Guide to

Faculty Research Interests

Class of 2022 Senior Thesis and Junior Independent Work

PRINCETON
School of Engineering and Applied Sciences
Research Interests

Financial market microstructure, high frequency trading, high frequency data analysis, optimal trade execution, and machine learning methods for all of the above.
Research Interests

- **Optimization**: algebraic methods in optimization, semidefinite programming, polynomial optimization.

- **Computational aspects of dynamics and control**: Optimization-based Lyapunov theory for verification of dynamical systems.

- **Control-oriented learning**: Learning dynamical systems from trajectory observations subject to side information.

- **Algorithms and complexity**: Computational complexity in numerical optimization, convex relaxations in combinatorial optimization.

- I am also interested in applications of these tools to semialgebraic problems in systems theory, machine learning, robotics, and economics.
Research Interests

Stochastic analysis, stochastic control and stochastic games, especially mean field games, reinforcement learning, even if it has to be a machine. High Frequency markets, environmental finance and energy and commodity markets.
**On Leave Spring 2022**
Professor Matias Cattaneo
Room 227, Sherrerd Hall - Ext. 8825
cattaneo@princeton.edu

**Research Interests**

Econometric theory and mathematical statistics; program evaluation and treatment effects; machine learning, nonparametric and semiparametric methods; high-dimensional inference; applications to social and behavioral sciences.
Research Interests

Finance, machine learning, statistics, portfolio choices, financial risks, computational biology, among others.

Junior Independent Work/Senior Thesis Topics

1. Statistical data analysis
2. Finance (Portfolio selection, asset pricing, financial risks, high-frequency trading)
3. Machine learning (deep learning, networks, topic modeling)
4. Biostatistics and genomics
Professor Boris Hanin  
bhanin@princeton.edu

Research Interests

Theory of Deep Learning: Neural networks are machine learning models that have achieved state of the art in a variety of practical tasks. I am interested in elucidating the principles by which they work, understanding how to make them better, and extracting the general lessons they hold for statistics and machine learning. Here are two big questions I'm excited about:

1. At finite but large depth and width, what is the statistical behavior of neural networks at initialization? This concerns inductive bias, aligning priors to datasets, and setting hyperparameters such as architecture, learning rate, non-linearity, batch size, etc in a principled way.

2. What can we say about the generalization ability of neural networks? What is the effect of the data distribution, initialization and optimization on the quality of loss minimizer we find?

In addition to such probabilistic/statistical questions, I am also fascinated by understanding the roles of data augmentation and large vs. small learning rates. The questions can involve a range of tools from random matrix theory to combinatorics, the renormalization group, and high dimensional probability.

Semiclassic Analysis: Bohr's correspondence principle says that in the limit of large quantum numbers, quantum mechanics should begin to resemble classical mechanics. For example, by studying a quantum mechanical system in the limit where Planck's constant $h$ tends to zero (the so-called semiclassical limit), the high energy behavior of quantum states should resemble the long-time behavior of the underlying classical system. I am interested in various questions that try to make such claims precise. Usually, they involve studying the zero set and/or density of states for randomized wavefunctions at high frequency/energy. The tools involved range from Riemannian geometry to Gaussian processes, geometric measure theory, spectral theory for self-adjoint operators, and special functions.

Random Polynomials: The location and behavior of the zeros of polynomials are a classical topic in math and science. It is a famous theorem that for polynomials of degree five and higher in one variable there is no formula for their zeros as a function of their coefficients. Nonetheless, a lot can be said about the zeros of a polynomial when it is chosen at random. For example, a result due to Gauss is that if $p_n$ is a degree $n$ polynomial of one complex variable, then its critical points (solutions to $p_n'(z)=0$) are inside the convex hull of its zeros. It turns out that much more is true with high probability if $p_n$ is chosen at random. Namely, most of its zeros have a unique nearby critical point. Amazingly, this rather surprising fact was not known until a few years ago, and much remains to be understood! The tools here involve probability and complex analysis.
Research Interests

I spent most of my professional life as a finance industry practitioner, including many years as a Goldman Sachs partner. Currently, my industry work focuses on start-up companies, where I invest, advise and serve on boards of technology and finance companies. I have developed an ORFE elective course on FinTech in consumer lending, where companies using novel data and ML to increase fairness and inclusion must navigate well-established competition and substantial regulation.

While a good number of my thesis advisees have pursued topics related to industry sectors I’ve worked in, e.g. FinTech, markets, investment banking and investing, I’ve also very much enjoyed learning about new areas and topics as an advisor, e.g. online review systems and e-sports. As an advisor, I encourage students to explore how problems can be approached from different angles, thinking widely about relevant data sets, computational methods and mathematical modeling approaches.

