# FEM-BASED DISCRETIZATION-INVARIANT MCMC METHODS FOR PDE-CONSTRAINED BAYESIAN INVERSE PROBLEMS

Tan Bui-Thanh*

Department of Aerospace Engineering and Engineering Mechanics
Institute for Computational Engineering & Sciences
The University of Texas at Austin
Austin, TX 78712, USA

Quoc P. Nguyen

Department of Petroleum and Geosystems Engineering
The University of Texas at Austin
Austin, TX 78712, USA

(Communicated by Jari Kaipio)

Abstract. We present a systematic construction of FEM-based dimension-independent (discretization-invariant) Markov chain Monte Carlo (MCMC) approaches to explore PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces. In particular, we consider two frameworks to achieve this goal: Metropolize-then-discretize and discretize-then-Metropolize. The former refers to the method of discretizing function-space MCMC methods. The latter, on the other hand, first discretizes the Bayesian inverse problem and then proposes MCMC methods for the resulting discretized posterior probability density. In general, these two frameworks do not commute, that is, the resulting finite dimensional MCMC algorithms are not identical. The discretization step of the former may not be trivial since it involves both numerical analysis and probability theory, while the latter, perhaps "easier", may not be discretization-invariant using traditional approaches. This paper constructively develops finite element (FEM) discretization schemes for both frameworks and shows that both commutativity and discretization-invariance are attained. In particular, it shows how to construct discretize-then-Metropolize approaches for both Metropolis-adjusted Langevin algorithm and the hybrid Monte Carlo method that commute with their Metropolize-then-discretize counterparts. The key that enables this achievement is a proper FEM discretization of the prior, the likelihood, and the Bayes' formula, together with a correct definition of quantities such as the gradient and the covariance matrix in discretized finite dimensional parameter spaces. The implication is that practitioners can take advantage of the developments in this paper to straightforwardly construct discretization-invariant discretize-then-Metropolize MCMC for large-scale inverse problems. Numerical results for one- and two-dimensional elliptic inverse problems with up to 17899 parameters are presented to support the proposed developments.

1. **Introduction.** Solving large-scale ill-posed inverse problems that are governed by partial differential equations (PDEs), though tremendously challenging, is of great practical importance in science and industry. Classical deterministic inverse methodologies, which provide point estimates of the solution, are not capable of rigorously accounting for the uncertainty in the inverse solution. The Bayesian formulation provides a systematic quantification of uncertainty by posing the inverse problem as one of statistical inference. The Bayesian framework for inverse problems proceeds as follows: given observational data $\mathbf{d} \in \mathbb{R}^K$ and their uncertainty, the governing forward problem and its uncertainty, and a prior probability density function describing uncertainty in the parameters $\mathbf{u} \in \mathbb{R}^N$, the solution of the inverse problems is the posterior probability distribution $\pi(\mathbf{u}|\mathbf{d})$ over the parameters. Bayes' Theorem explicitly gives the posterior density as

$$\pi(\mathbf{u}|\mathbf{d}) \propto \pi_{\text{like}}(\mathbf{d}|\mathbf{u}) \times \pi_{\text{prior}}(\mathbf{u})$$

which updates the prior knowledge $\pi_{\text{prior}}(\mathbf{u})$ using the likelihood $\pi_{\text{like}}(\mathbf{d}|\mathbf{u})$. The prior encodes any knowledge or assumptions about the parameter space that we may wish to impose before any data are observed, while the likelihood explicitly represents the probability that a given set of parameters $\mathbf{u}$ might give rise to the observed data $\mathbf{d}$.

Even when the prior and noise probability distributions are Gaussian, the posterior need not be Gaussian, due to the nonlinearity embedded in the likelihood. For large-scale inverse problems, exploring non-Gaussian posterior in high dimensions to compute statistics is a grand challenge since evaluating the posterior at each point in the parameter space requires a solution of the forward PDEs. Numerical quadrature to compute the mean and covariance matrix, for example, is generally infeasible in high dimensions. Usually, the method of choice for computing statistics is Markov chain Monte Carlo (MCMC), which judiciously samples the posterior distribution, so that sample statistics can be used to approximate the exact ones.

Perhaps the Metropolis-Hastings (MH) algorithm, first developed by Metropolis *et al.* [30] and then generalized by Hastings [23], is the most popular MCMC method. Its popularity and attractiveness come from the ease of implementation and minimal requirements on the target density and the proposal density [21, 37]. One of the simplest instances of the MH algorithm is the random walk Metropolis-Hastings (RWMH) which can be considered as the Euler-Maruyama discretization of a stochastic ordinary differential equation with step size $\Delta t = \sigma^2$. The RWMH method is simple except for a small detail: how to choose the optimal step size $\sigma^2$?

Choosing the time step, also known as the proposal variance, $\sigma^2$ optimally is vital since it determines the mixing time which is approximately the number of steps to explore the stationary distribution. If $\sigma^2$ is too small, it is most likely that all the proposed moves are accepted, but the chain explores $\pi(\mathbf{u}|\mathbf{d})$ very slowly since the proposed jump is small. On the other hand, if the proposal variance is large, it is most likely that the proposed move is in low probability regions, and hence rejected. This case also leads to slow mixing since the chain virtually does not move at all. As a result, the proposal variance should be in between these extremes, and this is known as the Goldilocks principle [40].

It turns out that the proposal variance for RWMH must scale like $\sigma^2 = \ell^2 N^{-1}$, with some constant $\ell$, for the average acceptance rate to be bounded away from zero as the dimension $N$ approaches infinity [18, 38]. Similarly, one can show that the optimal scaling for the Metropolis-adjusted Langevin algorithms (MALA) [39]

is $\sigma^2 = \ell^2 N^{-1/3}$, and $\sigma^2 = \ell^2 N^{-1/4}$ for hybrid Monte Carlo (HMC) methods [1]. Clearly, all these methods suffer from vanishingly small proposal variance as the parameter dimension increases. As such they are *unlikely to be useful* for Bayesian posteriors governed by PDEs in which mesh refinement, and hence increasingly higher parameter dimension, is often needed to resolve important physical features of the problem under consideration.

Meanwhile, efforts in designing MCMC methods to explore probability measures in infinite dimensional spaces have been greatly successful, including the creation of random walk Metropolis-Hasting methods [3, 14], Metropolis-adjusted Langevin algorithms [3, 14], and hybrid Monte Carlo methods [2, 33] in function spaces. We refer to these as *function-space MCMC methods*. The appealing feature of these methods is that, *if discretized properly*, their performance, particularly the acceptance rate, is independent of the parameter dimension, and hence the mesh size. Thus, they are perhaps one of *the viable MCMC options* for large-scale PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces.

To be amenable to computer implementations, function-space MCMC methods on infinite dimensional spaces need to be discretized. Finite difference method together with Karhunen-Loève truncation [3, 14, 2, 33] has been proposed to discretize function-space MCMC methods. Care must be taken in designing discretization schemes so that they preserve/inherit the *dimension-independent, also called discretization-invariant, property. While discretization of forward PDE using FEM is widespread, a systematic FEM discretization of function-space MCMCs is not popular, and this limits the potential impact of these scalable MCMC approaches on PDE-constrained inverse problems. On the other hand, a more straightforward approach is to first discretize the PDE and the parameter space, form a Bayesian posterior for discretized inverse problems and then propose a finite dimensional MCMC method. Naive discretization schemes may lead to MCMC methods whose performance deteriorates as the parameter dimension increases, e.g. when the mesh is refined, [3, 14, 2, 33]. How to constructively develop scalable MCMC methods for this "easier" route in the context of FEM methods has attracted little attention and this may prevent practitioners from constructing dimension-independent MCMC approaches for large-scale problems.*

The *main objective* of this paper is to address the aforementioned issues. In particular, the goal is to construct FEM-based discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces. To that end, we first present an inverse problem governed by elliptic PDEs in section 2 together with a well-defined infinite dimensional Bayesian setting with prior Gaussian measure. The task at hand is to explore the Bayesian inverse problem using MCMC techniques. To accomplish this, we consider two frameworks: *Metropolize-then-discretize* and *discretize-then-Metropolize*. The former refers to the approach of first proposing a function-space MCMC method for the Bayes' posterior measure in infinite dimensional spaces and then discretizing both of them. The latter, on the other hand, first discretizes the Bayesian inverse problem and then proposes MCMC methods for the resulting discretized posterior probability density in finite dimensional spaces.

In general, Metropolize-then-discretize and discretize-then-Metropolize do not commute, that is, the resulting finite dimensional algorithms are not identical. As shall be shown, the discretization step of the former may not be trivial since it involves both numerical analysis and probability theory, while the latter may not

be dimension-independent using standard approaches. In this paper, we develop dimension-independent finite element discretizations for both frameworks. In particular, we show how to construct discretize-then-Metropolize approaches for both the Metropolis-adjusted Langevin algorithm and the hybrid Monte Carlo method that *commute with their Metropolize-then-discretize counterparts, and hence are discretization-invariant.* The key that facilitates our developments is a proper FEM discretization of the prior, the likelihood, and the Bayes' formula, together with correct definitions of quantities such as the gradient and the covariance matrix in discretized finite dimensional parameter spaces. We next summarize some important details of each section of the paper.

Section 3 presents a Metropolize-then-discretize (MTD) framework for both function-space Metropolis-adjusted Langevin algorithm (denoted as FMALA) and function-space hybrid Monte Carlo (denoted as FHMC) method. In particular, we first summarize FMALA in section 3.1 and FHMC in section 3.2. Since either of the methods requires the gradient of the negative logarithm of the likelihood, we present an adjoint approach to compute the gradient efficiently in section 3.3. We next present in detail the discretization of the prior, the likelihood, the posterior, the FMALA, and the FHMC in section 3.4. At the heart of our discretization scheme is the combination of the Galerkin finite element method, the matrix transfer technique, the Karhunen-Loève expansion, and some functional analysis. We then discuss at length the "easier" route, namely, the discretize-then-Metropolize approach in section 4. Specifically, we present an $\mathbb{R}^N$-view discretization of the posterior measure in section 4.1, finite element gradient in section 4.2, finite dimensional MALA algorithms in section 4.3, and finite dimensional HMC algorithms in section 4.4. *The key finding in section 4 is that a proper construction of finite dimensional MCMC methods taking into account the mathematical structure of the problem under consideration are discretization-invariant while standard approaches may not be.* Various numerical results to validate our developments are presented in section 5 and conclusions together with future work are discussed in section 6. Finally, we close the paper with an appendix on the Hessian, its various definitions, and its computation, which is useful for Hessian-based discretization-invariant MCMC methods.

2. **Problem statement.** In this paper, though our proposed framework is valid for Bayesian inverse problems governed by any partial differential equations (PDEs), we restrict ourselves to those with elliptic PDEs for simplicity. For concreteness, we consider heat conduction problem governed by the following elliptic PDE in an open and bounded domain $\Omega \subset \mathbb{R}^d, d = \{1, 2\}$:

$$
\begin{aligned}
-\nabla \cdot (e^u \nabla w) &= 0 && \text{in } \Omega \\
-e^u \nabla w \cdot \mathbf{n} &= Bi\, w && \text{on } \partial\Omega \setminus \Gamma_R, \\
-e^u \nabla w \cdot \mathbf{n} &= -1 && \text{on } \Gamma_R,
\end{aligned}
$$

where $w$ is the forward state, $u$ the logarithm of distributed thermal conductivity on $\Omega$, $\mathbf{n}$ the unit outward normal on $\partial\Omega$, and $Bi$ the Biot number. Here, $\Gamma_R$ is a portion of the boundary $\partial\Omega$ on which the inflow heat flux is 1. The rest of the boundary is assumed to have Robin boundary condition.

In the forward problem, the task is to solve for the temperature distribution $w$ given a description of distributed parameter $u$. In the inverse problem, the task is to reconstruct $u$ given some available observations, e.g, temperature observed

at some parts/locations of the domain $\Omega$. We choose to cast the inverse problem into the framework of PDE-constrained optimization. To begin, let us consider the following additive noise-corrupted pointwise observation model

$$(1) \qquad\qquad d_j := w\left(\mathbf{x}_j\right) + \eta_j, \quad j = 1, \ldots, K,$$

where $K$ is the total number of observation locations, $\{\mathbf{x}_j\}_{j=1}^{K}$ the set of points at which $w$ is observed, $\eta_j$ the additive noise, and $d_j$ the actual noise-corrupted observations. In this paper we work with synthetic observations and hence there is no model inadequacy in (1). Concatenating all the observations, one can rewrite (1) as

$$(2) \qquad\qquad \mathbf{d} := \mathbf{F}\left(u\right) + \boldsymbol{\eta},$$

with $\mathbf{F} := \left[w\left(\mathbf{x}_1\right), \ldots, w\left(\mathbf{x}_K\right)\right]^T$ denoting the map (known as the forward or para- meter-to-observable map) from the distributed parameter $u$ to the noise-free obser- vations, $\boldsymbol{\eta}$ being random numbers normally distributed by $\mathcal{N}\left(0, \mathbf{L}\right)$ with bounded covariance matrix $\mathbf{L}$, and $\mathbf{d} = \left[d_1, \ldots, d_K\right]^T$. For simplicity, we take $\mathbf{L} = \sigma^2 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For notational convenience, we use boldface letters for vectors and matrices.

Our inverse problem can be now formulated as

$$(3) \qquad\qquad \min_u \mathcal{J}\left(u, \mathbf{d}\right) := \frac{1}{2} \left\|\mathbf{d} - \mathbf{F}\left(u\right)\right\|_{\mathbb{R}_{\mathbf{L}}^K}^2 = \frac{1}{2\sigma^2} \sum_{j=1}^{K} \left[w\left(\mathbf{x}_j\right) - d_j\right]^2$$

subject    to    the forward problem

$$(4\text{a}) \qquad\qquad -\nabla \cdot \left(e^u \nabla w\right) = 0 \qquad\qquad\qquad \text{in } \Omega,$$

$$(4\text{b}) \qquad\qquad -e^u \nabla w \cdot \mathbf{n} = Bi\,u \qquad\qquad \text{on } \partial\Omega \setminus \Gamma_R,$$

$$(4\text{c}) \qquad\qquad -e^u \nabla w \cdot \mathbf{n} = -1 \qquad\qquad \text{on } \Gamma_R, .$$

where $\left\|\cdot\right\|_{\mathbb{R}_{\mathbf{L}}^K} := \left\|\mathbf{L}^{-\frac{1}{2}} \cdot\right\|_{\mathbb{R}^K}$ denotes the weighted Euclidean norm induced by the canonical inner product $\left\langle\cdot, \cdot\right\rangle_{\mathbb{R}^K}$ in $\mathbb{R}^K$. This optimization problem is however ill- posed. An intuitive reason is that the dimension of observations $\mathbf{d}$ is much smaller than that of the parameter $u$ (typically infinite before discretization), and hence they provide limited information about the distributed parameter $u$. As a result, the null space of the Jacobian of the parameter-to-observable map $\mathbf{F}$ is non-empty. In particular, for a class of inverse problems, we have shown that the Gauss-Newton approximation of the Hessian (which is the square of the Jacobian, and is also equal to the full Hessian of the misfit $\mathcal{J}$ with noise-free data evaluated at the optimal parameter) is a compact operator [5, 6, 7], and hence its range space is effectively finite-dimensional.

One way to overcome the ill-posedness is to use *Tikhonov regularization* (see, e.g., [48]), which proposes to augment the cost functional (3) with a quadratic term, i.e.,

$$(5) \qquad\qquad \tilde{\mathcal{J}} := \frac{1}{2} \left\|\mathbf{d} - \mathbf{F}\left(u\right)\right\|_{\mathbb{R}_{\mathbf{L}}^K}^2 + \frac{\alpha}{2} \left\|R^{1/2} u\right\|^2,$$

where $\alpha$ is a regularization parameter, $R$ some regularization operator, and $\left\|\cdot\right\|$ some appropriate norm. This method is a representative of deterministic inverse solution techniques that typically do not take into account the randomness due to measurements and other sources, though one can equip the deterministic solution with a confidence region by post-processing (see, e.g., [47] and references therein).

