Workshop on Mathematics and Artificial Intelligence Abstracts

02 Feb 2026–06 Feb 2026

January 30, 2026

Learn to Synthesize Data in Imaging

Chenglong Bao

Yau Mathematical Sciences Center
Tsinghua University, China

Generative models are transforming scientific imaging by enabling the synthesis of realistic measurements that effectively compensate for scarce, noisy, or geometrically biased data. In this talk, I will begin with a motivating case from cryo-electron microscopy, where combining AI-generated data with experimentally acquired particles helps overcome the long-standing preferred-orientation challenge. Building on this insight, we abstract the data-synthesis process into a general formulation in which unpaired datasets are leveraged to recover the underlying joint distribution from their respective marginals. We further present a VAE based approach for solving this inverse problem.

AI for Mathematics: From Digitization to Intelligentization

Bin Dong

Beijing International Center for Mathematical Research
Peking University, China

Abstract: Artificial intelligence offers potential solutions to the efficiency bottlenecks in mathematical research. Crucially, the goal of "AI for Mathematics" is not the mere application of existing tools, but the enhancement of AI capabilities to genuinely support mathematical inquiry. This talk outlines the challenges necessitating this integration and reviews recent technical advances. We argue that formalization—the digitization of mathematics—is essential for bridging the gap between statistical models and rigorous reasoning. Finally, we present the research roadmap of the Peking University team, highlighting progress in formal tools, automated reasoning systems, and benchmark construction.

Hybrid Neural Modelling: Theory and Applications

Thomas Gaskin

The London School of Economics and Political Science, UK

Integrating neural networks into mechanistic, equation-based models is a field of growing scientific interest, yet it lacks consistent terminology and a unified mathematical framework. In this talk, I will systematise existing approaches and establish new connections between hybrid modelling, differential equations theory, and deep learning. I will discuss the role of inference, prediction, and generalisation in hybrid models, showing how neural components can capture mathematical structures within and across datasets, and connect these questions to structural identifiability theory. By further interpreting hybrid models through a manifold-learning lens, I demonstrate how parametric spaces of missing information can be learned, thereby providing an alternative approach to uncertainty quantification. Finally, some applications in computational social science are briefly presented.

A Geometric Perspective on Polynomially Solvable Convex Maximization

Shaoning Han

National University of Singapore, Singapore

Convex maximization arises in many applications but is generally NP-hard, even for low-rank objectives. This paper introduces a set of broadly applicable conditions that certify when such problems are polynomially solvable. Our main condition is a new property of the feasible set, which we term co-monotonicity. We show that this property holds for two important families: matroids and permutation-invariant sets. Under co-monotonicity and mild additional assumptions, we develop a geometric framework that generates polynomially many candidate solutions, one of which is optimal, which yields a polynomial-time algorithm.

Advances in Self-Supervised Image Denoising: From Gaussian Noise to Real-World Noise

Hui Ji

National University of Singapore, Singapore

Image denoising is a fundamental task in image restoration and a key component in solving inverse imaging problems. Deep learning has achieved remarkable success in denoising, particularly through supervised learning; however, its reliance on ground-truth images for network training limits its broader applicability in real-world scenarios. Recent research has increasingly shifted towards truth-free learning paradigms, where models are trained directly on noisy data without requiring clean references. In this talk, I will present a sequence of works on self-supervised image denoising, progressively addressing more complex noise models, from i.i.d. Gaussian noise to pixel-wise independent heteroscedastic noise, and ultimately to pixel-wise correlated noise observed in real-world data. Built upon data augmentation and mask-and-predict training strategies, our approaches enable networks to learn denoising solely from noisy observations. Experimental results demonstrate that the proposed self-supervised denoisers achieve performance comparable to supervised methods, offering a practical solution for data-limited imaging applications.

