

Our background

Our team

Overview of the project

Understanding matrix factorisation and Mercer's theorem

Algebraic Geometric Approach on the Manifold Hypothesis and Algebraic PCA

Neural Networks and Tropical Geometry: An Adapted Gradient Descent

Restricted Boltzmann machines

Overfitting in over-parametrised situation

Impact

# 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 and his collaborators have been able to tackle long-standing problems in complex dynamical systems using tropical methods.

- ▶ 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 and his collaborators have been able to tackle long-standing problems in complex dynamical systems using tropical methods.

# 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.

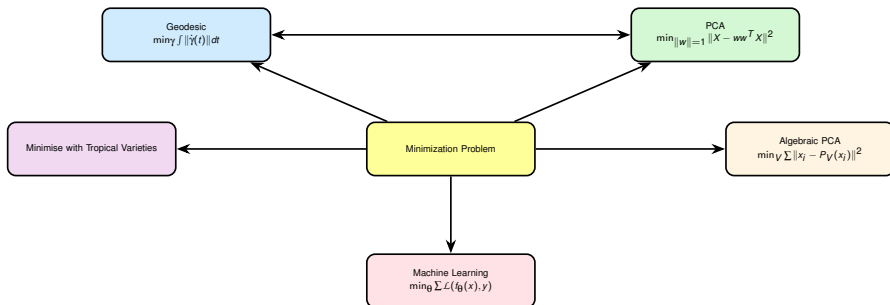
# Overview of the project

- ▶ 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 different minimization problems encountered in machine learning, as shown below.

# 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 different minimization problems encountered in machine learning, as shown below.

# Minimization Problems Across Fields



# Understanding matrix factorisation and Mercer's theorem

- ▶ We would like to investigate tropical versions of 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, which itself is connected to random matrix theory.

# Understanding matrix factorisation and Mercer's theorem

- ▶ We would like to investigate tropical versions of 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, which itself is connected to random matrix theory.



# 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:

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

where  $\text{Gr}(n,k)$  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:

$$\min_{V \in \mathcal{V}} \|x - P_V(x)\|_2^2$$

where  $\mathcal{V}$  is a family of algebraic varieties, and  $P_V$  is a projection function, to be defined, and which vanishes of  $V$ .

# 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.

# 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.

# 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 be published in world-leading journals and add to the toolbox for practitioners.

# 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 be published in world-leading journals and add to the toolbox for practitioners.

Our background

Our team

Overview of the project

Understanding matrix factorisation and Mercer's theorem

Algebraic Geometric Approach on the Manifold Hypothesis and Algebraic PCA

Neural Networks and Tropical Geometry: An Adapted Gradient Descent

Restricted Boltzmann machines

Overfitting in over-parametrised situation

Impact

# Thank you for your attention!