Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_07-005-cd

ISGSR 2022

Assessment of HMC Parameter Updates for Piping Zone Boundary Detection

Michael C. Koch[1], Kazunori Fujisawa[1] and Akira Murakami[1]

[1]Graduate School of Agriculture, Kyoto University, Kyoto, Japan
E-mail: koch.michaelconrad.5w@kyoto-u.ac.jp

Abstract: Estimation of the extent of piping or internal erosion in an earthen embankment/ levee can provide valuable information for the serviceability of such geotechnical structures. Working in a Bayesian framework, different parameter updates are discussed for the update of domain geometry, using a statistically efficient gradient based Markov Chain Monte Carlo (MCMC) algorithm called Hamiltonian Monte Carlo (HMC). Although, the hydraulic conductivity spatial random field is generally uncertain, it is assumed to be known exactly in this study (for simplicity), and the updates presented only consider uncertainty in domain boundary. The parameter updates are discussed for volume integral method based discretization of domains e.g. finite element method. The main challenge in such shape detection problems is to ensure a high quality of mesh as the boundary is updated. As such, the parameter update methods can be classified into two groups, one without remeshing and the other with remeshing. Additionally, in HMC, parameter updates have to be designed in a manner that finite element nodal coordinate functions are differentiable w.r.t the parameters. For the class of updates that do not involve remeshing, methods that maintain mesh quality even under large distortions, such as the mesh moving method are considered. The reversibility aspects of such updates and their implications on the inverse analysis are highlighted. A method to compute nodal coordinate gradients w.r.t parameters, in the remeshing case is also discussed. Finally, the merits and demerits of the two classes of methods are highlighted and an extension of the remeshing class to trans-dimensional parameter updates is briefly introduced.

Keywords: Inverse problems; HMC; boundary detection; piping zone; mesh moving method; remeshing.

1 Introduction

Piping zone boundary detection is an example of a geometric inverse problem where the target is the determination of the boundary of the domain. Accurate solution of the inverse problem naturally requires accurate computation of the forward problem also. This is dependent on the quality of mesh used for the forward analysis. Maintaining a high quality of mesh, free from distortions, with uniform-sized elements throughout the domain during inversion procedures where gradients of the forward problem are required, is a challenging task. Methods that continuously move a mesh with fixed connectivities (Stein et al. 2003; Koch et al. 2020) are one way to maintain a decent mesh quality. However, these methods provide no guarantee of a good mesh quality and can produce meshes that are severely non-uniform or in some cases folded too, especially around fixed observation points.

Remeshing is seen as a natural alternative to solve these issues. However, it isn't clear how the gradients of the nodal coordinate functions can be evaluated if naïve remeshing is done. These gradients are essential to all gradient based inversion algorithms, including the statistically efficient Markov Chain Monte Carlo (MCMC) algorithm called Hamiltonian Monte Carlo (HMC) (Neal 2011). Shape sensitivity analysis in such problems ultimately requires the computation of nodal coordinate gradients w.r.t to the shape (boundary) parameters. Additionally, as a numerical approximation is being made, it is almost certain that the Hamiltonian will not be constant before and after remeshing. In other words, remeshing acts like an event-driven discontinuity (Afshar, M.H. and Domke, 2015) at the current point in parameter space. Standard HMC isn't applicable in such cases.

This paper revisits the problem of analytical calculation of shape sensitivities when remeshing is done during the parameter update. First the basics of probabilistic inverse problems are highlighted in Section 2. Building upon the arbitrary reference mesh based geometry update of Koch et al. (2020), a method is highlighted in Section 3 that respects the reversibility criterion of all MCMC algorithms and can incorporate remeshing while allowing for gradients to be calculated. For simplicity of presentation, the hydraulic conductivity spatial field is assumed to be known and the only unknowns are the parameters defining the piping zone boundary in the numerical implementation in Section 4. Comments on applicability of such updates to trans-dimensional problems are highlighted in Section 5.

