

**Speaker: Yuchen WU**

**Title: Posterior Sampling from the Spiked Models via Diffusion Processes**

**Abstract:** Sampling from the posterior is a key technical problem in Bayesian statistics. Rigorous guarantees are difficult to obtain for Markov Chain Monte Carlo algorithms of common use. In this paper, we study an alternative class of algorithms based on diffusion processes. The diffusion is constructed in such a way that, at its final time, it approximates the target posterior distribution. The stochastic differential equation that defines this process is discretized (using a Euler scheme) to provide an efficient sampling algorithm. Our construction of the diffusion is based on the notion of observation process and the related idea of stochastic localization. Namely, the diffusion process describes a sample that is conditioned on increasing information. An overlapping family of processes was derived in the machine learning literature via time reversal. We apply this method to posterior sampling in the high-dimensional symmetric spiked model. We observe a rank-one matrix  $\theta\theta^T$  corrupted by Gaussian noise, and want to sample  $\theta$  from the posterior. Our sampling algorithm makes use of an oracle that computes the posterior expectation of  $\theta$  given the data and the additional observation process. We provide an efficient implementation of this oracle using approximate message passing. We thus develop the first sampling algorithm for this problem with approximation guarantees.