Application to Prob-AI Hub - Discipline Hopping Awards

Joseph Najnudel (PI) and Farhad Babaee (Co-PI)

University of Bristol, UK

25 July 2025

Our background

- My research covers a broad range of topics in probability theory, with a focus on stochastic analysis, probabilistic number theory and random matrix theory.
- I have authored more than 40 research articles, some published in top-tier mathematical journals such as Inventiones Mathematicae and the Duke Mathematical Journal.
- In collaboration with my former PhD student, we authored three papers on the loss surfaces of neural networks, using techniques from random matrix theory.

Our background

- My research covers a broad range of topics in probability theory, with a focus on stochastic analysis, probabilistic number theory and random matrix theory.
- I have authored more than 40 research articles, some published in top-tier mathematical journals such as Inventiones Mathematicae and the Duke Mathematical Journal.
- In collaboration with my former PhD student, we authored three papers on the loss surfaces of neural networks, using techniques from random matrix theory.

Our background

- My research covers a broad range of topics in probability theory, with a focus on stochastic analysis, probabilistic number theory and random matrix theory.
- I have authored more than 40 research articles, some published in top-tier mathematical journals such as Inventiones Mathematicae and the Duke Mathematical Journal.
- In collaboration with my former PhD student, we authored three papers on the loss surfaces of neural networks, using techniques from random matrix theory.

- Farhad Babaee is an algebraic geometer, and has recently worked on dynamical method to study tropicalisation, i.e. the degeneration of algebraic varieties into tropical varieties.
- Farhad has collaborations with Prof. June Huh (Fields Medal, 2020) and Prof. Karim Adiprasito (EMS Prize 2020), where they have tackled long-standing problems in complex dynamics.

- Farhad Babaee is an algebraic geometer, and has recently worked on dynamical method to study tropicalisation, i.e. the degeneration of algebraic varieties into tropical varieties.
- Farhad has collaborations with Prof. June Huh (Fields Medal, 2020) and Prof. Karim Adiprasito (EMS Prize 2020), where they have tackled long-standing problems in complex dynamics.

Our team

- Farhad and I have been colleagues since 2018, and regularly discuss mathematical questions, but didn't previously work on a shared project, due to our distinct research areas.
- The interest in ReLU activation functions in neural networks provides a natural connection with tropical geometry.
- On the other hand, the study of neural networks can be enriched by introducing random matrix theory methods in this setting, in particular, in order to model situations where noisy data is involved.

Our team

- Farhad and I have been colleagues since 2018, and regularly discuss mathematical questions, but didn't previously work on a shared project, due to our distinct research areas.
- The interest in ReLU activation functions in neural networks provides a natural connection with tropical geometry.
- On the other hand, the study of neural networks can be enriched by introducing random matrix theory methods in this setting, in particular, in order to model situations where noisy data is involved.

Our team

- Farhad and I have been colleagues since 2018, and regularly discuss mathematical questions, but didn't previously work on a shared project, due to our distinct research areas.
- The interest in ReLU activation functions in neural networks provides a natural connection with tropical geometry.
- On the other hand, the study of neural networks can be enriched by introducing random matrix theory methods in this setting, in particular, in order to model situations where noisy data is involved.

- This is why our collaboration presents a unique and timely opportunity. By combining our complementary expertise, we are well positioned to explore connections between the two points of view above.
- Our projects can be the starting point for what we envision as a broad and impactful area of research into the mathematical foundations of AI.

- This is why our collaboration presents a unique and timely opportunity. By combining our complementary expertise, we are well positioned to explore connections between the two points of view above.
- Our projects can be the starting point for what we envision as a broad and impactful area of research into the mathematical foundations of AI.

Overview of the project

- Our projects explore the intersection of algebraic geometry, probability theory, and AI, with a focus on developing new mathematical tools to analyse and improve neural network architectures.
- We plan to connect these architectures to tropical geometry, a variant of algebraic geometry in which addition and multiplication are replaced by maximum and addition, respectively. This framework is directly related to ReLU activation functions.
- Further, spectral properties of neural networks can be analysed using random matrix theory and free probability theory.

