Generative models with singular score functions. Application to speckle formation and laser light communication

Guillaume Bal, The University of Chicago

Generative models form a natural class of algorithms to sample complex probability distributions. This is done via a map to a simpler, approximately Gaussian distribution, whose inversion requires learning a singular score function when the initial distribution is supported on a low dimensional manifold. We identify the singular structure of the score function allowing us to optimally learn a modified, smoother, score function. We show how the modified algorithm outperforms generative models that do not account for such singularities. The algorithm is applied to the problem of laser light propagation through turbulent atmospheres used in free space communication. The objective is to classify incident signals from noisy measurements. We use the generative model to augment the number of available realizations and demonstrate improved classification with our modified generative model. This is joint work with Yubin Lu, Anjali Nair, Aokun Wang, and Zhongjian Wang.

Exact Lipschitz Regularization of Convex Optimization Problems

Amir Beck, Tel Aviv University

We consider the class of convex composite minimization problems which consists of minimizing the sum of two nonsmooth extended valued convex functions, with one which is composed with a linear map. Convergence rate guarantees for first order methods on this class of problems often require the additional assumption of Lipschitz continuity of the nonsmooth objective function composed with the linear map. We introduce a theoretical framework where the restrictive Lipschitz continuity of this function is not required. Building on a novel dual representation of the so-called Pasch-Hausdorff envelope, we derive an exact Lipshitz regularization for this class of problems. We then show how the aforementioned result can be utilized in establishing function values-based rates of convergence in terms of the original data. Joint work with Marc Teboulle.

Martingale deep neural networks for stochastic optimal controls

Wei Cai, Southern Methodist University

In this talk, we will present a highly parallel and derivative-free martingale neural network method, based on the probability theory of Varadhan's martingale formulation of PDEs, to solve Hamilton-Jacobi-Bellman (HJB) equations arising from stochastic optimal control problems (SOCPs), as well as general quasilinear parabolic partial differential equations (PDEs). In both cases, the PDEs are reformulated into a martingale problem such that loss functions will not require the computation of the gradient or Hessian matrix of the PDE solution, and can be computed in parallel in both time and spatial domains. Moreover, the martingale conditions for the PDEs are enforced using a Galerkin method realized with adversarial learning techniques, eliminating the need for direct computation of the conditional expectations associated with the martingale property. For SOCPs, a derivative-free implementation of the maximum principle for optimal controls is also introduced. The numerical results demonstrate the effectiveness and efficiency of the proposed method, which is capable of solving HJB and quasilinear parabolic PDEs accurately and fast in dimensions as high as 100,000.

Learning multiscale coarse space

Eric Chung, CUHK

We will present a learning algorithm for multiscale coarse space. This problem is motivated by solving Darcy flow problem in highly heterogeneous media. When an iterative solver is applied, one needs an efficient preconditioner. The aim of our work is to approximate this preconditioner efficiently. We will demonstrate the results using some numerical examples.

A global structure-preserving kernel method for the learning of Hamiltonian systems

Juan-Pablo Ortega Lahuerta, Nanyang Technological University

In this talk, I will present a structure-preserving kernel ridge regression method that allows the recovery of globally defined, potentially high-dimensional, and nonlinear Hamiltonian functions on Poisson manifolds out of datasets made of noisy observations of Hamiltonian vector fields. The proposed method is based on finding the solution of a non-standard kernel ridge regression where the observed data is generated as the noisy image by a vector bundle map of the differential of the function that one is trying to estimate. Additionally, it is shown how a suitable regularization solves the intrinsic non-identifiability of the learning problem due to the degeneracy of the Poisson tensor and the presence of Casimir functions. A full error analysis is conducted that provides convergence rates using fixed and adaptive regularization parameters. The good performance of the proposed estimator is illustrated with several numerical experiments. This talk is joint work with Jianyu Hu and Daiying Yin.

On the diameter of subgradient sequences in o-minimal structures

Lexiao Lai, HKU

We study subgradient sequences of locally Lipschitz functions definable in a polynomially bounded o-minimal structure. We show that the diameter of any subgradient sequence is related to the variation in function values, with error terms dominated by a double summation of step sizes. Consequently, we prove that bounded subgradient sequences converge if the step sizes are of order 1/k. The proof uses Lipschitz L-regular stratifications in o-minimal structures to analyze subgradient sequences via their projections onto different strata.

