Inverse analysis in geomechanical problems using Hamiltonian Monte Carlo

Koch Michael Conrad

2020

1 Introduction

Determination of probabilistic solutions can enable engineers make more informed decisions in practical inverse problems through point solutions and confidence intervals. Numerically, this amounts to the estimation of a parameter vector that represents all the unknowns in the system and the problem is known as a parameter identification problem. In this context, inversion is attempted on two classes of parameters commonly encountered in geomechanics, spatial distribution (elastic parameters, hydraulic conductivity etc.) parameters and geometrical (solid-void interface) parameters. Probabilistic inversion has ordinarily been executed in a Bayesian sense through Markov Chain Monte Carlo or MCMC algorithms. Owing to the pioneering work of Metropolis [6], the most famous among this class of algorithms is the Metropolis Hastings algorithm. Central to the algorithm is an aperiodic and irreducible transition kernel that maintains reversibility and guarantees invariance of the underlying probability distribution. Under these conditions, the Ergodic Theorem guarantees that the sample average converges to the true average of the parameters as the number of samples goes to infinity. However, such algorithms have typically been plagued by poor acceptance rates, due to the failure in designing proper proposal densities, a problem that is compounded further by the concentration of measure as the dimensionality of the parameter vector increases. This forces the user to use smaller proposal step sizes and the algorithm takes impractical amounts of time to explore the invariant probability distribution.

Hamiltonian Monte Carlo (HMC) [7, 1] is a modern MCMC algorithm where deterministic proposals are made through Hamiltonian dynamics. Such proposals (if calculated properly) yield samples on level sets of probability and enable movement to distant regions in parameter space at reasonable acceptance rates. This yields sampling efficiencies far superior to that observed in conventional MCMC algorithms. However, an important component of numerically applied Hamiltonian dynamics (Leapfrog method) is the gradient w.r.t the parameters, the evaluation of which forms the computational bottleneck of HMC. The ability to compute this gradient and the efficiency with which it can be computed, determine the practical usage of the algorithm. This forms the motivation behind this thesis. In particular, three inversion problems are attempted in the HMC framework:

1. Identification of spatial distribution of elastic modulus from elastic wave propagation data.
2. Explicit identification of solid-void interface parameters from elastic wave propagation data.
3. Simultaneous estimation of spatial distribution of hydraulic conductivity and piping zone interface from seepage flow data.

2 Hamiltonian Monte Carlo

The invariant density of the parameter vector $\boldsymbol{\theta} \in \mathbb{R}^K$ given the observations $\mathbf{y}_{1:n} = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$, in this thesis is of the form $p(\boldsymbol{\theta}|\mathbf{y}_{1:n})$, and is also known as the posterior and computed as

$$p(\boldsymbol{\theta}|\mathbf{y}_{1:n}) = \frac{p(\mathbf{y}_{1:n}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\mathbf{y}_{1:n})}. \tag{1}$$

The first term in the numerator represents the likelihood which itself can be calculated as $p(\mathbf{y}_{1:n}|\boldsymbol{\theta}) = \prod_{k=1}^{n} \mathbb{N}(\mathbf{y}_k|\mathbf{H}\mathbf{x}_k(\boldsymbol{\theta}), \mathbf{R}_k)$, for observation models of the form

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

Here, $\mathbf{x}_k \in \mathbb{R}^D$ is the state vector, $\mathbf{H}$ is the observation model matrix and $\mathbf{y}_k \in \mathbb{R}^m$ are the observations containing some Gaussian noise $\mathbf{r}_k \sim \mathbb{N}(\mathbf{0}, \mathbf{R}_k)$. The second term is the prior density over the parameters usually considered to be a Gaussian as $p(\boldsymbol{\theta}) = \mathbb{N}(\boldsymbol{\theta}|\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is the prior covariance matrix. The sum of the negative log of the likelihood and the prior simplify computation and define a term called Potential energy $\varphi(\boldsymbol{\theta}) : \mathbb{R}^K \to \mathbb{R}$ given as

$$\begin{aligned}\varphi(\boldsymbol{\theta}) &\equiv -\log p(\mathbf{y}_{1:n}|\boldsymbol{\theta}) - \log p(\boldsymbol{\theta}) \\ &= \sum_{k=1}^{n} \frac{1}{2} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k(\boldsymbol{\theta}))^{\mathrm{T}} \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k(\boldsymbol{\theta})) + \frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\theta} + \text{const.}.\end{aligned} \tag{3}$$

Hamiltonian Monte Carlo starts with an augmentation of the parameter space $\boldsymbol{\theta} \in \mathbb{R}^K$ with auxiliary momentum variables $\mathbf{p} \in \mathbb{R}^K$ and the definition of a joint density

$$p(\boldsymbol{\theta}, \mathbf{p}|\mathbf{y}_{1:n}) = p(\mathbf{p}|\boldsymbol{\theta}, \mathbf{y}_{1:n})\, p(\boldsymbol{\theta}|\mathbf{y}_{1:n}). \tag{4}$$

The term $p(\mathbf{p}|\boldsymbol{\theta}, \mathbf{y}_{1:n})$ is assumed to be Gaussian with a fixed covariance matrix $\mathbf{M}_H$ such that $p(\mathbf{p}|\boldsymbol{\theta}, \mathbf{y}_{1:n}) = p(\mathbf{p}) = \mathbb{N}(\mathbf{p}|\mathbf{0}, \mathbf{M}_H)$. The negative log from of Eqn. (4), then defines the Hamiltonian

$$H(\boldsymbol{\theta}, \mathbf{p}) \equiv -\log p(\boldsymbol{\theta}, \mathbf{p}) = K_H(\mathbf{p}) + \varphi(\boldsymbol{\theta}), \tag{5}$$

where $K_H(\mathbf{p}) = -\log p(\mathbf{p})$ is known as the Kinetic energy. The introduction of the Hamiltonian allows for the definition of proposals through Hamiltonian dynamics which are numerically applied through the leapfrog steps

$$\mathbf{p}\left(t + \frac{\epsilon}{2}\right) = \mathbf{p}(t) - \frac{\epsilon}{2}\frac{\partial \varphi(\boldsymbol{\theta}(t))}{\partial \boldsymbol{\theta}} \tag{6}$$

$$\boldsymbol{\theta}(t + \epsilon) = \boldsymbol{\theta}(t) + \epsilon \mathbf{M}_H^{-1} \mathbf{p}\left(t + \frac{\epsilon}{2}\right) \tag{7}$$

$$\mathbf{p}(t + \epsilon) = \mathbf{p}\left(t + \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\frac{\partial \varphi(\boldsymbol{\theta}(t + \epsilon))}{\partial \boldsymbol{\theta}}. \tag{8}$$