Sample of advisee topics:

- **Twitter Sentiment Analysis with Interaction Meta-Data**, exploring diverse ML ‘sentiment models’ and overlying interactions (likes/favorites) to distinguish trends. (Motivated by the student’s interests.)
- **Information Cascades and Herding in Online Reviews**, with work spanning data analysis of Yelp reviews, simulation and mathematical theory. (Motivated by the student’s active interest in online review systems, as an avid user.)
- **Wisdom of Crowds vs. Identifying Experts: Extracting Information from Industry Surveys**, which focused on application of ML methods to a proprietary dataset and included theory related to wisdom of crowds and detecting expert from crowds. (Motivated by the student’s summer internship work.)
- **Earnings Calls, Does Anything Really Happen?**, applied natural language processing methods to company earnings calls to identify information content relative to other financial and markets data. (Motivated by a popular article that focused on humor.)
- **Looking through the GlassDoor**, included a cross-sectional analysis of reviews to identify real outliers using NLP and other ML techniques. (Motivated by the student’s interest in start-ups and how employee satisfaction can help VC’s direct capital. I helped source data from an alternative data startup ThinkNum.)
- **Who gets the money? Analysis of ETF launches that succeeded, or failed, to grow their assets**, considering breadth of data related to fund attributes and performance. (Motivated by student’s curiosity about the high growth in this sector and many, many fund launches.)
- **Optimal Team Choice for competitive electronic gaming using ML methods** (Motivated and shaped by the student’s hobby interests.)
- **Beyond Prediction: Bayesian vs. Penalized Regression**, an algorithm rather than application-focused senior project (motivated by the student’s interest in this class of problems.)
**Research Interests:**

The research in Applied Mathematics I have been conducting with my co-authors during the past few years, has been motivated so far by two main words: interactions and incentives. These two terms have to be related in the context of my work to the Economics field, and in particular when focusing on the behaviour of economic actors. The term interaction thus refer to interaction between consumers, workers, more generally called Agents, through their decisions, their actions, the price they pay for a service, or even through an external parameter. When looking at interactions, the aim is generally to find an equilibrium in this game between the Agents, so that no one has an interest in deviating. This so-called Nash equilibrium can be extended in the case of a continuum of Agents, and then reaches a more recent theory borrowed from Physics, the Mean-Field Games. The term incentives also has an economic content, and refer to the Contract Theory, and more precisely to Principal-Agents problems. In this case, an Agent (he) is delegated to act on behalf of a Principal (she). The Principal's purpose is to find appropriate incentives, in the form of a contract, to encourage the Agent to act in her interest. When considering these issues from the mathematical point of view, the behaviour of Agents, in an uncertain environment and in continuous-time, can be modelled as a stochastic control problem. My research so far is thus a continuous path oscillating between Principal-Agent problems, Nash equilibria and Mean-Field Games, using recently introduced and state of the art tools in stochastic control. We study applications to Energy, Epidemiology and Finance.
Research Interests:

I am broadly interested in statistical and algorithmic aspects of various problems that arise in machine learning. Specifically, my research seeks to describe the tension between flexibility, computability, and accuracy. Some specific questions that interest me include: How do overparameterized neural networks generalize well to unseen data? What are the approximation properties of neural networks when the dimensionality is large? Can decision trees be used to identify important variables in a predictive model? Is it possible to precisely characterize the tradeoff between type I and type II errors for variable selection in a high dimensional linear model?
Research Interests

Development and application of operations research and other analytical techniques in various aspects of Autonomous Vehicles, aka "SmartDrivingCars", including

- the fundamental design of computer vision techniques for the rapid classification and identification of the driving environment, especially “deep learning convolutional neural networks”,

- analysis and classification of collision-free driving scenarios,

- quantification of accident risk and the investigation, formulation and design of "pay-as-you-drive, pay-as-the-car-drives" insurance,

- Investigation and creative design of the human-computer interfaces for SmartDrivingCars

- operational and feasibility analyses of autonomousTaxi (aTaxi) systems

Junior Independent Work/Senior Thesis Topics

Any of the above
Research Interests

The applications and theory for the efficient design of resource sharing services and systems. Motivating examples include communication networks and healthcare systems. Research interests include queueing theory with time varying rates, stochastic networks, dynamical systems and their optimal control, Monte-Carlo simulation, and time-inhomogeneous Markov processes.
Research Interests

Large-scale stochastic optimization models, algorithms, and applications, especially financial planning and wealth management. Multi-period financial planning applications for large insurance companies, hedge funds, global FinTech firms, and individuals. Apply novel methods in machine learning to financial planning systems. Combining advanced mathematical financial models with deep neural networks to address transaction costs and overcome the exponential growth in model size.

Conducting research with Ant Financial (Alibaba) on enterprise risk management for a global FinTech firm, and several large banks in San Francisco and New York City, and a multi-manager hedge fund in Austin Texas.

Junior Independent Work/Senior Thesis Topics


2. Modeling alternative asset categories. Over the past two decades, institutional investors have shifted capital to the so-called alternative asset categories, including private equity, hedge funds, leveraged loans, real assets, and others. These securities are generally hybrid in nature, with several embedded risk factors. We explore these topics, employing advanced topics in machine learning and stochastic optimization.

