Tianyuan Mathematics Research Center

Interdisciplinary Research Between Dynamical Systems and Artificial Intelligence

8 March-14 March 2026

Conference Organizing Committee:

Zhihong Xia (Great Bay University)

Hongkun Zhang (Great Bay University)

Jianyu Chen (Soochow University)

Talk Schedule

	March 9 （ Mon ）	March 10 （ Tue ）	March 11 （ Wed ）	March 12 （ Thu ）	March 13 （ Fri ）
9:00 - 9:40	Yuan Yao	Kai Ming Ting	Xijun Hu	Yanni Xiao	Jiancheng Jiang
9:45-10:25	Jian Sun	Zenglin Xu	Jianwei Shen	Junxiong Jia	Xiang Zhou
10:25-10:55	Tea Break				
10:55-11:35	Discussion	Jun Shu	Ruxin Wang	Jian Zu	Zhihong Xia
11:30-14:00	Lunch Break				
14:00-14:40	Jianping Shi	Jinchao Feng	Discussion	Feng Jiao	Discussion
14:40-15:20	Xubin Zheng	Siku Yang		Gangnan Yuan	
15:20-15:50	Tea Break	Tea Break		Tea Break	
15:50-16:30	Nenggan Zheng	Yuan Lin		Zhe Pu	
16:30-17:10	Discussion	Silu Yan		Mingkun Li	

Conference Schedule

time	content
All-day registration for the conference on Sunday, March 8	
Monday, 9 March	
8:50-9:00	opening ceremony
9:00-9:40	Speaker: Yuan Yao (The Hong Kong University of Science and Technology) Title: Revisiting Smale's 18th Problem in the Age of Artificial Intelligence
9:40-10:20	Speaker: Jian Sun (Xi'an Jiaotong University) Title: Mathematical Foundation and Cross-Application of Controllable Generation Modeling
10:20-10:50	Tea Break
10:50-11:30	Discussion
11:30-14:00	Lunch Break
14:00-14:40	Speaker: Jianping Shi (Kunming University of Science and Technology) Title: PD-Outsourced PINN - A Novel PINN Method for Solving PDEs
14:40-15:20	Speaker: Xubin Zheng (Great Bay University) Title: Multiscale and Multimodal Fusion Technology in Biomedicine
15:20-15:50	Tea Break
15:50-16:40	Speaker: Nenggan Zheng (Zhejiang University) Title: Motion Behavior Control of Animal Robots
16:40-19:00	Dinner

Tuesday, 10 March	
9:00-9:40	Speaker: Kai Ming Ting (Nanjing University) Title: Why You Should Consider Using Isolation Kernel and Isolation Distributional Kernel
9:40-10:20	Speaker: Zenglin Xu (Fudan University) Title: Long-Term Sequence Prediction: Challenges and Prospects
10:20-10:50	Tea Break
10:50-11:30	Speaker: Jun Shu (Xi'an Jiaotong University) Title: Representation and Prediction Method of Continuous Spectrum Dynamic System Based on Koopman Theory
11:30-14:00	Lunch, nap
14:00-14:40	Speaker: Jinchao Feng (Great Bay University) Title: Data-driven Discovery of Asymmetric Interacting Particle Systems
14:40-15:20	Speaker: Sikun Yang (Great Bay University) Title: Numerically Efficient Methods for Solving High-dimensional PDE and Probabilistic Modeling
15:20-15:50	Tea Break
15:50-16:20	Speaker: Yuan Lin (Shanxi University) Title: Reconstructing and Predicting Stochastic Dynamic Systems Using Probabilistic Deep Learning
16:20-16:50	Speaker: Silu Yan (Great Bay University) Title: Rhythm-ONet: A Query-Conditioned Neural Operator for High-Fidelity Emulation of Multiscale Rhythm Transitions in Neurological Disorder Modeling
16:50-19:00	Dinner