Reversible proposals are designed by applying the Leapfrog integrator for L steps, each with step size ϵ and then flipping the momentum at the end of the trajectory. This yields deterministic proposal densities, which for transitions of the type $(\boldsymbol{\theta}, \mathbf{p}) = (\boldsymbol{\theta}_0, \mathbf{p}_0) \to (\boldsymbol{\theta}', \mathbf{p}') = (\boldsymbol{\theta}_L, -\mathbf{p}_L)$ are given

as $q\left((\boldsymbol{\theta},\mathbf{p}),(\boldsymbol{\theta}',\mathbf{p}')\right) = q\left((\boldsymbol{\theta}_0,\mathbf{p}_0),(\boldsymbol{\theta}_L,-\mathbf{p}_L)\right) = \delta\left(\boldsymbol{\theta}'-\boldsymbol{\theta}_L\right)\delta\left(\mathbf{p}'+\mathbf{p}_L\right)$. This then yields acceptance probabilities of the form

$$A\left((\boldsymbol{\theta},\mathbf{p}),(\boldsymbol{\theta}',\mathbf{p}')\right) = A\left((\boldsymbol{\theta}_0,\mathbf{p}_0),(\boldsymbol{\theta}_L,-\mathbf{p}_L)\right) = \min\left\{1, \frac{p\left(\boldsymbol{\theta}_L,-\mathbf{p}_L\right)\delta\left(\boldsymbol{\theta}_0-\boldsymbol{\theta}_0\right)\delta\left(\mathbf{p}_0-\mathbf{p}_0\right)}{p(\boldsymbol{\theta}_0,\mathbf{p}_0)\delta\left(\boldsymbol{\theta}_L-\boldsymbol{\theta}_L\right)\delta\left(-\mathbf{p}_L+\mathbf{p}_L\right)}\right\}$$

$$= \min\left\{1, \exp\left(-H\left(\boldsymbol{\theta}_L,-\mathbf{p}_L\right)+H(\boldsymbol{\theta}_0,\mathbf{p}_0)\right)\right\}. \tag{9}$$

It is clear that $q\left((\boldsymbol{\theta}_L,\mathbf{p}_L),(\boldsymbol{\theta}_0,\mathbf{p}_0)\right)$ must be positive for this exercise to be meaningful. HMC proposals which satisfy this property are highlighted in Section **4**.

3 Adjoint HMC for elastic modulus estimation through elastodynamic inversion

Consider the existence of a sample space Ω_ω and a bounded domain $\Omega \subset \mathbb{R}^N$. In this section, the spatial distribution of the elastic modulus $E(\mathbf{z},\omega) : \Omega\times\Omega_\omega \to \mathbb{R}$ is considered as a random field, with mean function $\overline{E}(\mathbf{z})$ and symmetric, positive definite covariance function $C\left(\mathbf{z}_1,\mathbf{z}_2\right)$ defined on $\Omega\times\Omega$. Any meaningful discretization of the random field for computational purposes will result in a high dimensional elastic modulus vector $\mathbf{E} = \left(E\left(\mathbf{z}_1\right),\ldots,E\left(\mathbf{z}_a\right)\right)^{\mathrm{T}}$ such that $a \gg 0$. This poses computational challenges to HMC, especially in gradient computation. As such, it is desired to define an efficient update and is the main focus of this section.

3.1 Karhunen-Loève (KL) expansion and the Adjoint method

If λ_q and $\boldsymbol{\Phi}_q$ are the eigenvalues and eigenvectors corresponding to the covariance matrix $\mathbf{C}$ and the eigen values are arranged so that $\lambda_1 > \lambda_2 > \ldots > \lambda_K$, then the K-term KL expansion [4] for the discretized random field $\mathbf{E}(\omega)$ can be expressed as

$$\mathbf{E} = \overline{\mathbf{E}} + \sum_{q=1}^{K} \sqrt{\lambda_q}\theta_q(\omega)\boldsymbol{\Phi}_q, \tag{10}$$

where, $\boldsymbol{\Phi}_q = \left(\phi_q\left(\mathbf{z}_1\right),\ldots,\phi_q\left(\mathbf{z}_a\right)\right)$ and $\theta_q(\omega)$ is an unknown random variable with zero mean and unit variance. When $K\to\infty$, it can be shown that the KL expansion converges to the true random field in a mean square sense. In most practical problems the eigen values decay exponentially and the KL expansion can be truncated at K-terms such that $K \ll a$. This reduction in dimensionality of the problem significantly speeds up HMC and limits the number of parameters w.r.t to which the gradient needs to be calculated.

The spatially discretized governing equation for elastic wave propagation in the domain is the usual second order ODE, which is now parameterized by the parameter vector $\boldsymbol{\theta}$, i.e.

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}(\boldsymbol{\theta})\dot{\mathbf{u}}(t) + \mathbf{K}(\boldsymbol{\theta})\mathbf{u}(t) = \mathbf{f}(t). \tag{11}$$

This equation is discretized in time by one-step [3] methods. A large class of implicit, unconditionally stable time integration algorithms such as Newmark-β, HHT-α, CH-α etc. fall in this class and can be written in a general form for state vectors $\mathbf{m}_k = (\mathbf{u}_k, \dot{\mathbf{u}}_k, \ddot{\mathbf{u}}_k)^{\mathrm{T}}$ as

$$\mathbf{m}_k = \mathbf{L}(\boldsymbol{\theta})\mathbf{m}_{k-1} + \mathbf{F}_k(\boldsymbol{\theta}), \quad k \in \{2, \ldots, n\}. \tag{12}$$

$\mathbf{L}(\boldsymbol{\theta})$ is the amplification matrix and $\mathbf{F}_k(\boldsymbol{\theta})$ contains contributions from the external loads $\mathbf{f}(t)$.

It is further possible to improve computational efficiency by eliminating the repeated computations of the state vector w.r.t to the parameter vector. This is achieved through the Adjoint method where an adjoint variable is introduced into Eqn. (3) in the form

$$\begin{aligned}\varphi(\boldsymbol{\theta}) = \sum_{k=1}^{n} \frac{1}{2} \left(\mathbf{y}_k - \mathbf{H}\mathbf{m}_k(\boldsymbol{\theta})\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(\mathbf{y}_k - \mathbf{H}\mathbf{m}_k(\boldsymbol{\theta})\right) + \frac{1}{2}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\theta} \\ + \sum_{k=2}^{n} \boldsymbol{\psi}_k^{\mathrm{T}} \left(\mathbf{m}_k - \mathbf{L}(\boldsymbol{\theta})\mathbf{m}_{k-1} - \mathbf{F}_k(\boldsymbol{\theta})\right).\end{aligned} \tag{13}$$

Taking a simple derivative w.r.t $\boldsymbol{\theta}$ and collecting all coefficients of $\frac{\partial \mathbf{m}_k}{\partial \theta_i}$ and making them go to zero yield the adjoint equation

$$\boldsymbol{\psi}_k = \mathbf{L}(\boldsymbol{\theta})^{\mathrm{T}}\boldsymbol{\psi}_{k+1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}\mathbf{m}_k\right), \tag{14}$$

which must be solved along with

$$\frac{\partial \varphi(\boldsymbol{\theta})}{\partial \theta_i} = -\sum_{k=2}^{n} \boldsymbol{\psi}_k^{\mathrm{T}} \left(\frac{\partial \mathbf{L}(\boldsymbol{\theta})}{\partial \theta_i}\mathbf{m}_{k-1} + \frac{\partial \mathbf{F}_k(\boldsymbol{\theta})}{\partial \theta_i}\right) + \frac{1}{2}\frac{\partial\left(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\theta}\right)}{\partial \theta_i}, \tag{15}$$

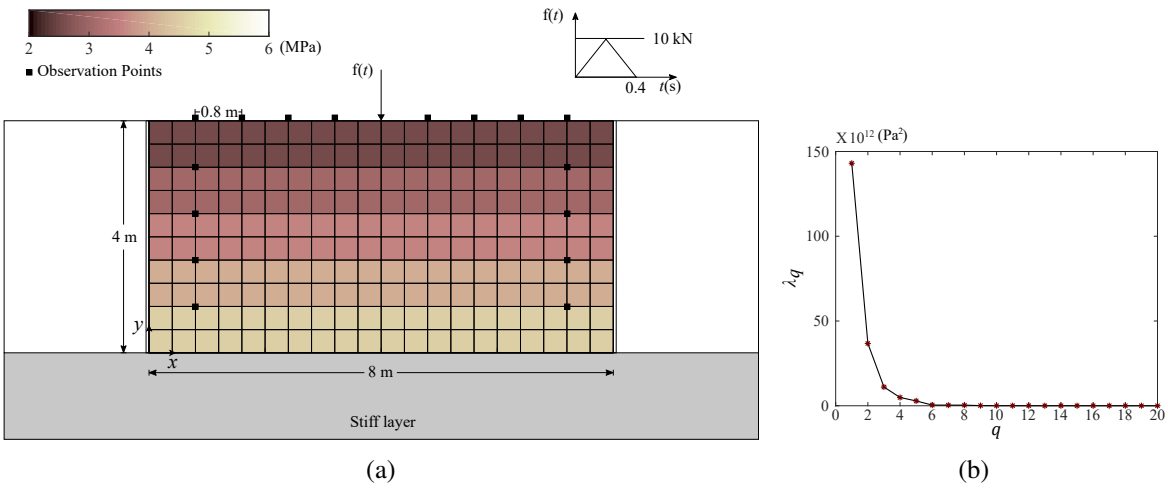
to yield the gradient. This process eliminates the need to compute $\frac{\partial \mathbf{m}_k}{\partial \theta_i}$ and further saves computation time. The KL expansion and Adjoint method are incorporated into the HMC algorithm for parameter estimation.

