Pure and Applied Functional Analysis
Volume 8, Number 2, 2023, 713–750

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

MULTILEVEL MARKOV CHAIN MONTE CARLO FOR BAYESIAN INVERSE PROBLEMS FOR LINEAR ELASTICITY EQUATIONS IN THE HELLINGER-REISSNER MIXED AND DUAL MIXED FORMS WITH GAUSSIAN PRIOR

JUNTAO YANG AND VIET HA HOANG

Dedicated to Professor Ronald DeVore

Abstract. We consider the Bayesian inverse problem for forward isotropic elasticity problems in the Hellinger-Reissner mixed and dual mixed forms to infer the Lamé parameters, under the Gaussian prior probability. These mixed forms are particularly convenient when observations on the stress tensor are available. The Lamé parameters are of the log-normal form. Their natural logarithms are expressed as a countable linear expansion of pairwise independent normal random variables, as in the Kahunen-Loève expansion. Assuming a polynomial decaying rate for the sup norm of the coefficient functions of these expansions, we deduce an approximation for the posterior probability measure by taking into account only a finite number of terms in these expansions. The resulting truncated forward problem is solved by mixed finite elements (FEs). An approximation of the posterior measure is derived from the FE solution of the forward equation, with an error estimate in the Hellinger distance in terms of the FE mesh size. For the Hellinger-Reissner mixed problem, we use the stable $\mathbb{P}_0-\mathbb{P}_1$ FE pair. For the dual mixed problem, due to the difficulty arising from constructing symmetric basis functions for the stress tensor in the $H(\text{div})$ space, we use a modified mixed form which is solved by the PEERS FEs. Plain application of the Markov Chain Monte Carlo method for sampling the posterior probability measure results in a prohibitively high level of complexity. We develop the Multilevel Markov Chain Monte Carlo method which requires an essentially optimal total number of degrees of freedom to produce an approximation for the posterior expectation of quantities of interests within a prescribed level of accuracy. The method essentially relies on the ones developed for Bayesian inverse problems of forward elliptic equations in V. H. Hoang, Ch. Schwab and A. M. Stuart, Inverse problems, vol. 29, 2013 for the uniform prior and V. H. Hoang, J. H. Quek and Ch. Schwab, Inverse problems, vol. 36, 2020 for the Gaussian prior. Numerical experiments confirm the theoretical error estimates.

1. Introduction

Quantifying uncertainty in partial differential equations governed physical models with random inputs attracts much attention from the scientific community recently

2020 *Mathematics Subject Classification.* 65N21, 65F15, 65C05, 65N30 .
Key words and phrases. Multilevel Markov Chain Monte Carlo, Bayesian inversion, Linear elasticity, Mixed finite elements, Optimal complexity.

due to its essential importance in many practical and engineering applications. Forward uncertainty quantification problems aim to find the statistics of the solutions of the partial differential equations given random inputs of the forcing and/or of the coefficients (see, e.g., [7–9, 27]), while inverse uncertainty quantification problems recover the uncertain coefficients of the equations given noisy observations on the solution (see, e.g, [24, 30]).

We consider Bayesian inverse problems to recover the elastic moduli of an isotropic elastic material. We particularly consider the case where the forward elasticity equation is written in the mixed form where the stress tensor forms a part of the solution. Considering the mixed setting of the forward equation is particularly beneficial when observations on the stress tensor are available. In the sampling Markov Chain Monte Carlo (MCMC) process, solving the mixed forward equation provides directly the stress tensor corresponding to a particular sample of the unknown elastic moduli. We consider the Hellinger-Reissner mixed form and the Hellinger-Reissner dual mixed form in this paper. The Lamé parameters of the underlying isotropic material are assumed to be of the log-normal form, i.e. their natural logarithms follow a Gaussian distribution and are written in the Kahunen-Loève expansion in the form of a linear combination of pairwise independent Gaussian random variables. To sample the posterior probability, we use MCMC where the forward mixed elasticity equation is solved by a stable mixed FE scheme. For the Hellinger-Reissner mixed forward problem, the $\mathbb{P}_0 - \mathbb{P}_1$ stable FE pair is used for the stress tensor and displacement (σ, u), while for the Hellinger-Reissner dual mixed forward problem, we use the PEERS FE to resolve the difficulty arising from the symmetry of the stress tensor in the $H(\mathrm{div})$ space (see, [5]).

It is well known that for Bayesian inverse problems of partial differential equations, a plain application of MCMC is prohibitively expensive as a large number of realizations of the forward equation needs to be solved with high accuracy, leading to a high level of complexity (see [22]). Hoang et al. [22] develop the multilevel Markov Chain Monte Carlo (MLMCMC) method which approximates the posterior expectation of quantities of interest within a prescribed level of accuracy with an essentially optimal complexity level, which is equivalent to that for solving only one realization of the forward equation to obtain an approximation for the solution with an equivalent level of accuracy. However, Hoang et al. [22] only shows rigorously the convergence of the method for the case of a uniform prior distribution where the solution of the forward equation is uniformly bounded for all the realizations. For the case of the Gaussian prior as considered in this paper, the solution to the forward equation is no longer uniformly bounded. Failing to take into account the unboundedness of the solution to the forward equation may result in a highly inaccurate approximation of the posterior expectation of quantities of interest by multilevel algorithms as shown numerically in Hoang et al. [20]. Hoang et al. [20] develop a novel MLMCMC method for approximating posterior expectation of quantities of interest for Bayesian inverse problems under the Gaussian prior, essentially modifying the method for the uniform prior in [22]. They show mathematically rigorously the optimal convergence of the method. The methods are applied to Baysian inverse problems for inferring the forcing and the initial condition of the forward Navier-Stokes equation, using mixed FEs to solve the forward equation, in [32]. There has been active research on reducing computation cost of MCMC sampling of posterior probability measures for Bayesian inverse problems of forward partial differential equations. We mention exemplarily the references [14], [13], [23], and the survey papers [16] and [17] and the references therein for other multilevel methods, though to the best of our knowledge, the method developed in [20] is the only one that has been mathematically rigorously justified for the log-Gaussian prior. High dimensional quadrature rules can also be employed for approximating integrals in the high dimensional parameter spaces in computing the posterior expectation of a quantity of interest, though certain levels of smoothness dependence of the integrands on the parameters are required, see, e.g., [28], [12].

