Mathematical Foundations & Methodologies for Artificial Intelligence & Data-Driven Scientific Computing
Workshop on Mathematical Approaches to Modern AI Challenges
Abstracts

31 Aug 2026–4 Sep 2026

September 2, 2026

Abstracts

The Effects of Feature-Space Dimensionality on Activation Steering

Laziz Abdullaev

National University of Singapore, Singapore

We observe empirically that a sufficiently large activation dimension is essential for activation steering as it enables better separated representations of contrastive concepts inside the pretrained model. We then propose new steering methods that utilize high dimensional feature maps and the kernel trick that enables the steering function to advantage from higher order structural separations in contrastive concept representations. Our methods show improved steering performance on standard benchmarks.

Agentic Reinforcement Learning

Bo An

Nanyang Technological University, Singapore

LLM-based autonomous agents are playing an increasingly important role in many fields, including general computer control, software engineering, scientific discovery, and social simulation. Early implementations of autonomous agents largely relied on manually constructed workflows based on prompt engineering. However, such static workflows often struggle to generalize effectively to out-of-domain tasks, lack the ability to maintain high-quality, continuous interaction with environments and users, and are unable to achieve self-improvement during long-term operation. As a result, a growing body of research has begun to introduce reinforcement learning, enabling agents to continuously optimize their policies through interaction and thereby systematically improve their generalization, interaction, and tool-use capabilities. This talk will discuss major challenges encountered in agentic RL and some recent progress.

Challenges and Opportunities of Self-Consuming Loops in AI

Richard Baraniuk
Rice University, USA

There are many arguments for training deep learning models on synthetic data, including 1) with the advent of trillion parameter models, we are simply running out of training data; 2) synthetic training data is much cheaper to source than real data; 3) some important domains lack large training data sets. Moreover, even if you think you are training your model on clean, real data, it is likely that your dataset has been polluted inadvertently with synthetic data. Our preliminary research in this space has shown that synthetic data training creates a feedback loop that, over generations of models, can amplify artifacts, increase biases, and reduce the models' quality and diversity (aka, "model collapse" or "MADness"). The potential negative ramifications of model collapse reach across the entire spectrum of machine learning applications, from generative models to decision making system to signal and image processing systems. It is imperative that we design new machine learning models and training paradigms that are at least robust and at best immunized against synthetic data. In this talk, we show that this is indeed possible.

Real-time Digital Twins: Challenges and Opportunities for Scientific Machine Learning

Tan Bui-Thanh
The University of Texas at Austin, USA

Digital twins (DTs) are high-fidelity virtual representations of physical systems and processes. At their foundation lie mathematical and physical models that describe system behavior across multiple spatial and temporal scales. A central purpose of DTs is to enable "what-if" analyses through hypothetical simulations, supporting lifecycle monitoring, parameter calibration against observational data, and systematic uncertainty quantification (UQ). For DTs to serve as a reliable basis for real-time forecasting, optimization, and decision-making, they must reconcile two traditionally competing requirements: mathematical rigor and physical fidelity, and computational efficiency at scale. This has motivated a new generation of approaches that combine classical tools from numerical analysis, partial differential equations, inverse problems, and optimization with the expressive power of Scientific Machine Learning (SciML). In this talk, I will outline a principled pathway from traditional computational mathematics to rigorously grounded SciML. I will then present recent Scientific Deep Learning (SciDL) methods for forward modeling, inverse and calibration problems, and uncertainty quantification, emphasizing mathematical structure, stability, and generalization. Both theoretical results and numerical demonstrations will be shown for representative problems governed by transport, heat, Burgers, Euler (including transonic and hypersonic regimes), and Navier–Stokes equations.