3.2 Numerical Implementation

A simple numerical example of estimation of elastic modulus in a rectangular domain (density = 2000 Kg/m^3 and Poisson's ratio = 0.25) from elastic wave propagation data measured at observation points marked in black is shown in Fig. 1(a). The profile shown represents the target profile. The parameter estimation problem is one over 200 parameters representing the elastic modulus of each element in the domain. A Gaussian covariance kernel is chosen with known scale ($\upsilon = 10^{12}\mathrm{m}^2$) and correlation length ($l = 4\mathrm{m}$) and is given as

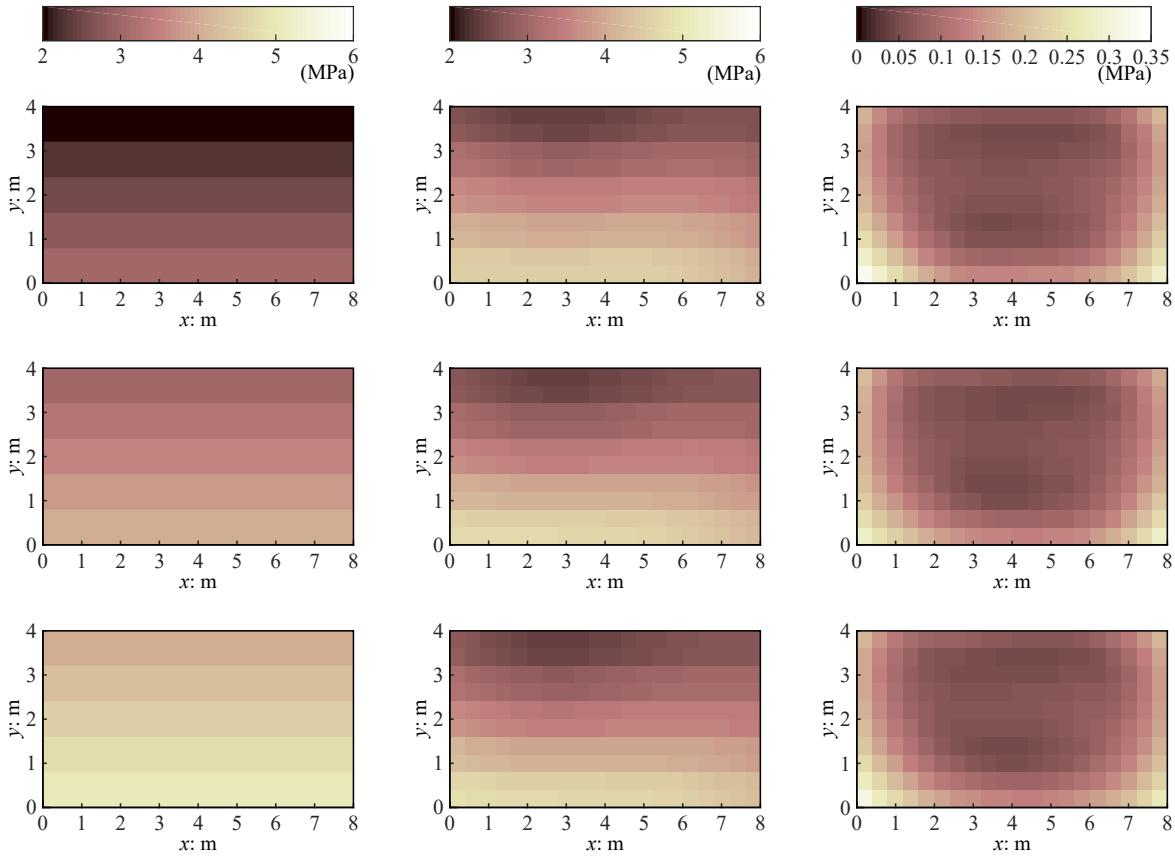
$$C\left(\mathbf{z}_i, \mathbf{z}_j\right) = \upsilon \exp\left(\frac{-\left|\mathbf{z}_i - \mathbf{z}_j\right|^2}{2l^2}\right). \tag{16}$$

The decay of the eigen values of the covariance matrix is shown in Fig. 1(b), based on which $K = 12$ terms are chosen in the KL expansion. 10000 samples are drawn using the Adjoint HMC algorithm and the first 1000 are discarded as burn-in samples, for three different starting points shown in Fig. 2.



(a) (b)

Fig. 1 (a) Discretized finite element domain showing variation of elastic modulus, observation points and point of application of triangular load. (b) Decay of eigen values of the covariance matrix



(a) (b) (c)

Fig. 2 Profiles of elastic modulus for (a) three starting points (left) and corresponding (b) Mean and (c) Standard Deviation.

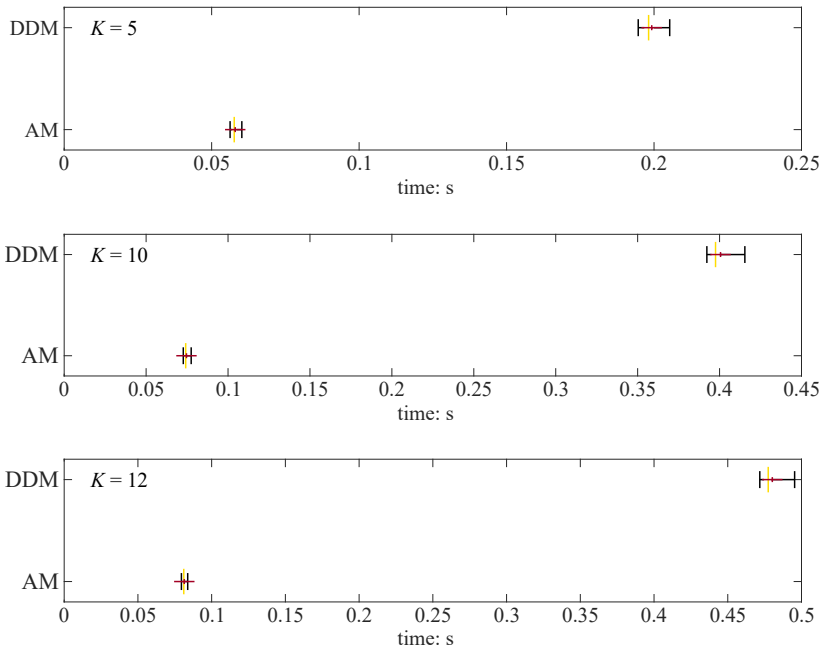


Fig. 3 Comparison of the Direct Differentiation Method (DDM) and Adjoint Method (AM) in terms of time taken for computation. Grey vertical bars represent the 5% and 95% quantiles, the yellow vertical bar represents the median and the red '+' sign represents the mean.