Origin of Quasiperiodic Interfaces

Kai Jiang

School of Mathematics and Computational Science
Xiangtan University, China

Perfect single-crystalline materials do not exist. Interfaces are prevalent in crystalline materials and exert a significant influence on material properties. The study of interface structures has long been recognized as one of the core issues in materials science. Most investigations focus on periodic interfaces; by contrast, the understanding of quasiperiodic interfaces—despite their ubiquity—remains limited. In this talk, we elucidate the origin of quasiperiodic order in interfaces by employing a unified theoretical and computational framework. Within the proximal coincidence point set theory we developed, quasiperiodicity arises intrinsically from cut-and-project constructions with visually indistinguishable perturbations, and is encoded in the corresponding Fourier-Bohr spectral structures. These spectral characteristics underpin our efficient numerical framework, which integrates the conserved Landau-Brazovskii model with the projection method to enable accurate entire-domain simulations. In this framework, quasiperiodicity is governed by the intensity distribution in reciprocal space and manifested in the resulting interfacial ordering. We report compelling quasiperiodic signatures across tilt grain boundaries, twist grain boundaries, and general interfaces, including finite local complexity, repetitivity, emergent generalized Fibonacci sequences, as well as 12-fold and 8-fold interfacial quasicrystal structures.

Blind Separation of Non-Stationary Multi-Component Signals: Enhanced SST/Chirplet Methods and Their Applications

Qingtang Jiang

School of Mathematical Sciences
Zhejiang Normal University, China

In nature, real-world phenomena that exist in the form of signals (or data acquired in the form of time series) are usually affected by various factors, ultimately manifesting as multi-component signals (or time series) that overlap with each other in the time dimension. To gain a deeper understanding of such phenomena and facilitate subsequent signal processing, we need to extract the unknown components of the target multi-component signals from blind source data. This is an inverse problem with a long research history. However, for non-stationary signals—i.e., signals whose component frequencies change over time—there has been a lack of rigorous mathematical processing methods until the research on the synchrosqueezing transform (SST) proposed by Ingrid Daubechies and her collaborators more than a decade ago. Despite the efforts of numerous researchers, SST can only be applied under strict conditions.

In this talk, we will introduce two improvement measures for SST. Firstly, time-varying parameters are introduced into the original SST to eliminate some constraints, especially those related to frequency separation, thereby achieving more robust separation of signal components. Secondly, we will discuss the newly developed chirplet transform, which can even separate multi-component signals with intersecting instantaneous frequency curves. Unlike the traditional two-dimensional time-frequency space, the chirplet transform maps signals into a three-dimensional time-frequency-chirp rate space. In addition, we will address a key challenge of this chirplet transform-based method: the slow decay problem along the chirp rate direction. Subsequently, we will present recent research results—overcoming this challenge by merging the X-ray transform with the chirplet wavelet transform. Furthermore, we will discuss the application of SST and chirplet transform to rotating machinery fault diagnosis and real-world depression screening in social media vlogs.

A New Variational model for Simulating Solid-state Dewetting problems

Wei Jiang

School of Mathematics and Statistics
Wuhan University, China

Solid-state dewetting is a ubiquitous phenomenon in materials science, and it describes the agglomeration of solid thin films into arrays of isolated particles on a substrate. In recent years, solid-state dewetting has found wide applications in modern technology, and many efforts have been devoted to understanding this important phenomenon. One of the effective tools for modeling and simulations of solid-state dewetting was the sharp-interface model proposed by my group [see Phys. Rev. B, 91:045303, 2015; SIAM J. Appl. Math., 80:1654, 2020]. However, these sharp-interface models belong to free boundary problems, which explicitly include surface diffusion and contact line migration. It is very difficult to design efficient numerical algorithms for solving the sharp-interface models. On the other hand, lots of pinch-off events will occur when a long island film evolves by solid-state dewetting, but the previous sharp-interface models can not automatically handle with topological events. In order to tackle these difficulties, we propose a new sharp-interface model for simulating solid-state dewetting by removing the singularity of the triple point among the film, vapor and substrate phases. The new model is only needed to solve in a fixed domain, and it can automatically capture "topological events". Numerical results demonstrate the high performance of the new model.

