



# Optimization and Machine Learning Seminar

## Non-convex Bayesian Learning via Stochastic Gradient Markov Chain Monte Carlo

**Dr. Wei Deng**

Machine Learning Researcher at Morgan Stanley

### Abstract

The training of modern deep neural networks (DNNs) boils down to a non-convex Bayesian learning problem. A standard tool to handle the problem is Langevin Monte Carlo, which, however, can be arbitrarily slow and often fails to explore the multi-modal posterior given a limited time.

As a result, advanced techniques are still required.

In this talk, we start with the replica exchange Langevin Monte Carlo (also known as parallel tempering), which is a Markov jump process that proposes appropriate swaps between exploration and exploitation to achieve accelerations. However, the naive extension of swaps to big data problems leads to a large bias, and bias-corrected swaps are required. Such a mechanism leads to few effective swaps and insignificant accelerations. To alleviate this issue, we first propose a control variates method to reduce the variance of noisy energy estimators and show a potential to accelerate the exponential convergence. We also present the population-chain replica exchange and propose a generalized deterministic even-odd scheme to track the non-reversibility and obtain an optimal round trip rate.

Date:	February 18, 2022 (Friday)
Time:	10:00 – 11:00am (Hong Kong Time)
Venue:	ZOOM: <a href="https://hku.zoom.us/j/">https://hku.zoom.us/j/</a>
	Meeting ID: 940 0962 9889
	Password: 286660

*All are welcome*