Narrative Models for Interpretable Use of LLMs

Larry Carin

*A*STAR, Singapore*

Pending

Emulating and Calibrating Antarctic Ice-Sheet Dynamics with Diffusion Models and Deep Feature Matching

Won Chang

Seoul National University, S. Korea

The Antarctic ice sheet is a major source of uncertainty in future sea-level projections, and physical simulators such as the PSU-3D model are essential for studying its long-term evolution. Calibrating these models against observations is challenging: both the simulator outputs and the observed ice-thickness fields are high-dimensional, spatially dependent, and semi-continuous, with a large point mass at zero denoting ice-free regions. These features make conventional Gaussian-process emulation and likelihood-based calibration ill-suited and computationally infeasible at full spatial resolution. We propose a fully neural framework for emulation and calibration tailored to this setting. For emulation, we develop a conditional diffusion model that jointly generates the binary ice presence–absence pattern and the continuous thickness field within a single architecture, using complementary global and local conditioning mechanisms to represent how the input parameters shape the output across space. Because the emulator induces an intractable likelihood, we introduce a likelihood-free calibration scheme that replaces the hand-chosen distance and tolerance of approximate Bayesian computation with a probabilistic acceptance rule learned by an iteratively retrained Siamese network, together with a data–model discrepancy adjustment. Applied to the West Antarctic Ice Sheet, our approach matches the accuracy of state-of-the-art Gaussian-process calibration at a small fraction of its computational cost and scales to the full-resolution domain, where existing methods are intractable.

Learning Rates, Restart and Adaptive Acceleration

Alexandre d'Aspermont

Ecole Normale Supérieure, France

We study restart schemes in stochastic optimization problems where the function to minimize is weakly convex and satisfies a Kurdyka-Lojasiewicz inequality. We show that the combination of restarts and KL inequalities yields improved rates of convergence, with acceleration depending explicitly on the KL exponent. Furthermore, optimal restart schemes on SGD define an explicit learning rate schedule akin to Polyak steps. While none of the regularity constants are observed in practice, we prove that restart schemes tend to be robust to a significant misspecification of these constants. We detail numerical experiments on both toy problems where the KL exponent is controlled and LLMs.

Joint work with Ali Elhishi and Christophe Roux

On Preventing Model Collapse in Generative Models

Bahman Gharesifard

Queen's University, Canada

Recursive training on synthetic data can lead to model collapse, a degenerative feedback loop in which models progressively lose information about the underlying data distribution. Incorporating fresh human-generated data alongside synthetic data offers a natural way to mitigate this risk. In this talk, I use a simple generative model to establish theoretical guarantees for the minimum rate of human data needed to prevent model collapse under recursive training. In particular, I characterize the human-to-synthetic data ratio required to maintain training stability by leveraging the information-geometric structure of the probability simplex and analyzing the resulting dynamics under the Fisher-Rao metric.

Learning Structured Relevance Networks

Alfred Hero

University of Michigan, USA

Structured relevance networks are probabilistic graphical models optimized for quantifying statistical dependency patterns between multiple covariates. They have been incorporated into AI to reveal groups of influential variables and to identify sources of algorithmic bias. Certain types of structural constraints on relevance network inference can facilitate computation, reduce sample complexity, reflect physical generative mechanisms, or penalize undesirable properties of the model. We will describe how such structural constraints can be used for parsimonious relevance network representation in multiway (spatio-temporal) analysis and in clustering with algorithmic fairness guarantees.

Sobolev IPM with Graph Metric: Regularization, Scalability, and Applications

Tam Le

The Institute of Statistical Mathematics, Japan

We study the Sobolev IPM problem for probability measures supported on a graph metric space. Sobolev IPM is an important instance of integral probability metrics (IPM), and is obtained by constraining a critic function within a unit ball defined by the Sobolev norm. Particularly, it has been used to compare measures and is crucial for several theoretical works in machine learning. However, there are no efficient algorithmic approaches to compute Sobolev IPM effectively, which hinders its practical applications. We establish a relation between Sobolev norm and weighted Lp-norm, and leverage it to propose a novel regularization for Sobolev IPM. By exploiting the underlying graph structure, we demonstrate that the regularized Sobolev IPM provides a closed-form expression for fast computation. This advancement addresses long-standing computational challenges, and paves the way to apply Sobolev IPM for large-scale applications. Moreover, building upon the insights of tree systems, we further introduce tree-sliced Sobolev IPM (TS-Sobolev), a tree-sliced metric that aggregates regularized Sobolev IPMs over random tree systems for applications with measures on Euclidean spaces or hyperspheres. Notably, TS-Sobolev admits the tree-sliced Wasserstein (TSW) as its special case. We empirically evaluate the proposed approaches on various tasks, e.g., document classification and topological data analysis for graph-based measures; as well as gradient flows, self-supervised learning, generative modeling, and text topic modeling for measures on Euclidean spaces and hyperspheres.