The mean from the three starting points shows a good match with the target distribution of elastic modulus. In the left and right lower corners the estimation isn't as accurate but this is highlighted in the larger standard deviation observed in these regions. This shows the usefulness of getting a statistical solution. The performance of the Adjoint Method is compared with the Direct Differentiation Method (plain derivative of Eqn. (3)), where the Adjoint Method clearly outperforms the DDM and the time taken for evaluation of gradients in the AM case is independent of the number of terms (K) considered.

4 Parameter update for solid-void interface detection through elasto-dynamic inversion

The second inversion problem considers the explicit identification of solid-void interfaces parameterized by the vector $\boldsymbol{\theta}$. Tomographic inversion studies based on identifying the spatial distribution of parameters such as elastic modulus, shear wave velocity profile etc. are rarely able to identify the exact location of interfaces and thereby accurately resolve the position of voids or cavities [9] such as pipelines, tunnels, sinkholes, cracks etc. Hence, the problem is changed to one of explicitly identifying the geometry of the interface. Inversion of this kind in an HMC framework requires the definition of a new parameter update that satisfies the following three requirements

1. Proposal densities must be reversible.
2. Good mesh quality must be maintained during each parameter update.

3. Update should not introduce any discontinuities ion the Hamiltonian and gradient computation should be possible.

4.1 Novel parameter update

The update is derived for a general setting of a void/cavity $\Omega_v \subset \Omega$ where $\Omega \subset \mathbb{R}^{n_{sd}}$ is the spatial domain and n_{sd} is the number of spatial dimensions. Let the node coordinates formed from a discretization of the solid domain $\Omega \backslash \Omega_v$ with $\tilde{d}$ nodes, be represented as $\mathbf{Z} \in \mathbb{R}^{\tilde{d} \times n_{sd}}$, where $\mathbf{Z} = (\mathbf{z}_1, \ldots, \mathbf{z}_d)^{\mathrm{T}}$ and $\mathbf{z}_j = \left(z_{1j}, \ldots, z_{n_{sd}j}\right)$. This discretization also produces a set of node coordinates $\mathbf{Z}_v(\boldsymbol{\theta}) = \left(\mathbf{z}_{v_1}(\boldsymbol{\theta}), \ldots, \mathbf{z}_{v_{\hat{d}}}(\boldsymbol{\theta})\right)^{\mathrm{T}}$, where $\mathbf{z}_{v_j} : \mathbb{R}^K \rightarrow \mathbb{R}^{n_{sd}}$, on the solid-void interface Γ_v, where $\hat{d}$ represents the total number of nodes at the interface.

4.1.1 Reversible proposals

For an HMC parameter update of the form $\boldsymbol{\theta}(t) \rightarrow \boldsymbol{\theta}(t+\epsilon)$, the associated solid-void interface nodal coordinate update is given by $\mathbf{Z}_v\left(\boldsymbol{\theta}(t)\right) \rightarrow \mathbf{Z}_v\left(\boldsymbol{\theta}(t+\epsilon)\right)$. This defines an update on the interface nodes. The update on the interior nodes is defined from a fixed reference mesh with nodal coordinates $\mathbf{Z}^{ref}\left(\boldsymbol{\theta}^{ref}\right) \in \mathbb{R}^{\tilde{d} \times n_{sd}}$. The update on the interface nodes is now defined through a prescribed displacement matrix

$$\mathbf{u}_v^{ref} = \mathbf{Z}_v\left(\boldsymbol{\theta}(t+\epsilon)\right) - \mathbf{Z}_v^{ref}\left(\boldsymbol{\theta}^{ref}\right). \tag{17}$$

Clubbing all other prescribed displacements in the domain to get the total displacement matrix $\mathbf{u}^{ref} \in \mathbb{R}^{\tilde{d} \times n_{sd}}$ and rearranging it as $\widetilde{\mathbf{u}}^{ref} \in \mathbb{R}^{d}$ where, $d = \tilde{d} \times n_{sd}$, results in a prescribed displacement problem on the interior nodes given by

$$\widetilde{\mathbf{u}}_{int}^{ref} = -\left(\mathbf{K}_{int}^{ref}\right)^{-1} \mathbf{K}_{pre}^{ref} \widetilde{\mathbf{u}}_{pre}^{ref}. \tag{18}$$

Here, $\widetilde{\mathbf{u}}_{int}^{ref}$ refers to the displacement on the interior nodes, $\widetilde{\mathbf{u}}_{pre}^{ref}$ the total prescribed displacements and $\mathbf{K}_{int}^{ref}$ and $\mathbf{K}_{pre}^{ref}$ represent the internal and prescribed degrees of freedom part of the global stiffness matrix $\mathbf{K}^{ref}$ on the reference mesh. The elemental reference mesh stiffness matrix is given as

$$\mathbf{K}_e^{ref} = \int_{\Omega_e} \mathbf{B}^{ref\,\mathrm{T}} \mathbf{D}^{ref} \mathbf{B}^{ref} \left| \boldsymbol{J}_e^{ref} \right| \bar{t} \rho d\Omega_e, \tag{19}$$

where all terms have their usual meanings. The complete mesh update is then given by

$$\mathbf{Z}(\boldsymbol{\theta}(t+\epsilon)) = \mathbf{Z}^{ref}\left(\boldsymbol{\theta}^{ref}\right) + \mathbf{u}^{ref}, \tag{20}$$

and is shown in Fig. 4. Proposals of this kind are reversible and the deterministic reverse proposal density $q\left((\boldsymbol{\theta}(t+2\epsilon), -\mathbf{p}(t+2\epsilon)), (\boldsymbol{\theta}(t), \mathbf{p}(t))\right) = 1$.

4.1.2 Mesh Moving Method

To eliminate anticipated mesh distortions in nodal coordinate updates from the reference mesh, a Mesh Moving method [8] is used in which the elastic modulus on the reference domain E_e^{ref} is scaled as

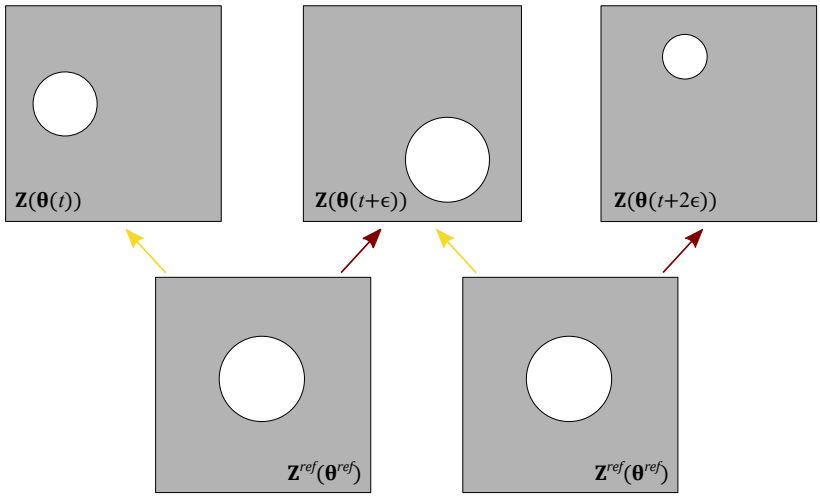


Fig. 4 Reversible update of the current mesh updated from a reference mesh.

$$E_e^{ref} = E_e \left(\frac{J^0}{\left| \boldsymbol{J}_e^{ref} \right|} \right)^{\chi}, \tag{21}$$

where J^0 is an arbitrary scaling parameter For $\chi = 1$, this results in the dropping of the Jacobian in Eqn. (19) and has the effect that small elements become more rigid and large elements deform more freely. Placement of small elements near the void/cavity where large displacements are anticipated then minimizes distortions in the domain.