Provable Diffusion Posterior Sampling for Bayesian Inversion

Yuling Jiao
School of Artificial Intelligence
Wuhan University, China

This paper proposes a novel diffusion-based posterior sampling method within a plug-and-play (PnP) framework. Our approach constructs a probability transport from an easy-to-sample terminal distribution to the target posterior, using a warm-start strategy to initialize the particles. To approximate the posterior score, we develop a Monte Carlo estimator in which particles are generated using Langevin dynamics, avoiding the heuristic approximations commonly used in prior work. The score governing the Langevin dynamics is learned from data, enabling the model to capture rich structural features of the underlying prior distribution. On the theoretical side, we provide non-asymptotic error bounds, showing that the method converges even for complex, multi-modal target posterior distributions. These bounds explicitly quantify the errors arising from posterior score estimation, the warm-start initialization, and the posterior sampling procedure. Our analysis further clarifies how the prior score-matching error and the condition number of the Bayesian inverse problem influence overall performance. Finally, we present numerical experiments demonstrating the effectiveness of the proposed method across a range of inverse problems.

Schrödinger-Föllmer Diffusion: Sampling, Optimization, Generative Learning

Lican Kang

School of Mathematics and Statistics
Wuhan University, China

Sampling from probability distributions is an important problem in statistics and machine learning, especially in Bayesian inference when integration with respect to posterior distribution is intractable and sampling from the posterior is the only viable option for inference. In this paper, we propose Schrodinger-Follmer sampler (SFS), a novel approach to sampling from possibly unnormalized distributions.The proposed SFS is based on the Schrodinger-Follmer diffusion process on the unit interval with a time-dependent drift term, which transports the degenerate distribution at time zero to the target distribution at time one. Compared with existing Markov chain Monte Carlo samplers that require ergodicity, SFS does not depend on the property of ergodicity. Computationally, SFS can be easily implemented using the Euler-Maruyama discretization. In theoretical analysis, we establish non-asymptotic error bounds for the sampling distribution of SFS in the Wasserstein distance under reasonable conditions. We conduct numerical experiments to evaluate the performance of SFS and demonstrate that it is able to generate samples with better quality than several existing methods.

Structure-Preserving Construction of Collision Operators for Kinetic Equations from Molecular Dynamics

Huan Lei

Michigan State University, USA

We introduce a data-driven approach to learn generalized collision operators from molecular dynamics. Unlike conventional models (e.g., Landau), the present operator takes a symmetry-breaking non-stationary form that depends not only on the relative velocity but also on the average velocity of the collision pair, capturing heterogeneous energy transfer arising from collective interactions with the environment. The constructed model strictly preserves the frame-indifference, conservation laws, and physical constraints such as the H-theorem. To enable efficient numerical evaluation, we develop a fast spectral separation method that represents the kernel as a low-rank tensor product of univariate basis functions. This formulation admits an O(N log N) algorithm and structure-preserving discretization. Numerical results demonstrate that the proposed model accurately captures plasma dynamics in the moderately coupled regime beyond the standard Landau model while maintaining high computational efficiency and structure-preserving properties.

Learning, Approximation and Control

Qianxiao Li

National University of Singapore, Singapore

In this talk, we discuss some interesting problems and recent results on the interface of deep learning, approximation theory and control theory. Through a dynamical system viewpoint of deep residual architectures, the study of model complexity in deep learning can be formulated as approximation or interpolation problems that can be studied using control theory, but with a mean-field twist. In a similar vein, training deep architectures can be formulated as optimal control problems in the mean-field sense. We provide some basic mathematical results on these new control problems that so arise, and discuss some applications in improving efficiency, robustness and adaptability of deep learning models.

