COMPARISON THEORY FOR MCMC ALGORITHMS

(with application to Subsampling MCMC)

Supervisor: Dr. Sam Power (sam.power@bristol.ac.uk)

1. Introduction

Bayesian statistics is a popular framework for carrying out statistical inference wherein uncertainty about parameters is encoded in a probability distribution called the *posterior distribution*. In all but the simplest cases, the posterior distribution does not lie in a standard parametric family; as such, in order to answer inferential questions, it is necessary to make use of numerical methods.

Among these methods, perhaps the most widely-used is *Markov chain Monte Carlo* (MCMC), a class of stochastic algorithms for generating approximate samples from the posterior distribution through an iterative approach (see e.g. [9] for some introduction). Given approximate samples of sufficient quality, one can then use these to form { statistical estimators, uncertainty sets, etc. }. It is thus of key practical and theoretical interest to identify algorithms in this class which can converge effectively, with rigorous complexity bounds depending favourably on problem parameters (e.g. dimension, number of observations, etc.).

For 'nice' statistical models in which covariates and observations are fully-observed, and data sets are of moderate size, there are several 'default' MCMC approaches for which convincing complexity analyses are now available, see e.g. [4] for some overview. Nevertheless, in the face of large-scale data and high-dimensional models, the practicality of these naïve approaches faces serious challenges. This has necessitated the development of more refined and scalable MCMC solutions, for which quantitative theoretical validation has thus far been more limited.

2. Project Proposal

The plan for this project would be to develop a theoretical framework for the quantitative analysis of scalable and 'exact' MCMC algorithms which are based around the use of subsampling strategies. In contrast to cruder 'stochastic gradient MCMC' strategies (see e.g. [6, 10, 3]), these algorithms employ certain ingenious constructions which allow for issues of asymptotic bias to be side-stepped, while retaining low per-iteration cost and favourable convergence properties. A number of such procedures have been developed in recent years (see e.g. [8, 5, 11]), and this project would seek to obtain robust quantitative guarantees for some of these, facilitating rigorous comparisons between competing methods.

Recent work [1] has demonstrated that for MCMC algorithms developed for a related class of intractable models (so-called 'pseudo-marginal' MCMC), the technique of *Markov chain comparison theorems* offers an attractive route to proving robust and interpretable estimates on their convergence behaviour, and follow-up works (e.g. [2, 7]) have demonstrated the broad applicability of this framework. For the present project, we would hope to adapt and extend these theoretical tools to handle large-scale inference tasks, and the range of MCMC algorithms which have been designed for that setting.

References

[1] Christophe Andrieu, Anthony Lee, Sam Power, and Andi Q Wang. Comparison of markov chains via weak poincaré inequalities with application to pseudo-marginal mcmc. *The Annals of Statistics*, 50(6):3592–3618, 2022.

[2] Christophe Andrieu, Anthony Lee, Sam Power, and Andi Q Wang. Weak poincar\'e inequalities for markov chains: theory and applications. *arXiv preprint arXiv:2312.11689*, 2023.

[3] Tianqi Chen, Emily Fox, and Carlos Guestrin. Stochastic gradient hamiltonian monte carlo. In *International conference on machine learning*, pages 1683–1691. PMLR, 2014.

[4] Sinho Chewi. Log-concave sampling. *Book draft available at https://chewisinho.github.io*, 2022.

[5] Khue-Dung Dang, Matias Quiroz, Robert Kohn, Minh-Ngoc Tran, and Mattias Villani. Hamiltonian monte carlo with energy conserving subsampling. *Journal of machine learning research*, 20(100):1–31, 2019.

[6] Christopher Nemeth and Paul Fearnhead. Stochastic gradient markov chain monte carlo. *Journal of the American Statistical Association*, 116(533):433–450, 2021.

[7] Sam Power, Daniel Rudolf, Björn Sprungk, and Andi Q Wang. Weak poincar\'e inequality comparisons for ideal and hybrid slice sampling. *arXiv preprint arXiv:2402.13678*, 2024.

[8] Matias Quiroz, Robert Kohn, Mattias Villani, and Minh-Ngoc Tran. Speeding up mcmc by efficient data subsampling. *Journal of the American Statistical Association*, 2018.

[9] Christian P Robert and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

[10] Max Welling and Yee W Teh. Bayesian learning via stochastic gradient langevin dynamics. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 681–688. Citeseer, 2011.

[11] Ruqi Zhang, A Feder Cooper, and Christopher M De Sa. Asymptotically optimal exact minibatch metropolis-hastings. *Advances in Neural Information Processing Systems*, 33:19500–19510, 2020.