4.1.3 Gradient computation

Gradients of the potential energy function $\varphi(\boldsymbol{\theta})$ involve derivatives of the mass and stiffness matrices $\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \mathbf{K}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$. On an elemental level these can be calculated as

$$\frac{\partial \mathbf{K}_e}{\partial \boldsymbol{\theta}} = \int_{\Omega_e} \left(\frac{\partial \mathbf{B}^{\mathrm{T}}}{\partial \boldsymbol{\theta}} \mathbf{D} \mathbf{B} \left| \boldsymbol{J}_e \right| + \mathbf{B}^{\mathrm{T}} \mathbf{D} \frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}} \left| \boldsymbol{J}_e \right| + \mathbf{B}^{\mathrm{T}} \mathbf{D} \mathbf{B} \frac{\partial \left| \boldsymbol{J}_e \right|}{\partial \boldsymbol{\theta}} \right) \bar{t} \rho d\Omega_e, \tag{22}$$

$$\frac{\partial \mathbf{M}_e}{\partial \boldsymbol{\theta}} = \int_{\Omega_e} \mathbf{N}^{\mathrm{T}} \mathbf{N} \frac{\partial \left| \boldsymbol{J}_e \right|}{\partial \boldsymbol{\theta}} \bar{t} \rho d\Omega_e, \tag{23}$$

where $\mathbf{K}_e$ and $\mathbf{M}_e$ are the elemental stiffness and mass matrices respectively and other terms have their usual meaning. The terms $\frac{\partial \mathbf{B}}{\partial \boldsymbol{\theta}}$ and $\frac{\partial |\boldsymbol{J}_e|}{\partial \boldsymbol{\theta}}$ are evaluated with ideas from shape optimization [2] and turn out to be functions of $\frac{\partial \mathbf{Z}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ which can readily be evaluated if $\frac{\partial \mathbf{Z}_v(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is differentiable. This defines a continuous reversible parameter update for HMC.

4.2 Numerical implementation

The novel parameter update is tested on a simple numerical example of identification of a target elliptical cavity as shown in Fig. 5, parameterized as $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^{\mathrm{T}} = \left[z_1^{\mathrm{center}}, z_2^{\mathrm{center}}, \tan\phi, b_1, b_2\right]^{\mathrm{T}} = [5, 5, 1/\sqrt{3}, 5, 2]^{\mathrm{T}}$ where, $\phi = 30°$ is the angle of inclination (measured anticlockwise) of the major axis with respect to the z_1 axis and b_1 and b_2 are the lengths of the major and minor axis respectively.

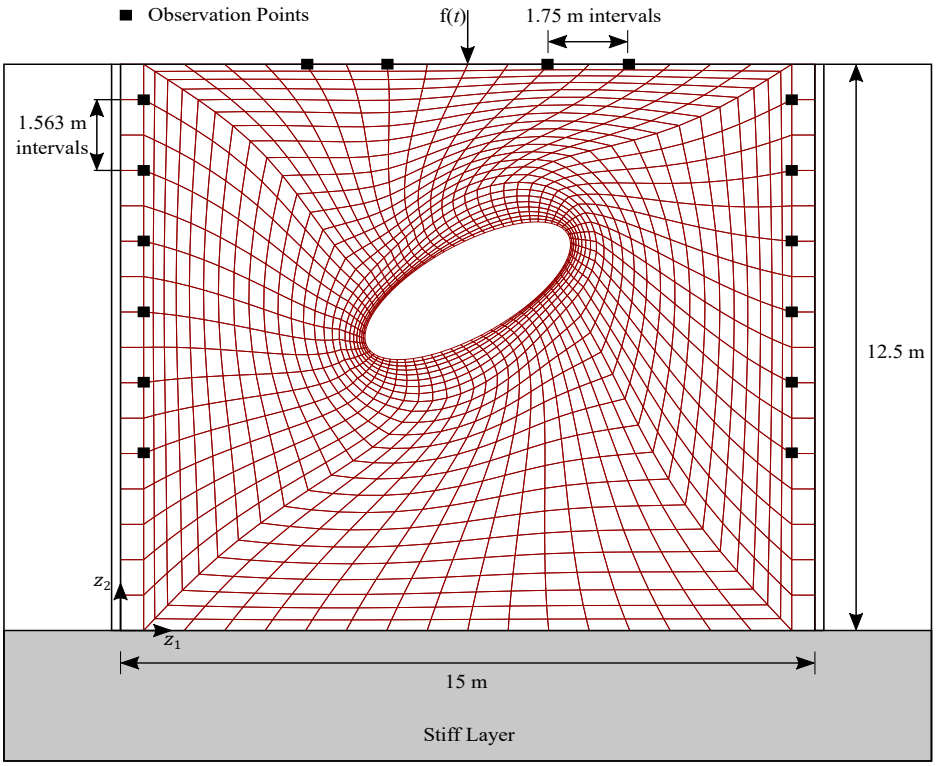


Fig. 5 Mesh showing position, size and orientation of the target elliptical solid-void interface.

HMC analysis is carried out for 2000 steps, out of which the first 200 are eliminated as burn-in, from the starting point $\boldsymbol{\theta} = [8, 5, 2-\sqrt{3}, 3, 1]^{\mathrm{T}}$ (see Fig. 6(a)). An arbitrary reference mesh is chosen with parameters $\boldsymbol{\theta}^{ref} = [5, 5, 0, 1, 1]^{\mathrm{T}}$ in Fig. 6(b). The parameterization of the elliptical interface is done such that the interface node coordinates matrix $Z_v(\boldsymbol{\theta})$ components can be defined as

$$\mathbf{z}_{v_j}(\boldsymbol{\theta}) = \begin{pmatrix} \theta_1 + \frac{\theta_4}{2}\cos\alpha_{v_j}\cos\left(\tan^{-1}\theta_3\right) - \frac{\theta_5}{2}\sin\alpha_{v_j}\sin\left(\tan^{-1}\theta_3\right) \\ \theta_2 + \frac{\theta_4}{2}\cos\alpha_{v_j}\sin\left(\tan^{-1}\theta_3\right) + \frac{\theta_5}{2}\sin\alpha_{v_j}\cos\left(\tan^{-1}\theta_3\right) \end{pmatrix}. \tag{24}$$

These equations are easily differentiable w.r.t $\boldsymbol{\theta}$ and enable computation of $\frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and $\frac{\partial \mathbf{K}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$.

The mesh at the 2000^{th} step of HMC is shown in Fig. 6(d). The good mesh quality is a clear indicator of the good performance of the mesh moving method where χ was chosen to be equal to 1. Normalized bin counts show the statistical nature of the solution. Even though the mean solution doesn't match the target solution, the target solution is covered by the normalized bin counts indicating that there is some probability for the target solution to occur.