Wednesday, 11 March	
9:00-9:40	Speaker: Xijun Hu (Shandong University) Title: Twist Dynamics in the Spatial Isosceles Three-body Problem
9:40-10:20	Speaker: Jianwei Shen (North China University of Water Resources and Electric Power) Title: Identifying Phase Transition Types in Epidemic Network Dynamics via Deep Learning
10:20-10:50	Tea Break
10:50-11:30	Reporter: Ruxin Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences) Title: Causal Condition-Driven Learning for Visual-Language Models in Biomedical Applications
11:30-14:00	Lunch Break
14:00-17:00	Discussion
17:00-19:00	Dinner

Thursday, 12 March	
9:00-9:40	Speaker: Yanni Xiao (Xi'an Jiaotong University) Title: Modelling Heterogeneities in Disease Transmission Dynamics
9:40-10:20	Speaker: Junxiong Jia (Xi'an Jiaotong University) Title: Nonlinear Transformation Based Infinite-Dimensional Variational Inference for Statistical Inverse Problems
10:20-10:50	Tea Break
10:50-11:30	Speaker: Jian Zu (Northeast Normal University) Title: Graph Hamiltonian Network: A Robust Model for Learning Particle Interactions in Lattice Systems
11:30-14:00	Lunch Break
14:00-14:40	Speaker: Feng Jiao (Guangzhou University) Title: Stochastic Dynamics Inference of Gene Expression Based on Interactive Signal Pathways
14:40-15:20	Speaker: Gangnan Yuan (Yunnan University) Title: From Approximation to Disentanglement: Trainable Gaussian Mixture Modules for Modern Neural Architectures
15:20-15:50	Tea Break
15:50-16:20	Speaker: Zhe Pu (Great Bay University) Title: Temporal Convolution-Based Physics-Informed Neural Networks for Solving Forward and Inverse Problems of Time Fractional Partial Differential Equations
16:20-16:50	Speaker: Mingkun Li (Great Bay University) Title: TS-PINNs: Physics-Informed Neural Networks in Temporal Sobolev Spaces
17:20-19:00	Dinner

Friday, 13 March	
9:00-9:40	Speaker: Jiancheng Jiang (Great Bay University) Title: How to Achieve Cross-Model Learning?
9:40-10:20	Speaker: Xiang Zhou (City University of Hong Kong) Title: Weak Generative Sampler
10:20-10:50	Tea Break
10:50-11:30	Speaker: Zhihong Xia (Great Bay University) Title: Mathematical Methods in Data Science
11:30-14:00	Lunch Break
14:00-17:00	Discussion
17:00-19:00	Dinner

Talk title and abstract

(Alphabetical order by surname)

Why you should consider using Isolation Kernel and Isolation Distributional Kernel

Kai Ming Ting (Nanjing University)

Abstract: The talk presents some successful applications of Isolation Kernel (IK) and Isolation Distributional Kernel (IDK). IK measures the similarity between two points/vectors. IK is better than the commonly used Euclidean distance and Gaussian kernel because of three unique characteristics. First, IK measures two points in a sparse region to be more similar than two points of the same inter-point distance in a dense region. This enables similarity to be compared more effectively in a dataset with varied densities. Second, IK is the only measure that has been proven to have broken the curse of dimensionality. This enables IK to deal with high-dimensional datasets more effectively than other measures. Third, IK has a finite-dimensional feature map, and this allows it to be computed more efficiently than other kernels. IDK is derived from IK to measure the similarity between two distributions, and it inherits the above characteristics of IK.