Average orders of Automorphism groups and Average-case complexity of Tensor Isomorphism problems

Yinan Li

School of Artificial Intelligence
Wuhan University, China

The Tensor Isomorphism problem asks whether two tensors lie in the same orbit under the natural action of general linear groups. In quantum information theory, this corresponds to determining whether two multipartite pure states are SLOCC-equivalent. In the setting of finite fields, many fundamental isomorphism problems—including those of graphs, groups, algebras, and polynomials—reduce to testing isomorphism between structural tensors.

In this talk, I will first survey recent progress on tensor isomorphism from a complexity-theoretic perspective. I will then present a proof that the average order of automorphism groups of 3-tensors is constant. This result yields nearly tight upper bounds on the number of p-groups of Frattini class 2 and resolves several open questions posed by Blackburn, Neumann, and Venkataraman. I will also describe average-case polynomial-time algorithms for structural tensor isomorphism problems, which in turn suggest potential attacks on certain post-quantum digital signature schemes based on isomorphism assumptions. The main mathematical tools involved include fixed-point analysis and random matrix theory over finite fields.

Approximation error for Holder class with Transformers

Xiliang Lu

School of Mathematics and Statistics
Wuhan University, China

In this talk we provide the approximation error upper and lower bounds for Holder class with Transformers. Specifically, we derive a new approximation upper bound for the standard Transformer architecture equipped with Softmax operators, ReLU activation functions, and residual connections. We prove that a Transformer network composed of at most $\mathcal{O}(\varepsilon^{-d_0/\alpha})$ blocks can approximate any bounded Hölder function with d_0-dimensional input and smoothness index $\alpha \in (0, 1]$ under any accuracy $\varepsilon > 0$. In the case of approximation lower bounds, leveraging the VC-dimension upper bound, we rigorously prove that Transformers demand for at least $\mathcal{O}(\varepsilon^{-d_0/(4\alpha)})$ blocks to achieve the ε approximation accuracy. As a final step, we extend the derived results for standard Transformers to a general regression task involving the non-negative scalar Hölder class, thereby establishing the corresponding excess risk rates. $\mathcal{O}(\varepsilon^{-d_0/\alpha})$

Understanding Weight Space Symmetries in Contemporary Deep Learning Architectures

Tan Minh Nguyen

National University of Singapore, Singapore

The parameter space of a neural network is often used as a proxy for the space of functions it represents, yet this correspondence is typically non-injective: distinct parameter configurations may realize the same function due to underlying symmetries. While such functional equivalence has been well studied in classical architectures, its role in modern models remains far less understood. In this talk, I will present our recent progress on the symmetry structure of modern neural architectures such as Transformers and Mixture-of-Experts. I will then discuss applications of these findings to equivariant metanetwork design and linear mode connectivity.

Diffusion Models for Inverse Problems: From Pretrained Priors to Posterior Sampling

Tongyao Pang

Yau Mathematical Sciences Center
Tsinghua University, China

Diffusion models have recently emerged as powerful generative priors for solving high-dimensional inverse problems. In this talk, I will present how pretrained diffusion models can be adapted from unconditional generation to posterior inference conditioned on observations. We revisit diffusion-based sampling through the lens of stochastic differential equations and probability flow dynamics, and discuss how data fidelity can be incorporated by modifying the underlying sampling process. This perspective connects diffusion models with classical Bayesian inference, stochastic control, and optimal transport, and aims to provide a principled understanding of their empirical success in both generative modeling and inverse problems.

Learning Sparse Representations with Symmetries

Yong Sheng Soh

National University of Singapore, Singapore

Symmetries and invariances feature in wide range of problems. For instance, translation invariance underpins the construction of convolutional neural networks used in vision tasks. In this talk, we will consider the task of learning data representations that obey some pre-specified notion of symmetry (e.g. translations or rotations). We will specifically focus on the task of learning sparse representations – or, equivalently, learning regularisers – for a dataset. We describe a fairly generic end-to-end recipe for learning such representations that can be instantiated for a fairly general class of groups. Important ideas that feature are (i) the representation theory of (compact) groups, which allows us to express these symmetries in terms of matrix groups, (ii) harmonic analysis, which allows us to leverage powerful ideas from Fourier analysis to facilitate computation, and (iii) convex geometry, which allows us to formulate these representation learning tasks as instances of structured conic programs