Overview of the project

- Our projects explore the intersection of algebraic geometry, probability theory, and AI, with a focus on developing new mathematical tools to analyse and improve neural network architectures.
- We plan to connect these architectures to tropical geometry, a variant of algebraic geometry in which addition and multiplication are replaced by maximum and addition, respectively. This framework is directly related to ReLU activation functions.
- Further, spectral properties of neural networks can be analysed using random matrix theory and free probability theory.

Overview of the project

- Our projects explore the intersection of algebraic geometry, probability theory, and AI, with a focus on developing new mathematical tools to analyse and improve neural network architectures.
- We plan to connect these architectures to tropical geometry, a variant of algebraic geometry in which addition and multiplication are replaced by maximum and addition, respectively. This framework is directly related to ReLU activation functions.
- Further, spectral properties of neural networks can be analysed using random matrix theory and free probability theory.

Understanding the Matrix Factorization Theorem

- We would like to investigate the algebraic and tropical versions of matrix factorisation and Mercer's theorem, and study the relations with the ReLU networks.
- Moreover, the properties of the eigenvalues provided by Mercer's theorem can also be better understood by comparing them to the spectrum of operators given by random kernels.
- Random matrix theory can be used to see how the eigenvalues of the matrix of data are perturbed by possible noise.

Understanding the Matrix Factorization Theorem

- We would like to investigate the algebraic and tropical versions of matrix factorisation and Mercer's theorem, and study the relations with the ReLU networks.
- Moreover, the properties of the eigenvalues provided by Mercer's theorem can also be better understood by comparing them to the spectrum of operators given by random kernels.
- Random matrix theory can be used to see how the eigenvalues of the matrix of data are perturbed by possible noise.

Understanding the Matrix Factorization Theorem

- We would like to investigate the algebraic and tropical versions of matrix factorisation and Mercer's theorem, and study the relations with the ReLU networks.
- Moreover, the properties of the eigenvalues provided by Mercer's theorem can also be better understood by comparing them to the spectrum of operators given by random kernels.
- Random matrix theory can be used to see how the eigenvalues of the matrix of data are perturbed by possible noise.

Algebraic Geometric Approach on the Manifold Hypothesis and Algebraic PCA

- Linear PCA can be viewed as the following minimization problem for the given data points x in $\mathbb{R}^n$ averaging to the origin:

$$\min_{L \in \mathrm{Gr}(n,k)} ||x - P_L(x)||_2^2$$

where $\mathrm{Gr}(n, k)$ is the Grassmannian of all linear subspaces of dimension k in $\mathbb{R}^n$, and P_L is the orthogonal projection onto L.

- This formulation can be generalized to algebraic varieties as follows:

$$\min_{V \subset V(I)} ||x - P_V(x)||_2^2$$

where I is an ideal in $\mathbb{C}[z_1, \ldots, z_n]$, and $V(I)$ denotes the algebraic variety defined by I.

- We aim to consider various alternatives for defining the projection function based on the generators of the ideals of algebraic varieties.

- This formulation can be generalized to algebraic varieties as follows:

$$\min_{V \subset V(I)} ||x - P_V(x)||_2^2$$

  where I is an ideal in $\mathbb{C}[z_1, \ldots, z_n]$, and $V(I)$ denotes the algebraic variety defined by I.
- We aim to consider various alternatives for defining the projection function based on the generators of the ideals of algebraic varieties.

Neural Networks and Tropical Geometry: An Adapted Gradient Descent

- We would like to analyse an adapted gradient descent method that aligns with the combinatorial structures inherent to tropical geometry.
- We also plan to study how these combinatorial structures interact when randomness is introduced, for example by considering the stochastic gradient descent method.
- We plan to study of the distribution of critical points of loss functions of smoothed ReLU network, which is related to random matrix theory.

