Tuning Hamiltonian Monte Carlo in waveform inversion

Jinji Li* and Kristopher A. Innanen
li.jinji@ucalgary.ca

ABSTRACT

Tuning the Hamiltonian Monte Carlo (HMC) stands as a crucial endeavor, especially within the realm of geophysical optimization and inversion where the problems are complex and high-dimensional. In the context of this report, we offer an in-depth examination of the functionalities associated with the tunable parameters, such as the integration length, time step, and mass matrix. The minimum-model sampling experiment has shown that assigning the second-order information to the mass matrix can help the sampler reach the target distribution more efficiently, and the HMC-FWI experiment has shown that an adaptive tuning strategy can significantly accelerate the convergence and thus produce more precise solutions within limited sampling attempts. This work contributes to a deeper understanding of the role of tuning in HMC applications and can serve as a valuable reference for future utilization.

TUNABLE PARAMETERS IN HMC

- Integration length L: The total length of the trajectory represents the number of leapfrog steps taken during the HMC simulation and impacts the acceptance probability of proposed states and the efficiency of the sampling process.
- Time step Δt: An inappropriate choice of time step can lead to biased or inaccurate samples, and too large a step size might result in poor exploration of the target distribution.
- Mass matrix $\mathbf{M}$: Analytically, the mass matrix's elements influence the energy landscape's curvature and the system's characteristic time scales.

SHOWCASE BY a 100D SAMPLING PROBLEM

$$U(\boldsymbol{m}) = \frac{1}{2}\boldsymbol{m}^T C_M^{-1}\boldsymbol{m} \qquad C_M = \begin{bmatrix} 0.01 & 0 & \cdots & 0 \\ 0 & 0.02 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

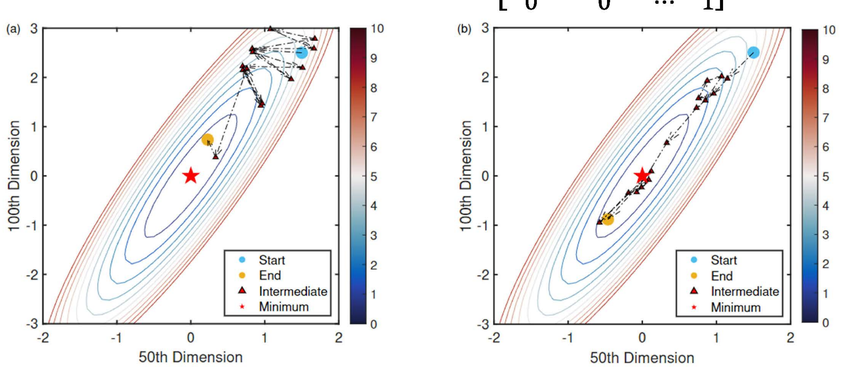

Figure 1. Initial 20 MC and HMC walker trajectories projected onto the 50th and 100th dimensions. (a) Trajectories of HMC when $\mathbf{M} = \boldsymbol{I}$. (b) Trajectories of HMC when matrix is set to Hessian.

