

# Department Lecture Series



## Stochastic Gradient and Mirror Descent: Minimax Optimality and Implicit Regularization

**Dr. Babak Hassibi**

**Mose and Lillian S. Bohn Professor  
Dept. of Electrical Engineering  
California Institute of Technology**

**Thursday, October 11th**

**4:30 pm 6142 Scott Hall**

**ECE Seminar Committee**

Aswin Sankaranarayanan  
[saswin@ece.cmu.edu](mailto:saswin@ece.cmu.edu)

Maysam Chamanzar  
[mchamanz@andrew.cmu.edu](mailto:mchamanz@andrew.cmu.edu)

Swarun Kumar  
[swarun@cmu.edu](mailto:swarun@cmu.edu)

Giulia Fanti  
[gfanti@andrew.cmu.edu](mailto:gfanti@andrew.cmu.edu)

### **Abstract:**

Stochastic descent methods (of the gradient and mirror varieties) have become increasingly popular in optimization. In fact, it is now widely recognized that the success of deep learning is not only due to the special deep architecture of the models, but also due to the behavior of the stochastic descent methods used, which play a key role in reaching "good" solutions that generalize well to unseen data. In an attempt to shed some light on why this is the case, we revisit some minimax properties of stochastic gradient descent (SGD)---originally developed for linear models in the context of H-infinity control in the 1990's---and extend them to general stochastic mirror descent (SMD) algorithms for general nonlinear models. These minimax properties can be used to explain the convergence and implicit-regularization behavior of the algorithms when the linear regression problem is over-parametrized (In what is now being called the "interpolating regime"). In the nonlinear setting, exemplified by training a deep neural network, we show that when the setup is "highly over-parametrized", stochastic descent methods have similar minimax optimality and implicit-regularization properties. This observation gives some insight into why deep networks exhibit such powerful generalization abilities. We shall also make some connections to online learning.

### **Bio:**

Babak Hassibi is the inaugural Mose and Lillian S. Bohn Professor of Electrical Engineering at the California Institute of Technology, where he has been since 2001, From 2011 to 2016 he was the Gordon M Binder/Amgen Professor of Electrical Engineering and during 2008-2015 he was Executive Officer of Electrical Engineering, as well as Associate Director of Information Science and Technology. Prior to Caltech, he was a Member of the Technical Staff in the Mathematical Sciences Research Center at Bell Laboratories, Murray Hill, NJ. He obtained his PhD degree from Stanford University in 1996 and his BS degree from the University of Tehran in 1989. His research interests span various aspects of information theory, communications, signal processing, control and machine learning. He is an ISI highly cited author in Computer Science and, among other awards, is the recipient of the US Presidential Early Career Award for Scientists and Engineers (PECASE) and the David and Lucille Packard Fellowship in Science and Engineering.

**(RECEPTION FOLLOWING: SCOTT HALL ATRIUM)**