2 Probabilistic inversion using HMC

The forward problem of steady state fluid flow is defined by the Laplace equation which is considered on a domain with a pipe, parameterized by the vector $\boldsymbol{\theta} = (l, w)$, at the top-right of the domain Ω, with standard Dirichlet and Neumann boundary conditions. The PDE is solved by the finite element method with varying boundary conditions at different time instants $k \in \{1, \ldots, n\}$, yielding discretized governing equations of the form

$$\mathbf{K}(\boldsymbol{\theta})\mathbf{h}_k = \mathbf{q}_k, \tag{1}$$

where $\mathbf{K}$ is the hydraulic conductivity matrix and $\mathbf{h}_k$ and $\mathbf{q}_k$ are the hydraulic head and nodal flux vectors respectively. To solve the inverse problem, a Gaussian observation model which connects the observations $\mathbf{y}_k$ to the state variables $\mathbf{m}_k = (\mathbf{h}_k, \mathbf{q}_k)$ is also considered and is given as

$$\mathbf{y}_k = \mathbf{H}\mathbf{m}_k(\boldsymbol{\theta}) + \mathbf{r}_k, \tag{2}$$

where $\mathbf{r}_k \sim \mathrm{N}(\mathbf{0}, \mathbf{R}_k)$ is the Gaussian observation error with zero mean and covariance matrix $\mathbf{R}_k$. $\mathbf{H}$ is a matrix mapping the state vector at all the nodes to the state vector at the observation nodes.

The unique feature of HMC is the association of a momentum $\mathbf{p}$ with every point in $\boldsymbol{\theta}$-space such that a joint probability distribution $p(\boldsymbol{\theta}, \mathbf{p})$ can be defined where $p(\boldsymbol{\theta}, \mathbf{p}) = p(\mathbf{p})p(\boldsymbol{\theta}|\mathbf{y}_{1:n})$. Considering a Gaussian momentum, $\mathbf{p} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\mathcal{M}})$, with a user defined covariance, the Hamiltonian (H) can then be defined as

$$H(\boldsymbol{\theta}, \mathbf{p}) = -\log p(\mathbf{p}) - \log p(\boldsymbol{\theta}|\mathbf{y}_{1:n}) + \text{const.}. \tag{3}$$

The potential energy or negative-log of the posterior distribution is $\varphi(\boldsymbol{\theta}) = -\log p(\boldsymbol{\theta}|\mathbf{y}_{1:n})$ and the remaining part (excluding the constants) is called the Kinetic energy. Ignoring the constants, considering a Gaussian prior $\boldsymbol{\theta} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ the potential energy is written as

$$\varphi(\boldsymbol{\theta}) = \sum_{k=1}^{n} \frac{1}{2}\left(\mathbf{y}_k - \mathbf{H}\mathbf{m}_k(\boldsymbol{\theta})\right)^{\mathrm{T}} \mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}\mathbf{m}_k(\boldsymbol{\theta})\right) + \frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\boldsymbol{\theta}. \tag{3}$$

HMC then explores level sets of H through numerical solutions of the Hamiltonian dynamics equations. One such numerical scheme, called the leapfrog method, is written as

$$\mathbf{p}\left(t + \frac{\varepsilon}{2}\right) = \mathbf{p}(t) - \frac{\varepsilon}{2}\frac{\partial\varphi(\boldsymbol{\theta}(t))}{\partial\boldsymbol{\theta}}, \boldsymbol{\theta}(t+\varepsilon) = \boldsymbol{\theta}(t) + \varepsilon\boldsymbol{\mathcal{M}}^{-1}\mathbf{p}\left(t + \frac{\varepsilon}{2}\right), \mathbf{p}(t+\varepsilon) = \mathbf{p}\left(t + \frac{\varepsilon}{2}\right) - \frac{\varepsilon}{2}\frac{\partial\varphi(\boldsymbol{\theta}(t+\varepsilon))}{\partial\boldsymbol{\theta}}. \tag{4}$$

These equations are solved L times with a step-size ε to move to the next sample in the HMC chain. Each transition $(\boldsymbol{\theta}, \mathbf{p}) \rightarrow (\boldsymbol{\theta}^*, \mathbf{p}^*)$ is exactly reversible, provided the gradients $\frac{\partial\varphi}{\partial\boldsymbol{\theta}}$ in Eq. 4 are one-to-one, and the detailed balance condition is obeyed through acceptance probabilities of the kind $\boldsymbol{A} = \min\left\{1, \exp\left(\frac{-H(\boldsymbol{\theta}^*, \mathbf{p}^*)}{-H(\boldsymbol{\theta}, \mathbf{p})}\right)\right\}$.

