Discussion on “Quasi-stationary Monte Carlo and the ScaLE Algorithm” by M. Pollock, P. Fearnhead, A.M. Johansen and G.O. Roberts

Ryan Chan, Hongsheng Dai

June 2020

We congratulate the authors for a stimulating and interesting paper which introduces a new class of algorithms that differs from traditional Markov chain Monte Carlo (MCMC) methods, in that the approach is based on the *quasi-stationary* distribution of an appropriately constructed diffusion process. A particularly impressive contribution of the proposed method is that it can be applied in big data contexts whilst remaining exact by adopting a sub-sampling approach. Interestingly, the ScaLE algorithm has connections to another method for tackling large data in the Bayesian framework, namely the *Monte Carlo Fusion* algorithm proposed by Dai et al. [2019]. In particular, both algorithms utilise methodology for the exact simulation of diffusions (Beskos et al. [2006], Beskos et al. [2008]) and use the Langevin diffusion in their mathematical construction (although it is not explicitly used in ScaLE). Further, the Monte Carlo Fusion algorithm uses the function $\phi : \mathbb{R}^d \to \mathbb{R}$ (defined in Section 2). However, the use of subsampling ideas were not explored in [Dai et al., 2019]. The unbiased estimators for ϕ outlined in Section 4 in this paper is a contribution which might be employed in the Monte Carlo Fusion algorithm.

Divide-and-conquer methods (for instance, Scott et al. [2016], Wang and Dunson [2013], Neiswanger et al. [2013], Dai et al. [2019]) have been proposed in order to adapt MCMC for reducing the computational cost of the algorithm. In these approaches, the data set is split into disjoint subsets and then standard MCMC methods are used for each subset. Inference is then combined into a single inference. In this framework, the target is of the form

$$f(\boldsymbol{x}) \propto \prod_{c=1}^{C} f_c(\boldsymbol{x}) \qquad (1)$$

where each *sub-posterior* $f_c(\boldsymbol{x})$ is a density (up to multiplicative constant) representing one of the C distributed inferences we wish to unify.

As noted in the introduction of the paper, the primary weakness of these methods thus far is that the recombination of the separately conducted inferences is inexact and involves some approximation of the sub-posteriors. However, the Monte Carlo Fusion algorithm is exact and is the first exact fusion inference method that allows for perfect sampling from (1). This is achieved by constructing a rejection sampling scheme on an extended space with the main difficulty being computing an intractable acceptance probability which requires the auxiliary simulation of collections of Brownian bridges. An advantage of the fusion approach is that it can be conducted in a distributed setting and one can exploit large clusters of computing cores. In contrast, the QSMC algorithm detailed in this paper and more traditional MCMC methods are single core algorithms. It would be interesting to see if the authors have any ideas for parallel implementations of ScaLE in the future.

References

Beskos, A., Papaspiliopoulos, O., and Roberts, G. O. (2008). A factorisation of diffusion measure and finite sample path constructions. *Methodology and Computing in Applied Probability*, 10(1):85–104.

Beskos, A., Papaspiliopoulos, O., Roberts, G. O., and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):333–382.

Dai, H., Pollock, M., and Roberts, G. (2019). Monte Carlo Fusion. *Journal of Applied Probability*, 56(1):174–191.

Neiswanger, W., Wang, C., and Xing, E. (2013). Asymptotically exact, embarrassingly parallel MCMC. *arXiv preprint arXiv:1311.4780*.

Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I., and McCulloch, R. E. (2016). Bayes and big data: The consensus Monte Carlo algorithm. *International Journal of Management Science and Engineering Management*, 11(2):78–88.

Wang, X. and Dunson, D. B. (2013). Parallelizing MCMC via Weierstrass Sampler. *arXiv e-prints*, page arXiv:1312.4605.