Generative AI: Mathematical Foundations and Applications

Jian Sun

School of Mathematics and Statistics
Xi'an Jiaotong University, China

Generative artificial intelligence is an important direction in the development of general artificial intelligence. It primarily focus on learning multimodal, high-dimensional, and complex data distributions, and new samples can be generated by the learned AI models. It serves as a methodological foundation for current AI applications such as automatic question answering, cross-modal generation, and AI for science. This talk introduces the underlying basis of generative AI in mathematics and statistics, presents optimal transport-based methods for building controllable/conditional generative AI models, with applications in medical image generation, multimodal image-text alignment, and molecular structure generation. Finally, it summarizes and provides an outlook on the development and future prospects of generative AI.

Adam-family Methods with Decoupled Weight Decay in Deep Learning

Kim Chuan Toh

National University of Singapore, Singapore

We investigate the convergence properties of a general class of Adam-family methods for minimizing quadratically regularized nonsmooth nonconvex optimization problems, especially in the context of training nonsmooth neural networks with weight decay. Motivated by AdamW, we propose a novel framework for Adam-family methods with decoupled weight decay. Within our framework, the estimators for the first-order and second-order moments of stochastic subgradients are updated independently of the weight decay term. Under mild assumptions and with non-diminishing stepsizes for updating the primary optimization variables, we establish the convergence properties of our proposed framework. In addition, we show that our framework encompasses a wide variety of well-known Adam-family methods, hence offering convergence guarantees for these methods in the training of nonsmooth neural networks. As a practical application of our framework, we propose a method named AdamD (Adam with Decoupled Weight Decay). Numerical experiments demonstrate that AdamD outperforms Adam and is comparable to the popular AdamW, in both the aspects of generalization performance and efficiency. [Based on joint work with Kuangyu Ding and Nachuan Xiao]

Diffusion models for High Dimensional Distributions

Xin Tong

National University of Singapore, Singapore

Diffusion model is a popular tool to generate new data samples. However, rigorous understanding of diffusion model is still lacking. One issue is how to train these models for high dimensional problems as score function estimation is subject to the curse of dimension. Another issue is how to avoid the memorization effect, where the diffusion model is bound to generate an exact copy from the training data. We will provide solutions to the first issue by focusing on high dimensional distributions with sparse dependence. We will leverage the sparse dependence to provide a local estimation of the score functions. As for the second issue, we will modify the diffusion model in the final stage and generate new samples close to the same manifold where the training data is originated.

Quantitative estimates on Convergence rates of Kinetic dynamics for Sampling

Lihan Wang

National University of Singapore, Singapore

Sampling has important applications in a wide range of areas, including molecular dynamics, Bayesian statistics and machine learning. Many of the most widely used dynamics for sampling have a kinetic structure, the two most prominent examples of which are underdamped Langevin dynamics and Hamiltonian Monte Carlo. We discuss the quantitative long-time convergence behavior of these kinetic dynamics for sampling, under different assumptions of the potential. For convex potentials, we show that kinetic sampling dynamics accelerate the convergence rate by a square root factor of the Poincaré constant, compared to the overdamped Langevin dynamics, which is the sampling analog of Nesterov acceleration in convex optimization. We also show in the weakly confining setting, how the growth rate of the potential impacts the convergence rates via weak variants of the Poincaré inequality. The analysis is inspired by the Armstrong-Mourrat variational framework for hypocoercivity, which combines a Poincaré-Lions inequality in time-augmented state space and an L^2 energy estimate.