3. Model comparison. All too often, operations researchers and decision analysts believe that their models are pure representations of the "scientific method.” Yet two modelers will likely employ two different models with varying recommendations when confronted with a decision problem. These differences are often due to behavioral bias and value considerations. Discover a real-world application and study its ramifications.

4. Educational aspects of financial planning. Most individuals are woefully under prepared for making significant financial decisions. There are many behavioral biases and unfortunate habits that lead to under performance. These issues are becoming more significant with the rise in automated financial planning systems (robo-advisors).
Professor Miklos Racz  
Room 204, Sherrerd Hall - Ext. 8281  
mracz@princeton

Research Interests:

My research focuses on probability, statistics, and its applications. I am interested in statistical inference problems in complex systems, in particular on random graphs and in genomics. I am also interested more broadly in applied probability, combinatorial statistics, information theory, control theory, interacting particle systems, and voting.
Research Interests

My research is in high-dimensional probability. It involves interactions with random matrix theory, numerical linear algebra, and mathematical data science. An overarching theme of my research is to study the structure of large high-dimensional objects in the presence of randomness and use this understanding to develop randomized algorithms that efficiently process complex data.

For example, my recent and current directions include
1. Non-asymptotic random matrix theory (a variety of theoretical math questions as well as some applications, e.g., to signal processing on graphs),
2. Low-rank matrix and tensor methods (factorization, dimension reduction, and recovery methods) and their applications to interpretable machine learning (dynamic topic modeling and beyond),
3. Randomized iterative methods for solving linear systems (e.g., robust methods for corrupted systems) and more general optimization problems.
Professor Mykhaylo Shkolnikov  
Room 202, Sherrerd Hall - Ext. 1044  
mykhaylo@princeton

Research Interests

1. Stochastic portfolio theory
2. Optimal Investment
3. Stochastic analysis
4. Interacting particle systems
5. Random matrix theory
6. Random Growth Models
Research Interests

Financial Mathematics & Engineering; stochastic models, especially for market volatility; optimal investment and hedging strategies; analysis of financial data; credit risk; dynamic game theory and oligopoly models; energy and commodities markets; reliability of the electricity grid under increased use of solar and wind technologies.
Research Interests

Mathematical theory of optimal control and decisions under uncertainty, and applications of stochastic optimization techniques in economics, financial economics and quantitative finance,
And high-dimensional computational problems. Current topics are:

2. McKean-Vlasov optimal control problems.
3. Model-independent and robust finance. In particular, pricing and hedging of complex financial instruments using market data with little or no model assumptions.
5. Hedging in markets with frictions such as illiquidity or transaction costs.
Research Interests:

Development of theory and data-driven computational tools for mathematical optimization, machine learning and optimal control. Applications include control of fast dynamical systems, transportation, finance, robotics, and autonomous vehicles.

Mathematical optimization
- Convex large-scale and embedded optimization.
- First-order methods for sparse, low-rank, and combinatorial optimization.
- Differentiable optimization in deep learning, i.e., optimization problems as layers in neural networks.

Machine learning for optimization
- Learning heuristics to accelerate combinatorial optimization algorithms.
- Learning solutions of optimization problems with varying data.

Control systems
- Learning control policies for continuous and hybrid systems.
- High-speed online optimization for real-time control systems.
Research Interests

My research interests are in stochastic calculus and mathematical finance. More specifically, my work in stochastic analysis aims at developing the understanding of stochastic control and stochastic differential games using probabilistic arguments. Such control and differential games arise in a variety of optimal decision applications, including optimal investment, economics, engineering or biology. Some of the main issues here concern the analysis of optimal decision policies in diffusion models, the study of representations that enable their efficient numerical computations, and similar questions for games with a large number of players.

In mathematical finance, I focus on developing the theory of quantitative risk management. Here, I put a particular emphasis on computational aspects, and develop tools needed to achieve efficient computation. In fact, financial agents (especially banks and insurance companies) are eager to evaluate the riskiness of their decision and often evaluate associated risks periodically. Therefore, it is essential to understand computational issues: When using a particular model, how to select it? Can we use the specificities of a particular model to improve the estimation? How do we account for model uncertainty? It is possible to build model-free methods? Or statistical estimations entirely data-driven? Can we derive theoretical guarantees and convergence rates? This line of work includes (and is not limited to) elements of probability theory, stochastic calculus, as well a numerical and data analysis.
Research Interests

Interior-point methods for linear and nonlinear optimization, semi-definite programming, robust optimization, probabilistic potential theory, optimization models for high contrast imaging, the search for exosolar planets, stable periodic solutions to the n-body problem, and the stability of Saturn-like ring systems.

Junior Independent Work/Senior Thesis Topics

1. Trajectory optimization problems to find orbits in celestial mechanics.
2. Fast Fourier Optimization with applications in optics, acoustics, digital filtering, etc.
3. Parametric simplex method with applications in image analysis.
4. Gerrymandering
5. Climate Change
Research Interests

My primary interests lie in several areas of (mostly pure) mathematics: probability theory, analysis, geometry, and their interactions. I am particularly fascinated by the development of principles and methods that explain the common structure in a variety of pure and applied mathematical problems.