is compared against benchmark LES simulations with standard SGS closure models. Important observations regarding the effect of the DL-SGS model architecture on the stability of the DL-LES simulation will also be discussed.

Principled Hyperparameter Transfer Across Depth and Width in Neural Networks

Boris Hanin

Princeton University

Thursday, February 22, 2024 @ 2:00 PM

Successfully deploying deep neural networks often requires significant experimentation to find good settings of hyperparameters that determine architecture (e.g. depth, width, residual blocks) and optimizer (e.g. learning rate schedule, weight decay, etc). As models and datasets continue to grow in complexity naive hyperparameter sweeps become prohibitively expensive. The purpose of this talk is to present several novel theoretical results that give reliable and extensively validated principles for hyperparameter transfer, in which optimal hyperparameters for very small models provably transfer to much larger models. Joint work with Blake Bordelon, Lorenzo Noci, Mufan Li, and Cengiz Pehlevan.

A Neural Network Warm-Start Approach for Inverse Scattering Problems

Jiequn Han

Flatiron Institute

Thursday, February 22, 2024 @ 2:40 PM

Inverse scattering problems play a crucial role in numerous applications across various fields. However, the widely-used optimization formulation is challenging to solve due to the computationally expensive evaluation of the objective function and the problem's highly nonlinear, non-convex, and ill-posed nature. In this talk, we introduce a neural network warm-start approach to effectively tackle these challenges while maintaining high precision. We will discuss the benefits of our approach, its implications, and how it can contribute to enhancing current practices in the field. The underlying philosophy of this method has the potential to be applied to a broader range of scientific challenges in the field of PDEs that demand low computational costs and high accuracy.

On the Stochastic (Variance-Reduced) Proximal Gradient Method for Regularized Expected Reward Optimization

Ling Liang

University of Maryland

Thursday, February 22, 2024 @ 3:45 PM

We consider a regularized expected reward optimization problem in the non-oblivious setting that covers many existing problems in reinforcement learning. In order to solve such an optimization problem, we apply and analyze the vanilla stochastic proximal gradient method. In particular, the method has shown to admit an $O(\epsilon^{-4})$ sample complexity to an ϵ-stationary point, under standard conditions. Since the variance of the vanilla stochastic gradient estimator is typically large which slows down the convergence, we also apply an efficent stochastic variance-reduce proximal gradient method with an importance sampling based ProbAbilistic Gradient Estimator (PAGE). To the best of our knowledge, the application of this method represents a novel approach in addressing the general regularized reward optimization problem. Our analysis shows that the sample complexity can be improved from $O(\epsilon^{-4})$ to $O(\epsilon^{-3})$ under additional conditions. Our results on the stochastic (variance-reduced) proximal gradient method match the sample complexity of their most competitive counterparts under similar settings for discrete-time and discounted Markov decision processes.

A Finite Expression Method for Solving High-Dimensional Committor Problems

Zezheng Song

University of Maryland

Thursday, February 22, 2024 @ 4:10 PM

Transition path theory (TPT) is a mathematical framework for quantifying rare transition events between a pair of selected metastable states A and B. Central to TPT is the committor function, which describes the probability to hit the metastable state B prior to A from any given starting point of the phase space. Once the committor is computed, the transition channels and the transition rate can be readily found. The committor is the solution to the backward Kolmogorov equation with appropriate boundary conditions. However, solving it is a challenging task in high dimensions due to the need to mesh a whole region of the ambient space. In this work, we explore the finite expression method (FEX, Liang and Yang (2022)) as a tool for computing the committor. FEX approximates the committor by an algebraic expression involving a fixed finite number of nonlinear functions and binary arithmetic operations. The optimal nonlinear functions, the binary operations, and the numerical coefficients in the expression template are found via reinforcement learning. The FEX-based committor solver is tested on several high-dimensional benchmark problems. It gives comparable or better results than neural network-based solvers. Most importantly, FEX is capable of correctly identifying the algebraic structure of the solution which allows one to reduce the committor problem to a low-dimensional one and find the committor with any desired accuracy.