Learning-Based Algorithms for Solving Combinatorial Optimization

Zaiwen Wen

Beijing International Center for Mathematical Research
Peking University, China

This talk explores learning-based optimization paradigms that deeply integrates data, models, and algorithms for combinatorial optimization. For binary integer programming, a Monte Carlo policy gradient method is employed to achieve efficient sampling and search within the feasible region. For the quadratic assignment problem, an integrated framework of offline training and online fine-tuning is constructed, balancing solution accuracy and scalability. In complex constrained path planning and directed acyclic graph-based heterogeneous scheduling, lazy mask decoding and heuristic search are synergistically combined. Extensive experiments demonstrate that the learning-based algorithms exhibit significant advantages in solution quality, feasibility, and robustness.

Variational Learning of Open Quantum Dynamics from Sparse and Noisy Data

Pinchen Xie

Lawrence Berkeley National Labs, USA

Modeling "open quantum systems" (quantum subsystems of interest that interact with a vast, unobserved environment) poses a significant computational challenge because the dimension of the environment's Hilbert space grows exponentially with its size. This exponential growth renders exact integration of the joint system's dynamics intractable. While traditional approximation methods force a trade-off between computational efficiency and physical accuracy, we introduce Data-Informed Quantum-Classical Dynamics (DIQCD) as a variational framework that learns an effective low-dimensional master equation of quantum evolution from data. Instead of simulating the full environment, DIQCD models the system's evolution using a time-dependent Lindblad equation, which is a matrix differential equation describing the density operator, driven by flexible classical stochastic processes optimized to mimic environmental fluctuations. By training this parameterized equation of motion on sparse, noisy time-series data, we demonstrate that DIQCD can predict complex phenomena, such as entanglement dynamics in Calcium Fluoride molecular qubits and carrier mobility in organic semiconductors. The method achieves accuracy comparable to expensive tensor-network methods at a fraction of the computational cost.

[P. Xie et al., Phys. Rev. Lett. (in press); arXiv:2508.17170 (2025)]

Towards Large Scientific Learning Models with In-Context Operator Networks (ICON)

Liu Yang

National University of Singapore, Singapore

Can we build a single large model for a wide range of scientific problems? We proposed a new framework for scientific machine learning, namely "In-Context Operator Learning" and the corresponding model "In-Context Operator Networks" (ICON). A distinguishing feature of ICON is its ability to learn operators from numerical prompts during the inference phase, without weight adjustments. A single ICON model can tackle a wide range of tasks involving different operators, since it is trained as a generalist operator learner, rather than being tuned to approximate a specific operator. This is similar to how a single Large Language Model can solve a variety of natural language processing tasks specified by the language prompt. We will show how a single ICON model (without fine-tuning) manages multiple distinct problem types, encompassing forward and inverse ODE, PDE, and mean-field control problems. Through a case study on 1D conservation laws, we will show ICON's strong generalization capability to new PDEs, as well as its advantage compared with classic operator learning methods, e.g., Fourier neural operator (FNO). We will also show the application of ICON in 2D fluid problems, where a single model can make predictions for incompressible or compressible fluids, with different viscosity.

A Complete Error Analysis for Deep Ritz Method

Zhijian Yang

School of Mathematics and Statistics
Wuhan University, China

It is widely known that the error analysis for deep learning involves approximation, statistical, and optimization errors. However, it is challenging to combine them together due to overparameterization. In this presentation, we address this gap by providing a comprehensive error analysis of the Deep Ritz Method (DRM). Specifically, we investigate a foundational question in the theoretical analysis of DRM under the overparameterized regime: given a target precision level, how can one determine the appropriate number of training samples, the key architectural parameters of the neural networks, the step size for the projected gradient descent optimization procedure, and the requisite number of iterations, such that the output of the gradient descent process closely approximates the true solution of the underlying partial differential equation to the specified precision.

