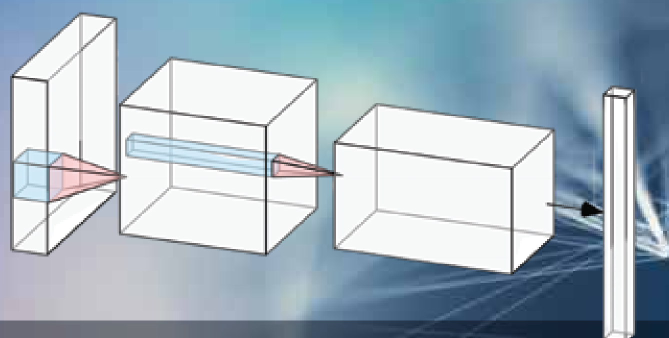
BRIN MATHEMATICS
RESEARCH CENTER

Statistical Theory of Deep Neural Network Models

November 7 - 9, 2024



About the Workshop

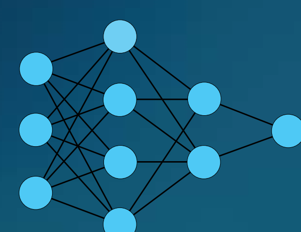
As deep learning has achieved breakthrough performance in a variety of application domains, a significant effort has been made to understand theoretical foundations of Deep Neural Network (DNN) models. Statisticians have devoted to understanding for example why deep neural networks models outperform classic nonparametric estimates and providing theoretical underpinning of DNN models from the lens of statistical theory. This workshop aims to bring together researchers in the field to discuss the pecent progress in statistical theory and foundations of DNN models, and chart possible future research directions.

Speakers

Arash Amini, *UCLA*
Peter Bartlett, *Berkeley*
Ismael Castillo, *Sorbonne*
Minwoo Chae, *Postech*
David Dunson, *Duke*
Jian Huang, *PolyU*
Yongdai Kim, *SNU*
Sophie Langer, *UT*
Wenjing Liao, *Georgia Tech*
Lu Lu, *Yale*
Rong Tang, *HKUST*
Matus Telgarsky, *NYU*
Jane-Ling Wang, *UC Davis*
Yixin Wang, *Michigan*
Yuting Wei, *U Penn*
Tong Zhang, *UIUC*
Yiqiao Zhong, *Wisconsin*

Organizers

Lizhen Lin, *Maryland*
Vince Lyzinski, *Maryland*
Yun Yang, *Maryland*



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

DEPARTMENT OF
MATHEMATICS

Contents

Schedule at a Glance

Workshop Overview

As deep learning has demonstrated breakthrough performance across various application domains, considerable efforts have been directed toward understanding the theoretical foundations of deep neural network (DNN) models. Statisticians, in particular, have focused on providing the statistical underpinnings of these models, investigating why DNNs outperform classical nonparametric estimators, and offering theoretical explanations for their practical success from the lens of statistical theory. This workshop aims to convene researchers in the field to discuss recent advances in the statistical theory and foundations of DNN models, and to identify and chart potential research directions.

Organizing committee

Lizhen Lin, University of Maryland

Vince Lyzinski, University of Maryland

Yun Yang, University of Maryland

Workshop Schedule

Thursday, November 7, 2024

8:30 - 9:00	Breakfast
9:00 - 9:10	Doron Levy (University of Maryland/Director, Brin MRC) *Opening*
9:10 - 10:00	Peter Bartlett (University of California, Berkeley) *Gradient Optimization Methods: Large Step-Size and the Edge of Stability*
10:00 - 10:50	Ismael Castillo (Sorbonne University) *Statistical Adaptation via Bayesian Deep GPs and Neural Networks*
10:50 - 11:20	Coffee Break
11:20 - 12:10	David Dunson (Duke University) *Bayesian Inferences in Interpretable Neural Network Models*
12:10 - 2:00	Lunch
2:00 - 2:50	Arash Amini (University of California, Los Angeles) *Polynomial Graph Neural Networks: Theoretical Limits and Graph Noise Impact*
2:50 - 3:40	Sophie Langer (University of Twente) *Image Classification - A Novel Statistical Approach*
3:40 - 4:10	Coffee Break
4:10 - 5:00	Jane-Ling Wang (University of California-Davis) *Hypothesis Testing for the Deep Cox Model*
6:30 - 8:30	Conference Dinner in North Bethesda

