Hamiltonian Monte Carlo in waveform inversion

Jinji Li* and Kristopher A. Innanen

li.jinji@ucalgary.ca

ABSTRACT

Classical Markov Chain Monte Carlo (MCMC) methods, while widely used for Bayesian inference, often suffer from computational demands and inefficiencies, particularly when dealing with high-dimensional parameter spaces. However, the Hamiltonian Monte Carlo (HMC) approach represents a notable advancement in the field of Monte Carlo (MC) methods. By simulating Hamiltonian dynamics through numerical integration and incorporating a Metropolis acceptance step, HMC avoids the limitations of random walks typically associated with MCMC. This results in a decent acceptance rate, enabling more efficient parameter space exploration. Additionally, the ability to generate plausible model candidates during the integration process opens up access to the null space, which can be particularly valuable in inversion problems such as full waveform inversion (FWI) where the model space is complex and multidimensional. In this report, we delve into the fundamental workings of HMC, shedding light on its mechanics and advantages. We present results from numerical experiments that showcase the special features of HMC. Our findings suggest that HMC holds significant promise as a robust tool for improving uncertainty quantification in applications of FWI, where accurately characterizing uncertainties is crucial for obtaining reliable model estimates.

HMC WORKFLOW

Algorithm 1: HMC iterations

for $M = 1$ to nsamples **do**
- Randomly generate $\mathbf{p}_{cur} \sim \mathcal{N}(0, \mathbf{M})$;
- $\mathbf{p}_{new} = \mathbf{p}_{cur}$
- $\mathbf{m}_{new} = \mathbf{m}_{cur}$
- $U(\mathbf{m}_{cur}) = -log\rho_{m_{cur}}(\mathbf{m}_{cur}|\mathbf{d})$
- $K(\mathbf{p}_{cur}) = \frac{\mathbf{p}_{cur}^T \mathbf{M}^{-1} \mathbf{p}_{cur}}{2}$
- $H(\mathbf{p}_{cur}, \mathbf{m}_{cur}) = U(\mathbf{m}_{cur}) + K(\mathbf{p}_{cur})$
- $\mathbf{p}_{new} = \mathbf{p}_{new} - \delta t \frac{\nabla U(\mathbf{m}_{new})}{2}$
- **for** $N = 1$ to $L-1$ **do**
  - $\mathbf{m}_{new} = \mathbf{m}_{new} + \delta t \nabla K(\mathbf{p}_{new})$
  - $\mathbf{p}_{new} = \mathbf{p}_{new} - \delta t \nabla U(\mathbf{m}_{new})$
- **end for**
- $\mathbf{m}_{new} = \mathbf{m}_{new} + \delta t \nabla K(\mathbf{p}_{new})$
- $\mathbf{p}_{new} = \mathbf{p}_{new} - \delta t \frac{\nabla U(\mathbf{m}_{new})}{2}$
- $\mathbf{p}_{new} = -\mathbf{p}_{new}$
- $U(\mathbf{m}_{new}) = -log\rho_{m_{new}}(\mathbf{m}_{new}|\mathbf{d})$
- $K(\mathbf{p}_{new}) = \frac{\mathbf{p}_{new}^T \mathbf{M}^{-1} \mathbf{p}_{new}}{2}$
- $H(\mathbf{p}_{new}, \mathbf{m}_{new}) = U(\mathbf{m}_{new}) + K(\mathbf{p}_{new})$
- **if** randnum $\leq exp(-(H_{new} - H_{cur}))$ **then**
  - $\mathbf{m}_{cur} = \mathbf{m}_{new}$
- **else**
  - $\mathbf{m}_{cur} = \mathbf{m}_{cur}$
- **end if**

end for

HMC-FWI CONSTRUCTION

$$U(\boldsymbol{m}) = -\log \rho_m(\boldsymbol{m}|\boldsymbol{d}) = \frac{1}{2}\left(\boldsymbol{d}_{obs} - \boldsymbol{d}_{syn}\right)^T C_D^{-1}\left(\boldsymbol{d}_{obs} - \boldsymbol{d}_{syn}\right). \quad (1)$$

$$K(\boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^T \mathbf{M}^{-1} \boldsymbol{p}. \quad (2)$$

In our inversion process, we consider the physical parameter to be inverted, namely the P-wave velocity, as independently and identically distributed (IID) following a Gaussian distribution. In addition, we assume that the data covariance matrix is identical. The gradient is calculated by the adjoint approach. 30,000 samples are generated.

