Large sample limit theorems for the Zig-Zag process

Sanket Agrawal
University of Warwick
sanket.agrawal@warwick.ac.uk

Coauthor(s): Gareth Roberts and Joris Bierkens

Sampling algorithms based on piecewise deterministic Markov processes (PDMPs) have incited much interest in the Monte Carlo community as an alternative to traditional Markov chain Monte Carlo methods. Their non-reversible nature, and the ability to admit sub-sampling strategies without introducing bias, offer much promise for efficient and scalable Monte Carlo simulation. The Zig-Zag sampler [1], based on the multi-dimensional Zig-Zag process has been shown to exhibit nice theoretical properties e.g. geometric ergodicity, large deviations, high dimensional scaling limits, to name a few. In settings of Bayesian inference (consisting of n observations, say), it admits sub-sampling versions which offer significant benefits, reducing computational effort per independent sample from $O(n)$ to $O(1)$ in some situations.

However, sub-sampling itself costs a positive excess switching rate of $O(n)$. This has two problems. First, the computational bounds needed for Poisson thinning may increase causing more proposed switches per unit time, thus increasing the algorithmic complexity of the sampler. Second, a positive excess rate induces diffusivity in the underlying Zig-Zag process. This causes the process to slow down and negatively affects its mixing properties.

We address the second problem by investigating the affects of sub-sampling on process dynamics as the data size n goes to infinity. We study the behaviour of Zig-Zag with sub-sampling (ZZ-SS) and sub-sampling using control variates (ZZ-CV). In the transient phase, by averaging out the fluctuations in the velocity process, we establish weak convergence of the position process to the solution of a deterministic ODE. In the stationary phase, after appropriate re-scaling, we prove the convergence of ZZ-SS and ZZ-CV to an Ornstein-Uhlenbeck process and to a PDMP respectively. Our results show that in the stationary phase, ZZ-CV mixes as fast as the canonical Zig-Zag and ZZ-SS mixes no slower than a factor $n^{1/2}$. In the transient phase, ZZ-SS and ZZ-CV have the same speed with ZZ-SS performing better for heavy tailed models.

[1] Bierkens, J., Fearnhead, P., and Roberts, G. O. (2019). The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data. *The Annals of Statistics*, 47(3):1288–1320.