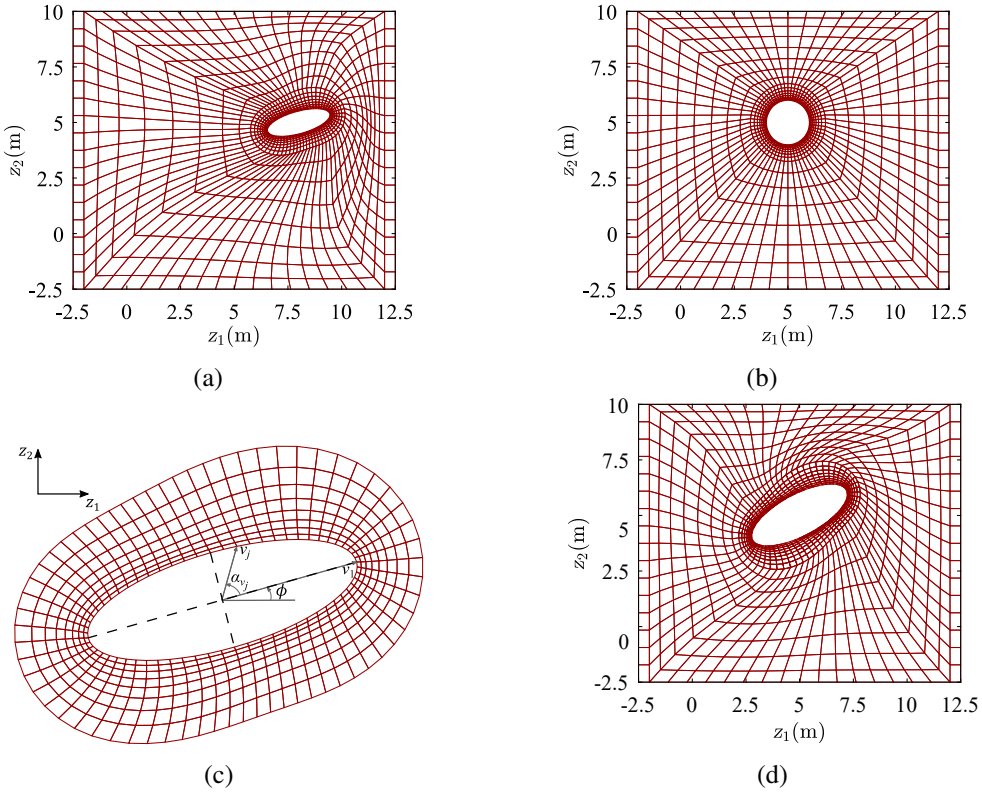


Fig. 6 (a) Initial and (b) Reference mesh for HMC. (c) Zoomed version of the fine mesh around the elliptical solid-void interface. (d) Mesh at 2000$^{\text{th}}$ step of HMC.

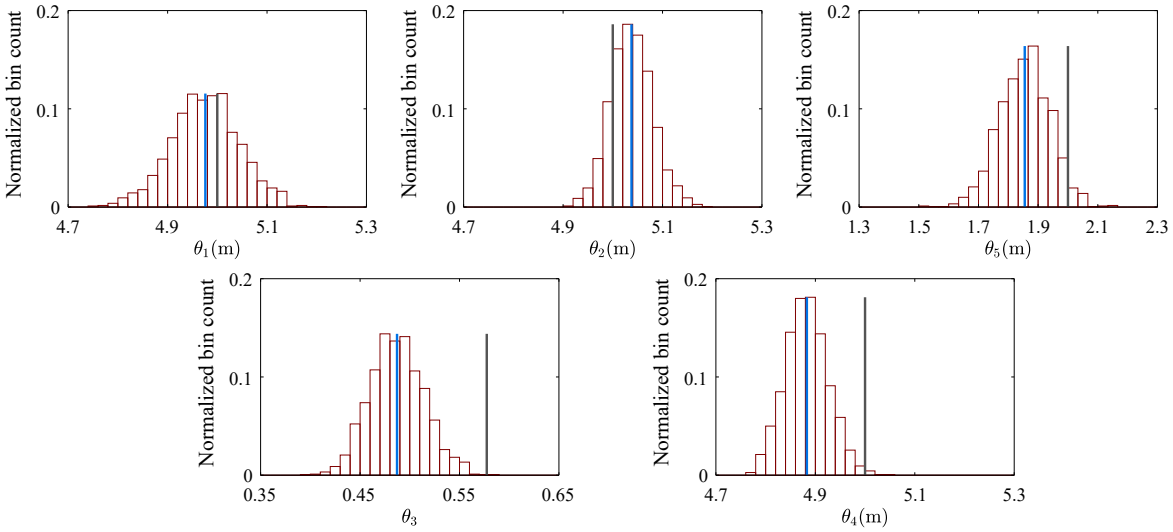


Fig. 7 Normalized bin counts (J_{bin}/J_{total}) for $\boldsymbol{\theta}$. J_{bin} and J_{total} represent the number of samples in a bin and the total number of samples respectively. Grey and blue lines represent the target and HMC mean parameter values respectively.

5 Exact and approximate implementations of HMC for simultaneous estimation of spatial distribution of hydraulic conductivity and piping zone interface

The third inversion problem combines the aims of the previous two sections for simultaneous estimation of spatial distribution of parameters and the interface of a domain. Specifically, the seepage flow forward problem is considered in a domain containing a predefined piping zone as shown in Fig. 8. The seepage flow experiment was performed by fixing the input hydraulic head at the top and the exit head at the bottom of the domain consisting of Silica sand #5. Hydraulic head was recorded at the observation points $2, \ldots, 6$ and discharge rate was measured at the exit after the seepage flow was allowed to reach steady state. This process was repeated 14 times with increasing input hydraulic head measured at point 1 on the top of the domain. The discretized governing equation for the problem is given as

$$\mathbf{K}(\boldsymbol{\theta})\mathbf{h}_k = \mathbf{q}_k, \tag{25}$$

where $\mathbf{h}_k$ and $\mathbf{q}_k \in \mathbb{R}^d$ are the global discretized hydraulic head vector and nodal flux vector respectively and $\mathbf{K}$ is the global hydraulic conductivity matrix. The vector $\boldsymbol{\theta} \in \mathbb{R}^K$ now contains both spatial field and interface parameters, which can be updated in two ways: exact and approximate.

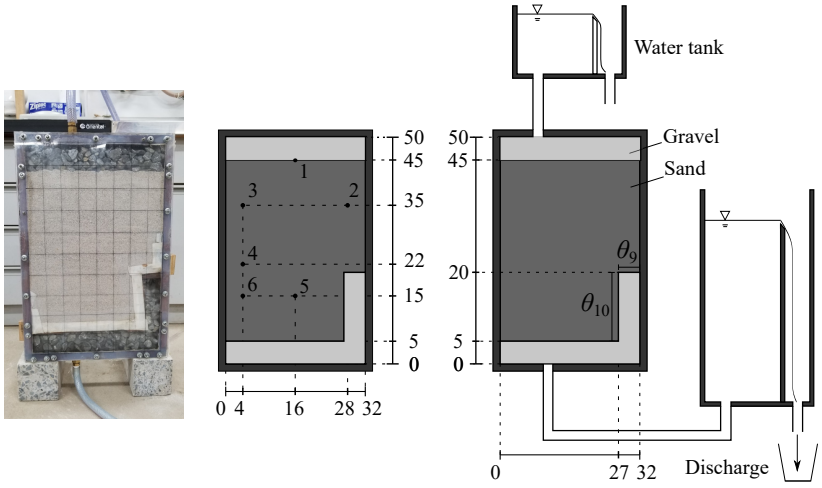


Fig. 8 Actual test apparatus (left) and corresponding illustration showing positions of observation points at which hydraulic head is recorded (middle) and piping zone dimensions (right). (unit: cms)

5.1 Exact parameter update

First the parameter vector is partitioned into two parts i.e. $\boldsymbol{\theta} = \left({}^1\boldsymbol{\theta}, {}^2\boldsymbol{\theta}\right)^{\mathrm{T}}$, where ${}^1\boldsymbol{\theta} \in \mathbb{R}^{K_1}$ correspond to the parameters defining the spatial distribution of hydraulic conductivity and ${}^2\boldsymbol{\theta} \in \mathbb{R}^{K_2}$ to those defining the solid-void interface. While the combination of the methods in the previous sections would appear to be straightforward, a new complication arises due to the update of the domain at

every step due to the update in the interface parameters. If the hydraulic conductivity is treated as a random field on a domain $\Omega\left({}^2\boldsymbol{\theta}\right) \subset \mathbb{R}^N$, then the accompanying covariance kernel $\boldsymbol{C}_{\Omega\left({}^2\boldsymbol{\theta}\right)}$ defined on $\Omega\left({}^2\boldsymbol{\theta}\right) \times \Omega\left({}^2\boldsymbol{\theta}\right)$ is now a function of the interface parameters.