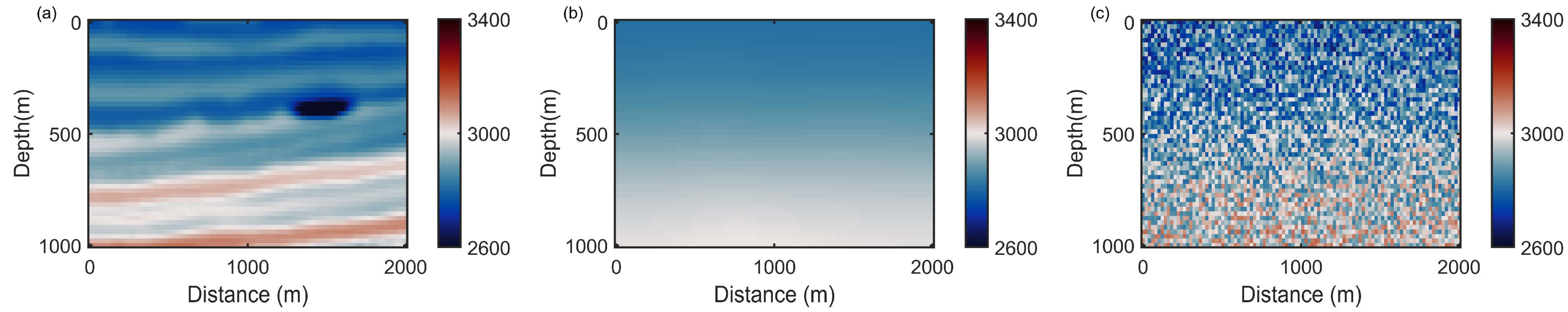

Figure 1. V_P models for HMC-FWI. (a) true model. (b) smoothed model. (c) initial model from $U(\boldsymbol{m}_{initial} \pm 150\ m/s)$.

RESULTS

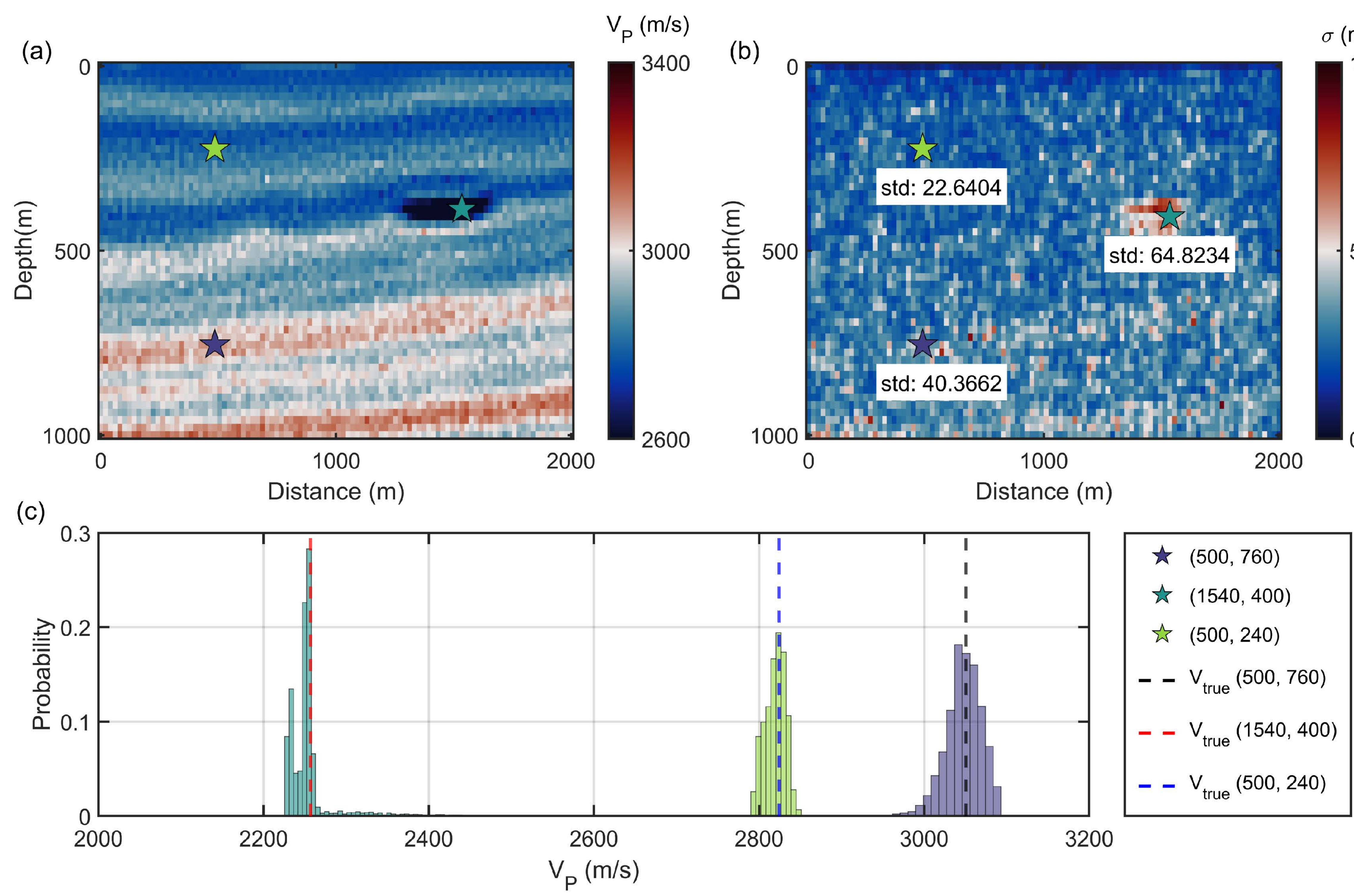

Figure 2. HMC-FWI results. The white star denotes the model position at (1540, 400), the blue star denotes the model position at (500, 760), and the red star denotes the model position at (500, 240). (a) mean of the accepted models. (b) the standard deviation of the accepted models. (c) probability distribution of models at positions denoted by the three stars, and the dashed line shows the true values at these positions..

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