It should be pointed out that if the regularization term is replaced by the Cameron-Martin norm of $u$ (the second term in (10) below), the Tikhonov solution is in fact identical to the maximum a posteriori point in (10). However, such a point estimate is insufficient for the purpose of fully taking the randomness into account.

In this paper, we choose to tackle the ill-posedness using a *Bayesian* framework [17, 27, 26, 43, 34]. We seek a statistical description of all possible $u$ that conforms to some prior knowledge and at the same time is consistent with the observations. The Bayesian approach does this by reformulating the inverse problem as a problem in *statistical inference*, incorporating uncertainties in the observations, the forward map $\mathbf{F}$, and prior information. This approach is appealing since it can incorporate most, if not all, kinds of randomness in a systematic manner. To begin, we postulate a Gaussian measure $\mu := \mathcal{N}(u_0, \mathcal{C})$ with mean function $u_0$ and covariance operator $\mathcal{C}$ on $u$ in $L^2(\Omega)$ where

$$(6) \qquad \mathcal{C} := \alpha^{-1}(I - \Delta)^{-s} =: \alpha^{-1}\mathcal{A}^{-s}$$

with the domain of definition of $\mathcal{A}$ defined as

$$(7) \qquad D(\mathcal{A}) := \left\{ u \in H^2(\Omega) : \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega \right\}.$$

Here, $H^2(\Omega)$ is the usual Sobolev space. Assume that the mean function $u_0$ lives in the Cameron-Martin space of $\mathcal{C}$, then one can show (see, e.g., [43] for the details) that the prior measure $\mu$ is well-defined when $s > d/2$ ($d$ is the spatial dimension), and in that case, any realization from the prior distribution $\mu$ almost surely resides in the Hölder space $X := C^{0,\beta}(\Omega)$ with $0 < \beta < 1 \wedge (s - d/2)$. That is, $\mu(X) = 1$, and the Bayesian posterior measure $\nu$ satisfies the Radon-Nikodym derivative

$$(8) \qquad \frac{\partial\nu}{\partial\mu}(u|\mathbf{d}) \propto \exp(-\mathcal{J}(u,\mathbf{d})) = \exp\left( -\frac{1}{2}\|\mathbf{d} - \mathbf{F}(u)\|_{\mathbb{R}^K_{\mathsf{L}}}^2 \right),$$

if $\mathbf{F}$ is a continuous map from $X$ to $\mathbb{R}^K$. Note that the Radon-Nikodym derivative is proportional to the the likelihood defined by

$$(9) \qquad \pi_{\text{like}}(\mathbf{d}|u) \propto \exp(-\mathcal{J}(u,\mathbf{d})).$$

The maximum a posteriori (MAP) point (see, e.g., [43, 15] for the definition of the MAP point in infinite dimensional settings) is defined as

$$(10) \qquad u^{MAP} := \arg\min_u \tilde{\mathcal{J}}(u,\mathbf{d}) := \frac{1}{2}\|\mathbf{d} - \mathbf{F}(u)\|_{\mathbb{R}^K_{\mathsf{L}}}^2 + \frac{1}{2}\|u - u_0\|_{\mathcal{C}}^2,$$

where $\|\cdot\|_{\mathcal{C}} := \left\|\mathcal{C}^{-\frac{1}{2}}\cdot\right\|_{L^2(\Omega)}$ denotes the weighted $L^2(\Omega)$ norm induced by the $L^2(\Omega)$ inner product $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$; in fact $\|\cdot\|_{\mathcal{C}}$ is the Cameron-Martin norm. We shall also use $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ to denote the duality pairing on $L^2(\Omega)$. Note that, by definition, the MAP point $u^{MAP}$ is a function in the Cameron-Martin space.

It should be pointed out that the last term in (10) can be considered as a regularization. Indeed, the MAP point can be considered as a solution to a deterministic inverse problem. However, the Bayesian approach provides a more complete picture of the solution. In particular, the posterior, i.e. the Bayesian solution, encodes all possible solutions with associated confidence (probability). The task at hand is to provide not only a solution but also its associated uncertainty. Perhaps the most general tool to accomplish this task is the Markov chain Monte Carlo (MCMC) framework [37] that we now discuss.

3. **Metropolize-then-discretize framework.** In this section we present a Metropolize-then-discretize framework in which Markov chain Monte Carlo (MCMC) methods are first introduced on function spaces for the Bayesian posterior (8) and then discretized using finite element method. We restrict ourselves to two classes of MCMC methods: the preconditioned Metropolis-adjusted Langevin algorithm, and the hybrid Monte Carlo (also known as Hamiltonian Monte Carlo) method.

3.1. **Function-space Metropolis-adjusted Langevin algorithm (FMALA).** To explore the posterior given by the Radon-Nikodym derivative (8) with nonlinear misfit functional $\mathcal{J}(u)$ we can use the preconditioned Crank-Nicholson Langevin MCMC algorithm developed in [3, 14]. The key feature of function-space MCMC methods is that, if properly discretized, the acceptance rate of resulting discretized methods is dimension-independent [3], and hence practical for large-scale inverse problems in which the mesh needs to be refined or adapted. Finite dimensional MCMC methods such as those in [38, 39] whose proposal variances (and hence acceptance rate) depend on the dimension of a discretization under consideration is unlikely to be useful for our target large-scale applications. Indeed, existing works [38, 39] show that the proposal variance must vanish as the parameter dimension approaches infinity for the acceptance rate to be bounded away from zero. The detailed derivation and analysis of the function-space Langevin MCMC method can be consulted in [3]. Let us summarize some important points that are needed in this paper.

To define a Metropolis-Hastings MCMC algorithm for the posterior measure $\nu$ on $L^2(\Omega)$, following [45], let us start by denoting $q(u, dv)$ as the transition (also known as proposal) measure on $L^2(\Omega)$ given $u$. Furthermore, let

$$\text{(11)} \qquad \eta(du, dv) := \nu(du)\, q(u, dv)$$

be a joint measure on the product space $L^2(\Omega) \times L^2(\Omega)$. The transpose measure is then defined as $\eta^T(du, dv) := \eta(dv, du)$. The standard Metropolis-Hastings MCMC [37] can be extended to function space [45] with the following acceptance probability

$$\text{(12)} \qquad a(u, v) := \min\left\{1, \frac{d\eta^T}{d\eta}(u, v)\right\},$$

where $\frac{d\eta^T}{d\eta}(u, v)$ is the Radon-Nikodym derivative of $\eta^T$ with respect to $\eta$. What remains to be defined are the transition measure $q(u, dv)$ and the explicit form of the acceptance probability (12). Care must be taken in designing $q(u, dv)$ so that $\eta$ is not singular with respect to $\eta^T$ since on infinite dimensional spaces even Gaussian probability measures are prone to be mutually singular due to the Feldman-Hajek theorem [35]. Following [22] we consider the following stochastic PDE

$$\text{(13)} \qquad \frac{du}{dt} = -u - \mathcal{C}\nabla\mathcal{J}(u) + \sqrt{2}\frac{db}{dt},$$

with $\nabla\mathcal{J}(u)$ being the Fréchet derivative of $\mathcal{J}(u)$, and $b$ being the nuclear $\mathcal{C}-$Wiener process [35] on $L^2(\Omega)$. Note that the "product" $\mathcal{C}\nabla\mathcal{J}(u)$ makes sense since $\mathcal{C}$ is an operator from the dual space $\left[L^2(\Omega)\right]'$ (the space of linear continuous functionals on $L^2(\Omega)$) to $L^2(\Omega)$, and $\nabla\mathcal{J}(u)$ is a member of $\left[L^2(\Omega)\right]'$. *We conventionally write $\nabla\mathcal{J}(u; z)$ to denote the action of $\nabla\mathcal{J}(u)$ on any $z \in L^2(\Omega)$.* One can show (see, e.g., [22, 35]) that $\nu$ is the invariant measure of (13). Now discretizing (13) with a preconditioned simplified Crank-Nicholson method [3] with time step $\Delta t$ we

obtain the proposal $v$ given $u$ as

$$(14) \qquad v = \frac{2 - \Delta t}{2 + \Delta t} u - \frac{2\Delta t}{2 + \Delta t} \mathcal{C}\nabla\mathcal{J}(u) + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\xi,$$

where $\xi$ is distributed by the prior Gaussian measure $\mu$. We define the proposal measure $q(u, dv)$ by (14). Then the work in [14] shows that

$$(15) \qquad \frac{d\eta^T}{d\eta}(u, v) = \exp\left[\rho(u, v) - \rho(v, u)\right],$$

where $\rho(u, v)$ is defined as

$$(16) \quad \rho(u, v) = \mathcal{J}(u) + \frac{1}{2}\nabla\mathcal{J}(u; v - u) + \frac{\Delta t}{4}\nabla\mathcal{J}(u; v + u)$$
$$+ \frac{\Delta t}{4}\left\langle \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}(u), \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}(u)\right\rangle_{L^2(\Omega)}.$$

The expressions (14), (15), and (16) constitute the preconditioned function-space Metropolis-adjusted Langevin algorithm (FMALA) that was proposed in [14]. Clearly, if we ignore terms involving $\nabla\mathcal{J}$, the preconditioned Crank-Nicholson random walk [14] is recovered.

3.2. **Function-space hybrid Monte Carlo method (FHMC).** In this section, we recall the hybrid Monte Carlo, also known as Hamiltonian Monte Carlo, method. In particular, we follow the Hilbert space setting presented in [2]. To begin, let us denote by $\Delta t$ the time step, $L$ the number of time steps, $t$ the artificial time, and $\vartheta$ the artificial velocity. Consider the following Hamiltonian

$$(17) \qquad \mathcal{H}(u, \vartheta) := \frac{1}{2}\left\langle \vartheta, \mathcal{C}^{-1}\vartheta\right\rangle_{L^2(\Omega)} + \mathcal{J}(u) + \frac{1}{2}\left\langle u, \mathcal{C}^{-1}u\right\rangle_{L^2(\Omega)},$$

and the following discrete dynamic

$$u\left(t + \frac{\Delta t}{2}\right) = u(t)$$

$$\vartheta\left(t + \frac{\Delta t}{2}\right) = \vartheta(t) - \frac{\Delta t}{2}\mathcal{C}\nabla\mathcal{J}(u(t)),$$

$$u\left(t + \frac{3\Delta t}{2}\right) = u\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t) + \vartheta\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t)$$

$$\vartheta\left(t + \frac{3\Delta t}{2}\right) = -u\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t) + \vartheta\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t)$$

$$u(t + 2\Delta t) = u\left(t + \frac{3\Delta t}{2}\right)$$

$$(18) \qquad \vartheta(t + 2\Delta t) = \vartheta\left(t + \frac{3\Delta t}{2}\right) - \frac{\Delta t}{2}\mathcal{C}\nabla\mathcal{J}\left(u\left(t + \frac{3\Delta t}{2}\right)\right)$$

which corresponds to integrating the splitting Hamiltonian dynamics [2]

$$(19) \qquad \begin{array}{l} \frac{du}{dt} = 0, \\ \frac{d\vartheta}{dt} = -\mathcal{C}\nabla\mathcal{J}(u), \end{array} \quad \text{and} \quad \begin{array}{l} \frac{du}{dt} = \vartheta, \\ \frac{d\vartheta}{dt} = -u. \end{array}$$

Since the discrete dynamics (18) only approximates the continuous counterpart (19), the Hamiltonian (17) is not preserved, though volume-preserving and reversibility are guaranteed. Thus, one needs to equip the HMC approach with an

acceptance/rejection rule to ensure the convergence of the Markov chain. Following [2], we define the acceptance probability as

$$(20) \qquad a\left(\left(u^1, \vartheta^1\right), \left(u^L, \vartheta^L\right)\right) := \min\left\{1, \exp\left(-\Delta\mathscr{H}\left(\left(u^1, \vartheta^1\right), \left(u^L, \vartheta^L\right)\right)\right)\right\},$$

where

$$\Delta\mathscr{H}\left(\left(u^1, \vartheta^1\right), \left(u^L, \vartheta^L\right)\right) = \mathcal{J}\left(u^L\right) - \mathcal{J}\left(u^1\right) - \Delta t \sum_{i=2}^{L-1} \left\langle \nabla\mathcal{J}\left(u^i\right), \vartheta^i \right\rangle_{L^2(\Omega)}$$

$$(21)$$

$$-\frac{\Delta t}{2}\left(\left\langle \nabla\mathcal{J}\left(u^1\right), \vartheta^1 \right\rangle_{L^2(\Omega)} + \left\langle \nabla\mathcal{J}\left(u^L\right), \vartheta^L \right\rangle_{L^2(\Omega)}\right)$$

$$+\frac{\Delta t^2}{8}\left(\left\langle \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}\left(u^1\right), \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}\left(u^1\right) \right\rangle_{L^2(\Omega)} - \left\langle \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}\left(u^L\right), \mathcal{C}^{\frac{1}{2}}\nabla\mathcal{J}\left(u^L\right) \right\rangle_{L^2(\Omega)}\right),$$

with $u^i, \vartheta^i$ as the intermediate steps of the integration scheme (18) defined in Algorithm 1. Note that this algorithm is used to generate one state/sample of the HMC chain, that is, it needs to be executed $N_s$ times if $N_s$ samples are required. Here, $u(0)$ is chosen as the last sample or as the MAP point for the first sample computation.

---

**Algorithm 1** FHMC Algorithm

---

Set $t = 0$
Initialize $u(0)$
Draw $\vartheta(0) \sim \mu$
**for** $i = 1, \ldots, L-1$ **do**
   1. Set $u^i := u(2(i-1)\Delta t)$
   2. Set $\vartheta^i := \vartheta(2(i-1)\Delta t)$
   3. Simulate (18)
**end for**
Set $u^L := u(2L\Delta t)$
Set $\vartheta^L := \vartheta(2L\Delta t)$
Compute the acceptance probability (20)
**Accept** $u^L$ with probability $a\left(\left(u^1, \vartheta^1\right), \left(u^L, \vartheta^L\right)\right)$

---

Note that the Hamiltonian difference in (21) is not in the standard form. The reason is that the term $\left\langle u, \mathcal{C}^{-1}u \right\rangle_{L^2(\Omega)}$ (and $\left\langle \vartheta, \mathcal{C}^{-1}\vartheta \right\rangle_{L^2(\Omega)}$) in the formal definition of the Hamiltonian (17) is almost surely infinite. It is only finite for $u$ in the Cameron-Martin space of $\mathcal{C}$, whose measure is zero under the prior Gaussian measure. Expression (21) takes into account the cancellation of almost surely infinite terms on the continuous level before discretization. As shall be demonstrated by numerical results in section 5, this is critical to the success of HMC methods in high dimensional parameter spaces.