Sharp error estimates for target measure diffusion maps with application to the committor problem

Shashank Sule

University of Maryland

Thursday, February 22, 2024 @ 4:35 PM

We obtain asymptotically sharp error estimates for the Target Measure Diffusion map (TMDmap) (Banisch et al. 2020), a variant of diffusion maps featuring importance sampling and hence allowing input data drawn from an arbitrary density. The derived error estimates include the bias error and the variance error. The resulting convergence rates are consistent with the approximation theory of graph Laplacians. The key novelty of our results lies in the explicit quantification of all the prefactors on leading-order terms. An important application of TMDmap is the analysis of rare events in systems governed by overdamped Langevin dynamics using the framework of transition path theory (TPT). The cornerstone ingredient of TPT is the solution of the committor problem, a boundary value problem for the backward Kolmogorov PDE. We derive an error estimate for the solution to the committor problem obtained with the TMDmap, showing how the consistency error transfers over to the solution error. Remarkably, the TMDmap algorithm is particularly suited as a meshless solver to the committor problem due to the cancellation of several error terms. Furthermore, significant improvements in bias and variance errors occur when using quasi-uniform sampling density. Our numerical experiments show that these improvements in accuracy are realizable in practice when using δ-nets as spatially uniform inputs to the TMDmap algorithm.

Optimal control for sampling the transition path process and estimating rates

Margot Yuan

University of Maryland

Thursday, February 22, 2024 @ 5:00 PM

Many processes in nature such as conformal changes in biomolecules and clusters of interacting particles, genetic switches, mechanical or electromechanical oscillators with added noise, and many others are modeled using stochastic differential equations with small white noise. The study of rare transitions between metastable states in such systems is of great interest and importance. The direct simulation of rare transitions is difficult due to long waiting times. Transition path theory is a mathematical framework for the quantitative description of rare events. Its crucial component is the committor function, the solution to a boundary value problem for the backward Kolmogorov equation. The key fact exploited in this work is that the optimal controller constructed from the committor leads to the generation of transition trajectories exclusively. We prove this fact for a broad class of stochastic differential equations. Moreover, we demonstrate that the committor computed for a dimensionally reduced system and then lifted to the original phase space still allows us to construct an effective controller and estimate the transition rate with reasonable accuracy. Furthermore, we propose an all-the-way-through scheme for computing the committor via neural networks, sampling the transition trajectories, and estimating the transition rate without meshing the space. We apply the proposed methodology to four test problems: the overdamped Langevin dynamics with Mueller's potential and the rugged Mueller potential in 10D, the noisy bistable Duffing oscillator, and Lennard-Jones-7 in 2D.

Representation of symmetric and anti-symmetric functions

Jianfeng Lu

Duke University

Friday, February 23, 2024 @ 9:15 AM

Efficient representation of high-dimensional (totally) symmetric and anti-symmetric functions have been an important task in scientific machine learning for applications to many-body quantum mechanics. In this talk, we will discuss some recent results in trying to understand the representational power of various ansatz. In particular, we will discuss both positive and negative results regarding representation of anti-symmetric functions.
Based on joint works with Chongyao Chen, Ziang Chen, Huang Hang, and JM Landsberg.

Elliptic PDE learning is provably data-efficient

Alex Townsend

Cornell University

Friday, February 23, 2024 @ 9:55 AM

Can one learn a solution operator associated with a differential operator from pairs of solutions and righthand sides? If so, how many pairs are required? These two questions have received significant research attention in operator learning. More precisely, given input-output pairs from an unknown elliptic PDE, we will derive a theoretically rigorous scheme for learning the associated Green's function. By exploiting the hierarchical low-rank structure of Green's functions and randomized linear algebra, we will have a provable learning rate. Along the way, we will develop a more general theory for the randomized singular value decomposition and show how these techniques extend to parabolic and hyperbolic PDEs. This talk partially explains the success of operator networks like DeepONet in data-sparse settings.

