# Continuous-time MCMC and super-efficient Monte Carlo for big data*

**Paul Fearnhead**

## Abstract

Current MCMC methods are based upon simulating a discrete-time Markov chain. In this talk I will present an alternative approach to MCMC, which simulates a continuous-time Markov process, called a piecewise deterministic process. Originally such MCMC methods were motivated as these are non-reversible processes, and it is commonly accepted that non-reversible MCMC mixes better than (standard) reversible approaches.

The resulting continuous-time MCMC algorithms seems particularly well-adapted for use in big data settings. The dynamics of the MCMC process depend on the gradient of the log-posterior. This is something that is easy to estimate using sub-sampling. Furthermore, these continuous-time MCMC processes obey an "exact approximation" property. Namely we can replace the gradient of the log-posterior by an unbiased estimator of it, and the process will still sample from the true posterior. Using low-variance estimators, which involve sub-sampling the data and using control variates, can lead to an MCMC sampler whose computational cost per effective sample size does not increase with the number of data points.

This is joint work with Joris Bierkens and Gareth Roberts (`https://arxiv.org/abs/1607.03188`).

* Machine Learning External Seminar, Gatsby Unit, November 2, 2016.