3 Geometry updates

3.1 State of the art in geometry updates

The parameter $\boldsymbol{\theta}$ characterizes the unknown part of the boundary of the domain Γ_v (see Fig. 1). The main question in geometry updates is: how to update the mesh nodal coordinates $\mathbf{Z}$, corresponding to the parameter update $\boldsymbol{\theta}(t) \rightarrow \boldsymbol{\theta}(t+\varepsilon)$? It is possible to pre-design one-to-one explicit $\boldsymbol{\theta}$-parameterized functions $\mathbf{Z}_v(\boldsymbol{\theta})$ for a subset of the nodal coordinates $\mathbf{Z}$, representing the coordinates of the nodes on the boundary Γ_v. Designing such explicit functions for the interior nodes subset of $\mathbf{Z}$ is however a much more complicated task and except for simple geometries, is almost impossible. A simple Lagrangian displacement of the mesh is also not a valid option as large distortions may occur during the update and no guarantees can be made on the quality of the updated mesh. To this end, Koch et al. (2020) developed a parameter update that incorporates the mesh moving method of Stein et al. (2003) to maintain a good mesh quality. However, even though the mesh moving method can significantly eliminate mesh distortion, it does not eliminate it completely. A folding of the mesh may still occur around the fixed observation points, when large displacements take place. Moreover, the mesh can also deform to such an extent that some elements might become "excessively large" and accurate computations might not be possible in these regions of the domain. Furthermore, these updates are not amenable to trans-dimensional jumps in parameter space. This paper investigates these aspects of the geometry update further.

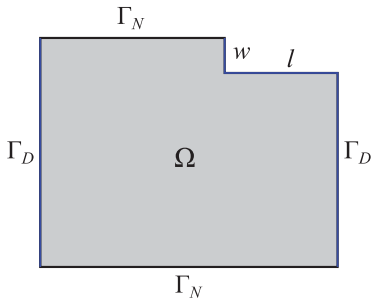

Figure 1. Domain schematic showing the pipe of length (l) and width (w). In this paper, $\Gamma_v = \Gamma_D \cup \Gamma_N$

An additional aspect of the parameter update that is unique to HMC is the question of reversibility. Sticking with the method described in Koch et al. (2020), the mesh moving method yields a displacement $\mathbf{u}^{ref}$ for all nodes in the mesh. Node coordinates at the next time-step $\mathbf{Z}\big(\boldsymbol{\theta}(t+\varepsilon)\big)$ can then be calculated as

$$\mathbf{Z}\big(\boldsymbol{\theta}(t+\varepsilon)\big) = \mathbf{Z}^{ref}(\boldsymbol{\theta}^{ref}) + \mathbf{u}^{ref}, \tag{4}$$

where $\mathbf{Z}^{ref}(\boldsymbol{\theta}^{ref})$ refers to a discretization of the reference domain i.e., a placeholder domain with node coordinates $\mathbf{Z}^{ref}$, with boundary defined by an arbitrary $\boldsymbol{\theta}^{ref}$. The subset of nodal displacements on the boundary Γ_v i.e. $\mathbf{u}_v^{ref}$ are known apriori because $\mathbf{Z}_v\big(\boldsymbol{\theta}(t+\varepsilon)\big)$ is known apriori and the displacements can be calculated as $\mathbf{u}_v^{ref} = \mathbf{Z}_v\big(\boldsymbol{\theta}(t+\varepsilon)\big) - \mathbf{Z}_v^{ref}(\boldsymbol{\theta}^{ref})$. The remaining subset of the interior nodes $\mathbf{u}_{int}^{ref}$ can then be obtained by solving an elastic deformation problem with prescribed displacements $\mathbf{u}_v^{ref}$. The mesh moving method is used in this step to ensure a good quality mesh and yields the total displacements of all the nodes $\{\mathbf{u}^{ref}\} = \{\mathbf{u}_v^{ref}\} \cup \{\mathbf{u}_{int}^{ref}\}$. Once the new position of all the nodes is known, the forward problem Eq. (1) can be solved and the potential energy (Eq. 3) an be computed. In this update, the gradient $\frac{\partial \varphi}{\partial \boldsymbol{\theta}}$ can be computed because the gradient of the mesh movement $\mathbf{u}^{ref}$ from the reference domain is a one-to-one function of $\boldsymbol{\theta}$. The derivative $\frac{\partial \mathbf{u}^{ref}}{\partial \boldsymbol{\theta}}$ is one-to-one as long as $\frac{\partial \mathbf{u}_v^{ref}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \mathbf{u}_{int}^{ref}}{\partial \boldsymbol{\theta}}$ are one-to-one.