Learning WENO for entropy stable schemes to solve conservation laws

Deep Ray

University of Maryland

Friday, February 23, 2024 @ 11:00 AM

Entropy-stable solvers for hyperbolic conservation laws ensure the selection of a physically relevant (weak) solution of the underlying PDE. Among such methods, the TeCNO schemes [Fjordholm et al, 2012] form a class of high-order finite difference-based solvers that utilize reconstruction algorithms satisfying a critical "sign-property" at the cell interfaces. However, only a handful of existing reconstructions are known to satisfy this property. In [Fjordholm & Ray, 2016], the first weighted essentially non-oscillatory (WENO) reconstruction satisfying the sign- property was developed. However, despite leading to provably entropy stable schemes, the numerical solutions using this reconstruction suffered from large under/overshoots near discontinuities. In this talk, we propose an alternate approach to constructing WENO schemes possessing the sign-property. In particular, we train a neural network to determine the polynomial weights of the WENO scheme, while strongly constraining the network to satisfy the sign-property. The training data comprises smooth and discontinuous data that represent the local solution features of conservation laws. Additional constraints are built into the network to guarantee the expected order of convergence (for smooth solutions) with mesh refinement. We present several numerical results to demonstrate a significant improvement over the existing variants of WENO with the sign- property.

Theoretical foundations for diffusion models

Holden Lee

Johns Hopkins University

Friday, February 23, 2024 @ 11:40 AM

Diffusion models are a highly successful approach for generative modeling, based on learning the score function (gradient of log-pdf) from data and then using it to simulate a stochastic process that transforms white noise into the data distribution. I'll discuss two fundamental questions for understanding diffusion models: (1) How do estimates of the score function translate into sampling guarantees? (2) For what families of distributions can we learn the score function efficiently? For the first question, I'll describe a framework for analyzing sampling error for general (non-smooth and multi-modal) distributions given only a L^2-accurate score estimate. This gives polynomial convergence for the SDE approach. However, in practice, it is important to reduce the dependence on the dimension d, as each step requires evaluating a large neural network, making diffusion models notoriously slow compared to other generative models. Towards this, I'll show that the ODE implementation (together with a corrector step) can reduce dimension dependence for smooth distributions to $O(\sqrt{d})$. For the second question, I'll show that the score function can be efficiently learned for mixtures of spherical gaussians and distributions supported on low-dimensional manifolds, giving an end-to-end result for learning these distributions.

The Brin Mathematics Research Center

The Brin Mathematics Research Center is a research center that sponsors activity in all areas of pure and applied mathematics and statistics. The Brin MRC was funded in 2022 through a generous gift from the Brin Family. The Brin MRC is part of the Department of Mathematics at the University of Maryland, College Park.

Activities sponsored by the Brin MRC include long programs, conferences and workshops, special lecture series, and summer schools. The Brin MRC provides ample opportunities for short-term and long-term visitors that are interested in interacting with the faculty at the University of Maryland and in experiencing the metropolitan Washington DC area.

The mission of the Brin MRC is to promote excellence in mathematical sciences. The Brin MRC is home to educational and research activities in all areas of mathematics. The Brin MRC provides opportunities to the global mathematical community to interact with researchers at the University of Maryland. The center allows the University of Maryland to expand and showcase its mathematics and statistics research excellence nationally and internationally.

List of Participants

Maria Cameron , University of Maryland
Ke Chen, University of Maryland
Ramani Duraiswami, University of Maryland
Jiequn Han, Flatiron Institute
Boris Hanin, Princeton University
Tom Hickling, University of Oxford
Yuehaw Khoo, University of Chicago
Samuel Lanthaler, California Institute of Technology
Holden Lee, Johns Hopkins University
Doron Levy, University of Maryland/Director, Brin MRC
Ling Liang, University of Maryland
Jianfeng Lu, Duke University
Mauro Maggioni , John Hopkins University
Reza Malek-Madani, Office Naval Research
James Murphy, Tufts University
Deep Ray, University of Maryland
Zezheng Song, University of Maryland
Shashank Sule, University of Maryland
Alex Townsend, Cornell University
Chunmei Wang, University of Florida
Rebecca Willett, University of Chicago
Jinchao Xu, Pennsylvania State University
Yunan Yang, Cornell University
Haizhao Yang, University of Maryland
Margot Yuan, University of Maryland