This presentation provides IK's successful applications to (a) improving t-SNE visualization and (b) retrieval in high-dimensional datasets. IDK has been successfully applied to (i) anomaly detection: creating a detector which is more effective than Isolation Forest---one of the most popular detectors used in industry and academia; and (ii) clustering: converting an NP-hard problem into a linear-time problem. A brief history of isolation-based methods can be found at https://github.com/IsolationKernel

Data-driven Discovery of Asymmetric Interacting Particle Systems

Jinchao Feng (Great Bay University)

Abstract: Interacting particle systems provide a powerful modeling framework for collective dynamics in nature and engineering. While prior methods have primarily addressed symmetric interactions using various learning techniques, many real-world systems exhibit asymmetric interactions, which demand more general and flexible modeling tools. In this talk, I will present a new Sparse Bayesian Learning (SBL) framework for identifying asymmetric interaction kernels in the Motsch – Tadmor model. By reformulating the nonlinear inverse problem as a subspace identification task, we establish identifiability guarantees and enable robust kernel recovery. Incorporating informative priors, the proposed SBL algorithm offers principled model selection and uncertainty quantification, achieving reliable inference from noisy trajectory data.

Twist Dynamics in the Spatial Isosceles Three-body Problem

Xijun Hu (Shandong University)

Abstract: We study dynamical constraints arising from Embedded Contact Homology (ECH) in the spatial isosceles three-body problem. For energies below the critical level, the dynamics on the energy surface is identified with a Reeb flow on the tight three-sphere. We obtain quantitative estimates for the Euler orbit, including monotonicity of its transverse rotation number and a strict inequality comparing its action with the contact volume. Combined with the ECH classification of Reeb flows on the tight three-sphere with two simple periodic orbits, these estimates rule out the two-orbit scenario, thus forcing every compact energy surface below the critical level to have infinitely many periodic orbits. The result admits a dynamical interpretation via disk-like global surfaces of section bounded by the Euler orbit. In this setting, the rotation number and the contact volume define a non-trivial twist interval which encodes the relative winding of periodic orbits.

Nonlinear Transformation Based Infinite-Dimensional Variational Inference for Statistical Inverse Problems

Junxiong Jia (Xi'an Jiaotong University)

Abstract: Inverse problems for PDEs are ubiquitous across scientific disciplines and can be formulated as statistical inference problems via Bayes’ theorem. For large-scale problems, the development of discretization-invariant algorithms is crucial, which can be achieved by formulating the methods in infinite-dimensional spaces. By restricting the variational family to be the pushforward of a nonlinear transformation of the prior measure, one obtains various classes of variational inference methods. By overcoming the singularity issues associated with probability measures defined on infinite-dimensional function spaces, we develop two such methods, termed infinite-dimensional Stein variational gradient descent (iSVGD) and functional normalizing flows (FNF). The transformations constructed in both iSVGD and FNF consist of a sequence of perturbations of the identity operator. In iSVGD, the perturbation mappings belong to an appropriate reproducing kernel Hilbert space, whereas in FNF, they are constructed using carefully designed neural network architectures. We apply these algorithms to an inverse problem governed by the steady-state Darcy flow equation. Numerical results validate the theoretical analysis, demonstrate the efficiency of the proposed algorithms, and confirm their discretization-invariant properties.

How to achieve cross-model learning?

Jiancheng Jiang (Great Bay University)

Abstract: Existing transfer learning methods for high-dimensional linear regression are constrained by the requirement of strict feature correspondence between source and target domains, which often proves impractical in real-world scenarios where feature representations differ (e.g., clinical scores versus sensor data) yet share predictive signals. To address this, we propose the Cross-Model Transfer Learning (CMTL) framework. By constructing data-adaptive

weight matrices that densely connect coefficients across all source and target models, CMTL enhances stable source signal transfer and suppresses noise through analytical means, eliminating the need for predefined feature mappings. Theoretically, CMTL achieves oracle efficiency under standard conditions. Empirically, it outperforms existing baseline methods in both simulation and real-world prediction tasks, demonstrating robustness when traditional approaches fail.