The choice of this placeholder reference domain and its discretization is up to the user and the simplest choice one can make is $\mathbf{Z}^{ref}(\boldsymbol{\theta}^{ref}) = \mathbf{Z}\big(\boldsymbol{\theta}(t)\big)$. This choice however produces irreversible updates because $\mathbf{u}_{int}^{ref}$ is obtained from an elastic deformation problem solved on a domain with coordinates $\mathbf{Z}\big(\boldsymbol{\theta}(t)\big)$ in the forward direction and $\mathbf{Z}\big(\boldsymbol{\theta}(t+2\varepsilon)\big)$ in the reverse direction. This means $\mathbf{u}_{int}^{ref}\big(\boldsymbol{\theta}(t)\big)$ is no longer one-to-one and by extension $\frac{\partial \varphi(\boldsymbol{\theta}(t))}{\partial \boldsymbol{\theta}}$ is not one-to-one. The solution to this problem proposed in Koch et al. (2020) is to consider $\mathbf{Z}^{ref}(\boldsymbol{\theta}^{ref}) = \mathbf{Z}^{fixed}(\boldsymbol{\theta}^{fixed})$, where the updates always take place from an arbitrary fixed reference configuration (see Fig. 2(a)). As the reference domain is always fixed, it is easy to see why this update is one-to-one. However, the problems mentioned earlier in this section remain and must be addressed.

3.2 New parameter updates

One method to ensure that a high-quality mesh exists at each time (t) is to remesh the domain. But this sets up an interesting question on how to compute the HMC gradients. Construction of one-to-one differentiable maps is impossible in such cases. Henceforth, a brief description is given on how remeshing is incorporated in the HMC framework.

3.2.1 Remeshing the reference domain

Two features are added to the reversible mesh moving proposal described above. It is apparent that the reversibility of this method comes from the fact that the mesh at each step (t) is updated from a fixed reference configuration. This can be generalized by allowing the reference configuration to change at every step (t), provided each configuration is kept in memory and the same corresponding configuration is used in a reverse move. Such an update would also guarantee reversibility. Mathematically, this can be written as $\mathbf{Z}_i^{ref}\big(\boldsymbol{\theta}_i^{ref}\big) = \mathbf{Z}_i^{fixed}\big(\boldsymbol{\theta}_i^{fixed}\big)$, where $i = (1, \dots, L)$. Each of these configurations can be set arbitrarily, but once set must be kept fixed in the forward and reverse moves. The second feature becomes apparent immediately. As each of these configurations can be arbitrarily set, different reference configurations corresponding to different $\boldsymbol{\theta}_i^{ref}$ can be formed by remeshing the reference domain (see Fig. 1(b)).

3.2.2 Selecting the reference configurations

Although remeshing is possible, it is not clear what advantage is to be gained by selecting different remeshed reference configurations in this problem. The need for a high-quality mesh (on which the forward problem is solved) at each time step motivates the goal for selecting a reference configuration i.e., there should little to no distortion in the elements in the mesh movement phase. This can be achieved by choosing the reference domain boundary at a particular time (t) to be as close as possible to the boundary of the domain (which is known apriori). As $\boldsymbol{\theta}(t) \to \boldsymbol{\theta}(t+\varepsilon)$, boundary nodal coordinates $\mathbf{Z}_v\big(\boldsymbol{\theta}(t+\varepsilon)\big)$ are available apriori. The remeshed arbitrary reference domain for the time step $(t+\varepsilon)$ can then be constructed with boundary nodal coordinates $\mathbf{Z}_{v1}^{ref}\big(\boldsymbol{\theta}_1^{ref}\big) = \mathbf{Z}_{v1}\big(\boldsymbol{\theta}(t+\varepsilon) + \boldsymbol{\delta}\big)$. The increment $\boldsymbol{\delta} \to \mathbf{0}$ is chosen so that the reference domain boundary is just a slightly perturbed version of the already known boundary at that time step. This then sets up an infinitesimal prescribed displacement problem to evaluate the HMC gradients which should ensure a high quality of mesh as displacements and distortions will be minimal.