Advances in Fast Nonconvex Algorithms for Low-Rank Hankel Matrix Recovery

Juntao You

School of Artificial Intelligence
Wuhan University, China

This talk studies the robust matrix completion problem for low-rank Hankel matrices. Low-rank Hankel matrix recovery provides a unified framework for spectral compressed sensing, which aims to reconstruct spectrally sparse signals from a limited number of randomly sampled time-domain observations. This problem arises frequently in signal processing and related areas. We give a partial overview of recent nonconvex approaches with provable guarantees, focusing on their sample complexity and convergence. In addition, to address practical challenges such as partial observations, impulsive noise, and ill-conditioning, we present a Newton-like algorithm designed to handle these issues in a unified manner.

A Gradient-Oriented Diffusion Sampling Method for Deep Partial Differential Equation Solvers

Haijun Yu

Academy of Mathematics and Systems Sciences
Chinese Academy of Sciences, China

In recent years, deep learning methods for the forward and inverse problems of partial differential equations have received increasing attention. The numerical errors of these methods can generally be decomposed into three parts: network approximation error, data sampling error, and training error. Among them, the data sampling error usually plays a dominant role in the total error. For cases where the smoothness of the solution is relatively uniform and the domain geometry is relatively simple, we can adopt Gaussian sampling to reduce sampling error. However, for complex domains, high-dimensional problems, and situations where the smoothness of the solution is highly non-uniform in spatial distribution, adaptive sampling is a mainstream approach to reducing sampling errors. Many scholars have focused on this issue and proposed a variety of adaptive sampling methods, but there is still no perfect solution currently. In this talk, we introduce a gradient-oriented diffusion sampling method. This method borrows the idea of moving mesh methods, uses the gradient information of the solution as an indicator for points movement, and moves the sampling points by solving a stochastic diffusion equation. This approach can avoid the large computational cost caused by optimizing the sampling points using the PINN's residual of the equation, and has good sampling efficiency for high-dimensional non-uniform problems. We have applied this method to elliptic steady-state problems, nonlinear parabolic problems (phase field equations), and the incompressible Navier-Stokes equations. The numerical results show that our method is significantly superior to the commonly used existing adaptive sampling methods.

On the Stabilization of PINNs

Cheng Yuan

School of Mathematics and Statistics
Wuhan University, China

In this talk, we demonstrate that by carefully selecting norms and regularization terms in the loss function, Physics-Informed Neural Networks (PINNs) can be effectively stabilized such that the solution error is rigorously bounded by the training loss itself. For forward problems, we introduce SPINNs, which replace the conventional L^2 penalties on initial and boundary conditions with H^1 norms when solving wave equations, achieving significantly faster and more robust convergence than vanilla PINNs. For inverse problems, we propose Tikhonov-PINNs, which incorporate Tikhonov regularization directly into the loss function to address inverse potential reconstruction, offering stability estimates and stochastic convergence guarantees even for nonlinear problems under unbounded noise. Through systematic theoretical analysis grounded in approximation and learning theory, supported by extensive numerical experiments, we illustrate the accuracy, stability, and convergence performance of these stabilized methods over standard PINNs.

SPIKE: Stable Physics-Informed Kernel Evolution Method for Solving Hyperbolic Conservation Laws

Lei Zhang

Beijing International Center for Mathematical Research
Peking University, China

We introduce the Stable Physics-Informed Kernel Evolution (SPIKE) method for numerical computation of inviscid hyperbolic conservation laws. SPIKE resolves a fundamental paradox: how strong-form residual minimization can capture weak solutions containing discontinuities. SPIKE employs reproducing kernel representations with regularized parameter evolution, where Tikhonov regularization provides a smooth transition mechanism through shock formation, allowing the dynamics to traverse shock singularities. This approach automatically maintains conservation, tracks characteristics, and captures shocks satisfying Rankine-Hugoniot conditions within a unified framework requiring no explicit shock detection or artificial viscosity. Numerical validation across scalar and vector-valued conservation laws confirms the method's effectiveness.