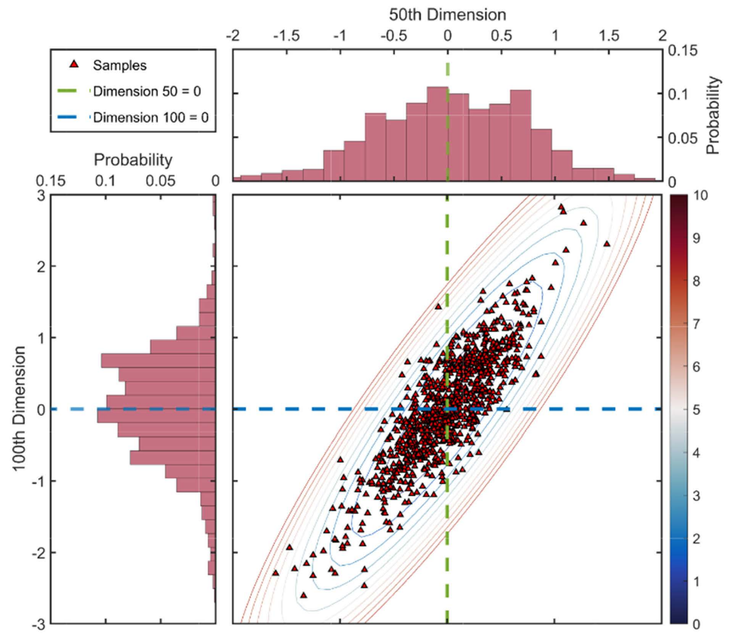

Figure 2. Accepted samples of MCMC projected onto the 50th and 100th dimensions, and the marginal distributions.

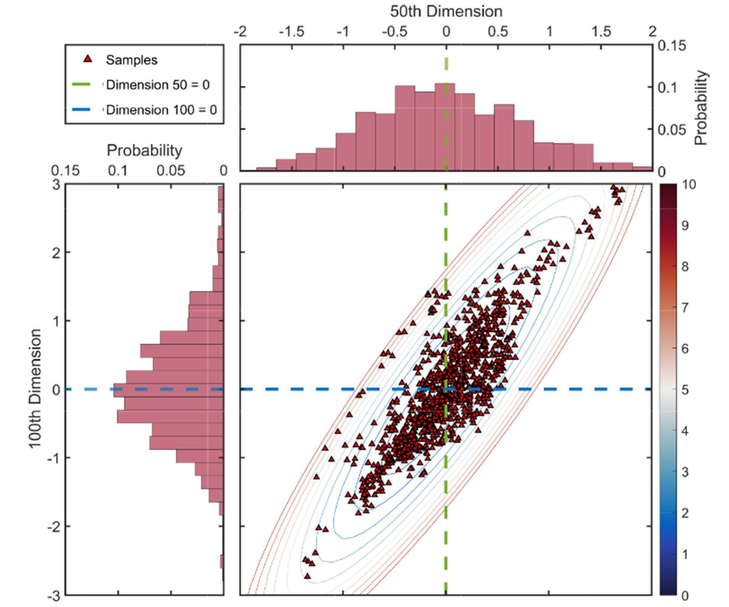

Figure 3. Accepted samples of HMC projected onto the 50th and 100th dimensions, and the marginal distributions.

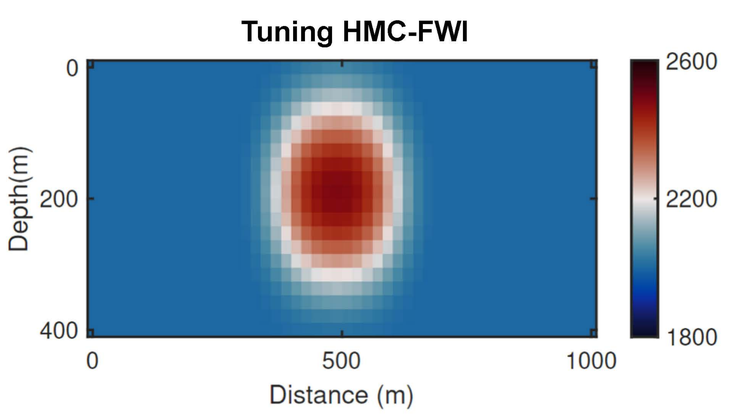

Figure 4. True model for HMC-FWI.