Friday, November 8, 2024

8:30 - 9:00	Breakfast
9:00 - 9:50	Yongdai Kim (Seoul National University) *ANOVA-NODE - An Identifiable Neural Network for Functional ANOVA*
9:50 - 10:40	Matus Telgarsky (New York University) *Logistic Regression with Arbitrary Step Sizes, and some Reflections on the Role of Theory*
10:40 - 11:10	Coffee Break
11:10 - 12:00	Lu Lu (Yale University) *Accurate, Efficient, and Reliable Learning of Deep Neural Operators for Multiphysics and Multiscale Problems*
12:00 - 1:30	Lunch
1:30 - 2:20	Jian Huang (The Hong Kong Polytechnic University) *Continuous Normalizing Flows for Learning Probability Distributions*
2:20 - 3:10	Minwoo Chae (Postech) *Nonparametric Structured Density Estimation Using Diffusion-Based Deep Generative Models*
3:10 - 3:40	Coffee Break
3:40 - 4:30	Rong Tang (Hong Kong University of Science and Technology) *Adaptivity of Diffusion Model to Manifold Structures*
4:30 - 5:20	Yuting Wei (University of Pennsylvania) *Towards Faster Non-Asymptotic Convergence for Diffusion-Based Generative Models*

Saturday, November 9, 2024

8:30 - 9:00	Breakfast
9:00 - 9:50	Yiqiao Zhong (University of Wisconsin) *Do LLMs Solve Novel Tasks? An Empirical Investigation of Out-Of-Distribution Generalization*
9:50 - 10:40	Yixin Wang (University of Michigan) *Towards Understanding Transformers and In-Context Learning: A Comparative Study with Word2Vec*
10:40 - 11:10	Coffee Break
11:10 - 12:00	Wenjing Liao (Georgia Tech) *Exploiting Low-Dimensional Data Structures and Estimating Scaling Laws for Transformer Neural Networks*
12:00 - 12:50	Tong Zhang (University of Illinois Urbana-Champaign) *Reinforcement Learning from Human Feedback: From Learning Theory to Reward Modeling and Policy Optimization*
12:50 - 1:00	Workshop Closing
1:00 - 2:00	Lunch

Abstracts of talks

Gradient Optimization Methods: Large Step-Size and the Edge of Stability

Peter Bartlett

University of California, Berkeley

Thursday, November 7, 2024 @ 9:10 AM

Optimization in deep learning relies on simple gradient descent algorithms. Although these methods are traditionally viewed as a time discretization of gradient flow, in practice, large step sizes—large enough to cause oscillation of the loss—exhibit performance advantages. This talk will review recent results on gradient descent with logistic loss with a step size large enough that the optimization trajectory is at the "edge of stability," and show the benefits of this initial oscillatory phase for linear functions and for two-layer networks.
Based on joint work with Yuhang Cai, Michael Lindsey, Song Mei, Matus Telgarsky, Jingfeng Wu and Bin Yu.

Statistical Adaptation via Bayesian Deep GPs and Neural Networks

Ismael Castillo

Sorbonne University

Thursday, November 7, 2024 @ 10:00 AM

Adaptation to structural parameters is an important topic in modern statistics. In regression models of possibly large input dimension, example of such parameters include smoothness of the regression function, intrinsic dimension of predictors and variables present or not in compositional structures. While from the Bayesian perspective adaptation can in principle be achieved by equipping each parameter, such as smoothness or the presence or not of a given variable, with a specific prior distribution, such as a gamma or a spike-and-slab prior respectively, implementing such hierarchies can be challenging computationally.

In this talk, we show that structural adaptation can be achieved through particularly simple deep learning priors, with posterior distributions that can be either simulated using MCMC or variational Bayes approximations.

A first contribution (joint work with Thibault Randrianarisoa, Toronto) concerns deep Gaussian processes, for which we show that a squared-exponential kernel combined with a well-chosen prior on lengthscale parameters leads to a (tempered) posterior distribution converging at near-optimal minimax rate, adaptive simultaneously in terms of smoothness and compositional structure. A second contribution (joint work with Paul Egels, Sorbonne University) concerns deep ReLU neural networks, for which we demonstrate that a specific class of heavy-tailed prior distributions placed on network weights leads to automatic and simultaneous adaptation to smoothness and compositional structure or intrinsic dimension of predictors.