Geometry-Preserving Numerical Schemes for Differential Equations on the Sphere

Tim Leung, HKUST

Differential equations with solutions constrained to the unit sphere appear in diverse fields. This talk presents recent advances in geometry-preserving computation for spherical-valued functions. We introduce the Spherical Essentially Non-Oscillatory (SENO) schemes, which generalize Bézier-type interpolation to the sphere and achieve high-order smoothness while suppressing spurious oscillations. We also present the Spherical Total-Variation-Diminishing Runge–Kutta (STVDRK) methods, which integrate differential equations directly on the sphere via the exponential map and spherical linear interpolation, maintaining geometric constraints without projection. Applications include spherical p-harmonic flows for color image denoising. These developments demonstrate how structure preserving discretizations enhance stability and accuracy.

Physics-Informed Neural Networks for Multiscale PDEs with Discontinuous High-Contrast Coefficients

Wing Tat Leung, City University of Hong Kong

Physics-informed neural networks (PINNs) are data-driven approaches to solving equations. It is successful in many applications; however, the accuracy of the PINN is not satisfactory when it is used to solve multiscale equations. In this paper, we propose a PINN to solve PDEs with a discontinuous high-contrast coefficient. To find such a solution using a neural network representation, we introduce an additional feature input to the network to retain the inherent solution properties. We train the network using the physics-informed framework in which the loss function comprises the residual of the differential equation. Using the specially designed residual, we can handle the discontinuity of the high-contrast coefficient. We conduct a series of numerical experiments to demonstrate the effectiveness and the accuracy of the present network model. Numerical results show that even with a one-hidden-layer (shallow) network and a moderate number of neurons, the present network model can achieve an accurate result using sufficient training data points.

Deep Neural Network Quasiconformal Approaches for Mapping Problems and Applications

Ronald Lui, CUHK

Mapping problems refer to the process of finding an optimal correspondence between different data domains. Such problems appear in many applications in image processing, computer vision, and computer graphics, including image registration, surface parameterization, and segmentation. Quasiconformal mappings play an important role in solving these mapping problems, as they provide a flexible way to control local geometric distortion. However, it becomes very challenging in high-dimensional and complex scenarios where optimization often gets stuck in local minima. In this talk, I will discuss several deep neural network-based quasiconformal approaches developed for solving these mapping problems. Applications in imaging, computer vision, and computer graphics will also be presented.

Localized Model Order Reduction via Multiscale Spectral Generalised Finite Elements

Robert Scheichl, Heidelberg University

Multiscale Spectral Generalized Finite Element Methods (MS-GFEM) are a powerful new discretisation method for general variational problems that satisfy a Gårding-type inequality, including strongly non-Hermitian problems. The construction of optimal approximation spaces is localised and requires no a priori regularity assumptions. The global approximation error is controlled by the local errors, which are rigorously shown to decay nearly exponentially. The optimality hinges on an SVD of the local restriction operator in a suitable, coefficient-dependent inner product on an oversampled patch. Compactness of this operator in the space of a-harmonic functions guarantees spectral accuracy akin to Weyl asymptotics for the Laplacian. As such, MS-GFEM can be seen as an "hp-version" of Localized Orthogonal Decomposition (LOD). Given the coefficient function the local approximation spaces can be constructed efficiently. In this talk, I will show how we can use Grassmannian interpolation of the resulting local approximation spaces on sparse grids in high dimensions to derive localized model order reduction methods that inherit the nearly exponential spatial convergence of MS-GFEM and parametric convergence of sparse grids. I will show some numerical experiments confirming the theoretical results in the context of elliptic problems.

Unveiling Hidden Dynamics: Reconstructing Systems from Time-Label-Free Data

Zuoqiang Shi, Tsinghua University

In many scientific fields, from molecular dynamics to single-cell genomics, we encounter a common data challenge: high-dimensional observations are collected, but their temporal labels are missing. This transforms sequential trajectory data into an unordered point cloud, breaking the standard paradigm for dynamical system identification. This talk presents two complementary frameworks for reconstructing the underlying dynamics and recovering the lost time labels. We begin with a method that treats the data as samples from a probability distribution. By formulating the recovery task as a distribution-matching problem, specifically minimizing the Sliced Wasserstein Distance between observed and simulated data, we can simultaneously infer the system's governing equations and assign temporal parameters to each data point. This approach leverages a two-phase algorithm that combines neural network surrogates with traditional solver-based optimization to ensure robustness and accuracy, even with noisy data and complex, chaotic systems. Building on this foundation, we then introduce BlinDNO, a novel distributional neural operator designed to tackle an even more general setting. When the dynamics are governed by stochastic or partial differential equations, BlinDNO learns a direct map from an unordered set of distribution snapshots to the system. By integrating a U-Net-based imaging operator with an attention-based feature fusion module, BlinDNO achieves permutation invariance and demonstrates high accuracy on challenging problems. Together, these works provide a comprehensive toolkit for extracting physical laws from modern, complex datasets where time information is lost or inaccessible.