3.3. **Adjoint computation of the derivative.** As can be seen in the previous sections, both FMALA and FHMC methods require the derivative $\nabla\mathcal{J}(u)$ of the cost functional (3). In this section, we present an adjoint technique to compute the derivative efficiently. We start by considering the weak form of the forward equation (4):

$$(22) \qquad \int_\Omega e^u \nabla w \cdot \nabla\hat{\lambda}\, d\Omega + \int_{\partial\Omega\backslash\Gamma_R} Bi\, w\hat{\lambda}\, ds = \int_{\Gamma_R} \hat{\lambda}\, ds,$$

with $\hat{\lambda}$ as the test function. Using the standard reduced space approach (see, e.g., a general discussion in [32] and a detailed derivation in [8]) one can show that the Fréchet derivative (assumed to exist) $\nabla \mathcal{J}(u; \cdot)$ of the cost functional $\mathcal{J}(u)$ acting in any direction $\tilde{u}$ is given by

$$(23) \qquad \nabla \mathcal{J}(u; \tilde{u}) = \int_\Omega \tilde{u} e^u \nabla w \cdot \nabla \lambda \, d\Omega,$$

where the adjoint state $\lambda$ satisfies the adjoint equation

$$(24) \qquad \int_\Omega e^u \nabla \lambda \cdot \nabla \hat{w} \, d\Omega + \int_{\partial\Omega \setminus \Gamma_R} Bi \, \lambda \hat{w} \, ds = -\frac{1}{\sigma^2} \sum_{j=1}^K \left( w(\mathbf{x}_j) - d_j \right) \hat{w}(\mathbf{x}_j),$$

with $\hat{w}$ as the test function. The procedure for computing the gradient acting on an arbitrary direction $\tilde{u}$ is now clear. One first solves the forward equation (22) for $w$, then solves the adjoint (24) for $\lambda$, and finally evaluates (23).

3.4. **Discretization using finite element method.** Clearly, for computer implementation, we need to discretize the prior, the likelihood, and hence the posterior. This section presents our contribution to the Metropolize-then-discretize approach, namely the discretization using the finite element method. In particular, we employ the standard $H^1(\Omega)$ finite element method (FEM) to discretize the forward and adjoint (the likelihood), and the operator $\mathcal{A}$ (the prior)[1]. For convenience, we further assume that the discretized state and parameter reside on the same finite element mesh. Since FEM approximation of elliptic operators is standard (see, e.g., [13]), we will not discuss it here.

3.4.1. *Finite element approximation of the prior.* In this section, we describe a combination of FEM and the matrix transfer technique (see, e.g. [25] and the references therein, for the detail) to discretize the truncated Karhunen-Loève (KL) expansions. Note that this is neither the unique nor the best way to discretize a Gaussian measure. In fact, it is chosen simply for convenience and the readers are referred to, for example, [10, 28] for a different approach. Let us start by defining $\mathcal{Q} := \mathcal{A}^{-s/2}$, then the eigenpairs $\left( \lambda_i, v^i \right)$ of $\mathcal{Q}$ define the KL expansion of the prior distribution as

$$(25) \qquad u = u_0 + \frac{1}{\sqrt{\alpha}} \sum_{i=1}^\infty a_i \lambda_i v^i,$$

where $a_i \sim \mathcal{N}(0, 1)$ (see, e.g., [35] for the details). The truncated KL expansion for $u$ with $N$ modes, *again denoted by $u$ for convenience*, reads

$$(26) \qquad u = u_0 + \frac{1}{\sqrt{\alpha}} \sum_{i=1}^N a_i \lambda_i v^i.$$

Clearly, we need to solve the following eigenvalue problem to determine $\left( \lambda_i, v^i \right)$:

$$\mathcal{Q} v^i = \lambda_i v^i,$$

or equivalently

$$(27) \qquad \mathcal{A}^{s/2} v^i = \frac{1}{\lambda_i} v^i.$$

---

[1]It should be pointed out that the Cameron-Martin space can be shown (see, e.g., [43]) to be a subspace of the usual fractional Sobolev space $H^s(\Omega)$, which is in turn a subspace of $H^1(\Omega)$ for spatial dimension greater than one. Thus, we are using a non-conforming FEM approach.

We choose to solve (27) using the matrix transfer technique (MTT). One of the main ideas of MTT is to indirectly discretize a fractional Laplacian using a discretization of the standard Laplacian. To that end, let us denote by $\mathbf{M}$ the mass matrix, and $\mathbf{K}$ the stiffness matrix resulting from the discretization of the Laplacian $\Delta$ using the FEM method. The representation of $\mathcal{A}$ (see, e.g., [42]) in the finite element space is given by

$$\mathbf{A} := \mathbf{M}^{-1}\mathbf{K} + \mathbf{I}.$$

We define $\left(\sigma_i, \mathbf{v}^i\right)$ as eigenpairs for $\mathbf{A}$, i.e,

$$\mathbf{A}\mathbf{v}^i = \sigma_i\mathbf{v}^i, \text{ or } \mathbf{A}\mathbf{V} = \mathbf{\Sigma}\mathbf{V}$$

where $\mathbf{\Sigma}$ is the diagonal matrix with entries $\sigma_i$ and the $i$th column of $\mathbf{V}$ is $\mathbf{v}^i$. It follows that

(28) $$\left(\mathbf{v}^i\right)^T \mathbf{M}\mathbf{v}^j = \delta_{ij}, \text{ or } \mathbf{V}^{-1} = \mathbf{V}^T\mathbf{M},$$

with $\delta_{ij}$ as the Kronecker delta function, i.e., $\mathbf{v}^i$ are $\mathbf{M}$-orthogonal. Note that $\mathbf{v}^i$ is in fact the FEM nodal vector of $v^i$. Since $\mathbf{A}$ is similar to $\mathbf{M}^{-\frac{1}{2}}\left(\mathbf{K} + \mathbf{M}\right)\mathbf{M}^{-\frac{1}{2}}$, a symmetric positive definite matrix, $\mathbf{A}$ has positive eigenvalues. Using MTT method, the matrix representation of (27) reads

$$\mathbf{A}^{s/2}\mathbf{v}^i = \frac{1}{\lambda_i}\mathbf{v}^i,$$

where

(29) $$\mathbf{A}^{s/2} := \mathbf{V}\mathbf{\Sigma}^{s/2}\mathbf{V}^{-1}.$$

It follows that

(30) $$\lambda_i = \sigma_i^{-s/2}.$$

We use the Lagrangian FEM to discretize the prior so that the FEM approximation of the parameter is given by the FEM ansatz

(31) $$u = \sum_{i=1}^{N} \mathbf{u}_i\varphi_i,$$

where $\mathbf{u}_i$ are the nodal values, $\varphi_i$ the nodal Lagrange basis functions, and $N$ the total number of FEM basis functions. *Note that the number of FEM basis functions does not need to be the same as the number of truncated modes in (26), but they are chosen to be the same for sake of convenience. This particular choice implies that more modes (higher parameter dimension) implies finer mesh and vice versa.* From now on to the end of the paper, let boldface symbol denote the corresponding vector of FEM nodal values, e.g., $\mathbf{u}$ is the vector containing all FEM nodal values, $\mathbf{u}_i$, of $\mathbf{u}$. Again, for ease in writing, we have used the same notation $u$ for both infinite dimensional quantity and its FEM approximation.

Substituting the FEM ansatz for both $u$ and $v^j$ into the truncated KL expansion (26), and using the Galerkin projection we obtain

(32) $$\mathbf{u} = \mathbf{u}_0 + \frac{1}{\sqrt{\alpha}}\sum_{i=1}^{N} a_i\lambda_i\mathbf{v}_i = \mathbf{u}_0 + \frac{1}{\sqrt{\alpha}}\mathbf{V}\mathbf{\Lambda}\mathbf{a},$$

after eliminating the mass matrix $\mathbf{M}$. Here, $\mathbf{\Lambda}$ is an $N \times N$ diagonal matrix with entries $\lambda_i$ and $\mathbf{a} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right)$ is a random vector whose components are $a_i$.

Since $u \in L^2(\Omega)$, $\mathbf{u}$ belongs to $\mathbb{R}_{\mathbf{M}}^N$, i.e. the Euclidean space with weighted inner product

$$\langle \cdot, \cdot \rangle_{\mathbb{R}_{\mathbf{M}}^N} := \langle \cdot, \cdot \rangle_{\mathbf{M}} := \langle \cdot, \mathbf{M} \cdot \rangle_{\mathbb{R}^N},$$

so that the induced norm $\|\mathbf{u}\|_{\mathbb{R}_{\mathbf{M}}^N}$ approaches $\|u\|_{L^2(\Omega)}$ as $N \to \infty$. Let us emphasize that the mass matrix $\mathbf{M}$, i.e. the discretized Riesz matrix in $L^2(\Omega)$, is the natural weight coming from the $L^2(\Omega)$-setting of the prior.

A question arises: what is the distribution of the FEM nodal vector $\mathbf{u}$? One can see from (32) that $\mathbf{u}$ is a Gaussian since $a_i$ are Gaussian. What remains to be done is to determine the mean and covariance matrix of $\mathbf{u}$. It is easy to see, by taking expectation of both sides of (32) with respect to $\mathbf{a}$ and using the fact that $\mathbb{E}_{\mathbf{a}}[a_i] = 0$, that the mean is $\mathbf{u}_0$. We find the covariance matrix $\mathbf{c}$ using the definition of random vectors in $\mathbb{R}_{\mathbf{M}}^N$, i.e.,

$$\langle \mathbf{z}, \mathbf{c}\mathbf{y} \rangle_{\mathbf{M}} := \mathbb{E}_{\mathbf{u}}\left[\langle \mathbf{u} - \mathbf{u}_0, \mathbf{z} \rangle_{\mathbf{M}} \langle \mathbf{u} - \mathbf{u}_0, \mathbf{y} \rangle_{\mathbf{M}}\right] = \frac{1}{\alpha} \mathbb{E}_{\mathbf{a}}\left[\mathbf{z}^T \mathbf{M}\mathbf{V}\mathbf{\Lambda}\mathbf{a}\mathbf{a}^T\mathbf{\Lambda}\mathbf{V}^T\mathbf{M}\mathbf{g}\right]$$

$$(33) \qquad = \frac{1}{\alpha}\mathbf{z}^T \mathbf{M}\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{M}\mathbf{y} = \left\langle \mathbf{z}, \frac{1}{\alpha}\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{M}\mathbf{y} \right\rangle_{\mathbf{M}},$$

where the expectation $\mathbb{E}_{\mathbf{u}}$ and $\mathbb{E}_{\mathbf{a}}$ are taken with respect to $\mathbf{u}$ and $\mathbf{a}$, respectively. Note that we have used (32) to obtain the second equality and the fact that $\mathbb{E}_{\mathbf{a}}\left[\mathbf{a}\mathbf{a}^T\right] = \mathbf{I}$ to reach the third equality. It follows that

$$(34) \qquad \mathbf{c} = \frac{1}{\alpha}\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{M} = \frac{1}{\alpha}\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^{-1},$$

where we have used (28) to obtain the second equality. A simple algebra manipulation shows that $\mathbf{c}$ is a self-adjoint operator from $\mathbb{R}_{\mathbf{M}}^N$ to $\mathbb{R}_{\mathbf{M}}^N$. The square root of the covariance matrix is thus given by

$$(35) \qquad \mathbf{c}^{\frac{1}{2}} := \frac{1}{\sqrt{\alpha}}\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1},$$

which is also a self-adjoint operator from $\mathbb{R}_{\mathbf{M}}^N$ to $\mathbb{R}_{\mathbf{M}}^N$. Again, using the identity (28) one can see that $\mathbf{c}^{\frac{1}{2}}\mathbf{c}^{\frac{1}{2}} = \mathbf{c}$. Now, the inverse of $\mathbf{c}$ reads

$$\mathbf{c}^{-1} = \alpha\mathbf{V}\mathbf{\Lambda}^{-2}\mathbf{V}^{-1} = \alpha\mathbf{V}\mathbf{\Lambda}^{-2}\mathbf{V}^T\mathbf{M},$$

whence the distribution of $\mathbf{u}$, i.e. *the FEM discretization of the prior*, is given by

$$\mathbf{u} \sim \mathcal{N}(\mathbf{u}_0, \mathbf{c}) \propto \exp\left[-\frac{1}{2}\left\langle \mathbf{u} - \mathbf{u}_0, \mathbf{c}^{-1}(\mathbf{u} - \mathbf{u}_0) \right\rangle_{\mathbf{M}}\right]$$

$$(36) \qquad = \exp\left[-\frac{\alpha}{2}(\mathbf{u} - \mathbf{u}_0)^T \mathbf{M}\mathbf{V}\mathbf{\Lambda}^{-2}\mathbf{V}^T\mathbf{M}(\mathbf{u} - \mathbf{u}_0)\right],$$

that is, the discretized prior is a Gaussian distribution in $\mathbb{R}_{\mathbf{M}}^N$ with mean $\mathbf{u}_0$ and covariance matrix $\mathbf{c}$.

3.4.2. *Finite element approximation of the likelihood.* The forward and adjoint equations (22) and (24) are second order elliptic PDEs. As such, they are (perhaps) best approximated with the standard FEM method, so that the forward solution, for example, can be approximated as

$$(37) \qquad w = \sum_{i=1}^{N} \mathbf{w}_i \varphi_i.$$

The FEM approximation for each equation results in a system of algebraic equations whose unknowns are the nodal values, e.g. $\mathbf{w}$ for the forward equation.

This standard process can be found in [13]. Once $\mathbf{w}$ is solved for, the misfit $\frac{1}{2\sigma^2} \sum_{j=1}^{K} [w(\mathbf{x}_j) - d_j]^2$, and hence the likelihood (9), is readily available.

We next focus on the discretization of the gradient. We begin with some conventions and definitions. Since $u \in L^2(\Omega)$, we follow [4] to define *the gradient* $\mathcal{G}(u) \in L^2(\Omega)$ *of the cost functional* (3) *as the Riesz representation of the gradient* (23) *with respect to the* $L^2(\Omega)$ *inner product*, i.e.,

$$(38) \qquad \langle \mathcal{G}(u), \tilde{u} \rangle_{L^2(\Omega)} := \nabla \mathcal{J}(u; \tilde{u}).$$

We can use FEM to approximate the gradient as

$$(39) \qquad \mathcal{G} = \sum_{i=1}^{N} \mathbf{g}_i \varphi_i,$$

with $\mathbf{g}_i$ as the nodal values of $\mathcal{G}$. Note that $\mathbf{g} = \mathbf{g}(u) = \mathbf{g}(\mathbf{u})$ but for simplicity we will ignore this dependency except where this could be ambiguous. Similar to $\mathbf{u}$, $\mathbf{g}$ belongs to $\mathbb{R}_{\mathbf{M}}^N$. It follows that when $u$ and $\tilde{u}$ are restricted to the finite element space we have

$$(40) \qquad \langle \mathcal{G}(u), \tilde{u} \rangle_{L^2(\Omega)} \overset{\text{FEM}}{=} \mathbf{g}^T \mathbf{M} \tilde{\mathbf{u}} = \langle \mathbf{g}, \tilde{\mathbf{u}} \rangle_{\mathbf{M}}.$$

We have defined the gradient $\mathcal{G}$ in $L^2(\Omega)$ (on which the Gaussian prior is postulated). We have also defined its counterpart, i.e. $\mathbf{g}$, in $\mathbb{R}_{\mathbf{M}}^N$, and their mutual relationships are dictated by the FEM ansatz (39). We shall derive the relationships between the gradient in $\mathbb{R}_{\mathbf{M}}^N$ and the FEM gradient defined in section 4.2.

3.4.3. *Finite element approximation of the posterior.* Since the infinite dimensional Bayes formula (8) becomes the standard one [43] when restricted to finite dimensional space, the FEM approximation of the posterior, particularly the posterior density, is therefore given by

$$(41) \quad \pi(\mathbf{u}|\mathbf{d}) \propto \exp\left( -\frac{1}{2} \|\mathbf{d} - \mathbf{F}(u)\|_{\mathbb{R}_{\mathbf{L}}^K}^2 \right) \times \exp\left( -\frac{1}{2} \langle \mathbf{u} - \mathbf{u}_0, \mathbf{c}^{-1}(\mathbf{u} - \mathbf{u}_0) \rangle_{\mathbf{M}} \right).$$

Since the posterior density is defined on finite dimensional space $\mathbb{R}_{\mathbf{M}}^N$, it is amenable to computer implementation of MCMC methods. However, care must be taken in designing MCMC methods so that their performance does not deteriorate as the dimension of the discretization increases (e.g. when the mesh is refined). This is desirable, especially for Bayesian inverse problems governed by PDEs in which mesh refinement is often needed to resolve important physical features of the problem under consideration. *One of the main objectives* of this paper is to constructively derive such FEM-based discretization-invariant MCMC methods.