Stochastic Dynamics Inference of Gene Expression Based on Interactive Signal Pathways

Feng Jiao (Guangzhou University)

Abstract: Our study on gene expression regulation via interactive signaling pathways originates from the physiological phenomenon of antimicrobial peptide gene expression in pathogen-regulated innate immune systems of mosquitoes, fruit flies, and humans. The core concept of single-cell transcription stochastic dynamics models involves introducing random switching between closed and open states of genes. Through the tandem structure of multiple gene states in the model, it accurately depicts the ordered biochemical reaction processes following gene activation. To address the inherent limitations of traditional models, we constructed an interactive signaling pathway transcription model with a multi-state parallel structure. By integrating large-scale experimental data, we revealed a universal transcriptional regulatory mechanism and established a novel parameter estimation and model screening methodology.

TS-PINNs: Physics-Informed Neural Networks in Temporal Sobolev Spaces

Mingkun Li (Great Bay University)

Abstract: We propose TS-PINNs, a novel class of physics-informed neural networks that incorporate temporal Sobolev regularization to enhance the accuracy, stability, and generalization of neural PDE solvers. Unlike standard PINNs that penalize residuals only at discrete points, TS-PINNs leverage the structure of Sobolev spaces to enforce smoothness in time by explicitly regularizing higher-order temporal derivatives of the solution. This approach improves numerical stability, mitigates error accumulation in long-time integration, and leads to physically consistent approximations for time-dependent differential equations.

The TS-PINN framework is implemented in a multi-stage training procedure, where earlier stages learn temporally coarse solutions, and later stages refine them using higher-order constraints derived from the governing PDE. We provide theoretical motivation based on functional approximation theory in Sobolev spaces, and we empirically validate the method on several benchmark problems, including the Burgers equation and nonlinear wave propagation. Across all tasks, TS-PINNs demonstrate improved accuracy, enhanced robustness to noise, and faster convergence relative to classical PINNs and recent variants. Our results show that temporal Sobolev regularization provides a principled and practical enhancement to physics-informed learning of time-evolving systems.

Reconstruction and Prediction of Stochastic Dynamic Systems Using Probabilistic Deep Learning

Yuan Lin (Shanxi University)

Abstract: Real-world dynamic systems are frequently subjected to stochastic noise disturbances. Such randomness disrupts the deterministic embedding topology, causing point-to-point mapping models based on deterministic delay embedding theory to fail in reconstruction or prediction when processing noisy data. The key challenge in this field is how to effectively characterize the phase-space reconstruction mapping of stochastic dynamic systems using deep learning methods. This study proposes a novel framework that integrates prior knowledge of stochastic dynamic systems with probabilistic deep learning. Based on the stochastic delay embedding theorem, we construct a deep stochastic delay embedding computational model. By employing variational approximation of stochastic processes, we reconstruct stochastic dynamic systems to achieve robust long-term prediction.

Temporal Convolution-Based Physics-Informed Neural Networks for Solving Forward and Inverse Problems of Time Fractional Partial Differential Equations

Zhe Pu (Great Bay University)

Abstract: In this talk, we focuses on establishing the efficient computational framework for sloving time fractional partial differential equations (PDEs) by combining physical information neural networks (PINNs). Since the automatic differentiation is not applicable for fractional derivatives, the numerical discretization scheme of fractional derivatives is introduced by Pang et al. (2019) called fPINNs. However, this approach faces serious efficiency bottlenecks in non local historical calculations, the standard fPINNs construction evaluates a discrete fractional operator through nested history summations, leading to $O(N^2)$ work per spatial sample over N time levels with high complexity and computational cost. We propose the fractional PINNs framework called Convolution-fPINNs based on convolutional structure for numerical discretization formula and Fourier Transform method, which can significantly reduces the computational complexity to $O(N\log N)$ while strictly maintaining the mathematical equivalence with the numerical discretization formula. We couple this accelerated framework with PINNs and demonstrate the effectiveness of such method on forward solution and inverse parameter identification tasks for time-fractional advection--diffusion equations with highly nonlinear operator. Numerical results show that the Convolution-fPINNs method can effectively solve both the forward and inverse problems of high-dimensional time fractional PDEs, and achieves significant computational efficiency improvement while ensuring the accuracy compared to fPINNs.

