Speaker: Inbar Seroussi

Title: **Stochastic Optimization in High Dimensional Structured Data**

Abstract: Stochastic optimization methods are fundamental in modern machine learning. However, understanding why these methods perform so well remains a major challenge. In this talk, I will present a theory for stochastic gradient descent (SGD) and its adaptive variant in high dimensions when the number of samples and problem dimensions are large. I will show that the dynamics of SGD applied to generalized linear models and multi-index problems with data possessing a general covariance matrix become deterministic in the large sample and dimensional limit. In particular, the limiting dynamics are governed by a set of low-dimensional ordinary differential equations (ODEs). Our setup encompasses many optimization problems, including linear regression, logistic regression, and two-layer neural networks as well as adaptive step size. In addition, it unveils the implicit bias inherent in SGD. For each of these problems, the deterministic equivalent of SGD enables us to derive a close approximation of the generalization error (with explicit and vanishing error bounds). Furthermore, we leverage our theorem to establish explicit conditions on the step size, ensuring the convergence and stability of SGD and a wide class of adaptive stochastic method. This is joint work with Elizabeth-Collins-Woodfin, Courtney Paquette, and Elliot Paquette, for more information see

https://arxiv.org/abs/2308.08977