Abstract: We consider the online version of the well-known Principal Component Analysis (PCA) problem. Let \( X \in \mathbb{R}^p \) be a random vector with zero mean and unknown covariance matrix \( \Sigma \). In the online setting, we receive i.i.d. copies of \( X \) one at a time and our goal is to update our estimate of the top eigenvector of \( \Sigma \) with each new data point. Krasulina (1969) suggested an online estimation procedure and proved its consistency in the standard setting where \( p \) is fixed and \( n \to \infty \). In this work, we derive non-asymptotic estimation bound for Krasulina’s method that holds valid in the high-dimensional setting \( p \gg n \).

Joint work with Jiangning Chen (Georgia Inst. of Technology)

Bio: Karim Lounici is a Professor of Statistics at Université de Nice-Sophia Antipolis. His main field of research concerns the estimation of various types of high-dimensional matrices with specific low complexity structures. His most recent work is about the theoretical analysis of standard PCA for high-dimensional data.