Identifying phase transition types in epidemic network dynamics via deep learning

Jianwei Shen(North China University of Water Resources and Electric Power)

Abstract: Traditional statistical indicators face challenges in identifying phase transition types in epidemic dynamics on complex networks due to strong structural dependence and limited generalization capability. To address this issue, this paper proposes a deep learning framework that integrates network topology with node-level dynamical information. Based on a generalized SIS model, we construct four typical propagation scenarios (pairwise-only transmission, higher-order synergistic transmission only, mixed transmission, and cases with exogenous driving) and systematically generate dynamical samples labeled as first-order phase transition, second-order phase transition, or no transition. To mimic real-world observations, all model inputs are strictly truncated to the pre-transition period using node-level residual sequences, ensuring no future information leakage. The model adopts an end-to-end architecture combining graph convolutional networks (GCN), attention pooling, and long short-term memory (LSTM) networks. It encodes network structure via the adjacency matrix and constructs a dual-channel temporal input consisting of node residuals and their first-order differences to enhance the extraction of precursory dynamical features. Experimental results on Watts – Strogatz small-world networks demonstrate that the proposed method effectively distinguishes different phase transition types, significantly outperforming a CNN-LSTM baseline built on 12 handcrafted spatiotemporal indicators, particularly in discriminating first-order from second-order transitions. Ablation studies further reveal that the closer the truncation point to the critical transition, the higher the classification accuracy, indicating that pre-transition dynamics contain rich discriminative information about the transition type. This study provides a feasible pathway for intelligent identification of phase transition types in epidemic dynamics on complex networks and offers methodological support for early warning and intervention strategy design under higher-order synergistic transmission mechanisms.

PD-Outsoured PINN-A New Method for Solving PDEs with PINN

Jianping Shi (Kunming University of Science and Technology)

Abstract: Solving partial differential equations (PDEs) has always been a challenging issue in scientific research, especially when there are fractional-order derivatives in the equations. This paper proposes a new method for solving α (α ≤1)-order PDEs based on the physical information neural network (PINN): Portion-Derivative Outsourced PINN (PD-Outsourced PINN) framework. This framework introduces a analytical module to handle the α -order differential term, enabling the PINN to focus on learning the remaining part of the equation. Furthermore, the results of the analytical module is incorporated into the loss function of the PINN for correcting the network parameters. Three (1+1)-dimensional and two (2+1)-dimensional equations are used to validate the effectiveness and reliability of this method. The results show that, given the initial and boundary values of the PDEs, by combining the constructed analytical basis function and the output solution of the PINN, this method can obtain the high-precision data-driven solutions of the original equation. The PD-Outsourced PINN provides a high precision, good convergence, and strong robustness approach for solving complex PDEs.

Representation and Prediction Method of Continuous Spectrum Dynamic System Based on Koopman Theory

Jun Shu (Xi'an Jiaotong University)

Abstract: Representing and predicting high-dimensional spatiotemporal chaotic dynamical systems remains a fundamental challenge in the fields of dynamical systems theory and machine learning. High-dimensional, nonlinear systems with continuous spectrum structures are ubiquitous in real-world scenarios such as climate evolution, turbulent flows, complex network propagation, and neurodynamics. Although current data-driven methods can achieve relatively accurate short-term predictions, they often lack stability, interpretability, and scalability in systems dominated by broad-spectrum or continuous spectrum behaviors. While Koopman theory provides a linearized perspective for characterizing and predicting nonlinear dynamics, existing approaches typically rely on finite-dimensional approximations, which often lead to performance degradation in high-dimensional scenarios. This report proposes a novel neural Koopman method that achieves structured representation of dynamical systems by separating reversible motion from irreversible dissipation. This approach not only enhances long-term prediction accuracy and stability but also helps reveal which aspects of chaotic behavior can be understood and learned.