Though the convergence analysis of our finite element approximation scheme is important, it is a subject of future work. On the other hand, it should be mentioned that a convergence analysis for finite dimensional posterior can be found in [43] using a spectral method and in [9] using a combination of the spectral method and the boundary element approach. A different FEM approximation of both prior and likelihood, and hence posterior, can be found in [10] for a particular case of $s = 2$ in a three dimensional spatial seismic inverse problem.

3.4.4. *Construction of a discretization-invariant MALA method in* $\mathbb{R}_{\mathbf{M}}^N$. The *goal* of this section is to derive a FEM-based discretization-invariant MALA algorithm. We shall achieve this objective by using the FEM discretization proposed in sections 3.4.1, 3.4.2, and 3.4.3 to discretize the FMALA discussed in section 3.1. *The task*

*at hand is to carry out a finite element approximation for* (14), (15), *and* (16). The question that needs to be addressed is how to cast (14), (15), and (16) in terms of quantities in $\mathbb{R}_{\mathbf{M}}^N$. The non-trivial term is $\mathcal{C}\nabla\mathcal{J}(u)$, namely, the action of the prior covariance operator on the gradient. We begin with the definition of the covariance operator acting on $z, \nabla\mathcal{J}(u) \in L^2(\Omega)$:

$$\langle z, \mathcal{C}\nabla\mathcal{J}(u)\rangle_{L^2(\Omega)} := \mathbb{E}_r\left[\langle z, r - u_0\rangle_{L^2(\Omega)} \nabla\mathcal{J}(u; r - u_0)\right]$$

$$= \mathbb{E}_r\left[\langle z, r - u_0\rangle_{L^2(\Omega)} \langle\mathcal{G}(u), r - u_0\rangle_{L^2(\Omega)}\right]$$

where we have used (38), and the expectation is taken with respect to the random function $r$ distributed by the prior measure $\mu$. Now, using the FEM method to discretize $z$ as in (31), $\mathcal{G}(u)$ as in (40), and $r$ as in (32) we obtain

(42)

$$\langle z, \mathcal{C}\nabla\mathcal{J}(u)\rangle_{L^2(\Omega)} \stackrel{\text{FEM}}{=} \mathbb{E}_{\mathbf{r}}\left[\mathbf{z}^T\mathbf{M}(\mathbf{r} - \mathbf{u}_0)(\mathbf{r} - \mathbf{u}_0)^T\mathbf{Mg}\right]$$

$$= \frac{1}{\alpha}\mathbb{E}_{\mathbf{a}}\left[\mathbf{z}^T\mathbf{MV\Lambda aa}^T\mathbf{\Lambda V}^T\mathbf{Mg}\right] = \frac{1}{\alpha}\mathbf{z}^T\mathbf{MV\Lambda}^2\mathbf{V}^T\mathbf{Mg} = \left\langle\mathbf{z}, \frac{1}{\alpha}\mathbf{V\Lambda}^2\mathbf{V}^T\mathbf{Mg}\right\rangle_{\mathbf{M}},$$

where we have used $\mathbb{E}_{\mathbf{a}}\left[\mathbf{aa}^T\right] = \mathbf{I}$. Clearly, if $z = \varphi_i$, i.e. the $i$th FEM nodal basis function, then $\mathbf{z}$ is the vector with $\mathbf{z}_i = 1$ and $\mathbf{z}_j = 0, \forall j \neq i$. The result in (42) implies the discretization of the square root $\mathcal{C}^{\frac{1}{2}}$, that is,

(43) $\quad \left\langle z, \mathcal{C}^{1/2}\nabla\mathcal{J}(u^*)\right\rangle_{L^2(\Omega)} \stackrel{\text{FEM}}{=} \frac{1}{\sqrt{\alpha}}\mathbf{z}^T\mathbf{MV\Lambda V}^T\mathbf{Mg} = \left\langle\mathbf{z}, \frac{1}{\sqrt{\alpha}}\mathbf{V\Lambda V}^T\mathbf{Mg}\right\rangle_{\mathbf{M}}.$

On the other hand, since $v \in L^2(\Omega)$ the proposal equation (14) is equivalent to

$$\langle\phi, v\rangle_{L^2(\Omega)} = \frac{2 - \Delta t}{2 + \Delta t}\langle\phi, u\rangle_{L^2(\Omega)} - \frac{2\Delta t}{2 + \Delta t}\langle\phi, \mathcal{C}\nabla\mathcal{J}(u)\rangle_{L^2(\Omega)} + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\langle\phi, \xi\rangle_{L^2(\Omega)},$$

for all $\phi \in L^2(\Omega)$. Now, discretizing the random function $\xi$ as in (32), restricting $v$ and $u$ in the FEM space as in (31), representing $\nabla\mathcal{J}(u)$ as in (38) and (40), taking $\phi$ as FEM nodal basis functions, and using (42) we arrive at a FEM discretization of (14):

(44) $$\mathbf{v} = \frac{2 - \Delta t}{2 + \Delta t}\mathbf{u} - \frac{2\Delta t}{2 + \Delta t}\frac{1}{\alpha}\mathbf{V\Lambda}^2\mathbf{V}^T\mathbf{Mg} + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\boldsymbol{\xi},$$

after eliminating the Riesz matrix $\mathbf{M}$. Here, $\boldsymbol{\xi}$ is a random vector given by (32), and hence distributed by the discretized prior (36). Next, using (34) we conclude

(45) $$\mathbf{v} = \frac{2 - \Delta t}{2 + \Delta t}\mathbf{u} - \frac{2\Delta t}{2 + \Delta t}\mathbf{cg} + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\boldsymbol{\xi}.$$

In order to compute the acceptance probability (12), we need to discretize (16). This can be done by first substituting the FEM ansatz (31) for $v, u$, second using definition (38) and FEM approximation (40) for $\nabla\mathcal{J}(u; v \pm u)$, and third employing (43) for $\mathcal{C}^{\frac{1}{2}}$. Doing so yields

$$\rho(\mathbf{u}, \mathbf{v}) = \mathcal{J}(\mathbf{u}) + \frac{1}{2}\langle\mathbf{g}, \mathbf{v} - \mathbf{u}\rangle_{\mathbf{M}} + \frac{\Delta t}{4}\langle\mathbf{g}, \mathbf{v} + \mathbf{u}\rangle_{\mathbf{M}}$$

$$+ \frac{\Delta t}{4\alpha}\langle\mathbf{V\Lambda V}^T\mathbf{Mg}, \mathbf{V\Lambda V}^T\mathbf{Mg}\rangle_{\mathbf{M}}.$$

Here, by $\rho(\mathbf{u}, \mathbf{v})$ and $\mathcal{J}(\mathbf{u})$, we mean the substitutions of FEM ansatz (37) for $u$ and $v$ into $\rho(u, v)$ and $\mathcal{J}(u)$. Using (28) and (35) yield

$$(46) \quad \rho(\mathbf{u}, \mathbf{v}) = \mathcal{J}(\mathbf{u}) + \frac{1}{2}\langle \mathbf{g}, \mathbf{v} - \mathbf{u}\rangle_{\mathbf{M}} + \frac{\Delta t}{4}\langle \mathbf{g}, \mathbf{v} + \mathbf{u}\rangle_{\mathbf{M}} + \frac{\Delta t}{4}\left\langle \mathbf{c}^{\frac{1}{2}}\mathbf{g}, \mathbf{c}^{\frac{1}{2}}\mathbf{g}\right\rangle_{\mathbf{M}}.$$

The acceptance rate (15) can be therefore approximated as

$$(47) \qquad \frac{d\eta^T}{d\eta}(\mathbf{u}, \mathbf{v}) \stackrel{\mathrm{FEM}}{=} \exp\left[\rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})\right].$$

Similar to the function space setting in section 3.1, if we ignore terms involving $\mathbf{g}$ in (45) and (46), we obtain a FEM approximation of the preconditioned Crank-Nicholson random walk.

We have shown how to constructively discretize the FMALA method using a Galerkin FEM method. In particular, formulas (14)–(16) for FMALA on $L^2(\Omega)$ have been cast into (45)–(47) for discretized FMALA on $\mathbb{R}^N_{\mathbf{M}}$. Recall that FMALA is a MALA method defined on the Hilbert space $L^2(\Omega)$. A closer comparison between (45)–(47) and (14)–(16) reveals an important result, namely, *the discretized MALA is nothing more than an application of the FMALA algorithm to the Hilbert space* $\mathbb{R}^N_{\mathbf{M}}$. To see this, we note that the results (14)–(16) are valid for any Hilbert space. Thus, they are certainly applicable for $\mathbb{R}^N_{\mathbf{M}}$ once all the corresponding quantities such as gradient $\mathbf{g}$, covariance matrix $\mathbf{c}$ (and its square root), etc, in $\mathbb{R}^N_{\mathbf{M}}$ are identified, and this has been one of the main focuses in sections 3.4.1 and 3.4.2. In particular, replacing $v, u, \mathcal{C}, \nabla \mathcal{J}(u)$ and $\xi$ with $\mathbf{v}, \mathbf{u}, \mathbf{c}, \mathbf{g}$ and $\boldsymbol{\xi}$ in (14) immediately gives (45). Similarly, using $\mathbf{u}, \mathbf{v}, \mathbf{g}$ and $\mathbf{c}^{\frac{1}{2}}$ in the place of $u, v, \nabla \mathcal{J}(u)$ and $\mathcal{C}^{\frac{1}{2}}$ in (16) and replacing the $L^2(\Omega)-$inner product with the $\mathbf{M}-$inner product we obtain exactly (46). Finally, replacing $u, v$ with $\mathbf{u}, \mathbf{v}$ we recover (47).

We have constructed a FEM-based MALA method on $\mathbb{R}^N_{\mathbf{M}}$. As can be seen, it is derived by either discretizing the FMALA on $L^2(\Omega)$ or applying the FMALA to $\mathbb{R}^N_{\mathbf{M}}$. We refer to this method as $\mathbb{R}^N_{\mathbf{M}}$-MALA, for which the proposal $\mathbf{v}$ given the current state $\mathbf{u}$ of the Markov chain is computed using (45) and the acceptance rate is computed via (47). Let us now provide a brief discussion on the discretization-invariant property of $\mathbb{R}^N_{\mathbf{M}}$-FMALA, and a detailed rigorous analysis is left for future work. The convergence of the distribution of the sequence of random functions with FEM nodal vector $\boldsymbol{\xi}$ to the prior distribution $\mu$ is clear by the convergence of FEM [13], of the matrix transfer technique [42], and of the Karhunen-Loève expansion. The convergence of $\mathbf{c}$ as a representation of $\mathcal{C}$ in $\mathbb{R}^N_{\mathbf{M}}$ is due to the same reasons. The convergence of the function sequences with FEM nodal vectors $\mathbf{g}$ to the gradient $\mathcal{G}$ is due to the interpolation theory on finite element spaces [13]. The convergence of the function sequences with FEM nodal vectors $\mathbf{u}$ to $u$ is due to the same reason. Consequently, the function sequences with FEM nodal vectors $\mathbf{v}$ in (45) converges to the proposal function $v$ in (14). The convergence of $\rho(\mathbf{u}, \mathbf{v})$ to $\rho(u, v)$, and hence the convergence of $\frac{d\eta^T}{d\eta}(\mathbf{u}, \mathbf{v})$ to $\frac{d\eta^T}{d\eta}(u, v)$, is due to similar reasons. As shall be shown in the numerical results in section 5, this is the key to ensuring the inheritance of dimension-independence, and thus mesh-independence, property of the $\mathbb{R}^N_{\mathbf{M}}$-MALA method.

3.4.5. *Construction of a discretization-invariant HMC method in* $\mathbb{R}^N_{\mathbf{M}}$. In this section, we construct a FEM-based discretization-invariant HMC method in $\mathbb{R}^N_{\mathbf{M}}$. Similar to section 3.4.4, we accomplish this goal using a FEM approximation of the FHMC defined on $L^2(\Omega)$ that is described in section 3.2. The presentation is

necessarily brief since we use similar Galerkin projection and MTT technique as discussed at length in section 3.4.4. In particular, the Galerkin FEM discretization of the dynamics (18) can be shown to be

$$\mathbf{u}\left(t + \frac{\Delta t}{2}\right) = \mathbf{u}(t)$$

$$\boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right) = \boldsymbol{\vartheta}(t) - \frac{\Delta t}{2}\mathbf{cg}(\mathbf{u}(t)),$$

$$\mathbf{u}\left(t + \frac{3\Delta t}{2}\right) = \mathbf{u}\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t) + \boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t)$$

$$\boldsymbol{\vartheta}\left(t + \frac{3\Delta t}{2}\right) = -\mathbf{u}\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t) + \boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t)$$

$$\mathbf{u}(t + 2\Delta t) = \mathbf{u}\left(t + \frac{3\Delta t}{2}\right)$$

$$(48) \qquad \boldsymbol{\vartheta}(t + 2\Delta t) = \boldsymbol{\vartheta}\left(t + \frac{3\Delta t}{2}\right) - \frac{\Delta t}{2}\mathbf{cg}\left(\mathbf{u}\left(t + \frac{3\Delta t}{2}\right)\right),$$

where, let us recall, except for $\mathbf{c}$, bold symbols denote vectors of FEM nodal values.

Similarly, a FEM approximation of the acceptance probability (20) can be written as

$$(49) \quad a\left(\left(\mathbf{u}^1, \boldsymbol{\vartheta}^1\right), \left(\mathbf{u}^L, \boldsymbol{\vartheta}^L\right)\right) := \min\left\{1, \exp\left(-\Delta\mathscr{H}\left(\left(\mathbf{u}^1, \boldsymbol{\vartheta}^1\right), \left(\mathbf{u}^L, \boldsymbol{\vartheta}^L\right)\right)\right)\right\},$$

where
(50)

$$\Delta\mathscr{H}\left(\left(\mathbf{u}^1, \boldsymbol{\vartheta}^1\right), \left(\mathbf{u}^L, \boldsymbol{\vartheta}^L\right)\right) = \mathcal{J}\left(\mathbf{u}^L\right) - \mathcal{J}\left(\mathbf{u}^1\right) - \Delta t\sum_{i=2}^{L-1}\left\langle\mathbf{g}^i, \boldsymbol{\vartheta}^i\right\rangle_{\mathbf{M}}$$

$$- \frac{\Delta t}{2}\left(\left\langle\mathbf{g}^1, \boldsymbol{\vartheta}^1\right\rangle_{\mathbf{M}} + \left\langle\mathbf{g}^L, \boldsymbol{\vartheta}^L\right\rangle_{\mathbf{M}}\right) + \frac{\Delta t^2}{8}\left(\left\langle\mathbf{c}^{\frac{1}{2}}\mathbf{g}^1, \mathbf{c}^{\frac{1}{2}}\mathbf{g}^1\right\rangle_{\mathbf{M}} - \left\langle\mathbf{c}^{\frac{1}{2}}\mathbf{g}^L, \mathbf{c}^{\frac{1}{2}}\mathbf{g}^L\right\rangle_{\mathbf{M}}\right).$$

Here, we have defined $\mathbf{g}^i := \mathbf{g}\left(\mathbf{u}^i\right), i = 1, \ldots, L$, and $\mathbf{u}^i$ is the FEM nodal vector of $u^i$ defined in Algorithm 1.