Bayesian Inferences in Interpretable Neural Network Models

David Dunson

Duke University

Thursday, November 7, 2024 @ 11:20 AM

In applying deep neural networks (DNNs) in scientific applications, it is critical for the results to be interpretable; black box predictive algorithms are insufficient. In addition, obtaining valid uncertainty quantification is of utmost importance in avoiding over-interpretation of the results. With these goals in mind, this talk focuses on recent developments in identifiable and interpretable Bayesian inferences in neural networks. In particular, we highlight the Bayesian pyramids framework and its applications in ecology and biomedicine, and describe related ideas for replicated network data with applications to neuroimaging.

Polynomial Graph Neural Networks: Theoretical Limits and Graph Noise Impact

Arash Amini

University of California, Los Angeles

Thursday, November 7, 2024 @ 2:00 PM

This talk examines the theoretical foundations of Graph Neural Networks (GNNs), focusing on polynomial GNNs (Poly-GNNs). We start with empirical evidence challenging the need for complex GNN architectures in semi-supervised node classification, showing simpler methods often perform comparably.

We then analyze Poly-GNNs within a contextual stochastic block model, addressing a key question: Does increasing GNN depth improve class separation in node representations?

Our results show that for large graphs, the rate of class separation remains constant regardless of network depth. We demonstrate how "graph noise" can overpower other signals in deeper networks, negating the benefits of additional feature aggregation. The analysis also reveals differences in noise propagation between even and odd-layered GNNs, providing insights for network architecture design and highlighting trade-offs between model complexity and performance in graph-based machine learning.

Image Classification - A Novel Statistical Approach

Sophie Langer

University of Twente

Thursday, November 7, 2024 @ 2:50 PM

Convolutional neural networks (CNNs) excel in image recognition, showcasing remarkable performance in face recognition, medical diagnosis, and autonomous driving. However, their reliability remains a concern due to the lack of robust theoretical foundations. Establishing a solid statistical framework is essential before we can fully analyse CNNs from a theoretical perspective.

Current interpretations treat image classification as a high-dimensional classification problem with each pixel value being an independent variable. Function recovery in high dimensions lead to slow convergence rates, known as the curse of dimensionality. Consequently, CNN classifiers show worse performance with increased pixel count, contradicting empirical observations.

In this talk, I will present a novel statistical approach that reconceptualizes images not as high-dimensional entities but as highly structured objects. Within one class, objects arise from different geometric deformations, including shifts, scales, and orientations. The goal of the classification rule is then to learn the uninformative deformations, resulting in faster convergence rates for higher dimensions, i.e., a higher resolution of the image.

This new perspective not only provides novel approximation and convergence guarantees for deep learning-based image classification, but also redefines our perception of image analysis, bridging theory with practice.

Hypothesis Testing for the Deep Cox Model

Jane-Ling Wang

University of California-Davis

Thursday, November 7, 2024 @ 4:10 PM

Deep learning has become enormously popular in the analysis of complex data, including event time measurements with censoring. To date, deep survival methods have mainly focused on prediction. Such methods are scarcely used in matters of statistical inference such as hypothesis testing. Due to their black-box nature, deep-learned outcomes lack interpretability which limits their use for decision-making in biomedical applications. Moreover, conventional tests fail to produce reliable type I errors due to the ability of deep neural networks to learn the data structure under the null hypothesis even if they search over the full space. This paper provides testing methods for the nonparametric Cox model – a flexible family of models with a nonparametric link function to avoid model misspecification. Here we assume the nonparametric link function is modeled via a deep neural network.

To perform hypothesis testing, we utilize sample splitting and cross-fitting procedures to get neural network estimators and construct the test statistic.

These procedures enable us to propose a new significance test to examine the association of certain covariates with event times.

We show that our test statistic converges to a normal distribution under the null hypothesis and establish its consistency, in terms of the Type II error, under the alternative hypothesis. Numerical simulations and a real data application demonstrate the usefulness of the proposed test.

Based on joint work with Qixian Zhong and Jonas Mueller.