Mathematical Foundation and Cross Application of Controllable Generation Modeling

Jian Sun (Xi'an Jiaotong University)

Abstract: Generative AI represents a pivotal direction in the development of general artificial intelligence. It primarily achieves learning from multimodal and high-dimensional complex sample distributions, as well as generating new samples, through the design of AI algorithms. This serves as the methodological foundation for current AI applications in automated question-answering, cross-modal generation, and AI for science. The underlying basis of generative AI is mathematics and statistics. This report primarily introduces the background of generative AI, its mathematical/statistical principles, and the challenges it faces. It further elaborates on methods for constructing controllable/conditional generation in AI based on optimal transport, and their applications in medical image generation, multimodal image-text alignment, and molecular structure generation. Finally, it summarizes and prospects the development and future of generative AI.

Causal conditional prompt learning for biomedical visual-linguistic models

Ruxin Wang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences)

Abstract: This report focuses on visual-language models for biomedical applications, exploring fine-tuning methods for large models based on prompt learning and their biomedical applications. Addressing the multimodal nature and sparse annotation of biomedical data, causal interventions are introduced to enhance the interpretability and generalization of prompt learning, effectively

mitigating data bias and false positives. Through efficient parameter fine-tuning on pre-trained models, deep alignment between visual features and clinical text knowledge is achieved, significantly improving model performance in tasks such as medical image analysis, molecular representation, and cross-modal retrieval.

Mathematical Methods in Data Science

Zhihong Xia (Great Bay University)

Abstract: We discuss some fundamental mathematical theory behind data science and AI, including approximation theory and a new complexity theory for data.

Modelling heterogeneities in disease transmission dynamics

Yanni Xiao (Xi'an Jiaotong University)

Abstract: Accurate prediction of epidemics is pivotal for making well-informed decisions for the control of infectious diseases, but modelling heterogeneity in the system becomes a challenge. In this talk, we propose a novel modelling framework integrating the spatio-temporal heterogeneity of susceptible individuals into homogeneous models, by introducing a continuous recruitment process for the susceptibles. Then, a general human heterogeneous disease model with mutation is proposed to comprehensively study the effects of human heterogeneity on basic reproduction number, final epidemic size and herd immunity. We show that human heterogeneity may increase or decrease herd immunity level, strongly depending on some convexity of the heterogeneity function. Finally, we illustrate how to link the deep learning to dynamic model to examine time-dependent transmission rate or the intensity of interventions.

Long-term Sequence Prediction: Challenges and Prospects

Zenglin Xu (Fudan University)

Abstract: In the field of multivariate time series prediction, accurate long-term trend forecasting plays a pivotal role across diverse domains ranging from climate science to financial markets. With the increasing complexity of dynamic systems and the surge in time-series data across various fields, there is an urgent need to address inherent complex dependencies and temporal uncertainties in time-series data, while elucidating the mechanisms underlying time-series dynamics. This report outlines our latest research efforts, tackling these challenges through several distinct approaches: decoupling time-series variables, modeling dependencies between time-series variables, exploring predictive mechanisms for multivariate time-series systems, and proposing to conceptualize temporal evolution in time-series analysis as a continuous flow analogous to physical processes described by partial differential equations. Finally, we delineate the boundaries between physics-driven and data-driven paradigms, and envision a future where long-term time-series prediction evolves from "black-box fitting" toward "explainable, generalizable, and mechanism-integrated" methodologies and technical pathways.

Rhythm-ONet: A query-conditioned neural operator for high-fidelity emulation of multiscale rhythm transitions in neurological disorder modeling

Silu Yan (Great Bay University)