Similar to section 3.4.4 we refer to the HMC method defined by (48)–(50) and Algorithm 1 with $u, \vartheta$ replaced by $\mathbf{u}, \boldsymbol{\vartheta}$ as $\mathbb{R}^N_{\mathbf{M}}$-HMC. Note that, due to the Galerkin FEM approximation, we draw $\boldsymbol{\vartheta}(0)$ using the truncated KL expansion in (32). As a result, there are two equivalent view points on $\mathbb{R}^N_{\mathbf{M}}$-HMC: it can be considered as an algorithm resulting from a systematic discretization of the FHMC method on $L^2(\Omega)$ using the FEM method, or it is an algorithm obtained by applying the FHMC method to the weighted Euclidean space $\mathbb{R}^N_{\mathbf{M}}$. Either of the views requires proper definitions/constructions of the prior, the gradient $\mathbf{g}$, and the covariance matrix $\mathbf{c}$ (and its square root) on $\mathbb{R}^N_{\mathbf{M}}$ to which sections 3.4.1 and 3.4.2 are devoted. Numerical results in section 5 show that $\mathbb{R}^N_{\mathbf{M}}$-HMC is discretization-invariant and the reasons are similar to those discussed at the end of section 3.4.4.

4. **Discretize-then-Metropolize approach.** We have presented FEM approximations of the prior, the likelihood, and the posterior in section 3.4. As pointed out, the nodal vector of the parameter, i.e. $\mathbf{u}$, should be considered as a vector in the weighted Euclidean space $\mathbb{R}^N_{\mathbf{M}}$. This is inherited from the $L^2(\Omega)$-setting of the prior. In practice, the $\mathbb{R}^N$-view of the nodal vector seems to be more natural and "easier" for practitioners. *The challenge here is how to construct a discretization-invariant*

*approximation procedure in $\mathbb{R}^N$ and whether it is equivalent to the $\mathbb{R}_{\mathbf{M}}^N$-view.* The *goal* of this section is to give detailed answers for these questions.

### 4.1. $\mathbb{R}^N$-discretization of the infinite dimensional Bayes' formula.

We begin with the formal density of the posterior with respect to the Lebesgue measure as

$$(51) \qquad \pi\left(u|\mathbf{d}\right) \propto \pi_{\text{like}}\left(\mathbf{d}|u\right) \times \pi_{\text{prior}}\left(u\right) \propto \pi_{\text{like}}\left(\mathbf{d}|u\right) \times \exp\left(-\frac{1}{2}\left\|u - u_0\right\|_{\mathcal{C}}^2\right),$$

and we wish to discretize the formal density $\pi\left(u|\mathbf{d}\right)$. To that end, we need to discretize the likelihood $\pi_{\text{like}}\left(\mathbf{d}|u\right)$ using the FEM method, and this was already done in section 3.4.2, since FEM approximation of the likelihood is independent of either of the view points. What remains to be done is to discretize the formal density of the prior, and it is sufficient to consider

$$\frac{1}{2}\left\|u - u_0\right\|_{\mathcal{C}}^2 := \frac{\alpha}{2}\left\|\mathcal{A}^{s/2}\left(u - u_0\right)\right\|_{L^2(\Omega)}^2 =: \frac{\alpha}{2}\left\|z\right\|_{L^2(\Omega)}^2,$$

where we have defined a function $z \in L^2\left(\Omega\right)$ as

$$(52) \qquad \mathcal{A}^{s/2}\left(u - u_0\right) = z.$$

Note that (52) is similar to (27) and for that reason we also use the MTT approach as in section 3.4.1 to discretize (52). Doing so casts (52) into

$$\mathbf{A}^{s/2}\left(\mathbf{u} - \mathbf{u}_0\right) = \mathbf{z},$$

where $\mathbf{u}, \mathbf{u}_0$ and $\mathbf{z}$ are the FEM nodal vectors of $u, u_0$ and $z$, respectively. Consequently, we have

$$\begin{aligned}
\frac{1}{2}\left\|u - u_0\right\|_{\mathcal{C}}^2 = \frac{\alpha}{2}\left\|z\right\|_{L^2(\Omega)}^2 &\overset{\text{FEM}}{=} \frac{\alpha}{2}\left\|\mathbf{A}^{s/2}\left(\mathbf{u} - \mathbf{u}_0\right)\right\|_{\mathbb{R}_{\mathbf{M}}^N}^2 \\
&= \frac{\alpha}{2}\left(\mathbf{u} - \mathbf{u}_0\right)^T\left(\mathbf{A}^{s/2}\right)^T\mathbf{M}\mathbf{A}^{s/2}\left(\mathbf{u} - \mathbf{u}_0\right) \\
&= \frac{\alpha}{2}\left(\mathbf{u} - \mathbf{u}_0\right)^T\mathbf{V}^{-T}\boldsymbol{\Sigma}^{s/2}\mathbf{V}^T\mathbf{M}\mathbf{V}\boldsymbol{\Sigma}^{s/2}\mathbf{V}^{-1}\left(\mathbf{u} - \mathbf{u}_0\right) \\
&= \frac{\alpha}{2}\left(\mathbf{u} - \mathbf{u}_0\right)^T\mathbf{M}\mathbf{V}\boldsymbol{\Sigma}^{s}\mathbf{V}^T\mathbf{M}\left(\mathbf{u} - \mathbf{u}_0\right) \\
(53) \qquad &= \frac{\alpha}{2}\left(\mathbf{u} - \mathbf{u}_0\right)^T\mathbf{M}\mathbf{V}\boldsymbol{\Lambda}^{-2}\mathbf{V}^T\mathbf{M}\left(\mathbf{u} - \mathbf{u}_0\right)
\end{aligned}$$

where we have used the FEM approximation (29) for $\mathbf{A}^{s/2}$ and (28) for simplifying the results[2]. As can be seen, (53) coincides with the negative logarithm of the discrete prior probability density of $\mathbf{u}$ in (36) using the $\mathbb{R}_{\mathbf{M}}^N$-view. Recall that we derived (36) using statistics/probability theory while we have used the MTT technique, a deterministic approach, to arrive at (53), and yet they are identical.

We conclude that the FEM approximation of the formal posterior density (51) is exactly the same as (41), i.e.,

$$(54)$$
$$\pi\left(\mathbf{u}|\mathbf{d}\right) \propto \exp\left(-\frac{1}{2}\left\|\mathbf{d} - \mathbf{F}\left(u\right)\right\|_{\mathbb{R}_{\mathbf{L}}^K}^2\right) \times \exp\left(-\frac{\alpha}{2}\left(\mathbf{u} - \mathbf{u}_0\right)^T\mathbf{M}\mathbf{V}\boldsymbol{\Lambda}^{-2}\mathbf{V}^T\mathbf{M}\left(\mathbf{u} - \mathbf{u}_0\right)\right).$$

---

[2]Note that the term $\frac{1}{2}\left\|u - u_0\right\|_{\mathcal{C}}^2$ is almost surely infinite for $u \sim \mu$ since it is only finite if $u$ is in the Cameron-Martin space, namely the range of operator $\mathcal{Q}$, which has zero measure under $\mu$ [35].

In other words, we obtain identical FEM discretizations of the infinite dimensional Bayes' formula using either the $\mathbb{R}_M^N$- or the $\mathbb{R}^N$-view.

It is clear from (54) that the prior distribution in $\mathbb{R}^N$ is a Gaussian $\mathcal{N}(\mathbf{u}_0, \mathbf{C})$ and our next task is to determine $\mathbf{C}$. From (54), we infer

$$\mathbf{C}^{-1} := \alpha \mathbf{M} \mathbf{V} \mathbf{\Lambda}^{-2} \mathbf{V}^T \mathbf{M},$$

from which we can deduce its inverse as

$$(55) \qquad \mathbf{C} := \left(\mathbf{C}^{-1}\right)^{-1} = \frac{1}{\alpha} \mathbf{V} \mathbf{\Lambda}^2 \mathbf{V}^T,$$

where we have used (28) to obtain the last equality. Comparing (34) and (55) we obtain

$$(56) \qquad \mathbf{c}\mathbf{M}^{-1} = \mathbf{C},$$

that is, the covariances in $\mathbb{R}^N$ and $\mathbb{R}_{\mathbf{M}}^N$ are related but different. It follows that $\mathbf{u} \sim \mathcal{N}(\mathbf{u}_0, \mathbf{C})$ in $\mathbb{R}^N$.

The next question that we like to address is how to generate random vectors from the prior $\mathcal{N}(\mathbf{u}_0, \mathbf{C})$. By inspection we see that, if $\mathbf{a} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ then a random vector $\mathbf{v} \in \mathbb{R}^N$ defined by

$$(57) \qquad \mathbf{v} = \mathbf{u}_0 + \frac{1}{\sqrt{\alpha}} \mathbf{V} \mathbf{\Lambda} \mathbf{a}$$

is distributed by the prior $\mathcal{N}(\mathbf{u}_0, \mathbf{C})$. As can be seen, (57) is exactly the same as (32). The difference here is the interpretation, namely, random vectors generated by (57) belong to $\mathbb{R}^N$ while those from (32) are understood as vectors in $\mathbb{R}_{\mathbf{M}}^N$.

4.2. **Gradient in** $\mathbb{R}^N$**.** In the following, we will present discretization-invariant MCMC methods in $\mathbb{R}^N$ to explore the posterior (54). To that end, we need to define the gradient in $\mathbb{R}^N$. We begin with the definition of the FEM gradient vector $\mathbf{G} \in \mathbb{R}^N$:

$$(58) \qquad \mathbf{G}_i := \nabla \mathcal{J}(\mathbf{u}) := \nabla \mathcal{J}(u; \varphi_i),$$

i.e. the derivative of the cost functional along direction spanned by the FEM basis $\varphi_i$, which is precisely the gradient of the cost functional $\mathcal{J}(\mathbf{u})$ with respect to $\mathbf{u}_i$, namely,

$$\mathbf{G}_i = \frac{\partial \mathcal{J}}{\partial \mathbf{u}_i}.$$

Again, $\mathcal{J}$ and $\mathbf{G}$ are functions of $\mathbf{u}$, and we omit their argument for simplicity, but we recover the dependency occasionally if there is a chance of misunderstanding. We will use this convention throughout the paper for other quantities as well.

We have the following simple result relating the gradients in $\mathbb{R}^N$ and $\mathbb{R}_{\mathbf{M}}^N$:

$$(59) \qquad \mathbf{M}\mathbf{g} = \mathbf{G},$$

which can be proved by taking $\tilde{u}$ in (38) as $\varphi_i$, for $i = 1, \ldots, N$, and using (39) and (58).

**Remark 1.** Note that the FEM gradient $\mathbf{G}$ is typically the output of a FEM computer code. It can be considered to be the gradient of the cost functional with respect to $\mathbf{u}$ if we view $\mathbf{u}$ as a vector in the standard Euclidean space $\mathbb{R}^N$. Identity (59) allows us to express the gradient in $\mathbb{R}_{\mathbf{M}}^N$ as a function of its counterpart in $\mathbb{R}^N$.

4.3. **Construction of a discretization-invariant MALA method in** $\mathbb{R}^N$. At this point, we have defined the posterior (54), the prior covariance $\mathbf{C}$, and the gradient $\mathbf{G}$ in $\mathbb{R}^N$. These provide sufficient ingredients to define MALA methods on $\mathbb{R}^N$. In particular, we are going to present two $\mathbb{R}^N$-MALA approaches to explore the posterior. Clearly, one can employ the original MALA algorithm developed in [39], but for the sake of fairness with comparing to the $\mathbb{R}^N_{\mathbf{M}}$-MALA let us use the same Crank-Nicholson discretization strategy proposed in [3, 14] for the posterior (54). Doing so yields

$$(60) \qquad \mathbf{v} = \frac{2 - \Delta t}{2 + \Delta t}\mathbf{u} - \frac{2\Delta t}{2 + \Delta t}\mathbf{CG} + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\boldsymbol{\zeta},$$

where $\boldsymbol{\zeta}$ is distributed by $\mathcal{N}(\mathbf{u}_0, \mathbf{C})$. In other words, $\boldsymbol{\zeta}$ is distributed exactly by (32), but is now understood as a random vector in $\mathbb{R}^N$. Thus, the proposal density of $\mathbf{v}$ given $\mathbf{u}$ reads

$$(61) \qquad q(\mathbf{u}, \mathbf{v}) = \mathcal{N}\left(\frac{2 - \Delta t}{2 + \Delta t}\mathbf{u} - \frac{2\Delta t}{2 + \Delta t}\mathbf{CG}, \frac{8\Delta t}{(2 + \Delta t)^2}\mathbf{C}\right).$$

Alternatively, *these results can be considered as an application of the function space MALA approach, i.e. FMALA, to the posterior* (54) *on the Hilbert space* $\mathbb{R}^N$.

What remains to be defined is the acceptance rate. At this point, there are two obvious approaches. In the first approach, we continue applying the FMALA approach presented in section 3.1 to $\mathbb{R}^N$. In particular, the acceptance rate is computed using expressions (12), (15), and (16). Nevertheless, similar to section 3.4.4, $\rho(\mathbf{u}, \mathbf{v})$ needs to be computed using appropriate quantities in $\mathbb{R}^N$, i.e. using the FEM gradient $\mathbf{G}$ and the covariance matrix $\mathbf{C}$. Doing so gives

$$(62) \quad \rho(\mathbf{u}, \mathbf{v}) = \mathcal{J}(\mathbf{u}) + \frac{1}{2}\langle\mathbf{G}, \mathbf{v} - \mathbf{u}\rangle_{\mathbb{R}^N} + \frac{\Delta t}{4}\langle\mathbf{G}, \mathbf{v} + \mathbf{u}\rangle_{\mathbb{R}^N} + \frac{\Delta t}{4}\langle\mathbf{G}, \mathbf{CG}\rangle_{\mathbb{R}^N}.$$

We next show that *the* $\mathbb{R}^N$-*MALA method consisting of* (60), (62), *and* (15) *is identical to the* $\mathbb{R}^N_{\mathbf{M}}$-*MALA approach* in section 3.4.4. Indeed, using identities (59) and (56) we can rewrite (45) as

$$\mathbf{v} = \frac{2 - \Delta t}{2 + \Delta t}\mathbf{u} - \frac{2\Delta t}{2 + \Delta t}\mathbf{CG} + \frac{\sqrt{8\Delta t}}{2 + \Delta t}\boldsymbol{\xi},$$

which is equivalent to (60) since both $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are random vectors distributed by (32). On the other hand, if we use identities (59) and (56), it is easy to see that (46) is identical to (62). Again, the difference here is the interpretation, namely, nodal vectors in the $\mathbb{R}^N_{\mathbf{M}}$-MALA method belong to $\mathbb{R}^N_{\mathbf{M}}$, but those from the $\mathbb{R}^N$-MALA counterpart reside in $\mathbb{R}^N$.

