Auto-tuning Hamiltonian Monte Carlo

Andreas Fichtner1, Lars Gebraad1, Christian Boehm1, and Andrea Zunino2
1ETH Zurich, Department of Earth Sciences, Zurich, Switzerland (andreas.fichtner@erdw.ethz.ch)
2Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark (zunino@nbi.dk)

Hamiltonian Monte Carlo (HMC) is a Markov chain Monte Carlo method that exploits derivative information in order to enable long-distance moves to independent models, even when the model space dimension is high (Duane et al., 1987). This feature motivates recent research aiming to adapt HMC for the solution of geophysical inverse problems (e.g. Sen & Biswas, 2017; Fichtner et al., 2018; Gebraad et al., 2020).

Here we present applications of HMC to inverse problems at variable levels of complexity. At the lowest level, we study linear inverse problems, including, for instance, linear traveltime tomography. Though this is not the class of problems for which Monte Carlo methods have been developed, it allows us to understand the important role of HMC tuning parameters. We then demonstrate that HMC can be used to obtain probabilistic solutions for two important classes of inverse problems: 2D nonlinear traveltime tomography and 2D elastic full-waveform inversion. In both scenarios, no super-computing resources are needed for model space dimensions from several thousand to ten thousand.

By far the most important, but also most complex, tuning parameter in HMC is the mass matrix, the choice of which critically controls convergence. Since manual tuning of the mass matrix is impossible for high-dimensional problems, we develop a new HMC flavour that tunes itself during sampling. This rests on the combination of HMC with a variant of the L-BFGS method, well-known from nonlinear optimisation. L-BFGS employs a few Monte Carlo samples to compute a matrix factorisation \( LL^T \) which dynamically approximates the local Hessian \( \mathbf{H} \), while the sampler traverses model space in a quasi-random fashion. The local curvature approximation is then used as mass matrix. Following an outline of the method, we present examples where the auto-tuning HMC produces almost perfectly uncorrelated samples for model space dimensions exceeding 105.

References