Abstract: Mechanistic neural mass models (NMMs) are pivotal for digital twins in neurological disorder modeling, offering the potential for diagnosis and treatment of abnormal rhythms associated with Alzheimer' s disease (AD) and epilepsy. However, their solutions can change sharply with synaptic connectivity, making numerical parameter sweeps computationally expensive, while standard neural operator surrogates struggle with spectral bias in resolving these multi-frequency dynamics. We introduce Rhythm-ONet, a query-conditioned neural operator with a time-conditioned coefficient generator in the branch pathway. Inspired by attention mechanisms, the design dynamically reweights branch features according to query points, mitigating spectral bias in operator learning. Combined with a Fourier-embedded trunk, it enables high-fidelity reconstruction of multiscale NMM dynamics. On two canonical NMMs, including the Jansen-Rit system for Alzheimer's-like activity and the Wendling-Chauvel system for seizure-like dynamics, Rhythm-ONet achieves substantially lower relative L2 error than DeepONet, Fourier-DeepONet, and DeepONet-Bt, yielding superior time-frequency agreement across synaptic connectivity regimes. Once trained, the emulator enables rapid mapping of synaptic connectivity dependent rhythm transitions, including in silico signatures of rhythm slowing in AD and ictal-like discharges within these mechanistic models.

Numerically Efficient Methods for Solving High-dimensional PDE and Probabilistic Modeling

Sikun Yang (Great Bay University)

Abstract: This report presents numerically efficient methods for solving high-dimensional partial differential equations (PDEs) and their applications in probabilistic modeling, with a focus on score-based generative models. We discuss Cauchy Networks (CauchyNets), a novel single-layer feedforward neural architecture inspired by the Cauchy's integral theorem from complex analysis. The Cauchy activation function provides a principled functional basis for approximating analytic functions in high-dimensional spaces with theoretical approximation guarantees of order O((1/m)ˆ k) for any integer K. We integrate CauchyNets with the deep backward stochastic differential equation (BSDE) framework to solve semilinear parabolic PDEs, including Hamilton-Jacobi-Bellman equations arising from stochastic control problems. Numerical experiments demonstrate that a CauchyNet-BSDE solver with only 10 hidden units achieves competitive performance compared to traditional deep BSDE approaches requiring larger architectures. The methodology extends to score-based generative modeling through an important connection: the Fokker-Planck equation governing diffusion processes via Hopf-Cole transform to a Hamilton-Jacobi-Bellman equation, enabling PDE-based solutions for learning generative models. We parameterize the negative log-density using CauchyNets with time embedding, demonstrating effectiveness on high-dimensional Gaussian mixture densities. Finally, we address the computational challenge of sampling in score-based generative models by developing fast sampling strategies based on probability flow ODEs and fixed-point iteration. These approaches

reduce sampling steps by more than 50% (from 50 to 22-25 steps) in video generation models including Open-Sora 1.3 and 2.0, while maintaining generation quality.

Revisiting Smale's 18th Problem in the Age of Artificial Intelligence

Yuan Yao (The Hong Kong University of Science and Technology)

Abstract: In 1998, Steve Smale proposed his 18th mathematical problem for the 21st century, calling for a deeper understanding of the fundamental limitations of intelligence. Rapid advances in artificial intelligence over the past decade have lent renewed urgency to this challenge, underscoring the need for principled theoretical frameworks that characterize the capabilities and limitations of modern learning systems. The structural differences between human intelligence and contemporary artificial intelligence models shed light on central issues in the development of trustworthy AI, including robustness, interpretability, uncertainty quantification, and computational efficiency. In this talk, I examine these questions through through the lens of topology learning, dynamical systems, optimization algorithms, and statistical inference.

From Approximation to Disentanglement: Trainable Gaussian Mixture Modules for Modern Neural Architectures

Gangnan Yuan (Yunnan University)

Abstract: Neural networks in general, from MLPs and CNNs to attention-based Transformers, are constructed from layers of linear combinations followed by nonlinear operations such as ReLU, Sigmoid, or Softmax. Despite their strength, these conventional designs are often limited in introducing non-linearity by the choice of activation functions. In this work, we introduce Gaussian Mixture-Inspired Nonlinear Modules (GMNM), a new class of differentiable modules that draw on the universal density approximation Gaussian mixture models (GMMs) and distance properties (metric space) of Gaussian kernal.