In the second $\mathbb{R}^N$-MALA approach, instead of using (62) and (15) to compute the acceptance probability we use the standard Metropolization:

$$(63) \qquad a(\mathbf{u}, \mathbf{v}) := \min\left\{1, \frac{\pi(\mathbf{v}|\mathbf{d}) \times q(\mathbf{v}, \mathbf{u})}{\pi(\mathbf{u}|\mathbf{d}) \times q(\mathbf{u}, \mathbf{v})}\right\}.$$

As can be observed, compared to the first $\mathbb{R}^N$-MALA approach (or equivalently the $\mathbb{R}^N_{\mathbf{M}}$-MALA approach in section 3.4.4), the second $\mathbb{R}^N$-MALA approach is the same up to the proposal (60). However, unlike the former, the acceptance rate of the latter in (63) is purely from the $\mathbb{R}^N$-view and *does not take into account the fact that the finite dimensional posterior* (54) *is a discretization of an infinite dimensional problem.* In particular, while almost surely infinite terms are eliminated

in the former approach, they are still present in the latter one. Indeed, without loss of generality let us assume $\mathbf{u}_0 = \mathbf{0}$, one can show that

$$
(64) \quad \log \left[ \frac{\pi(\mathbf{v}|\mathbf{d}) \times q(\mathbf{v}, \mathbf{u})}{\pi(\mathbf{u}|\mathbf{d}) \times q(\mathbf{u}, \mathbf{v})} \right] = \rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})
$$
$$
+ \frac{1}{2}\mathbf{v}^T \mathbf{C}^{-1}\mathbf{v} + \frac{(2-\Delta t)^2}{16\Delta t}\mathbf{v}^T \mathbf{C}^{-1}\mathbf{v} - \frac{(2+\Delta t)^2}{16\Delta t}\mathbf{v}^T \mathbf{C}^{-1}\mathbf{v}
$$
$$
- \frac{1}{2}\mathbf{u}^T \mathbf{C}^{-1}\mathbf{u} - \frac{(2-\Delta t)^2}{16\Delta t}\mathbf{u}^T \mathbf{C}^{-1}\mathbf{u} + \frac{(2+\Delta t)^2}{16\Delta t}\mathbf{u}^T \mathbf{C}^{-1}\mathbf{u}.
$$

Note that terms involving $\mathbf{v}^T\mathbf{C}^{-1}\mathbf{v}$ and $\mathbf{u}^T\mathbf{C}^{-1}\mathbf{u}$ are very large, in fact approach infinity as the parameter dimension increases, (due to the presence of $\boldsymbol{\zeta}$ in (60)) if the finite element space does not conform in the Cameron-Martin space. Nevertheless, the last two lines in (64) vanish in exact arithmetic. While the first $\mathbb{R}^N$-MALA approach takes this into account, the second does not. Consequently, the last two lines are not zero (due to round-off errors) and may dominate $\rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})$. As will be shown in the numerical results, *it is these terms that make the acceptance rate of the second approach deteriorate as the dimension of $\mathbf{u}$ increases (e.g. when the mesh is refined)*.

4.4. **Construction of a discretization-invariant HMC method in $\mathbb{R}^N$.** In section 3.4.5 we have presented a FEM discretization of the FHMC method on $L^2(\Omega)$. In this section, we will first discretize the Hamiltonian, and then present two $\mathbb{R}^N$-HMC algorithms parallel to the $\mathbb{R}^N$-MALA methods in section 4.3. To begin, using the MTT technique as in (53) and the prior covariance (55) we can discretize the Hamiltonian (17) to obtain

$$
(65) \qquad \mathscr{H}(\mathbf{u}, \boldsymbol{\vartheta}) := \frac{1}{2}\left\langle \boldsymbol{\vartheta}, \mathbf{C}^{-1}\boldsymbol{\vartheta} \right\rangle_{\mathbb{R}^N} + \mathcal{J}(\mathbf{u}) + \frac{1}{2}\left\langle \mathbf{u}, \mathbf{C}^{-1}\mathbf{u} \right\rangle_{\mathbb{R}^N}.
$$

In the first $\mathbb{R}^N$-HMC algorithm, we apply the FHMC discrete dynamic (18) to the Hamiltonian (65) in the Hilbert space $\mathbb{R}^N$. Substituting appropriate quantities, e.g. $\mathbf{G}$ and $\mathbf{C}$, of $\mathbb{R}^N$ in (18) gives

$$
\mathbf{u}\left(t + \frac{\Delta t}{2}\right) = \mathbf{u}(t)
$$
$$
\boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right) = \boldsymbol{\vartheta}(t) - \frac{\Delta t}{2}\mathbf{C}\mathbf{G}(\mathbf{u}(t)),
$$
$$
\mathbf{u}\left(t + \frac{3\Delta t}{2}\right) = \mathbf{u}\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t) + \boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t)
$$
$$
\boldsymbol{\vartheta}\left(t + \frac{3\Delta t}{2}\right) = -\mathbf{u}\left(t + \frac{\Delta t}{2}\right)\sin(\Delta t) + \boldsymbol{\vartheta}\left(t + \frac{\Delta t}{2}\right)\cos(\Delta t)
$$
$$
\mathbf{u}(t + 2\Delta t) = \mathbf{u}\left(t + \frac{3\Delta t}{2}\right)
$$
$$
(66) \qquad \boldsymbol{\vartheta}(t + 2\Delta t) = \boldsymbol{\vartheta}\left(t + \frac{3\Delta t}{2}\right) - \frac{\Delta t}{2}\mathbf{C}\mathbf{G}\left(\mathbf{u}\left(t + \frac{3\Delta t}{2}\right)\right).
$$

The acceptance rate (20), together with (21), when applied to the discrete Hamiltonian (65) in $\mathbb{R}^N$ becomes

$$
(67) \quad a\left(\left(\mathbf{u}^1, \boldsymbol{\vartheta}^1\right), \left(\mathbf{u}^L, \boldsymbol{\vartheta}^L\right)\right) := \min\left\{1, \exp\left(-\Delta\mathscr{H}\left(\left(\mathbf{u}^1, \boldsymbol{\vartheta}^1\right), \left(\mathbf{u}^L, \boldsymbol{\vartheta}^L\right)\right)\right)\right\},
$$

where

(68)

$$
\Delta \mathscr{H} \left( \left( \mathbf{u}^1, \boldsymbol{\vartheta}^1 \right), \left( \mathbf{u}^L, \boldsymbol{\vartheta}^L \right) \right) = \mathcal{J} \left( \mathbf{u}^L \right) - \mathcal{J} \left( \mathbf{u}^1 \right) - \Delta t \sum_{i=2}^{L-1} \left\langle \mathbf{G}^i, \boldsymbol{\vartheta}^i \right\rangle_{\mathbb{R}^N}
$$

$$
- \frac{\Delta t}{2} \left( \left\langle \mathbf{G}^1, \boldsymbol{\vartheta}^1 \right\rangle_{\mathbb{R}^N} + \left\langle \mathbf{G}^L, \boldsymbol{\vartheta}^L \right\rangle_{\mathbb{R}^N} \right) + \frac{\Delta t^2}{8} \left( \left\langle \mathbf{G}^1, \mathbf{C}\mathbf{G}^1 \right\rangle_{\mathbb{R}^N} - \left\langle \mathbf{G}^L, \mathbf{C}\mathbf{G}^L \right\rangle_{\mathbb{R}^N} \right).
$$

Here, we have defined $\mathbf{G}^i := \mathbf{G} \left( \mathbf{u}^i \right), i = 1, \dots, L$, and $\mathbf{u}^i$ is the FEM nodal vector of $u^i$ defined in Algorithm 1. Now using identities (59) and (56) it is easy to see that (67)–(68) are *identical* to (49)–(50). That is, *the first $\mathbb{R}^N$-HMC algorithm is equivalent to the $\mathbb{R}^N_{\mathbf{M}}$-HMC method.*

Let us now construct the second $\mathbb{R}^N$-HMC approach by using the standard finite dimensional HMC approach for the Hamiltonian (65). For the sake of comparison, instead of considering the standard momentum-based HMC approach [16, 31], we use the same velocity-based HMC framework proposed in [2]. That is, the same numerical integrator (66) is employed, but the acceptance rate is based on the usual formula

(69)    $a^S \left( \left( \mathbf{u}^1, \boldsymbol{\vartheta}^1 \right), \left( \mathbf{u}^L, \boldsymbol{\vartheta}^L \right) \right) := \min \left\{ 1, \exp \left( \mathscr{H} \left( \left( \mathbf{u}^1, \boldsymbol{\vartheta}^1 \right) \right) - \mathscr{H} \left( \mathbf{u}^L, \boldsymbol{\vartheta}^L \right) \right) \right\}.$

As can be observed, the second $\mathbb{R}^N$-HMC approach is almost identical to the first one. Indeed the difference is only at the acceptance rate computation. The former suffers from large errors due to approximating terms with large value (in fact infinite as the number of mesh points, and hence parameter dimension, increases), while this does not happen for the latter because large-value terms are already eliminated on the continuous level. As shall be shown in the numerical results, the approximation errors turn out to be detrimental. In particular, they lead to vanishing small acceptance rates for the second $\mathbb{R}^N$-HMC approach.

Let us summarize one of the most important results in this paper about the commutativity of the Metropolize-then-discretize and the discretize-then-Metropolize approaches. Here, by *commutativity we mean that both approaches yield the same discretization-invariant finite dimensional MCMC methods.*

**Theorem 4.1.** *Suppose proper definitions of the gradients* $\mathbf{g}, \mathbf{G}$ *and covariance matrices* $\mathbf{c}, \mathbf{C}$ *are provided on* $\mathbb{R}^N_{\mathbf{M}}$ *and* $\mathbb{R}^N$*, respectively. If function-space MCMC methods are applied to the Hilbert space* $\mathbb{R}^N$*, the discretize-then-Metropolize approaches commute with the Metropolize-then-discretize counterparts. In particular,* $\mathbb{R}^N$-MALA *and* $\mathbb{R}^N$-HMC *algorithms are identical to* $\mathbb{R}^N_{\mathbf{M}}$-MALA *and* $\mathbb{R}^N_{\mathbf{M}}$-HMC *counterparts.*

Clearly the result also holds true for other function-space MCMC techniques [3, 14] though we limit ourselves to the MALA and HMC approaches. The key to the commutativity is to use the correct and well-defined quantities, e.g. gradient, covariance, and inner product, in the setting under consideration. *Defining these properly and systematically in the finite element context is one of the main contributions of this paper.* More importantly, *the commutativity allows one to constructively develop discretization-invariant finite dimensional MCMC methods, such as those in this paper, whose performance is independent of the parameter dimension, and hence the mesh refinement.*

**Remark 2.** It should be pointed out that the commutativity discussed in this paper is not the same as that of discretize-then-optimize and optimize-then-discretize approaches in the PDE-constrained optimization literature (see, e.g., [20]). The FEM method typically yields the latter for elliptic forward PDEs. The former is completely different, namely, it is solely based first on proper definitions of $\mathbf{g}, \mathbf{G}, \mathbf{c}$ and $\mathbf{C}$, and then on the application of function-space MCMC methods on the Hilbert spaces $\mathbb{R}^N_{\mathbf{M}}$ and $\mathbb{R}^N$. The former is thus guaranteed, irrespective of the discretization methods and PDE types (though we use the FEM method and the forward PDE happens to be elliptic). Indeed, the discretization of the forward PDE has no role in our commutativity results. On the other hand, even when the FEM method is used, our notion of commutativity may not hold and this is what happens for the second $\mathbb{R}^N$-MALA and the second $\mathbb{R}^N$-HMC approaches.

5. **Numerical results.** In this section we present numerical results to justify our developments in this paper. Recall that we have presented one Metropolize-then-discretize (MTD) method for MALA and one for HMC in sections 3.4.4 and 3.4.5, respectively, but two discretize-then-Metropolize (DTM) ones for either of the approaches in sections 4.3 and 4.4. Since the first DTM method is the same as the MTD $\mathbb{R}^N_{\mathbf{M}}$ one, we call either of them as the MTD method, and we refer to the second $\mathbb{R}^N$-approach as the DTM method.

We will present examples in one and two dimensional spatial spaces. In the first example, we consider the case with $\Omega = [0, 1]$ and $\Gamma_R = \{1\}$. We take $K = 65$ synthetic observations at $x_j = (j-1)/2^6$, $j = 1, \ldots, K$, and the noise is assumed to be i.i.d. with variance taken as $\sigma = 0.01 \max_j \{w(\mathbf{x}_j)\}$. Here, $w(\mathbf{x}_j)$ are synthetic observations with the ground truth parameter field $u(\mathbf{x}) = 0.1 \cos(2\pi\mathbf{x})$. The inversion task is to recover this ground truth given the noisy data $\mathbf{d}$ modeled as in (1)–(2). For the numerical results, we take $Bi = 0.1$ in (4), $\alpha = 8$, and $s = 0.9$ in (6). For all MCMC runs, we take the MAP point as the initial state. For all results, we by no means attempt to run MCMC chains until convergence. Instead, we simply take, say 5000, as a representative sample size for numerical illustrations.

We first compare the MTD MALA, the DTM MALA, and the standard MALA [39]. For this comparison, we consider three different mesh sizes $h = \{2^{-9}, 2^{-8}, 2^{-7}\}$, or equivalently $N = \{129, 257, 513\}$. For time step $\Delta t$, we choose the following values $\{0.1, 0.05, 0.02, 0.01, 0.005, 0.002\}$ for both MTD MALA and DTM MALA approaches, while it is $\{0.128, 0.064, 0.032, 0.016, 0.08\} \times 10^{-3}$ for the standard MALA method. For all MCMC runs, after discarding the first 100 samples as burn-ins we take the chain with 5000 samples from which the average acceptance rate is computed.

The results for MTD MALA, DTM MALA, and standard MALA are shown in Figures 1(a), 1(b), and 1(c) respectively. As expected, the acceptance rate of the MTD MALA is independent of the mesh size, and hence the parameter dimension. However, the DTM MALA approach is better than expected. Indeed, it is identical to the MTD MALA counterpart. *The reason is that the Cameron-Martin space is $H^s = H^{0.9}$ which is a superset of the $H^1$ finite element space. In other words, the finite element space conforms in the Cameron-Martin space, and hence all the terms involving $\mathbf{C}^{-1}$ in (64) are well-defined and their sum is negligible compared to $\rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})$. This implies that the first two methods are identical.* In contrast, the standard MALA approach suffers from mesh refinement, i.e., the acceptance rate
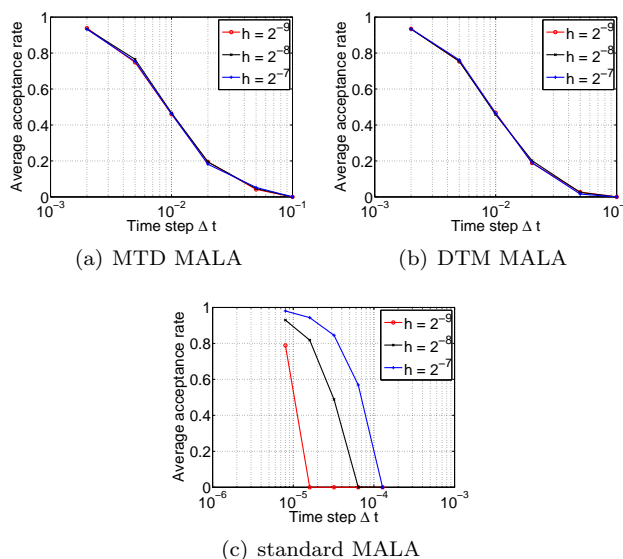
(a) MTD MALA

(b) DTM MALA

(c) standard MALA

FIGURE 1. Average acceptance rate for Metropolize-then-Discretize MALA, Discretize-then-Metropolize MALA, and standard MA LA as the mesh is refined.

decreases. This is consistent with existing results in [3, 14, 38, 39] (in which different discretizations are used). This highlights the need for discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces. Clearly, for all methods, the acceptance rate, as expected, decreases as the time step $\Delta t$, i.e. the proposal variance, increases.