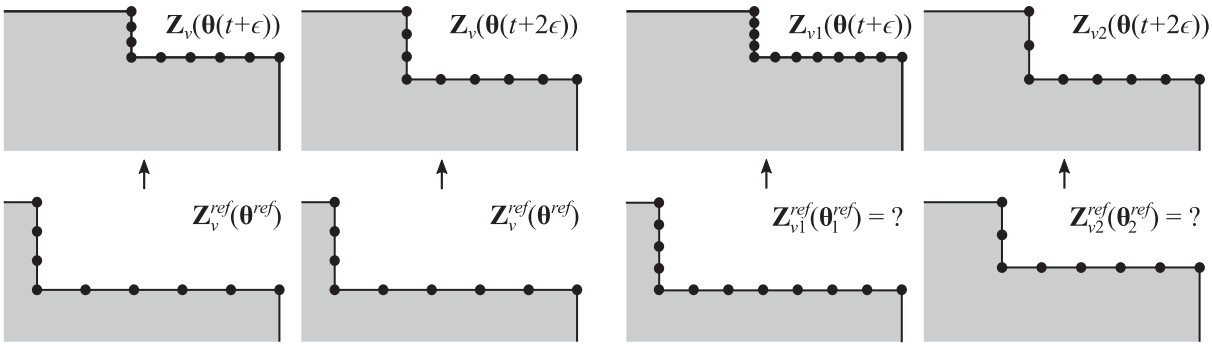

Figure 2. Schematic showing (a) update from a fixed reference configuration and (b) update from different reference configurations at different steps of HMC. Only the top-right part of the domain is shown.

4 Numerical implementation and results

4.1 Observation data

Synthetic observation data for inversion was obtained from a domain with $\boldsymbol{\theta}^{true} = (0.14, 0.04)$. The position of the observation points B,…,F are shown in Fig. 3 and the discretization was done with uniform sized elements of average size 1/4th of the smallest elements shown in Fig. 3. Dirichlet BCs were set such that steady state flow was prescribed from left to right. The Neumann boundaries were set as no-flow boundaries. For a fixed uniform hydraulic conductivity, hydraulic head data was measured at the five red observation points and the outward flow data was measured as the sum of total outward normal flow from the right boundary nodes shown in green. Ten sets ($k = 10$) of observations were obtained for different Dirichlet BCs. An observation noise was added to the recorded data in the proportion of 1% for $\mathbf{h}_k$ and 5% for $\mathbf{q}_k$.

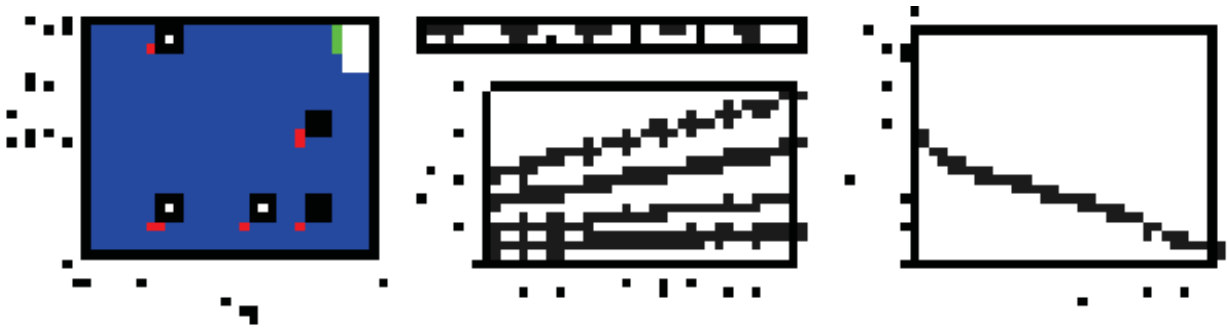

Figure 3. (left) Initial reference domain with $\boldsymbol{\theta}^{ref} = (0.05\text{ m}, 0.08\text{ m})$ showing position of observation points. (center) Hydraulic head observation data (diamonds represent input hydraulic head on left boundary, head on right boundary is fixed at zero). (right) Total outward flow data.

