

CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - October 16, 2020

**ACCELERATING CONVERGENCE OF REPLICA EXCHANGE
STOCHASTIC GRADIENT MCMC VIA VARIANCE REDUCTION**

Wei Deng, Purdue University

Replica exchange stochastic gradient Langevin dynamics (reSGLD) has shown promise in accelerating the convergence in non-convex learning; however, an excessively large correction for avoiding biases from noisy energy estimators has limited the potential of the acceleration. To address this issue, we study the variance reduction for noisy energy estimators, which promotes much more effective swaps. Theoretically, we provide a non-asymptotic analysis on the exponential acceleration for the underlying continuous-time Markov jump process; moreover, we consider a generalized Girsanov theorem which includes the change of Poisson measure to overcome the crude discretization based on the Grownall's inequality " and yields a much tighter error in the 2-Wasserstein (W_2) distance. Numerically, we conduct extensive experiments and obtain the state-of-the-art results in optimization and uncertainty estimates for synthetic experiments and image data