We now consider HMC methods for the same inverse problem but present the result differently. In particular, we choose a single time step of $\Delta t = 0.05$ with $L = 50$ time steps, and a wide range of mesh size $h = \left\{2^{-13}, 2^{-12}, 2^{-11}, 2^{-10}, 2^{-9}, 2^{-8}, 2^{-7}\right\}$. That is, the smallest parameter dimension is 129 and the largest is 8193. For all MCMC runs, after discarding the first 100 samples as burn-ins we take the chain with 5000 samples from which the average acceptance rate is computed. In Figure 2 are the average acceptance rates for the MTD HMC, DTM HMC, standard HMC [16, 31], and a prior-conditioned standard HMC methods. Note that the last method uses the inverse of the prior covariance matrix as the mass matrix in the kinetic energy term (see, e.g., [31] for the definition of the mass matrix). As can be seen, the acceptance rates for the first two and the last HMC methods are mesh-independent, while it is not for the standard HMC method. This is expected since standard HMC is purely based on the $\mathbb{R}^N$-view, and hence its performance deteriorates as the parameter dimension increases.

Note that the cost of the four methods are comparable since the action of the prior is negligible in this example. The results therefore seem to suggest that the first two and the last methods are the method of choice. Among these three methods, the dimension-independent behavior of the MTD HMC is expected and is consistent with [2]. However, the DTM HMC approach is better than expected. Indeed, it is identical to the MTD HMC one. *The reason is that the Cameron-Martin space is $H^s = H^{0.9}$ which is a superset of the $H^1$ finite element space. In other words,*

*the finite element space conforms in the Cameron-Martin one, and hence all the terms in the formal Hamiltonian* (17) *are well-defined and finite. This implies that the standard Hamiltonian difference used in* (69) *is (numerically) the same as the improved one in* (21). *Thus, the first (MTD HMC) and second (DTM HMC)* $\mathbb{R}^N$-*HMC methods behave the same in this case.* The prior-conditioned standard HMC is surprisingly good.
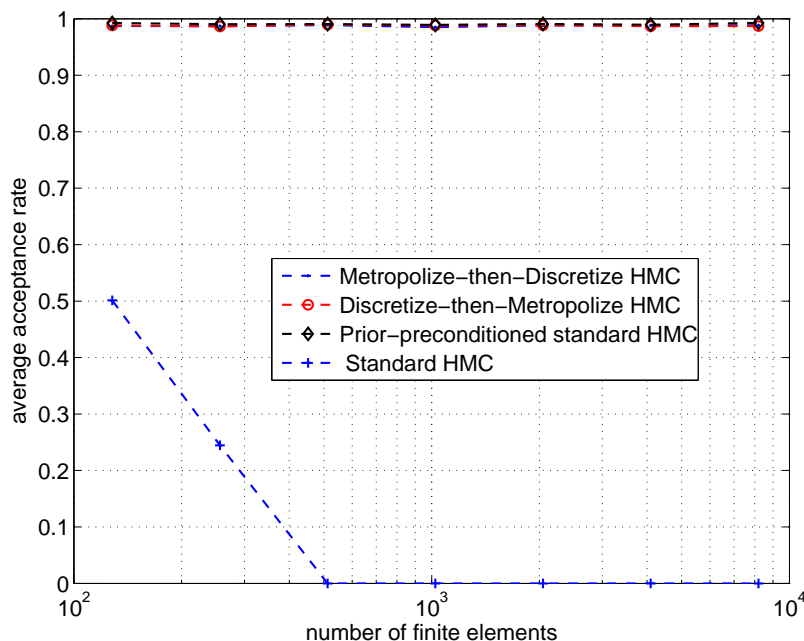


FIGURE 2. Average acceptance rate for various HMC approaches as a function of the mesh size (parameter dimension): Metropolize-then-discretize HMC (Hilbertian HMC with discretized infinite dimensional acceptance rate), discretize-then-Metropolize HMC (Hilbertian HMC with the standard finite dimensional acceptance rate), prior-preconditioned standard HMC, and the standard HMC with identity mass matrix.

To see if the above results for MTD and DTM approaches are carried over to higher dimensional spatial space, we consider an example in two dimensions. In this case, the forward equation (4) is still the same, but the domain $\Omega$ together with its dimensions is given in Figure 3(a). Red circles are the measurement/observation points. We will consider three mesh sizes $\{h, h/2, h/4\}$ and the coarsest mesh with size $h$ is shown in Figure 3(a). The coarsest mesh has 1333 parameters, medium 4760, and fine 17899. Here, $\Gamma_R = [-0.5, 0.5] \times \{0\}$, $s = 1.4$, $Bi = 0.1$, $\alpha = 5$, and $\sigma = 0.01 \max_j \{w(\mathbf{x}_j)\}$. The exact synthetic parameter field is a Gaussian blob given by

$$u = \frac{1}{0.3\sqrt{2\pi}} \exp\left[-\frac{1}{0.18}\left(x^2 + (y-2)^2\right)\right].$$

Let us first compare MTD MALA and DTM MALA. The time step is chosen to be $\Delta t = 0.00015$ for both approaches. In Table 1 are the average acceptance rates
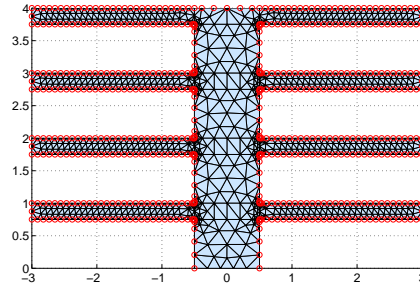
(a) Mesh and observations
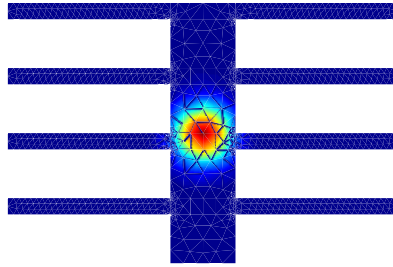


(b) Synthetic exact parameter field

FIGURE 3. Left: finite element coarsest mesh of the thermal fin and its dimensions. Red circles are the measurement/observation points. Right: the exact synthetic parameter field $u$

over 5000 samples for different mesh sizes. As can be seen, the acceptance rate of the former is mesh-independent while that of the latter is not. The reason is the following. For the Gaussian measure to be well-defined we have chosen $s = 1.4 > d/2 = 1$. Thus, the $H^1$ finite element space does not conform in the Cameron-Martin space $H^s = H^{1.4}$, and terms such as $\mathbf{u}^T\mathbf{C}^{-1}\mathbf{u}$, $\mathbf{v}^T\mathbf{C}^{-1}\mathbf{v}$ in (64) are very large. Round-off errors from these terms dominate $\rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})$, and this destroys the Metropolization step. To demonstrate this problem let us define

$$\kappa = \rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})$$

for the MTD MALA approach and

$$\kappa = \log\left[\frac{\pi(\mathbf{v}|\mathbf{d}) \times q(\mathbf{v}, \mathbf{u})}{\pi(\mathbf{u}|\mathbf{d}) \times q(\mathbf{u}, \mathbf{v})}\right]$$

for the DTM MALA counterpart. The difference between these two are terms involving $\mathbf{C}^{-1}$ in (64). Again, the difference is zero in exact arithmetic. Figure 4 shows the plot of $\kappa$ for the coarsest mesh case in Figure 3(a). As can be seen, $\kappa$ of the DTM MALA method is more than three orders of magnitude larger than that of the MTD MALA approach. In other words, the round-off errors from large terms involving $\mathbf{C}^{-1}$ entirely overwhelm $\rho(\mathbf{u}, \mathbf{v}) - \rho(\mathbf{v}, \mathbf{u})$. Consequently, the acceptance

rate in the third column of Table 1 is completely erratic. This is the reason why the DTM MALA method fails.

TABLE 1. Average acceptance rate of the Metropolize-then-discretize MALA and the discretize-then-Metropolize MALA as the mesh is refined.

|  | MTD MALA | DTM MALA |
|---|---|---|
| $h$ | 0.63098 | 0.01000 |
| $h/2$ | 0.63137 | 0.00235 |
| $h/4$ | 0.63627 | 0.00549 |
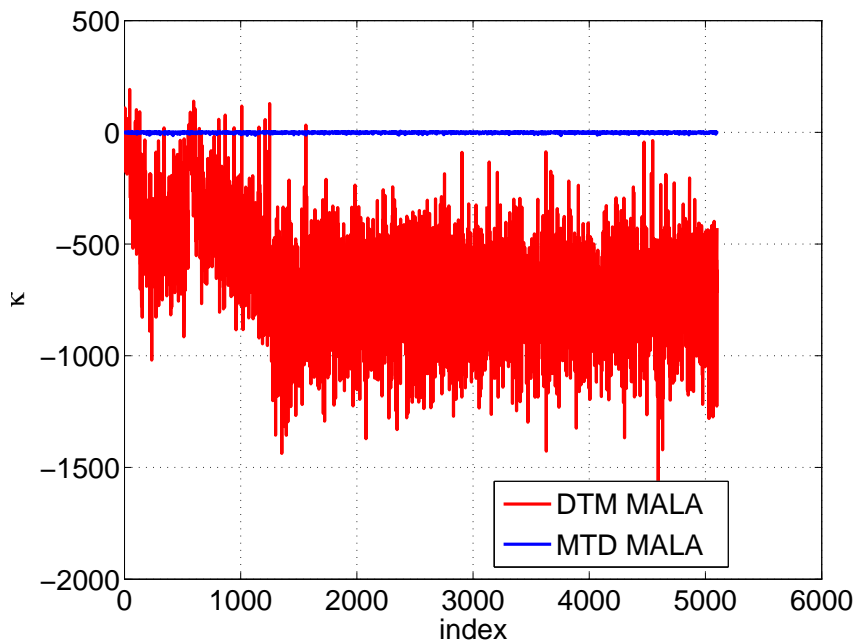


FIGURE 4. A plot of $\kappa$ for both MTD MALA and DTM MALA approaches on the coarsest mesh.

We next compare the average acceptance rate of MTD HMC, DTM HMC, and the prior-conditioned standard HMC approaches as the mesh is refined. To that end, we take $\Delta t = 0.015, L = 100$, and generate 1000 samples for each method and each mesh after discarding the first 100 samples as burn-ins. The results for three methods are shown in Table 2. As can be seen, the acceptance rate of the prior-preconditioned standard HMC is better than that of the DTM HMC, though it degrades as the mesh is refined. The DTM HMC seems to be mesh-independent, but the acceptance rate is vanishingly small. The MTD HMC is the best in the sense that it not only is independent of the mesh, but also has very high acceptance rate.

To gain further insights into each method, let us present the sample mean with 5000 samples. The result is shown in Figure 5 on the same color scale for all three

TABLE 2. Average acceptance rate of the Metropolize-then-discretize (MTD) HMC, the discretize-then-Metropolize (DTM) HMC, and the prior-conditioned standard HMC as the mesh is refined.
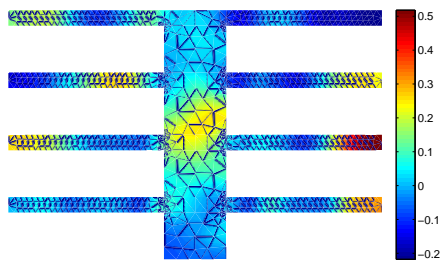
|  | MTD HMC | DTM HMC | prior-preconditioned standard HMC |
| --- | --- | --- | --- |
| $h$ | 0.82 | 0.017 | 0.511 |
| $h/2$ | 0.76 | 0.014 | 0.288 |
| $h/4$ | 0.79 | 0.018 | 0.279 |

methods on the coarsest mesh. As can be observed, the sample means of MTD HMC and prior-preconditioned standard HMC approaches in Figures 5(a) and 5(c), respectively, are similar, and they both reveal the position of the Gaussian blob at $(0, 2)$ (compared to the synthetic ground truth parameter field in Figure 3(b)). The DTM HMC method, however, does not seem to succeed in exploring the posterior, and its sample mean in Figure 5(b) is completely random.

Let us now explain why the DTM HMC approach fails in this case. For the Gaussian measure to be well-defined we have chosen $s = 1.4 > d/2 = 1$. Thus, the $H^1$ finite element space does not conform in the Cameron-Martin space $H^s = H^{1.4}$, and terms such as $\langle u, \mathcal{C}^{-1} u \rangle_{L^2(\Omega)}$ in the formal Hamiltonian (17) are almost surely infinite. Finite element approximation of these terms incurs large errors (which are supposed to be zero) that dominate the actual Hamiltonian difference, and this ruins the Metropolization step. To demonstrate this problem, let us plot, for the coarsest mesh case in Figure 5, the Hamiltonian difference using the standard and improved approaches in Figure 6. As can be seen, the standard Hamiltonian difference is more than three orders of magnitude larger than the improved one (the actual difference), and the decision on whether to accept a proposal is completely wrong.

We conclude that standard MCMC methods are not discretization-invariant. On the other hand, straightforward approaches such as the DTM methods considered in this section are corrupted by round-off/approximation errors due to theoretically infinite terms. The MTD approaches take this into account and provide methods that are independent of parameter dimensions.

6. **Conclusions and future work.** We have presented two FEM-based discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems in infinite dimensional parameter spaces. These result from a systematic FEM discretization of the prior, the likelihood, the posterior, and function-space MALA and HMC algorithms. We have shown that they can also be considered as the result of applying the function-space MALA and HMC methods to the Hilbert spaces $\mathbb{R}_{\mathbf{M}}^N$ and $\mathbb{R}^N$. The key that enables our achievements is the proper definition of quantities such as gradient and covariance matrix in these finite dimensional spaces, and we have provided step-by-step derivations using the standard finite element method. Here, by discretization-invariant MCMC methods we mean those whose behavior such as acceptance rate does not deteriorate as the parameter dimension increases. This is of paramount importance since mesh refinement, and hence increasingly higher parameter dimension, is often needed to resolve important features of physical phenomena modeled by partial differential equations.

(a) MTD HMC mean
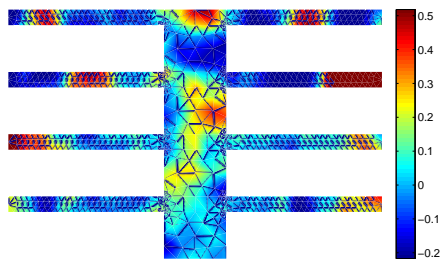


(b) DTM HMC mean
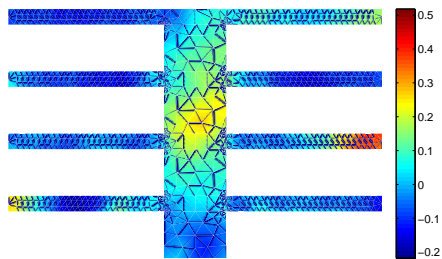


(c) Prior-preconditioned standard HMC mean

Figure 5. The sample mean of HMC methods with 5000 samples. Top: sample mean from the Metropolize-then-discretize (MTD) HMC method; Middle: sample mean from the discretize-then-Metropolize (DTM) HMC method; Bottom: sample mean from the prior-preconditioned standard HMC method.
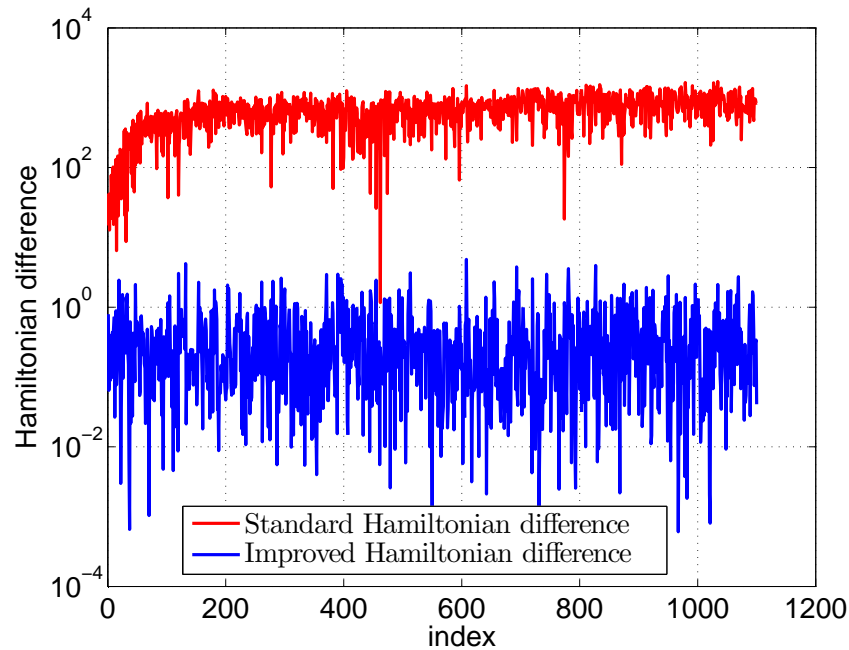
FIGURE 6. Standard versus improved Hamiltonian difference for two dimensional spatial problem.