Motion Behavior Control of Animal Robots

Nenggan Zheng (Zhejiang University)

Abstract: A direct connection between animal brains and computers can be established through implantable brain-computer interfaces, forming animal-based robots that integrate computer systems. The control of such robots' motor behaviors requires addressing challenges including neural modulation of animal brains, development of lightweight wearable brain-computer microsystems, and autonomous generation of motor control commands. Solving these issues may lead to the creation of animal robots with simple system structures, superior sensory-motor capabilities, low energy consumption, and excellent stealth performance. These robots are suitable for exploring confined spaces and can serve as research subjects and tools for fundamental scientific studies on sensory-motor neural mechanisms and swarm control. By investigating the

motor behavior control mechanisms of animal carriers and optimizing the composition of animal robot systems, we have achieved "wireless motor control of bearbee robots" and "drone navigation of rat robots."

Multiscale and multimodal fusion technology in biomedicine

Xubin Zheng (Great Bay University)

Abstract: Gliomas are tumors originating from glial cells in the brain, with glioblastoma (GBM) being the most aggressive type, exhibiting a 5-year mortality rate of 95%. Preoperative risk prediction for GBM aids in treatment decision-making. Clinical data for GBM involve multiple modalities and scales, including magnetic resonance imaging (millimeter-scale), cellular status (micrometer-scale), and genetic information (nanometer-scale). Integrating multi-modal and multi-scale data to achieve early non-invasive risk prediction for GBM remains an urgent challenge. Here, we present recent progress in the integration of multi-modal and multi-scale biomedical data. First, we integrated brain magnetic resonance imaging with gene expression using techniques such as prompt learning and image decoupling to construct an early non-invasive risk prediction model for GBM. However, the excessive granularity of genes resulted in a significant disparity between imaging and genetic scales. To address this, we proposed a bioinformatics neural network (BINN) to integrate functionally similar genes into gene functional clusters. Additionally, we employed capsule networks and transfer learning to represent single cells using gene functional clusters and to assist in disease diagnosis using human samples. Given that cells are also influenced by external environments, we developed a spatial multi-modal integration model, SMART, which achieves unified representation of multi-level omics and cellular environment, laying the foundation for future accurate risk stratification of GBM.

Weak Generative Sampler

Xiang Zhou (City University of Hong Kong)

Abstract: Sampling invariant distributions from an Ito diffusion process presents a significant challenge in stochastic simulation. The current deep learning-based method solves the stationary Fokker--Planck equation to determine the invariant probability density function in the form of deep neural networks, but they generally do not directly address the problem of sampling from the computed density function. In this work, we introduce a framework that employs a weak generative sampler (WGS) to directly generate independent and identically distributed (iid) samples induced by a transformation map derived from the stationary Fokker--Planck equation. Our proposed loss function is based on the weak form of the Fokker--Planck equation, integrating normalizing flows to characterize the invariant distribution and facilitate sample generation from a base distribution. Our randomized test function circumvents the need for min-max optimization in the traditional weak formulation. Our method necessitates neither the computationally intensive calculation of the Jacobian determinant nor the invertibility of the transformation map. A crucial component of our framework is the adaptively chosen family of test functions in the form of Gaussian kernel functions with centers related to the generated data samples.

Graph Hamiltonian network: A robust model for learning particle interactions in lattice systems

Jian Zu (Northeast Normal University)

Abstract: Addressing the challenges posed by nonlinear lattice models, which are vital across diverse scientific disciplines, in this talk, we introduce separable graph Hamiltonian networks (α-SGHN) that reveals complex interaction patterns between particles in lattice systems. Utilizing trajectory data, α-SGHN infers potential interactions without prior knowledge about particle coupling, overcoming the limitations of traditional graph neural networks that require predefined links. Furthermore, α-SGHN preserves all conservation laws during trajectory prediction. Experimental results demonstrate that α-SGHN, incorporating structural information, outperforms baseline models based on conventional neural networks in predicting lattice systems. This work is in collaboration with Ru Geng, Yixian Gao, Hong-Kun Zhang and Panayotis Kevrekidis.