Non-Hackable Confidence Reward Functions

Wee Sun Lee

National University of Singapore, Singapore

We consider the setting where large language models (LLMs) are trained using reinforcement learning (RL) to simultaneously improve reasoning accuracy and verbalize its confidence. Our reward scheme uses two functions for rewarding confidence verbalized by the LLM: one when the LLM is correct and a different one when the LLM is incorrect. With a poorly designed reward scheme, the LLM may be incentivized to answer incorrectly in order to maximize the expected reward, a phenomenon that we call confidence reward hacking. We demonstrate that selective confidence reward hacking can occur in practical datasets with reward schemes that are not designed to be non-hackable. We propose the concept of non-hackable confidence reward schemes and provide characterization for such schemes. Our characterization allows us to define a spectrum of such reward schemes for RL confidence calibration training in LLMs. When the approximator capacity is limited, different schemes allow different trade-offs between accuracy and calibration.

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.

Nipping the Butterfly Effect in the Bud: Self-Output Fine-Tuning for Autoregressive Weather Prediction

Hsuan-Tien Lin

National Taiwan University, Taipei

Long-horizon weather forecasting is a fundamental challenge in atmospheric science, for which autoregressive Deep Learning Weather Prediction (DLWP) has emerged as the primary paradigm. Although the autoregressive pipeline is highly scalable and flexible, its prediction errors grow rapidly over long forecasting horizons. In this work, we study this error growth phenomenon from both theoretical and empirical perspectives. Our analysis reveals that the growth is driven by a feedback loop between output errors and input distribution shifts. Specifically, the autoregressive process amplifies small initial output errors, which progressively corrupt subsequent input distributions, echoing the butterfly effect in atmospheric science and ultimately deteriorating forecasting accuracy over longer horizons. Furthermore, we show that this distributional shift originates at the earliest stage of inference, with out-of-distribution signatures detectable as early as the first autoregressive step. To mitigate this issue, we propose Self-Output Fine-Tuning (SOFT), a plug-and-play strategy that leverages the model's own one-step predictions to calibrate the biased input distribution encountered at the first step. Extensive experiments demonstrate that, despite its simplicity, SOFT achieves state-of-the-art performance on long-horizon forecasting tasks and substantially reduces both prediction errors and distributional discrepancy. The empirical success of SOFT highlights the importance of reexamining the fundamental pipeline of atmospheric science, representing a critical pipeline advance for atmospheric science.

Instances where Compositional Structure Helps Avoid the Curse of Dimensionality

Mauro Maggioni

Johns Hopkins University, USA

I will discuss instances of multiple problems where function composition can be exploited in statistical models in order to obtain efficient estimators, implemented by efficient algorithms, that avoid the curse of dimensionality:

1. a model of high-dimensional functions that depend only on the projection of the data on a low-dimensional manifold, for which we construct estimators that, under suitable assumptions, provably defeat the statistical and computational curse of dimensionality, and achieve nearly optimal learning rates.

2. the problem of estimating interaction laws in stochastic systems of interacting particles given trajectories, where again considering compositional structure imposed by modeling and symmetries allows one to provably avoid the curse of dimensionality.

3. an application of deep learning in the context of digital twins in cardiology, and in particular the use of operator learning architectures for predicting solutions of parametric PDEs, or functionals thereof, on a family of diffeomorphic domains — the patient-specific hearts – which we apply to the prediction of medically relevant electrophysiological features of heart digital twins.

4. recently-introduced Multiscale Markov Decision Processes (MDPs) in Reinforcement Learning, which enable the efficient solution of Markov Decision Processes (MDPs), where scales represent MDPs at different levels of abstraction, that also provide highly transferable skills to new MDPs (via compositionally).