ANOVA-NODE - An Identifiable Neural Network for Functional ANOVA

Yongdai Kim

Seoul National University

Friday, November 8, 2024 @ 9:00 AM

Interpretability for machine learning models is becoming more and more important as machine learning models become more complex. The functional ANOVA model which decomposes a high-dimensional function into a sum of lower dimensional functions (e.g. components), is one of the most popular tools for interpretable AI, and recently, various neural network models have been developed for estimating each component in the functional ANOVA model. However, such neural networks are highly unstable in estimation of each component since the components themselves are not identifiable. That is, there are multiple functional ANOVA decompositions for a given function. In this paper, we propose a new neural network model which guarantees a unique functional ANOVA decomposition and thus is able to estimate each component stably. The main idea is to put a constraint, the so called "sum-to-zero" condition, which is a popularly used approach in statistics, on a given neural network model. A critical problem of this approach is that a standard gradient descent method cannot be applied to learn a neural network due to the identifiability condition. Our proposed neural network resolves this problem since it always satisfies the identifiability condition without any constraint on the parameters and thus standard gradient descent algorithms can be directly applied. Theoretically, we prove that our neural network is complete in the sense that it approximates any smooth function well. In addition, we discuss relations of our neural network with existing methods for interpretable AI including SHAP.

Logistic Regression with Arbitrary Step Sizes, and some Reflections on the Role of Theory

MATUS TELGARSKY

New York University

Friday, November 8, 2024 @ 9:50 AM

This talk will have a philosophical part, covering the role of theory in the face of LLMs, and a technical part, where I will share joint work describing a large step size analysis of logistic regression. In the first part, I will try to distill key discussions from the current semester at the Simons Institute (one on LLMs, one on generalization), where both the role of theory and the role of humanity have been raised. In the second part I will prove that logistic regression can use an arbitrarily large step size, loosely following joint work with Jingfeng Wu, Bin Yu, and Peter Bartlett (https://arxiv.org/abs/2402.15926), however presenting a modified proof via direct analogy to the classical perceptron method.

Accurate, Efficient, and Reliable Learning of Deep Neural Operators for Multiphysics and Multiscale Problems

LU LU

Yale University

Friday, November 8, 2024 @ 11:10 AM

It is widely known that neural networks (NNs) are universal approximators of functions. However, a less known but powerful result is that a NN can accurately approximate any nonlinear operator. This universal approximation theorem of operators is suggestive of the potential of deep neural networks (DNNs) in learning operators of complex systems. In this talk, I will present the deep operator network (DeepONet) to learn various operators that represent deterministic and stochastic differential equations. I will also present several extensions of DeepONet, such as DeepM&Mnet for multiphysics problems, DeepONet with proper orthogonal decomposition or Fourier decoder layers, MIONet for multiple-input operators, and multifidelity DeepONet. I will demonstrate the effectiveness of DeepONet and its extensions to diverse multiphysics and multiscale problems, such as bubble growth dynamics, high-speed boundary layers, electroconvection, hypersonics, geological carbon sequestration, and full waveform inversion. Deep learning models are usually limited to interpolation scenarios, and I will quantify the extrapolation complexity and develop a complete workflow to address the challenge of extrapolation for deep neural operators.

Continuous Normalizing Flows for Learning Probability Distributions

Jian Huang

The Hong Kong Polytechnic University

Friday, November 8, 2024 @ 1:30 PM

Continuous normalizing flows (CNFs) are a generative method based on ordinary differential equations for learning probability distributions. This method has shown success in applications like image synthesis, protein structure prediction, and molecule generation. We present the CNF method and study its theoretical properties using a flow matching objective function. We establish non-asymptotic error bounds for the distribution estimator based on CNFs, in terms of the Wasserstein-2 distance, under the assumption that the target distribution has bounded support, is strongly log-concave, or is a mixture of Gaussian distributions. Our convergence analysis addresses errors due to velocity estimation, discretization, and early stopping. We also develop uniform error bounds with Lipschitz regularity control for deep ReLU networks approximating the Lipschitz function class. Our analysis provides theoretical guarantees for using CNFs to learn probability distributions from finite random samples.

Nonparametric Structured Density Estimation Using Diffusion-Based Deep Generative Models

Minwoo Chae

Postech

Friday, November 8, 2024 @ 2:20 PM

In recent years, diffusion-based deep generative models have achieved remarkable success in various applications. In this talk, we present statistical theories for diffusion models within the framework of nonparametric structured density estimation. To address the curse of dimensionality in nonparametric density estimation, we assume that the underlying density function factorizes into several low-dimensional components. Such factorizable densities are common in important examples, such as Bayesian networks and Markov random fields. We prove that an implicit density estimator constructed from diffusion models achieves the minimax optimal convergence rate with respect to total variation. Technically, we design a novel network architecture, which includes convolutional neural networks as a special case, to construct a minimax optimal estimator.