4.2 Numerical simulation and results

As the solution is expected to be correlated, the initial reference mesh was chosen to be highly uncorrelated i.e. $\boldsymbol{\theta}^{ref} = (0.05\text{ m}, 0.08\text{ m})$ as shown in Fig. 3. The number of leapfrog steps required to generate each sample was determined randomly from a normal distribution $\mathrm{N}(10,3)$. A total of 10,000 samples were obtained, and the first 2000 samples were considered as burn-in samples. Posterior 1-D and 2-D marginals, obtained from 8000 samples post burn-in are shown in Fig. 4. Dark grey contours indicate the 50% and 95% High Probability Density (HPD) regions such that 50% and 95% of the total marginal probability lie within these contours. The 2-D marginal is highly correlated (which is physically expected) and justifies the use of more advanced MCMC algorithms like HMC for the efficient solution of such problems. Additionally, the true solution marked as a diamond is well captured by both the 1-D and 2-D marginals, indicating the correctness of the method.

The effect of different choices of the reference configuration is shown in Fig. 5. The inverse analysis is carried out by choosing the reference configuration boundary to be infinitesimally perturbed, $\boldsymbol{\delta} = \left(0.001, \frac{0.001}{3}\right)$, from the actual boundary in Fig. 5(a). At the end of 11 leapfrog steps (arbitrarily chosen) in the forward direction, the momentum is reversed i.e. $\mathbf{p} = -\mathbf{p}$ and the leapfrog steps are again executed 11 times with the same step size. There is a perfect match between the forward and reverse trajectory at every step which indicates the correctness of the remeshed reference domain based reversible proposal. To show that the update from the previous step is not reversible, the remeshed reference domain boundary is chosen to be the boundary at the previous step. As shown

in Fig. 5(b), the reverse trajectory in red deviates from the forward trajectory and the condition of reversibility is not exactly met.

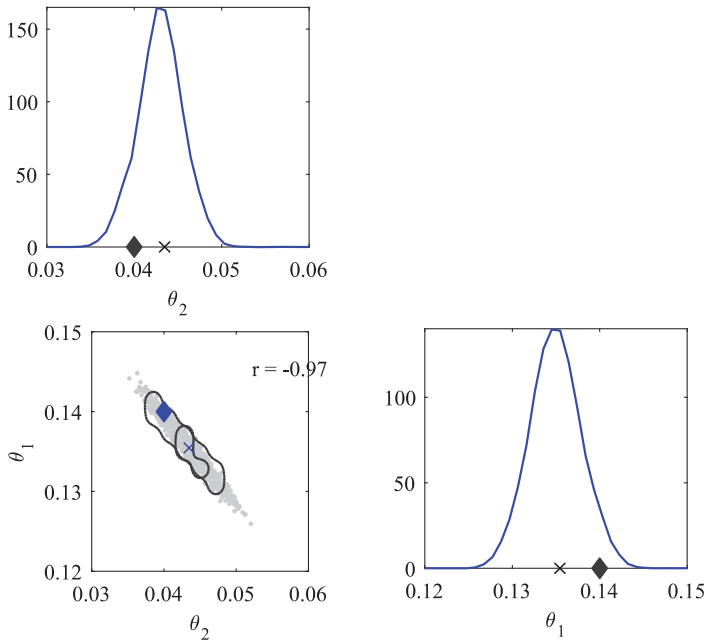

Figure 4. Marginal posterior distributions for different components of the parameter vector $\boldsymbol{\theta}$. The light grey circles represent thinned samples (every 5th sample) obtained after burn-in from HMCSISFD. Dark grey contours indicate the 50% and 95% High Probability Density (HPD) regions. The cross mark represents the HMC mean, and the blue diamond represents the true value of the parameters. Pearson correlation coefficient between representative parameters is denoted by r. (Magnitudes are in m)