Applied Mathematics Inspired by Irrational Numbers

Pingwen Zhang
Wuhan University, China

Real numbers underpin mathematical research and the understanding of natural phenomena. The real number set comprises rational numbers (a measure-zero set) and irrational numbers (a full-measure set), implying that systems governed by irrational numbers are more prevalent. However, current computers cannot exactly store or represent irrational numbers, rendering numerical computations for irrational-dominated systems susceptible to Diophantine errors—incurred when approximating irrationals with rationals. Such errors significantly compromise numerical results. In this talk, we analyze the mechanism of Diophantine error and its impact on numerical computations. We also develop an approximation theory for multi-dimensional quasiperiodic functions with Diophantine frequencies when approximated by periodic functions. To avoid Diophantine errors, we propose two novel algorithms: the projection method and the finite points recovery method, which extend high-precision numerical computation from the rational to the real domain. Leveraging the ergodic and arithmetic properties of irrational numbers, we establish the mathematical foundations of these algorithms. Furthermore, we present their applications to several irrational-dominated scientific problems to discover new physical phenomena, including quasicrystals and their phase transitions, interfacial structures in materials, and quasiperiodic Schrödinger operators.

Unsupervised Operator Learning Approach for Dissipative equations via Onsager principle

Xiaofei Zhao

School of Mathematics and Statistics
Wuhan University, China

In this talk, we propose the deep Onsager operator learning (DOOL) method, a novel unsupervised framework for solving dissipative equations. Rooted in the Onsager variational principle (OVP), DOOL trains a deep operator network by directly minimizing the OVP-defined Rayleighian functional, requiring no labeled data, and then proceeds in time explicitly through conservation/change laws for the solution.The external time stepping enables temporal extrapolation. Numerical experiments on typical dissipative equations validate the effectiveness of the DOOL method, and systematic comparisons with supervised DeepONet and MIONet demonstrate its enhanced performance. Extensions are made to cover the second-order wave models with dissipation that do not directly follow OVP.

Efficient Deep Learning Methods for Very High Dimensional Quasilinear Parabolic PDEs and HJB Equations

Tao Zhou

Academy of Mathematics and Systems Sciences
Chinese Academy of Sciences, China

Solving high-dimensional PDEs with deep learning methods is often computationally and memory intensive, primarily due to the need for automatic differentiation to compute large Hessian matrices. We propose a deep random difference method (DRDM) that addresses these issues by approximating the convection-diffusion operator using first-order random differences, avoiding explicit Hessian computation. When incorporated into a Galerkin framework, the DRDM eliminates the need for pointwise evaluation of expectations, resulting in very efficient training procedure. Rigorous error estimates for DRDM are presented for linear PDEs. We further extend the approach to the Hamilton-Jacobi-Bellman (HJB) equations in stochastic optimal control. Numerical experiments demonstrate the efficiency of DRDM for solving quasilinear parabolic PDEs and HJB equations in dimensions up to 100000.

Optimal PhiBE — A Model-Free PDE-Based Framework for Continuous-Time Reinforcement Learning

Yuhua Zhu

University of California, Los Angeles, USA

This talk addresses continuous-time reinforcement learning (RL) in settings where the system dynamics are governed by a stochastic differential equation but remain unknown, with only discrete-time observations available. Existing approaches face fundamental limitations: model-based PDE methods suffer from non-identifiability, while classical model-free RL algorithms suffer from large discretization error.

We introduce Optimal-PhiBE, an equation that integrates discrete-time information into a continuous-time PDE, combining the strengths of both RL and PDE formulations. In linear-quadratic control, Optimal-PhiBE can even achieve the accurate continuous-time optimal policy with only discrete-time information. In addition, we establish convergence guarantees under model misspecification. Unlike classical RL analyses, whose errors typically blow up as the sampling interval shrinks, the approximation error of PhiBE remains stable and independent of discretization by exploiting the smoothness intrinsic to continuous-time dynamics.

Finally, we develop model-free algorithms for solving Optimal-PhiBE that operate in both online learning and off-policy data settings, requiring only minimal modifications to standard RL methods.