RESULTS

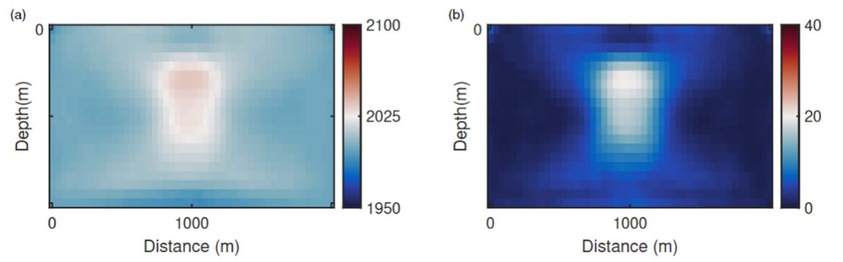

Figure 5. HMC-FWI results without adaptive tuning. (a) mean model. (b) the standard deviation of the accepted models.

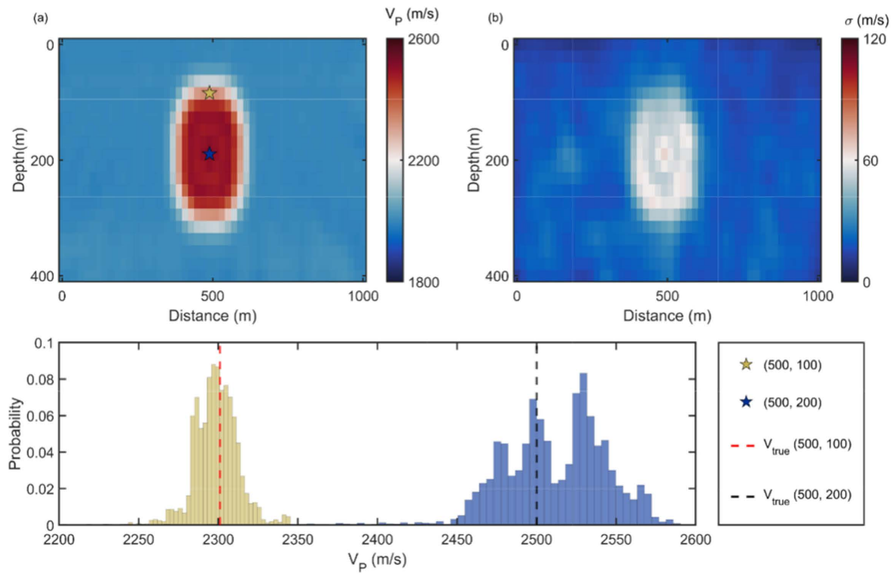

Figure 6. HMC-FWI results with adaptive tuning. The yellow star denotes the model position at (500, 100), and the blue star denotes the model position at () (a) mean of the accepted models. (b) the standard deviation of the accepted models. (c) probability distribution of models at (500, 100) and (500, 200), and the dashed line shows the true values at these positions.

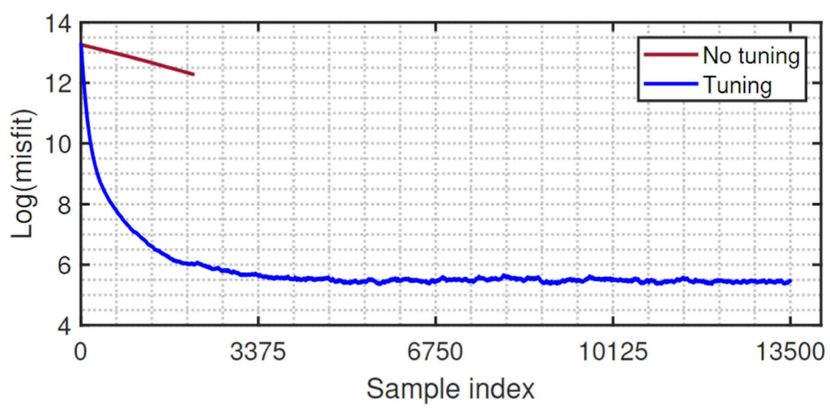

Figure 7. Log misfit's variation with iteration.

ACKNOWLEDGEMENTS

This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19. One of the authors of this report was supported by the CSEG Foundation.

NSERC CRSNG
UNIVERSITY OF CALGARY
FACULTY OF SCIENCE
Department of Earth, Energy, and Environment