Dimensionality reduction through the K-term KL expansion will result in a discretized hydraulic conductivity field $\mathbf{k}\left({}^1\boldsymbol{\theta}, {}^2\boldsymbol{\theta}\right) \in \mathbb{R}^a$ given as

$$\mathbf{k}\left({}^1\boldsymbol{\theta}, {}^2\boldsymbol{\theta}\right) = \overline{\mathbf{k}} + \sum_{q=1}^{K_1} \sqrt{\lambda_q\left({}^2\boldsymbol{\theta}\right)}\, {}^1\theta_q \boldsymbol{\Phi}_q\left({}^2\boldsymbol{\theta}\right), \tag{26}$$

where $\boldsymbol{\Phi}_q\left({}^2\boldsymbol{\theta}\right) = \left(\phi_q\left(\mathbf{z}_1^c\left({}^2\boldsymbol{\theta}\right)\right), \ldots, \phi_q\left(\mathbf{z}_a^c\left({}^2\boldsymbol{\theta}\right)\right)\right)^{\mathrm{T}}$ are the eigen vectors and $\lambda_q\left({}^2\boldsymbol{\theta}\right)$ are the eigen values of the covariance matrix $\left(\boldsymbol{C}\left({}^2\boldsymbol{\theta}\right)\right)$ computed on the domain $\Omega\left({}^2\boldsymbol{\theta}\right)$ according to Eqn. (16). The covariance matrix has to be updated at every step and eigen values and eigen vectors recomputed.

The gradient $\frac{\partial \varphi(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ necessitates the computation of the gradient of the state vectors i.e. $\frac{\partial \mathbf{m}_k}{\partial \boldsymbol{\theta}} = \left(\frac{\partial \mathbf{h}_k}{\partial \boldsymbol{\theta}}, \frac{\partial \mathbf{q}_k}{\partial \boldsymbol{\theta}}\right)^{\mathrm{T}}$. These can be obtained by a direct differentiation of Eqn. (15) given as

$$\frac{\partial \mathbf{K}}{\partial \boldsymbol{\theta}} \mathbf{h}_k + \mathbf{K} \frac{\partial \mathbf{h}_k}{\partial \boldsymbol{\theta}} = \frac{\partial \mathbf{q}_k}{\partial \boldsymbol{\theta}}. \tag{27}$$

Using standard boundary conditions in a seepage flow problem the equations can be solved for the unknowns. The derivatives $\frac{\partial \mathbf{K}}{\partial \boldsymbol{\theta}}$ comprise of the derivatives $\frac{\partial \mathbf{K}}{\partial {}^1\boldsymbol{\theta}}$ and $\frac{\partial \mathbf{K}}{\partial {}^2\boldsymbol{\theta}}$. Derivatives w.r.t to the field parameters are given by

$$\frac{\partial \mathbf{K}_e}{\partial {}^1\boldsymbol{\theta}} = \int_{\Omega_e} \mathbf{G}^{\mathrm{T}} \frac{\partial k_e\left({}^1\boldsymbol{\theta}\right)}{\partial {}^1\boldsymbol{\theta}} \mathbf{G} \left|\boldsymbol{J}_e\right| \bar{t} d\Omega_e, \tag{28}$$

where, k_e is the hydraulic conductivity of element e, $\mathbf{G}$ contains the derivatives of the elemental shape functions and other terms have the usual meaning. Following from the KL expansion in Eqn. (26), it is easy to compute the derivatives

$$\frac{\partial \mathbf{k}}{\partial {}^1\theta_q} = \sqrt{\lambda_q\left({}^2\boldsymbol{\theta}\right)} \boldsymbol{\Phi}_q\left({}^2\boldsymbol{\theta}\right). \tag{29}$$

The derivatives w.r.t the interface parameters are the same as mentioned in Section 4, with the exception that the derivatives $\frac{\partial \mathbf{k}}{\partial {}^2\boldsymbol{\theta}}$ now need to be considered. Using the same isoparametric formulation , the derivatives are given by

$$\frac{\partial \mathbf{K}_e}{\partial {}^2\boldsymbol{\theta}} = \int_{\Omega_e} \left(\frac{\partial \mathbf{G}^{\mathrm{T}}}{\partial {}^2\boldsymbol{\theta}} k_e \mathbf{G} \left|\boldsymbol{J}_e\right| + \mathbf{G}^{\mathrm{T}} k_e \frac{\partial \mathbf{G}}{\partial {}^2\boldsymbol{\theta}} \left|\boldsymbol{J}_e\right| + \mathbf{G}^{\mathrm{T}} k_e \mathbf{G} \frac{\partial \left|\boldsymbol{J}_e\right|}{\partial {}^2\boldsymbol{\theta}} + \mathbf{G}^{\mathrm{T}} \frac{\partial k_e}{\partial {}^2\boldsymbol{\theta}} \mathbf{G} \left|\boldsymbol{J}_e\right| \right) \bar{t} d\Omega_e. \tag{30}$$

Derivative of Eqn.(26), results in the following expression for $\frac{\partial \mathbf{k}}{\partial {}^2\boldsymbol{\theta}}$

$$\frac{\partial \mathbf{k}(\boldsymbol{\theta})}{\partial {}^2\boldsymbol{\theta}} = \sum_{q=1}^{K_1} \left(\frac{1}{2\sqrt{\lambda_q}} \frac{\partial \lambda_q}{\partial {}^2\boldsymbol{\theta}} \boldsymbol{\Phi}_q + \sqrt{\lambda_q} \frac{\partial \boldsymbol{\Phi}_q}{\partial {}^2\boldsymbol{\theta}} \right) {}^1\theta_q. \tag{31}$$

This necessitates the computation of the gradients of the eigen values and eigen vectors of the covariance matrix. These derivatives are given as [5]

$$\frac{\partial \lambda_q}{\partial^2 \boldsymbol{\theta}} = \boldsymbol{\Phi}_q^{\mathrm{T}} \frac{\partial \boldsymbol{C}_{\Omega(^2\boldsymbol{\theta})}}{\partial^2 \boldsymbol{\theta}} \boldsymbol{\Phi}_q \qquad , \qquad \frac{\partial \boldsymbol{\Phi}_q}{\partial^2 \boldsymbol{\theta}} = \left(\lambda_q \mathbf{I}_a - \boldsymbol{C}_{\Omega(^2\boldsymbol{\theta})}\right)^{\dagger} \frac{\partial \boldsymbol{C}_{\Omega(^2\boldsymbol{\theta})}}{\partial^2 \boldsymbol{\theta}} \boldsymbol{\Phi}_q, \tag{32}$$

where the gradient of the eigen vectors requires the computation of the Moore-Penrose inverse (represented by the dagger) of $\left(\lambda_q \mathbf{I}_a - \boldsymbol{C}_{\Omega(^2\boldsymbol{\theta})}\right)^{\dagger}$. Clearly the exact algorithm is expensive.

