

## DEPARTMENT OF STATISTICS

**Mohamed Ndaoud****Department of Decision Sciences  
ESSEC Business School, Paris, France***Minimax Supervised Clustering in the Anisotropic  
Gaussian Mixture Model: A new take on Robust  
Interpolation***Wednesday, November 3, 2021  
11:45 AM EST****Zoom Meeting: Meeting ID: 924 8283 7055  
Password: 378160**

<https://rutgers.zoom.us/j/92482837055?pwd=TktKVVU2Z1JzZnZQbm1TNUU3K0pmZz09>

**Abstract:** We study the supervised clustering problem under the two-component anisotropic Gaussian mixture model in high dimensions and in the non-asymptotic setting. We first derive a lower and a matching upper bound for the minimax risk of clustering in this framework. We also show that in the high-dimensional regime, the linear discriminant analysis (LDA) classifier turns out to be sub-optimal in the minimax sense. Next, we characterize precisely the risk of  $\ell_2$  regularized supervised least squares classifiers. We deduce the fact that the interpolating solution may outperform regularized classifiers, under mild assumptions on the covariance structure of the noise. Our analysis also shows that interpolation can be robust to corruption in the covariance of the noise when the signal is aligned with the "clean" part of the covariance, for the properly defined notion of alignment. To the best of our knowledge, this peculiar phenomenon has not yet been investigated in the rapidly growing literature related to interpolation. We conclude that interpolation is not only benign but can also be optimal, and in some cases robust.

This is joint work with Stanislav Minsker and Yiqiu Shen.

**Bio:** Mohamed Ndaoud is an Assistant Professor in the Department of Decision Sciences at ESSEC Business School in Paris. His research interests include variable selection, community detection and robust statistics in high dimensional settings. He received his Ph.D. degree in statistics from CREST-ENSAE Paris in 2019 under the supervision of Prof. Alexandre Tsybakov. He has also spent two years as an instructor in the Mathematics Department at University of Southern California.

