Analysis and applications of piecewise deterministic Markov Chain Monte Carlo methods.

Principal Investigator: Pierre Nyquist

A fundamental aspect of many machine learning methods is the need for efficient Markov chain Monte Carlo (MCMC) methods. For example, in the setting of latent variable models, such as variational auto-encoders and generative adversarial networks, the training involves evaluating expectations with respect to the latent variables. This evaluation in turn relies on MCMC methods. In the high-dimensional setting often encountered in machine learning, the underlying energy landscapes typically have complex geometries, making standard implementations of MCMC methods too slow for practical purposes. The development of new methods with improved convergence properties is therefore an important part of making more advanced latent variable models feasible from a computational standpoint.

A major development in MCMC methods in recent years is the introduction of piecewise-deterministic Markov processes for simulation. The two most prominent examples are the *zig-zag* and the *bouncy particle* samplers. The zig-zag sampler in particular has been shown to have traits that makes it suitable for the setting of large data sets: there is a natural process for sub-sampling the data whilst still approximating the correct (invariant) distribution.

Despite a large effort to understand the properties of the zig-zag and bouncy particle samplers, many questions about their efficiency remain. Ergodicity results exist, but so far convergence rates are missing. Recent results on spectral properties of the zig-zag process in one dimension is a step towards this, but there is currently no way to extend these results beyond the one-dimensional setting. Moreover, the spectral gap has been shown to have drawbacks as a measure of convergence for Markov processes.

The proposed project is focused on analysis and applications of piecewise-deterministic MCMC methods. Topics of interest include, but are not limited to:

I. *Measure-valued process formulation.* Measure-valued processes are used for analysing certain Markov processes, primarily in the setting of queueing systems. We propose that this framework is suitable for the study of simulation algorithms as well, specifically piecewise-deterministic processes as the corresponding dynamics in the space of measures takes a rather simple form.

II. *Large deviations for empirical measures of zig-zag and bouncy particle samplers.* Piecewise-deterministic MCMC methods correspond to irreversible process of jump type and thus the assumptions used in the well-established Donsker-Varadhan framework for large deviations of empirical measures do not hold. Recently we established the large deviation principle for the zig-zag process in one dimension. However for dimension ≥ 2 this remains an open problem.

III. *Qualitative and quantitative comparisons of convergence rates.* Building on topic II, we are interested in comparing the convergence properties of piecewise-deterministic MCMC methods with other common simulation methods; a comparison between zig-zag and bouncy particle is also of particular interest. The rate function associated with large deviations is one way to approach this problem, and empirical measure large deviations have become a useful tool for studying convergence in the MCMC setting.

IV. *Applications in machine learning.* Thus far piecewise-deterministic MCMC methods have mainly been studied in the context of Bayesian statistics. We are interested in implementation and analysis of the performance, both through empirical and theoretical studies, of such methods in the machine learning context, for example in the setting of training latent variable models.