Neural Networks and Tropical Geometry: An Adapted Gradient Descent

- We would like to analyse an adapted gradient descent method that aligns with the combinatorial structures inherent to tropical geometry.
- We also plan to study how these combinatorial structures interact when randomness is introduced, for example by considering the stochastic gradient descent method.
- We plan to study of the distribution of critical points of loss functions of smoothed ReLU network, which is related to random matrix theory.

Neural Networks and Tropical Geometry: An Adapted Gradient Descent

- We would like to analyse an adapted gradient descent method that aligns with the combinatorial structures inherent to tropical geometry.
- We also plan to study how these combinatorial structures interact when randomness is introduced, for example by considering the stochastic gradient descent method.
- We plan to study of the distribution of critical points of loss functions of smoothed ReLU network, which is related to random matrix theory.

Restricted Boltzmann machines

- We plan to study applications of restricted Boltzmann machines for learning probability distributions, which are models inspired by statistical physics.
- Such models have been used in estimating loss of portfolios in mathematical finance. We plan to extend the application to more general random data.

Restricted Boltzmann machines

- We plan to study applications of restricted Boltzmann machines for learning probability distributions, which are models inspired by statistical physics.
- Such models have been used in estimating loss of portfolios in mathematical finance. We plan to extend the application to more general random data.

Overfitting in over-parametrised situation

- Insights from the previous projects might help us to understand why overfitting does not sometimes occur in overparametrised architectures.
- In the analysis, it is crucial to distinguish which part of the data should be explained by the relevant parameters and which part corresponds only to perturbation or noise.
- For data in a high-dimensional space, in view of the manifold hypothesis, the relevant part is expected to lie on a submanifold of much lower dimension, whereas the noise is not expected to exhibit such structure.
- We plan to develop tools to detect when a large set of points lies close to a lower-dimensional manifold.

Overfitting in over-parametrised situation

- Insights from the previous projects might help us to understand why overfitting does not sometimes occur in overparametrised architectures.
- In the analysis, it is crucial to distinguish which part of the data should be explained by the relevant parameters and which part corresponds only to perturbation or noise.
- For data in a high-dimensional space, in view of the manifold hypothesis, the relevant part is expected to lie on a submanifold of much lower dimension, whereas the noise is not expected to exhibit such structure.
- We plan to develop tools to detect when a large set of points lies close to a lower-dimensional manifold.

Overfitting in over-parametrised situation

- Insights from the previous projects might help us to understand why overfitting does not sometimes occur in overparametrised architectures.
- In the analysis, it is crucial to distinguish which part of the data should be explained by the relevant parameters and which part corresponds only to perturbation or noise.
- For data in a high-dimensional space, in view of the manifold hypothesis, the relevant part is expected to lie on a submanifold of much lower dimension, whereas the noise is not expected to exhibit such structure.
- We plan to develop tools to detect when a large set of points lies close to a lower-dimensional manifold.

Overfitting in over-parametrised situation

- Insights from the previous projects might help us to understand why overfitting does not sometimes occur in overparametrised architectures.
- In the analysis, it is crucial to distinguish which part of the data should be explained by the relevant parameters and which part corresponds only to perturbation or noise.
- For data in a high-dimensional space, in view of the manifold hypothesis, the relevant part is expected to lie on a submanifold of much lower dimension, whereas the noise is not expected to exhibit such structure.
- We plan to develop tools to detect when a large set of points lies close to a lower-dimensional manifold.

Impact

- Our two first projects can be seen as promising methods for generalising PCA in cases where less linearity is involved.
- Since PCA is a fundamental tool in data analysis, any generalisation can have a considerable impact, and we hope to be able to publish our results in a world-leading journal, which in turn, will add to the toolbox for the data analysts and practitioners in the area.

Impact

- Our two first projects can be seen as promising methods for generalising PCA in cases where less linearity is involved.
- Since PCA is a fundamental tool in data analysis, any generalisation can have a considerable impact, and we hope to be able to publish our results in a world-leading journal, which in turn, will add to the toolbox for the data analysts and practitioners in the area.

Thank you for your attention!