How Many Different Outputs Can a Transformer Generate?

Maxime Meyer

National University of Singapore, Singapore

We show that a few architectural characteristics of a transformer are enough to closely predict how many different sentences it can generate, both qualitatively and quantitatively. We derive an upper bound that depends on the prompt length and show empirically that it is tight within a factor of 10 across architectures and model sizes. Our analysis also explains previously observed failures of transformers on simple sequence tasks such as copying and cramming. I will then discuss how the same framework extends to prompt tuning, information retrieval, and sliding-window attention.

Discriminants for Neural Network Verification

Guido Montufar
University of California, Los Angeles, USA

Neural network verification seeks to provide formal guarantees about the behavior of learned models. In this talk, I will present two recent projects on the mathematical foundations of neural network verification and the evaluation of verification systems. The first develops an algebraic framework based on metric algebraic geometry. We study the distance degree as a measure of the intrinsic complexity of the verification problem and introduce two discriminants, one associated with the input space and one with the parameter space. The second presents VeriStress-GT, a benchmark for neural network verification based on instances with known ground truth and controlled difficulty profiles for the systematic evaluation and comparison of verification algorithms. The first part is joint work with Yulia Alexandr and Hao Duan. The second is joint work with David Troxell, Yulia Alexandr, Sofia Hunt, and Stephanie Lei.

Numerical Approximation for McKean-Vlasov SDE with Super-linear Growth Coefficients

Hoang Long Ngo
Hanoi National University of Education, Vietnam

We consider the numerical approximation for McKean-Vlasov stochastic differential equations driven by Lévy processes. We propose a tamed-adaptive Euler-Maruyama scheme and consider its strong convergence in both finite and infinite time horizons when applied to certain classes of Lévy-driven McKean-Vlasov stochastic differential equations with non-globally Lipschitz-continuous, super-linear-growth drift and diffusion coefficients.

Missing Data Imputation using Neural Cellular Automata

Thanh Bình Nguyen

University of Science, Viet Nam National University Ho Chi Minh City, Vietnam

Missing data is a pervasive challenge when working with tabular data, and a wide range of imputation methods have been proposed, from classical discriminative approaches to modern deep generative models. However, existing methods typically rely on fixed relational structures, one-shot inference, or complex encoder–decoder architectures. In this paper, we propose Neural Imputation Cellular Automata (NICA), a novel imputation method inspired by Neural Cellular Automata (NCA). NICA iteratively grows missing values through residual updates, dynamically recomputing sample-wise similarities at every step to identify informative neighbors. This design yields a simple yet effective architecture that naturally captures evolving inter-sample relationships throughout the imputation process. Extensive experiments across diverse benchmark datasets demonstrate that NICA achieves the strongest overall performance among a wide range of baselines in terms of imputation error, downstream prediction performance, and robustness under different missing rates and missingness mechanisms.

Mixture of Experts Inside the Euler Equation

Tien Trinh Nguyen

Shanghai Jiao Tong University, China

The gating mechanism of a mixture of experts assigns each input a softmax-weighted combination of expert outputs, with an inverse-temperature parameter β controlling how decisively the router commits to a single expert. We run the same mechanism for a two-dimensional incompressible inviscid flow whose local swirl takes finitely many values, one per region. This regularizes the solution without mollification or diffusion, and preserves the clean interface structure of the flow at every $\beta > 0$.

Slowly Annealed Langevin Dynamics: Theory and Applications to Training-Free Guided Generation

Atsushi Nitanda

Nanyang Technological University, Singapore

We introduce Slowly Annealed Langevin Dynamics (SALD), a sampler for tracking a path of moving target distributions and approximating the terminal target through time slowdown. We establish non-asymptotic convergence guarantees via a KL differential inequality, showing that slowdown improves tracking through contraction of intermediate targets and the complexity of the path. Motivated by training-free guided generation with pretrained score-based generative models, we further introduce Velocity-Aware SALD (VA-SALD), which explicitly incorporates the underlying marginal distributions of the pretrained model and uses slowdown to correct the additional deviation induced by guidance. This yields a principled framework for training-free guided generation for diffusion-based and related generative model families, together with convergence guarantees that clarify the roles of intermediate functional inequalities and guidance bias.