5 Conclusions and discussions

To develop parameter updates for general geometric inverse problems, it is argued that remeshing is a necessity to maintain a good mesh quality. This paper highlights a method to include remeshing into HMC updates, while also ensuring continuity of the Hamiltonian and reversibility of the proposal. The method relies on choosing slightly perturbed reference domain boundaries from the already available boundaries at each step. These slightly perturbed

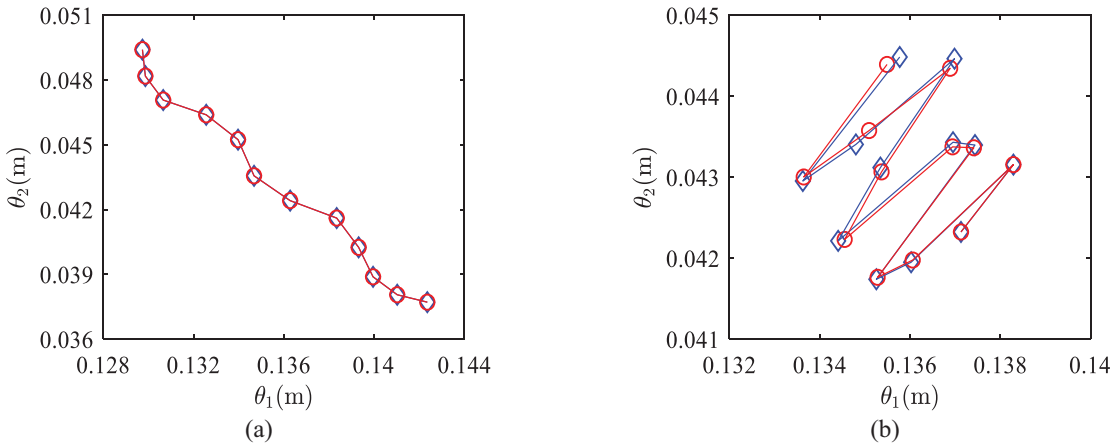

Figure 5. Leapfrog trajectories in $\theta_1 - \theta_2$ space when remeshed reference configuration is parameterized as (a) $\boldsymbol{\theta}^{ref} = \boldsymbol{\theta}(t+\varepsilon) + \boldsymbol{\delta}$ with $\boldsymbol{\delta} = \left(0.001, {}^{0.001}/_{3}\right)$ and (b) $\boldsymbol{\theta}^{ref} = \boldsymbol{\theta}(t)$. Blue line with diamond markers indicates the forward trajectory and the red line with circle markers indicates the reverse trajectory.

domains are then remeshed and the mesh moving method is used to perform infinitesimal displacements of the entire mesh to yield the HMC gradient. The performance of the method is validated on a synthetic example and the deterministic leapfrog trajectory is shown to be exactly reversible as opposed to that updated from the previous step. More studies need to be conducted to understand the effect of non-reversibility of the trajectory on the posterior. Initial results (not shown in this paper) indicate little to no change in the posterior for the two different cases shown in Fig. 5(a) and Fig. 5(b). This result is intriguing and will be explored in the future.

As the reference domain boundary can be set to be a slightly perturbed version of the domain boundary already available at the next time step, the remeshed reference domain-based method possesses significant potential for application to trans-dimensional problems (Green 1995), where the dimensionality of parameter space can change. New discontinuities can be added or removed from the domain easily, all while the Hamiltonian is kept continuous, enabling efficient computation of HMC gradients. This opens up an exciting avenue for research which will be pursued in the future.

Acknowledgments

The authors gratefully acknowledge the support provided by JSPS KAKENHI, Grant Number 21H02304.

References

Afshar, M.H. and Domke, J. (2015) Reflection, refraction, and Hamiltonian Monte Carlo, *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2, Advances in neural information processing systems, 28.*

Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika*, 82(4), 711–732.

Koch, M.C., Fujisawa, K. and Murakami, A. (2020) Novel parameter update for a gradient based MCMC method for solid-void interface detection through elastodynamic inversion, *Probabilistic Engineering Mechanics*, 62, 103097.

Neal, R. (2011) MCMC Using Hamiltonian Dynamics, *in Handbook of markov chain monte carlo. CRC Press.*

Stein, K., Tezduyar, T. and Benney, R. (2003) Mesh moving techniques for fluid-structure interactions with large displacements, *Journal of Applied Mechanics*, 70, 58–63.