# Two Stochastic Approximation Frameworks with Applications in Control and Reinforcement Learning

Sihan Zeng
Georgia Tech

**Abstract**: We propose two stochastic approximation (SA) framework, characterize their finite-time convergence, and discuss their applications to control and reinforcement learning.

First, we look at a decentralized variant of stochastic approximation over a network of agents, where the goal is to find the root of the aggregate of the local operators at the agents. In this method, each agent implements a local stochastic approximation using noisy samples from its operator while averaging its iterates with the ones received from its neighbors. Our model for the data observed at each agent is that it is sampled from a Markov process. Under mild assumptions we show that the convergence rate of the proposed method is essentially the same as if the samples were independent, differing only by a log factor that represents the mixing time of the Markov process. We finally present applications of the framework in decentralized robust system identification and decentralized Q-learning for solving multi-task reinforcement learning.

Second, we study a novel two-time-scale stochastic gradient method for solving optimization problems where the gradient samples are generated from a time-varying Markov random process parameterized by the underlying optimization variable. We consider a two-time-scale update scheme, where one scale is used to estimate the true gradient from the Markovian samples and the other scale is used to update the decision variable with the estimated gradient. While these two iterates are implemented simultaneously, the former is updated "faster" (using bigger step sizes) than the latter (using smaller step sizes). We derive the finite-time complexity of the method under different objective functions, namely, strong convexity, convexity, non-convexity under the PL condition, and general non-convexity. We apply this framework to study the performance of the online actor-critic methods in solving the linear-quadratic regulator and the standard MDP with finite state and action space.

**Acknowledgement:** This work was supported by ARL DCIST CRA W911NF-17-2-0181.