Deep Neural ODE Operator Networks for PDEs

Yongcun Song, NTU

Operator learning has emerged as a promising paradigm for developing efficient surrogate models to solve partial differential equations (PDEs). However, existing approaches often overlook the domain knowledge inherent in the underlying PDEs and hence suffer from challenges in capturing temporal dynamics and generalization issues beyond training time frames. This paper introduces a deep neural ordinary differential equation (ODE) operator network framework, termed NODE-ONet, to alleviate these limitations. The framework adopts an encoder-decoder architecture comprising three core components: an encoder that spatially discretizes input functions, a neural ODE capturing latent temporal dynamics, and a decoder reconstructing solutions in physical spaces. Theoretically, error analysis for the encoder-decoder architecture is investigated. Computationally, we propose novel physics-encoded neural ODEs to incorporate PDE-specific physical properties. Such well-designed neural ODEs significantly reduce the framework's complexity while enhancing numerical efficiency, robustness, applicability, and generalization capacity. Numerical experiments on nonlinear diffusion-reaction and Navier-Stokes equations demonstrate high accuracy, computational efficiency, and prediction capabilities beyond training time frames. Additionally, the framework's flexibility to accommodate diverse encoders/decoders and its ability to generalize across related PDE families further underscore its potential as a scalable, physics-encoded tool for scientific machine learning. This is joint work with Z. Li, K. Liu, H. Yue, and E. Zuazua.

ADAPTING NOISE TO DATA: GENERATIVE FLOWS FROM 1D PROCESSES

Gabriele Steidl, TU Berlin

We introduce a general framework for constructing generative models using one-dimensional noising processes. Beyond diffusion processes, we outline examples that demonstrate the flexibility of our approach, e.g. based on the Kac process and the related damped wave equation. Motivated by this, we propose a novel framework in which the 1D processes themselves are learnable, achieved by parameterizing the noise distribution through quantile functions that adapt to the data. Our construction integrates seamlessly with standard objectives, including Flow Matching and consistency models. Learning quantile-based noise naturally captures heavy tails and compact supports when present. Numerical experiments highlight both the flexibility and the effectiveness of our method. Joint work with J. Chemseddine, G. Kornhardt, R. Duong, P. Friz.

Proximal Gradient Methods in Nonsmooth Nonconvex Minimax: A Unified Convergence Analysis Framework

Marc Teboulle, Tel Aviv University

Nonconvex minimax problems are prevalent in modern applications. We focus on nonsmooth nonconvex minimax, thus departing from the more common weakly convex/concave and smooth models assumed in the recent literature. We present a unified framework for the convergence analysis of proximal gradient methods, for both parallel and alternating schemes. In contrast to the current literature which addresses the complexity of obtaining only "nearly approximate" stationary solutions, here we derive pointwise global convergence and refined complexity results. Furthermore, the flexibility of our approach allows us to expand the scope of minimax problems that can be addressed through the use of Non Euclidean proximal gradient steps, and to extend the convergence and complexity results to this broader setting. This is joint work with Eyal Cohen.

Overfitting has a limitation: a model-independent generalization error bound based on Rényi entropy

Atsushi Suzuki, HKU

Will further scaling up of machine learning models continue to bring success? A significant challenge in answering this question lies in understanding generalization error, which is the impact of overfitting. Understanding generalization error behavior of increasingly large-scale machine learning models remains a significant area of investigation, as conventional analyses often link error bounds to model complexity, failing to fully explain the success of extremely large architectures. This research introduces a novel perspective by establishing a model-independent upper bound for generalization error applicable to algorithms whose outputs are determined solely by the data's histogram, such as empirical risk minimization or gradient-based methods. Crucially, this bound is shown to depend only on the Rényi entropy of the data-generating distribution, suggesting that a small generalization error can be maintained even with arbitrarily large models, provided the data quantity is sufficient relative to this entropy. This framework offers a direct explanation for the phenomenon where generalization performance degrades significantly upon injecting random noise into data, where the performance degrade is attributed to the consequent increase in the data distribution's Rényi entropy. Furthermore, we adapt the no-free-lunch theorem to be data-distribution-dependent, demonstrating that an amount of data corresponding to the Rényi entropy is indeed essential for successful learning, thereby highlighting the tightness of our proposed generalization bound.