We have considered two frameworks: *Metropolize-then-discretize* and *discretize-then-Metropolize*. The former refers to the method of first proposing a function-space MCMC method for a PDE-constrained Bayesian inverse problem in infinite dimensional parameter space and then discretizing both of them. The latter, on the other hand, first discretizes the Bayesian inverse problem and then proposes MCMC methods for the resulting finite dimensional discretized Bayesian posterior. In general, these two frameworks do not commute, that is, the resulting finite dimensional algorithms are not identical. The discretization step of the former may not be trivial, while the latter, as has been shown, may not be dimension-independent using traditional or straightforward approaches. In this paper, we develop finite element discretization schemes for both frameworks so that both commutativity and discretization-invariance are attained. In particular, we show how to construct discretize-then-Metropolize approaches for both Metropolis-adjusted Langevin algorithm and hybrid Monte Carlo method that commute with their Metropolize-then-discretize counterparts. The implication is that practitioners can take advantage of the developments in this paper to straightforwardly construct discretization-invariant discretize-then-Metropolize MCMC for large-scale inverse problems. Numerical results for one- and two-dimensional elliptic inverse problems with up to 17899 parameters have been presented to support our developments.

Ongoing research is to develop mesh-independent Riemann manifold Hamiltonian MCMC methods. This class of methods has been shown to be every effective [11, 19] with very high acceptance rate and with almost uncorrelated samples. The work on the Hessian in the Appendix will be useful for these types of methods in which the

Hessian is the key player. We are also investigating rigorous convergence analysis for the proposed discretization-invariant MCMC methods using ideas from [43, 9].

**Appendix.** Though the second order information, namely the Hessian, is not required for MCMC methods considered in this paper, it is necessary for higher order methods such as Newton-type MCMC methods [41, 12, 44, 36, 29, 24, 49, 46] and preconditioned HMC algorithms [19, 16, 31]. Thus, for completeness, let us also derive the Hessian. In particular, the Hessian acting along directions $\tilde{u}$ and $\check{u}$ reads

$$(70) \quad \nabla^2 \mathcal{J}(u; \tilde{u}, \check{u}) = \int_\Omega \tilde{u} e^u \nabla w \cdot \nabla \check{\lambda} \, d\Omega + \int_\Omega \tilde{u} e^u \nabla \check{w} \cdot \nabla \lambda \, d\Omega + \int_\Omega \tilde{u} \check{u} e^u \nabla w \cdot \nabla \lambda \, d\Omega,$$

where the second order forward state $\check{w}$, i.e. the variation of $w$ with respect to $u$ in the direction $\check{u}$, obeys the second order forward equation

$$(71) \qquad \int_\Omega e^u \nabla \check{w} \cdot \nabla \hat{\lambda} \, d\Omega + \int_{\partial\Omega \setminus \Gamma_R} Bi \, \check{w} \hat{\lambda} \, ds = -\int_\Omega \check{u} e^u \nabla w \cdot \nabla \hat{\lambda} \, d\Omega,$$

and the second order adjoint state $\check{\lambda}$, i.e. the variation of $\lambda$ with respect to $u$ in the direction $\check{u}$, is governed by the second order adjoint equation

(72)

$$\int_\Omega e^u \nabla \check{\lambda} \cdot \nabla \hat{w} \, d\Omega + \int_{\partial\Omega \setminus \Gamma_R} Bi \, \check{\lambda} \hat{w} \, ds = -\frac{1}{\sigma^2} \sum_{j=1}^K \check{w}(\mathbf{x}_j) \hat{w}(\mathbf{x}_j) - \int_\Omega \check{u} e^u \nabla \lambda \cdot \nabla \hat{w} \, d\Omega.$$

Similar to the gradient computation, one can compute the Hessian acting on two arbitrary directions by first solving the second order forward equation (71) for $\check{w}$, then computing $\check{\lambda}$ from the second order adjoint equation (72), and finally evaluating (70).

Following section 3.4.2, let us present a FEM discretization of the Hessian. We define the Hessian $\mathcal{H}(u)$ as the Fréchet derivative of $\mathcal{G}(u)$, particularly, the Hessian acting on $\check{u}$ and $\tilde{u}$ reads

$$(73) \qquad \langle \mathcal{H}(u) \check{u}, \tilde{u} \rangle_{L^2(\Omega)} := \langle \nabla \mathcal{G}(u; \check{u}), \tilde{u} \rangle_{L^2(\Omega)} = \nabla^2 \mathcal{J}(u; \tilde{u}, \check{u}).$$

It follows that $\mathcal{H}(u) : L^2(\Omega) \to L^2(\Omega)$ is a linear and self-adjoint mapping.

Recall that $\mathbf{g} = \mathbf{g}(\mathbf{u})$ is a function of $\mathbf{u}$, and thus we can define its derivative, denoted by $\mathbf{h}$, as a linear map from $\mathbb{R}_{\mathbf{M}}^N$ to $\mathbb{R}_{\mathbf{M}}^N$ as

$$\mathbf{h}_{ij} := \frac{\partial \mathbf{g}_i}{\partial \mathbf{u}_j}.$$

From (73) and (39), we deduce the following relationship between $\mathcal{H}$ and $\mathbf{h}$ when restricted to finite element space:

$$(74) \qquad \mathcal{H}(u) \check{u} \stackrel{\text{FEM}}{=} \sum_{i=1}^N \sum_{j=1}^N \mathbf{h}_{ij} \check{\mathbf{u}}_j \varphi_i.$$

Similarly, we define the FEM Hessian $\mathbf{H}$ as

$$\mathbf{H}_{ij} := \nabla^2 \mathcal{J}(u; \varphi_i, \varphi_j),$$

that is, the second derivative of the cost functional along directions spanned by the FEM basis functions $\varphi_i$ and $\varphi_j$, and we deduce

$$\mathbf{H}_{ij} = \frac{\partial \mathbf{G}_i}{\partial \mathbf{u}_j} = \frac{\partial^2 \mathcal{J}}{\partial \mathbf{u}_i \partial \mathbf{u}_j}.$$

It follows that

$$(75) \qquad\qquad \mathbf{Mh} = \mathbf{H}.$$

**Remark 3.** Note that the FEM gradient $\mathbf{G}$ and Hessian $\mathbf{H}$ are typically the outputs of a FEM computer code. They can be considered as the gradient and Hessian of the cost functional with respect to $\mathbf{u}$ if $\mathbf{u}$ is viewed as a vector in the standard Euclidean space $\mathbb{R}^N$. Expressions (59) and (75) allow us to express the gradient and Hessian in $\mathbb{R}^N_{\mathbf{M}}$ as a function of their counterparts in $\mathbb{R}^N$.

## REFERENCES

[1] A. Beskos, N. Pillai, G. Roberts, J-M Sanz-Serna and A. Stuart, Optimal tuning of the hybrid Monte Carlo algorithm, *Bernoulli,* **19** (2013), 1501–1534.

[2] A. Beskos, F. J. Pinski, J. M. Sanz-Serna and A. M. Stuart, Hybrid Monte Carlo on Hilbert spaces, *Stochastic Processes and their Applications*, **121** (2011), 2201–2230.

[3] A. Beskos and A. M. Stuart, MCMC methods for sampling function space, in *Invited Lectures: Sixth International Congress on Industrial and Applied Mathematics, ICIAM 2007, R. Jeltsch and G. Wanner, eds., European Mathematical Society,* 2009, 337–364.

[4] A. Borzì and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, SIAM, 2012.

[5] T. Bui-Thanh and O. Ghattas, Analysis of the Hessian for inverse scattering problems. Part I: Inverse shape scattering of acoustic waves, *Inverse Problems,* **28** (2012), 055001, 32pp.

[6] ——, Analysis of the Hessian for inverse scattering problems. Part II: Inverse medium scattering of acoustic waves, *Inverse Problems*, **28** (2012), 055002.

[7] ——, Analysis of the Hessian for inverse scattering problems. Part III: Inverse medium scattering of electromagnetic waves. Inverse Problems and Imaging, 2013.

[8] ——, Randomized maximum likelihood sampling for large-scale Bayesian inverse problems, In preparation, 2013.

[9] ——, An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation, *SIAM Journal of Uncertainty Quantification*, **2** (2014), 203–222.

[10] T. Bui-Thanh, O. Ghattas, J. Martin and G. Stadler, A computational framework for infinite-dimensional Bayesian inverse problems Part I: The linearized case, with application to global seismic inversion, *SIAM Journal on Scientific Computing,* **35** (2013), A2494–A2523.

[11] T. Bui-Thanh and M. A. Girolami, Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo, *Inverse Problems, Special Issue,* **30** (2014), 114014, 23pp.

[12] F. P. Casey, J. J. Waterfall, R. N. Gutenkunst, C. R. Myers and J. P. Sethna, Variational method for estimating the rate of convergence of Marko-chain Monte Carlo algorithms, *Phy. Rev. E.,* **78** (2008), 046704.

[13] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, New York, 1978.

[14] S. L. Cotter, G. O. Roberts, A. M. Stuart and D. White, MCMC methods for functions: Modifying old algorithms to make them faster, *Statistical Science,* **28** (2013), 424–446.

[15] M. Dashti, K. J. H. Law, A. M. Stuart and J. Voss, MAP estimators and their consistency in Bayesian nonparametric inverse problems, *Inverse Problems,* **29** (2013), 095017, 27pp.

[16] S. Duane, A. D. Kennedy, B. Pendleton and D. Roweth, Hybrid Monte Carlo, *Phys. Lett. B,* **195** (1987), 216–222.

[17] J. N. Franklin, Well-posed stochastic extensions of ill–posed linear problems, *Journal of Mathematical Analysis and Applications,* **31** (1970), 682–716.

[18] A. Gelman, G. O. Roberts and W. R. Gilks, Efficient Metropolis jumping rules, in *Bayesian Statistics,* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds., Oxford Univ Press, **5** (1996), 599–607.

[19] M. Girolami and B. Calderhead, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **73** (2011), 123–214.

[20] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, 2003.

[21] H. Haario, M. Laine, A. Miravete and E. Saksman, DRAM: Efficient adaptive MCMC, *Statistics and Computing,* **16** (2006), 339–354.

[22] M. Hairer, A. M. Stuart and J. Voss, Analysis of SPDEs arising in path sampling. part II: The nonlinear case, *Annals of Applied Probability,* **17** (2007), 1657–1706.

[23] W. Keith Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika,* **57** (1970), 97–109.

[24] E. Herbst, Gradient and Hessian-based MCMC for DSGE models, (2010). Unpublished manuscript.

[25] M. Ilić, F. Liu, I. Turner and V. Anh, Numerical approximation of a fractional-in-space diffusion equation, *Frac. Calc. and A Anal.,* **8** (2005), 323–341.

[26] S. Lasanen, *Discretizations of Generalized Random Variables with Applications to Inverse Problems*, PhD thesis, University of Oulu, 2002.

[27] M. S. Lehtinen, L. Päivärinta and E. Somersalo, Linear inverse problems for generalized random variables, *Inverse Problems*, **5** (1989), 599–612.

[28] F. Lindgren, H. Rue and J. Lindström, An explicit link between gaussian fields and gaussian markov random fields: The stochastic partial differential equation approach, *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **73** (2011), 423–498.

[29] J. Martin, L. C. Wilcox, C. Burstedde and O. Ghattas, A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion, *SIAM Journal on Scientific Computing,* **34** (2012), A1460–A1487.

[30] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, Equation of state calculations by fast computing machines, *The Journal of Chemical Physics,* **21** (1953), 1087–1092.

[31] R. M. Neal, *Handbook of Markov Chain Monte Carlo*, Chapman & Hall / CRC Press, 2010, ch. MCMC using Hamiltonian dynamics.

[32] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Verlag, Berlin, Heidelberg, New York, second ed., 2006.

[33] M. Ottobre, N. S. Pillai, F. J. Pinski and A. M. Stuart, A function space HMC algorithm with second order langevin diffusion limit, *Bernoulli,* **22** (2016), 60–106, arXiv, arXiv:1308.0543 (2014).

[34] P. Piiroinen, *Statistical Measurements, Experiments, and Applications*, PhD thesis, Department of Mathematics and Statistics, University of Helsinki, 2005.

[35] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambidge University Press, 1992.

[36] Y. Qi and T. P. Minka, Hessian-based Markov chain Monte-Carlo algorithms, in First Cape Cod Workshop on Monte Carlo Methods, Cape Cod, MA, USA, September 2002.

[37] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods (Springer Texts in Statistics)*, Springer-Verlag, New York, 2004.

[38] G. O. Roberts, A. Gelman and W. R. Gilks, Weak convergence and optimal scaling of random walk Metropolis algorithms, *The Annals of Applied Probability,* **7** (1997), 110–120.

[39] G. O. Roberts and J. S. Rosenthal, Optimal scaling of discrete approximations to Langevin diffusions, *J. R. Statist. Soc. B,* **60** (1997), 255–268.

[40] J. Rosenthal, Optimal proposal distributions and adaptive MCMC, in *Handbook of Markov chain Monte Carlo, Steve Brooks, Andrew Gelman, Galin L. Jones, and Xiao-Li Meng, eds., CRC Press,* 2011, 93–112.

[41] P. J. Rossky, J. D. Doll and H. L. Friedman, Brownian dynamics as smart Monte Carlo simulation, *J. Chem. Phys.,* **69** (1978), p. 4268.

[42] D. P. Simpson, *Krylov subpsace methods for approximating functions of symmetric positive definite matrices with applications to applied statistics and anomalous Diffusion*, PhD thesis, School of Mathematical Sciences Qeensland University of Technology, 2008.

[43] A. M. Stuart, Inverse problems: A Bayesian perspective, *Acta Numerica,* **19** (2010), 451–559.

[44] A. M. Stuart, J. Voss and P. Wiberg, Conditional path sampling of SDEs and the Langevin MCMC method, *Communications in Mathematical Sciences,* **2** (2004), 685–697.

[45] L. Tierney, A note on Metropolis–Hastings kernels for general state spaces, *Annals of Applied Probability,* **8** (1998), 1–9.

[46] C. Vacar, J.-F. Giovannelli and Y. Berthoumieu, Langevin and Hessian with Fisher approximation stochastic sampling for parameter estimation of structured covariance, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* (2011), 3964–3967.

[47] B. Vexler, Adaptive finite element methods for parameter identification problems, *Model Based Parameter Estimation,* Volume 4 of *the series Contributions in Mathematical and Computational Sciences,* (2012), 31–54.

[48] C. R. Vogel, *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[49] Y. Zhang and C. Sutton, Quasi-Newton methods for Markov chain Monte Carlo, in Advances in Neural Information Processing Systems, J. Shawe-Taylor, R. S. Zemel, P. Bartlett, F. C. N. Pereira, and K. Q. Weinberger, eds., vol. 24, 2011.

*E-mail address:* tanbui@ices.utexas.edu

*E-mail address:* quoc_p_nguyen@mail.utexas.edu