Post-doc proposition
ANITI, MADLADS CHAIR

1 Context

The postdoc candidate is applying to work with the following team of researchers

- CHHAIBI Reda, Professor at Université Côté d'Azur,
- GAMBOA Fabrice, Professor at Université Toulouse III,
- LAGNOUX Agnès, Associate professor at Université Toulouse II,
- PELLEGRINI Clément, Associate professor at Université of Toulouse III.

The team specializes in the mathematics of artificial intelligence, and the focus on the project at hand is the use of innovative tools such as Random Matrix Theory.

We are interested in both proving new theorems in the field, for publication in optimization, probability or statistics journals, in practical implementations, for publication in AI conference proceedings.

If interested, send your application to clement.pellegrini@univ-toulouse.fr and reda.chhaibi@univ-cotedazur.fr along with

- a curriculum vitae.
- a motivation letter.
- at least one recommendation letter, directly sent by the supporter to the aforementioned address.

Deadline for application is the 29th of May 2025.

Position begins on 1st of September 2025 (2 years).

2 Some relevant literature on RMT and AI

RMT is a branch of mathematics that studies the statistical properties of large matrices with random entries. It has found applications in a wide range of fields, including physics, finance, and statistics.

Relevance for high dimensional statistics. In the context of covariance matrices, RMT can help us better understand the behavior of these matrices when the dimensionality of the data is high while the number of samples is small. One key result from RMT is the Marchenko-Pastur (MP) theorem, which describes the eigenvalue distribution of large covariance matrices with random entries. In particular, the MP theorem shows that when the dimensionality of the data is much larger than the number of samples, the eigenvalue spectrum of the covariance matrix converges to a limiting distribution that depends only on the ratio of the dimensionality to the number of samples. This limiting distribution is the MP distribution, which has a well-defined shape and can be used to estimate the noise level in the eigenvalues of the covariance matrix.

Relevant applications and their state of the art.

First, in the line of Pennington et al.'s work [8], RMT is a natural tool in order to control the spectra of the Jacobian of Neural Networks at the initialization. This approach is a more rich and quantitative version than the seminal paper by Glorot et al. [7] which was mainly based on variance considerations. In [6], we consider a computational solution to the metamodel built from RMT. We use a homotopy method based on the chaining of basins of attraction for the Newton-Raphson algorithm. Not only is the result guaranteed to be correct – unlike the solution previously proposed by Pennington et al., but the method is also very fast.

Second, it is well-known that PCA (Principal Component Analysis) is the initial step of many statistical methods, either for the purpose of dimensionality reduction, or for the detection of a signal of small rank. And PCA is about diagonalizing large covariance matrices, hence the relevance of RMT. In fact, there is a flurry of theoretical works analyzing "signal plus noise" models, where one is interested in the spectrum of a large matrix deformed by some noise. Existing mathematical results are under quite general hypotheses, and describe refined phenomenons such as the celebrated BBP (Baik-Benarous-Péché) phase transition [1]. In comparison, there are only a handful of papers giving practical computational solutions which leverage RMT to estimate the spectra of large covariance matrices. We can mention El Karoui's work [4] as a prime example.

Finally, in the context of kernel learning, it is now common to use random features whose behavior is controlled thanks to RMT. The most known example is certainly that of Random Fourier Features which is accelerated by performing products of structured matrices as in [5].

3 Research topics

Following the three paragraphs describing each a different topic, the candidate will join projects in

- the use of Free Probability Theory for studying Jacobians of Neural Networks. Basically, we shall aim for the natural follow-up to the paper [6];

- the problem of Free Deconvolution and cleaning large covariance matrices in regards to [2, 4];
- the use of RMT when generating random features for kernel learning as in [3, 5].

References

[1] Baik, Jinho, Gérard Ben Arous, and Sandrine Péché. "Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices." (2005): 1643-1697.

[2] Bun, Joël, Jean-Philippe Bouchaud, and Marc Potters. "Cleaning large correlation matrices: tools from random matrix theory." Physics Reports 666 (2017): 1-109.

[3] Demni, Nizar, and Hachem Kadri. "Orthogonal Random Features: Explicit Forms and Sharp Inequalities." arXiv preprint arXiv:2310.07370 (2023).

[4] El Karoui N. Spectrum estimation for large dimensional covariance matrices using random matrix theory (2008): 2757-2790.

[5] Le, Quoc, Tamás Sarlós, and Alex Smola. "Fastfood-approximating kernel expansions in loglinear time." Proceedings of the international conference on machine learning. Vol. 85. No. 8. 2013.

[6] Chhaibi, Reda, Tariq Daouda, and Ezechiel Kahn. "Free Probability for predicting the performance of feed-forward fully connected neural networks." Advances in Neural Information Processing Systems 35 (2022): 2439-2450.

[7] Glorot, Xavier, and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks." Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010.

[8] Pennington J., Schoenholz S., and Ganguli S. The emergence of spectral universality in deep networks. In International Conference on Artificial Intelligence and Statistics (2018): 1924–1932.