5.2 Inversion using experimental data

Inverse analysis is carried out where 10000 samples are drawn using the exact HMC algorithm and the first 1000 samples are rejected as burn-in. For this case $K_1 = 8$ such that ${}^1\boldsymbol{\theta} = (\theta_1, \ldots, \theta_8)$ while the true piping zone interface is parameterized by ${}^2\boldsymbol{\theta} = ({}^2\theta_1, {}^2\theta_2) = (\theta_9, \theta_{10}) = (0.05, 0.15)$. The inversion is carried out post an appropriate parameterization of the interface and the selection of an arbitrary reference mesh. The center figure in Fig. 9 shows the true interface is estimated well and lies well within the confidence intervals. Even the hydraulic conductivity is estimated well with confidence intervals lying in the range of 0.07 to 0.11 cm/s which matches well with actual samples where hydraulic conductivity lies in the range 0.10 to 0.17 cm/s. Comparison of simulated response using HMC identified parameters and those obtained from experiment show good agreement as shown in Fig. 10.

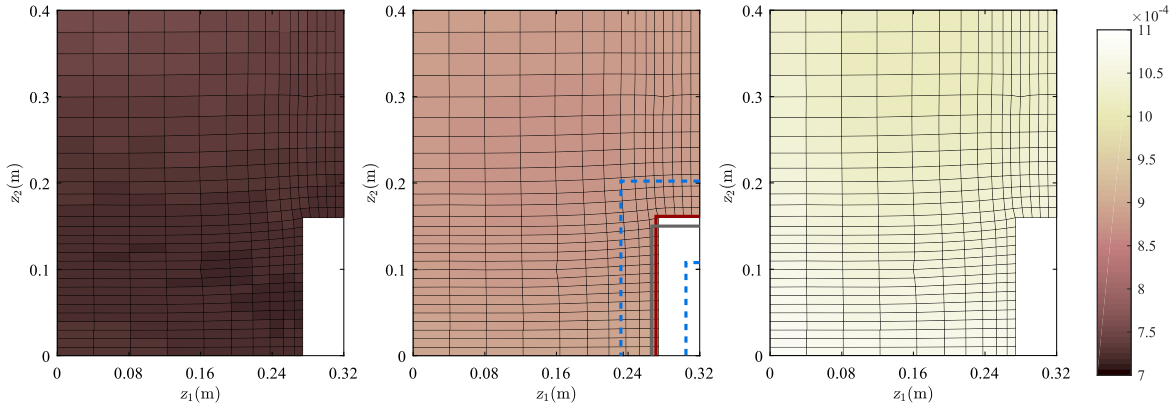


Fig. 9 Hydraulic conductivity profile of 0.01 (left) 0.5 (middle) and 0.99 (right) quantiles (Magnitudes are in m/s). Interface corresponding to 0.01 and 0.99 quantiles (blue dash lines) enclosing the mean interface (red solid line) is shown in the center figure. True interface is represented by the grey line.

5.3 Approximate parameter update

An approximate implementation of the algorithm was also attempted, where the covariance matrix ($\boldsymbol{C}_{\Omega^b}$) is fixed and approximated from a bounding domain that encloses all possible realizations from HMC. in this case, the hydraulic conductivity field is not a function of the interface parameters. Hence, the eigen value problem doesn't need to be solved at every step and the derivatives of the eigen values and eigen vectors don't need to be computed repeatedly. Comparison of the statistical performance of the two algorithms: approximate and exact through the autocorrelation function is shown in Fig. 11.

Surprisingly, the autocorrelations are almost the same for both cases, which naturally implies that the approximate algorithm will be computationally more efficient than the exact algorithm.

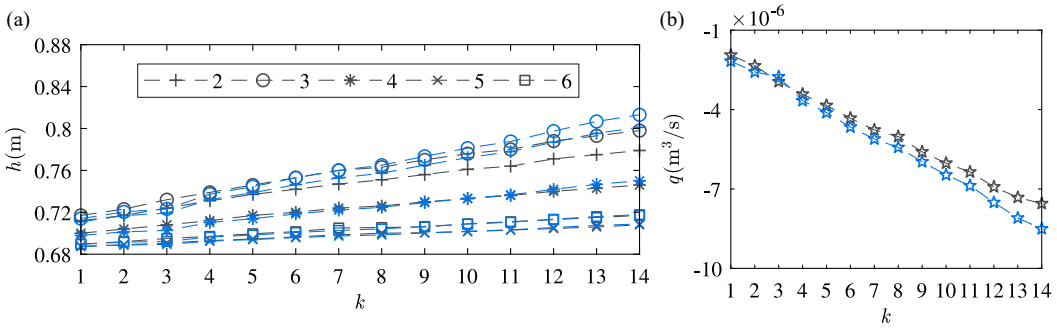


Fig. 10 Comparison of simulated response (shown in blue) and observed data (shown in grey) for (a) Hydraulic head $h_2, \ldots, h_6$ at manometers 2 to 6 and (b) Discharge rate q.

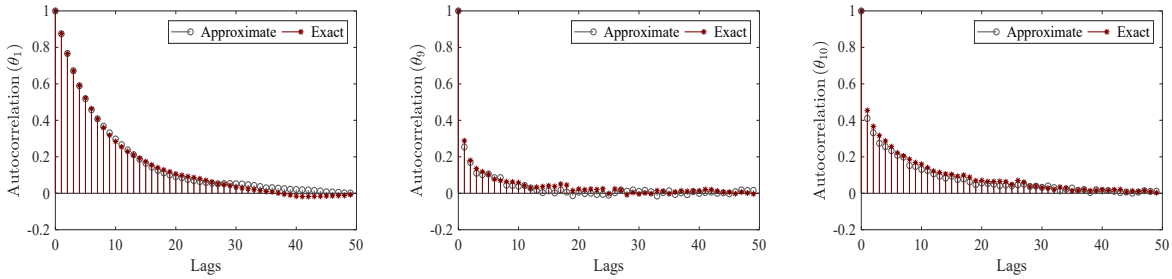


Fig. 11 Autocorrelation of θ_1 (left) and the two interface parameters θ_9 (center) and θ_{10} (right).

6 Conclusions

Three common inverse problems in geomechanics are solved in an HMC framework.

1. In the first problem, an Adjoint HMC algorithm is proposed that improves computational efficiency by reducing the dimensionality of the problem through the KL expansion and by eliminating the need to compute the gradients of large state vectors w.r.t parameters through the Adjoint Method.
2. In the second problem, a new update that maintains reversibility, maintains a good mesh quality and enables computation of gradients is proposed in an HMC framework for explicit identification of solid-void interfaces.
3. In the third problem, the two methods are combined in an exact sense, where the eigen value problem is solved and gradients are evaluated at every step of HMC. In this case the gradients of the eigen values and eigen vectors need to be computed. In the approximate sense, the covariance matrix is approximated from a fixed bounding domain and the eigen value problem and its gradients don't need to be evaluated. In this study, the approximate algorithm was found to be more computationally efficient, however, this comment cannot be generalized yet.

References

[1] Duane, S., Kennedy, A., Pendleton, B. J., and Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222.

[2] Haslinger, J. and Mäkinen, R. (2003). *Introduction to Shape Optimization- Theory, Approximation, and Computation*. SIAM, Philadelphia.

[3] Hughes, T. J. (1987). *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover, New York.

[4] Loève, M. (1978). *Probability Theory II*. Springer-Verlag New York.

[5] Magnus, J. R. (1985). On differentiating eigenvalues and eigenvectors. *Econometric Theory*, 1(2):179–191.

[6] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

[7] Neal, R. M. (2011). Handbook of Markov Chain Monte Carlo, chapter 5: MCMC Using Hamiltonian Dynamics. CRC Press.

[8] 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.

[9] Xia, J., Chen, C., Li, P. H., and Lewis, M. J. (2004). Delineation of a collapse feature in a noisy environment using a multichannel surface wave technique. *Geotechnique*, 54(1):17–27.