PODNO: Proper Orthogonal Decomposition Neural Operators

Zhongjian Wang, Nanyang Technological University

In this paper, we introduce Proper Orthogonal Decomposition Neural Operators (PODNO) for solving partial differential equations (PDEs) dominated by high-frequency components. Building on the structure of Fourier Neural Operators (FNO), PODNO replaces the Fourier transform with (inverse) orthonormal transforms derived from the Proper Orthogonal Decomposition (POD) method to construct the integral kernel. Due to the optimality of POD basis, the PODNO has potential to outperform FNO in both accuracy and computational efficiency for high-frequency problems. From analysis point of view, we established the universality of a generalization of PODNO, termed as Generalized Spectral Operator (GSO). In addition, we evaluate PODNO's performance numerically on dispersive equations such as the Nonlinear Schrödinger (NLS) equation and the Kadomtsev–Petviashvili (KP) equation.

A Multi-Agent Framework for Discovering Physical Laws

Tongqi Wen, HKU

In this talk, I will introduce a multi-agent artificial intelligence framework designed for the autonomous discovery of physical laws. The framework coordinates several specialized agents that perform data curation, feature construction, symbolic regression, and human-in-the-loop verification. By embedding domain knowledge into each stage of the workflow, the system ensures that the resulting equations remain interpretable and physically meaningful. We apply this framework to a range of problems in materials science, including the prediction of glass-forming ability in metallic glasses, Vickers hardness, and Young's modulus in multi-principal element alloys (MPEAs). Across these tasks, the discovered formulas exhibit strong out-of-distribution generalization and show close correspondence with established physical principles. Moreover, the learned laws can be used not only for prediction but also for design: when applied to property estimation in quaternary MPEAs, they achieve performance competitive with state-of-the-art atomistic models. More broadly, this work aims to move toward autonomous and interpretable AI systems capable of identifying concise, generalizable governing equations—an objective that resonates with both applied mathematics and the physical sciences.

Computing on spheres: From spherical designs to scattered and random data

Hao-Ning Wu, University of Georgia

Spherical t-designs provide a foundation for numerical integration and approximation on the sphere, offering spectral accuracy by ensuring exact integration of polynomials up to degree t. However, their computational construction becomes prohibitively difficult for large t, creating a significant practical barrier. To overcome this limitation, we introduce a framework that relaxes the stringent requirement of quadrature exactness. Leveraging the Marcinkiewicz--Zygmund inequality, we derive error bounds for hyperinterpolation that enable its use with point sets that are not strict spherical designs, thereby relaxing its original requirement for high quadrature exactness. This approach maintains provably good convergence rates even with non-optimal points. We further demonstrate the utility of this framework by applying hyperinterpolation to develop a spectral method for solving the Allen--Cahn equation on spheres. This method facilitates rigorous numerical analysis with general point sets, including random or scattered points. The efficacy of our approach is confirmed through numerical examples for both approximation and PDE solving.

Data Driven Modeling for Scientific Discovery and Digital Twins

Dongbin Xiu, The Ohio State University

We present a data-driven modeling framework for scientific discovery, termed Flow Map Learning (FML). This framework enables the construction of accurate predictive models for complex systems that are not amenable to traditional modeling approaches. By leveraging data and the expressiveness of deep neural networks (DNNs), FML facilitates long-term system modeling and prediction even when governing equations are unavailable.

FML is particularly powerful in the context of Digital Twins, an emerging concept in digital transformation. With sufficient offline learning, FML enables the construction of simulation models for key quantities of interest (QoIs) in complex Digital Twins, when direct mathematical modeling of the QoIs is infeasible. During the online execution of a Digital Twin, the learned FML model can simulate the QoIs without reverting to the computationally intensive Digital Twin simulation model. As a result, FML serves as an enabling methodology for real-time control and optimization for complex systems.

Machine-Learning Interatomic Potentials for Long-Range Systems

Zhenli Xu, Shanghai Jiao Tong University

Machine-learning interatomic potentials have emerged as a revolutionary class of force-field models in molecular simulations, delivering quantum-mechanical accuracy at a fraction of the computational cost and enabling the simulation of large-scale systems over extended timescales. However, they often focus on modeling local environments, neglecting crucial long-range interactions. We propose a Sum-of-Gaussians Neural Network (SOG-Net), a lightweight and versatile framework for integrating long-range interactions into machine learning force field. The SOG-Net employs a latent-variable learning network that seamlessly bridges short-range and long-range components, coupled with an efficient Fourier convolution layer that incorporates long-range effects. By learning sum-of-Gaussians multipliers across different convolution layers, the SOG-Net adaptively captures diverse long-range decay behaviors while maintaining close-to-linear computational complexity during training and simulation via non-uniform fast Fourier transforms. The method is demonstrated effective for a broad range of long-range systems.

