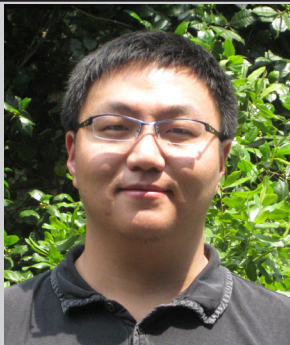


# Computer Science Seminar

## Local Dynamics: A Diffusion Approximation Framework on Nonconvex Stochastic Gradient Descent

Dr. Junchi Li, Princeton University



In this talk, I will develop a general framework for analyzing nonconvex statistical optimization problem using the stochastic gradient descent method. Using the theory of Morse's and Itô diffusion, we prove that under the presence of omnidirectional noise, the SGD iteration escapes a local minimum (resp. saddle point) in exponential time steps (resp. linear time steps) with respect to the inverse stepsize. We take deep neural network as an example. Our theory substantiate the numerical observations by a recent paper (Keskar et al., 2016), suggesting that large batch sizes for synchronous optimization leads to poor generalization error.

**Bio:** Dr. Junchi Li obtained his B.S. in Mathematics and Applied Mathematics at Peking University in 2009, and his PhD in Mathematics at Duke University under Professor Rick Durrett in 2014. He is currently working as a postdoctoral research associate at Department of Operations Research and Financial Engineering, Princeton University, with Professor Han Liu and Professor Tong Zhang. His research interests include statistical machine learning and optimization, stochastic algorithms for big data analytics, and stochastic dynamics on graphs and complex networks.

**Date: April 21, 2017**

**Time: 1:00 pm**

**216 Computer Science Building**

MISSOURI  
S&T