Adaptivity of Diffusion Model to Manifold Structures

Rong Tang

Hong Kong University of Science and Technology

Friday, November 8, 2024 @ 3:40 PM

Empirical studies have demonstrated the effectiveness of (score-based) diffusion models in generating high-dimensional data, such as texts and images, which typically exhibit a low-dimensional manifold nature. These empirical successes raise the theoretical question of whether score-based diffusion models can optimally adapt to low-dimensional manifold structures. We show that the forward-backward diffusion can adapt to the intrinsic manifold structure by showing that the convergence rate of the inducing distribution estimator depends only on the intrinsic dimension of the data, and demonstrate that the forward-backward diffusion can achieve the minimax optimal rate under the Wasserstein metric. We also extend our analysis by considering a class of conditional forward-backward diffusion models for conditional generative modeling, that is, generating new data given a covariate (or control variable). We extend our theory by allowing both the data and the covariate variable to potentially admit a low-dimensional manifold structure. In this scenario, we demonstrate that the conditional forward-backward diffusion model can adapt to both manifold structures, meaning that the derived estimation error bound (under the Wasserstein metric) depends only on the intrinsic dimensionalities of the data and the covariate.

Towards Faster Non-Asymptotic Convergence for Diffusion-Based Generative Models

Yuting Wei

University of Pennsylvania

Friday, November 8, 2024 @ 4:30 PM

Diffusion models, which convert noise into new data instances by learning to reverse a Markov diffusion process, have become a cornerstone in contemporary generative artificial intelligence. While their practical power has now been widely recognized, the theoretical underpinnings remain far from mature. In this work, we develop a suite of non-asymptotic theory towards understanding the data generation process of diffusion models in discrete time, assuming access to ℓ_2-accurate estimates of the (Stein) score functions. For a popular deterministic sampler (based on the probability flow ODE), we establish a convergence rate proportional to $1/T$ (with T the total number of steps), improving upon past results; for another mainstream stochastic sampler (i.e., a type of the denoising diffusion probabilistic model), we derive a convergence rate proportional to $1/\sqrt{T}$, matching the state-of-the-art theory. Imposing only minimal assumptions on the target data distribution (e.g., no smoothness assumption is imposed), our results characterize how ℓ_2 score estimation errors affect the quality of the data generation processes. Further, we design two accelerated variants, improving the convergence to $1/T^2$ for the ODE-based sampler and $1/T$ for the DDPM-type sampler, which might be of independent theoretical and empirical interest.

Do LLMs Solve Novel Tasks? An Empirical Investigation of Out-Of-Distribution Generalization

Yiqiao Zhong

University of Wisconsin

Saturday, November 9, 2024 @ 9:00 AM

Large language models (LLMs) such as GPT-4 sometimes appeared to be creative, solving novel tasks with a few demonstrations in the prompt. These tasks require the pre-trained models to generalize on distributions different from those from training data—which is known as out-of-distribution (OOD) generalization. For example, in symbolized language reasoning where names/labels are replaced by arbitrary symbols, yet the model can infer the names/labels without any finetuning.

In this talk, I will offer some new angles for understanding the emergent phenomena in LLMs, which hopefully provide empirical foundations for statistical theory for LLMs. By focusing on induction heads, which are a type of pervasive components within LLMs, I will show that learning the right compositional structure is a key to OOD generalization, and this learning process exhibits sharp transitions in training dynamics. Further, I propose the "common bridge representation hypothesis" as a compositional mechanism in Transformers, where a latent subspace in the embedding space acts as a bridge to align multiple attention heads across early and later layers.

Towards Understanding Transformers and In-Context Learning: A Comparative Study with Word2Vec

Yixin Wang

University of Michigan

Saturday, November 9, 2024 @ 9:50 AM

Large language models (LLMs) based on transformer architectures have demonstrated remarkable capabilities in natural language processing, including impressive in-context learning (ICL) abilities. However, the statistical underpinnings of these models and how these capabilities emerge from unsupervised training on unstructured text data remain unclear. Are these capabilities unique to transformers, or do they share principles with earlier generations of language models? This talk explores these questions through comparing transformers with classical language models like word2vec.