Ab Initio Many-body Quantum Chemistry: from Algorithm-driven to AI-driven Developments

Jun YANG, HKU

Solving Schrödinger equation with full many-body molecular Hamiltonian is prohibitive for complex molecules. By nailing down several theory and algorithm approximations [1], efficient ab initio quantum chemistry has been demonstrated to achieve a great success in making quantitative and accurate prediction of solids [2] and complex processes [3]. The core idea supporting these developments is the low-rank tensor factorization technique, which has also inspired us recently to develop emergent AI models in quantum chemistry. The computational complexities have been reduced by several orders of magnitude for macromolecules that were previously impossible. In this talk, I will discuss my past works [4-10] that push the limit of scalable tensor factorization schemes towards low-rank quantum chemistry and deep neural network surrogate models. I will introduce the theory formulation, technical implementation and applications of these low-rank techniques, and highlight the experiences and lessons that we learned from traditional theories to motivate low-dimensional AI schemes for quantum chemistry.

References

1. Yang, J. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2024, 14, e1706.
2. Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schuetz, M.; Chan, G.K.-L. Science, 2014, 345, 640.
3. Liang, Q.; Zhu, C.; Yang, J. J. Am. Chem. Soc. 2023, 145, 10159.
4. Liang, Q.; Yang, J. J. Chem. Theory Comput. 2021, 17, 6841.
5. Zhou, R.; Liang, Q.; Yang, J. J. Chem. Theory Comput. 2020, 16, 196.
6. Ng, W.-P.; Liang, Q.; Yang, J. J. Chem. Theory Comput. 2023, 19, 5439.
7. Ng, W.-P.; Zhang, Z.; Yang, J. J. Chem. Theory Comput. 2025, 21, 1602.
8. Liang, Q.; Yang, J. J. Chem. Theory Comput. 2025, 21, 3360.
9. Li, J.; Yang, J. J. Phys. Chem. Lett. 2022, 13, 10042.
10. Liang, Q.; Yang, J. J. Chem. Theory Comput. 2025, under preparation.

Unconditionally energy-stable SLM-DG methods for Navier-Stokes equations at high Reynolds number

Lina Zhao, City University of Hong Kong

In this talk, we present fully discrete schemes for the unsteady Navier–Stokes equations that simultaneously achieve global and local mass conservation, pressure robustness and Reynolds-semi-robustness. These schemes are based on a stabilized Lagrange multiplier (SLM) formulation with first- and second-order accuracy in time and H(div)-conforming discontinuous Galerkin (DG) methods in space, where the nonlinear convection term is treated explicitly. The unconditional H^1 stability for both the first- and second-order time-stepping schemes will be presented. Furthermore, the pressure-robust and Reynolds-semi-robust error estimates under mild conditions on the temporal and spatial step sizes in both two- and three-dimensional cases will be developed. Finally, numerical experiments are presented to validate the accuracy and efficiency of the proposed schemes.

Some Patterns in NS-KPP

Kun Zhao, Haerbing Engineering University

This presentation is built on recent findings in the study of the coupled Navier-Stokes-KPP equations, a simplified framework used to model the dynamics of biological fluid flows. We introduce several families of nontrivial steady-state solutions of the system on two-dimensional bounded domains, under specific parametric and boundary constraints. The stability of these solutions will also be discussed.

Time reversal and generation: the backward heat equation meets diffusion models

Enrique Zuazua, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Diffusion-based generative models, which have recently revolutionized AI image and data synthesis, are deeply connected to one of the classical equations of mathematical physics: the heat equation.

In this talk, we explore this connection from a mathematical perspective, focusing on the time-reversal (inverse design) problem for parabolic systems. Although the heat equation satisfies backward uniqueness, diffusion irreversibility makes reconstructing initial states from later observations a major challenge. We present a recent approach that leverages long-time moment expansions and representer theorems to efficiently recover atomic initial sources from diffused data. In parallel, using the Li–Yau differential inequality, we provide a quantitative framework to understand diffusion-based generative models and their role in modern AI.

The talk will combine analytical insight with numerical experiments and conclude with open questions at the interface of PDE analysis, inverse problems, and generative machine learning. The contents of this talk are inspired by recent joint work with Kang Liu (Université de Bourgogne, France).