AI meets Fourier: Towards Interpretability in Large Language Models

Kannan Ramchandran

University of California, Berkeley, USA

Interpreting the output produced by a Large Language Model (LLM) involves the complex question of attribution: which small set of interacting input components (features) primarily influences this output decision? In this talk, we describe our recent framework named SPEX for addressing this attribution challenge at much larger scale than possible with popular methods like Shapley values. SPEX exploits the underlying structure with its novel framing of masked attribution as learning a sparse spectral (Fourier) representation of the value function characterizing the model. This allows SPEX to leverage a rich set of foundational tools from signal processing and information/coding theory, which we will describe. We will overview an extension called ProxySPEX, which uses gradient boosted trees to better exploit a certain hierarchical structure revealed by SPEX.

Beyond Independence: Strongly Correlated Particle Systems in Sampling for ML

Hoang Son Tran

National University of Singapore, Singapore

The classical paradigm of randomness in the sciences is based on independent and identically distributed (i.i.d.) random variables. In this talk, we explore a complementary perspective, namely, tractable forms of dependence can enable major advances in ML applications. We will focus on determinantal point processes (DPPs), a class of strongly dependent particle systems originating in quantum physics that has emerged as a powerful toolbox for sampling.

A notable application is DPP-based minibatch sampling in stochastic gradient descent (SGD), where we will show how DPP minibatches can outperform i.i.d. minibatches while remaining computationally feasible. We will discuss how DPP samplers adapt to intrinsic structure and geometry latent in data, enabling a kernelized approach to sampling that provably outperforms independent samplers and achieves optimal guarantees in general settings. This builds upon connections to several classical ingredients, including orthogonal polynomials, wavelets, diffusions, and more.

References:

1. R. Bardenet, S. Ghosh, H. Simon-Onfroy, and H. S. Tran, “Small coresets via negative dependence: DPPs, linear statistics, and concentration.”

2. H. S. Tran, V. Petrovich, R. Bardenet, and S. Ghosh, “Negative Dependence as a toolbox for machine learning: review and new developments.”

3. H. S. Tran, P. Gupta, R. Bardenet, and S. Ghosh, “State-of-the-art minibatches via novel DPP kernels: discretization, wavelets, and rough objectives.”

4. H. S. Tran, P. Gupta, and S. Ghosh, “Fast determinantal sampling on general spaces and diffusion geometry.”

Homepage: `https://sites.google.com/view/hoangson-tran`

Symmetry in Deep Learning

Robin Walters

Northeastern University, USA

Symmetry is found across AI – in tasks, data, and models. I will discuss several approaches to leveraging various forms of invariance to improve model training and generalization. This includes networks with built in strict symmetry constraints, equivariant neural networks, but it also includes relaxed equivariance, symmetry discovery, and symmetry of model parameters. Lastly, I will discuss how understanding the degree of symmetry in the data distribution can be used to guide model selection and inform learning.

Impact of Data and Model Geometry on Learning Complexity in Deep Learning

Melanie Weber

Harvard University, USA

The utility of encoding geometric structure, such as known symmetries, into machine learning architectures has been demonstrated empirically in domains ranging from biology to computer vision.However, rigorous analysis of its impact on the learning complexity of neural networks remains largely unexplored. Recent advances in learning theory have shown that training shallow, fully connected neural networks, which are agnostic to data geometry, has exponential complexity in the statistical query model, a framework encompassing gradient descent. In this talk, we ask whether knowledge on data and model geometry is sufficient to alleviate the fundamental hardness of learning neural networks. We analyze two geometric settings: First, equivariant neural networks, a class of geometric architectures that explicitly encode symmetries; and second, learning neural networks under the manifold hypothesis, which assumes that high-dimensional input data lies on or near a low-dimensional manifold.