We begin by examining the core attention mechanism in transformers, showing its surprising similarity to the continuous bag of words (CBOW) architecture in word2vec. Specifically, we prove that fitting a single-layer, single-head bidirectional attention model is equivalent to fitting a CBOW model with mixture-of-experts (MoE) weights. Further, bidirectional attention with multiple heads and multiple layers is equivalent to stacked MoEs and a mixture of MoEs, respectively. These connections reveal the distinct use of MoE in bidirectional attention, explaining its effectiveness with heterogeneous data. We also leverage this connection to adapt transformers for categorical tabular data, which enables improved out-of-distribution generalization performance over existing tabular transformers.

Building on this foundation, we next investigate how the ICL capability of transformers emerges from unsupervised learning with unstructured data, again highlighting its tight connections to CBOW. We theoretically prove (and empirically validate) that ICL can arise simply by modeling co-occurrence information using classical language models like CBOW. Additionally, we demonstrate that positional encodings and certain noise structures are crucial for generalization to unseen data. Taken together, these findings underscore the close relationships between transformers and classical language models like CBOW, offering new perspectives on how ICL emerges from training on unstructured data

[1] https://arxiv.org/abs/2307.04057
[2] https://arxiv.org/abs/2406.00131

Exploiting Low-Dimensional Data Structures and Estimating Scaling Laws for Transformer Neural Networks

Wenjing Liao

Georgia Tech

Saturday, November 9, 2024 @ 11:10 AM

When training deep neural networks, a modelís generalization error is often observed to follow a power scaling law dependent on the model size and the data size. Perhaps the best-known example of such scaling laws is for transformer-based large language models (LLMs), where networks with billions of parameters are trained on trillions of tokens of text. One theoretical interest in LLMs is to understand why transformer scaling laws exist. To answer this question, we exploit low-dimensional structures in language datasets by estimating its intrinsic dimension, and establish statistical estimation and mathematical approximation theories for transformers to predict the scaling laws. By leveraging low-dimensional data structures, we can explain transformer scaling laws in a way which respects the data geometry. Furthermore, we test our theory with empirical observations by training LLMs on natural language datasets, and find strong agreement between the observed empirical scaling laws and our theoretical predictions.
This is a joint work with Alex Havrilla.

Reinforcement Learning from Human Feedback: From Learning Theory to Reward Modeling and Policy Optimization

Tong Zhang

University of Illinois Urbana-Champaign

Saturday, November 9, 2024 @ 12:00 PM

Reinforcement Learning from Human Feedback (RLHF) has emerged as a key method for aligning large language models with human preferences. In this talk, we will explore the fundamental formulation of RLHF as an entropy-regularized reinforcement learning problem.

From a theoretical perspective, we present new statistical convergence results for entropy-regularized RL, highlighting its advantages over the traditional RL formulation in terms of both sample and exploration complexity. In practice, RLHF often relies on data presented in pairs, annotated with human preference labels. A widely used approach is Direct Preference Optimization (DPO), which directly learns the language model from these preference pairs. However, we demonstrate the limitations of DPOís implicit reward structure and emphasize the importance of a dedicated reward modeling process.

We further show that incorporating a learned reward function and employing an iterative refinement process can significantly enhance DPO performance. The statistical theory and algorithmic techniques discussed are applicable to complex reasoning tasks, such as mathematical problem-solving.

The Brin Mathematics Research Center

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

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

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

List of Participants

Arash Amini, University of California, Los Angeles
Peter Bartlett, University of California, Berkeley
Ismael Castillo, Sorbonne University
Minwoo Chae, Postech
David Dunson, Duke University
Francesco Gafi, University of Maryland
Subhashis Ghosal, North Carolina State University
Jian Huang, The Hong Kong Polytechnic University
Yongdai Kim, Seoul National University
Kunwoong Kim, Seoul National University
Shivam Kumar, University of Notre Dame
Hyeok Kyu Kwon, Postech
Sophie Langer, University of Twente
Jeyong Lee, Postech
Kyeongwon Lee, University of Maryland
Doron Levy, University of Maryland/Director, Brin MRC
Wenjing Liao, Georgia Tech
Lizhen Lin, University of Maryland
Lu Lu, Yale University
Vince Lyzinski, University of Maryland
Seokhun Park, Seoul National University
Rong Tang, Hong Kong University of Science and Technology
Matus Telgarsky, New York University
Jane-Ling Wang, University of California-Davis
Yixin Wang, University of Michigan
Yuting Wei, University of Pennsylvania
Yun Yang, University of Maryland
Tong Zhang, University of Illinois Urbana-Champaign
Yiqiao Zhong, University of Wisconsin