We develop in this paper the optimal mixed FE MLMCMC for Bayesian inverse problems under the Gaussian prior for mixed elasticity problems with noisy observation on the stress tensors. The paper shows theoretically and numerically that the optimal convergence of the MLMCMC method for Bayesian inverse problems with the Gaussian prior developed for scalar elliptic problems in [22] works equally for Hellinger-Reissner mixed and dual mixed form elasticity forward equation, where observations are on the stress tensor. The paper is organized as follows. In the next section, we consider the Bayesian inverse problem for the Hellinger-Reissner mixed forward elasticity problem. In particular, in subsection 2.1, we introduce the forward parametric mixed elasticity problem, and establish the realization dependent bounds for its solution. In subsection 2.2, we set up the Bayesian inverse problem where observations are on the stress tensor. We show the existence of the posterior probability measure and its locally Lipschitz continuity with respect to the observations. In subsection 2.3, we approximate the posterior by considering only a finite number of terms in the expansions of the Lamé parameters. We show a convergence rate for this approximation in terms of the expansion's finite truncation level, given a decaying rate for the coefficient functions of the expansions. Mixed finite element approximation for the forward equation is considered in subsection 2.4. An approximation for the posterior probability measure is constructed from the FE solution of the forward equation, with an error estimate in the Hellinger distance in terms of the FE mesh size. In section 3, we consider the Hellinger-Reissner dual mixed forward problem. We introduce the parametric dual mixed problem in subsection 3.1. In subsection 3.2, we show the existence and well-posedness of the Bayesian inverse problem. Approximation of the posterior probability measure with respect to finitely truncating the expansion of the Lamé parameters is considered in subsection 3.3. FE approximation of the forward problem is studied in subsection 3.4. Due to the difficulty of constructing symmetric basis functions for the stress tensor in the $H(\mathrm{div})$ space, we employ the PEERS elements together with a modified form for the dual mixed problem. From the FE solution of the forward equation, we construct an approximation of the posterior probability measure with an error estimate in the Hellinger distance which depends on the FE convergence rate. In section 4, we present the MLMCMC method for approximating the posterior expectation of quantities of interests for the Bayesian inverse problem of mixed elasticity forward equation. The method essentially relies on the method for forward elliptic equations of [20]. In section 5 we present numerical examples on FE MLMCMC sampling for Bayesian inverse problems for elasticity forward equation of the Hellinger-Reissner mixed and dual mixed form. The numerical results support the theoretical error estimates. Appendix A contains proof of the error estimates for the approximation of the posterior in the Hellinger distance. We recapitulate the details of the construction of the MLMCMC method in appendix B.

Throughout the paper, by C and c we denote generic constants whose values may change between different appearances, and do not depend on the approximation parameters. Repeated indices indicate summation.

2. Linear elasticity problem in the Hellinger-Reissner mixed form

2.1. **Parametric forward problem.** Let D be a bounded Lipschitz domain in $\mathbb{R}^d$ ($d = 2, 3$). We consider the linear isotropic elasticity equation in D. We denote by $\mathcal{C}$ the stiffness tensor, ε and σ the strain tensor and the stress tensor, and u the displacement vector. Assuming that the material is isotropic, $\mathcal{C}$ can be represented in the following explicit form with the Lamé parameters μ and λ,

$$\mathcal{C}_{ijkl} = \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) + \lambda\delta_{ij}\delta_{kl}, \tag{2.1}$$

where δ_{ij} is the Kronecker symbol ($i, j, k, l = 1, ..., d$). The relationship between the stress and strain tensors and the displacement is governed by the following linear constitutive equations

$$\begin{aligned} &\varepsilon = \nabla^{(s)}u, \\ &\sigma = \mathcal{C}\varepsilon = \{\mathcal{C}_{ijkl}\varepsilon_{kl}\} = 2\mu\varepsilon + \lambda(tr(\varepsilon))I, \end{aligned} \tag{2.2}$$

where $\nabla^{(s)}$ is the symmetric gradient, which is defined as $\left(\nabla^{(s)}u\right)_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$ with $i, j = 1, ..., d$. We consider the following forward linear elasticity equation with Dirichlet boundary,

$$\begin{cases} \nabla \cdot \sigma = -f & \text{in } D, \\ \sigma = \mathcal{C}\nabla^{(s)}u & \text{in } D, \\ u = 0 & \text{on } \Gamma, \end{cases} \tag{2.3}$$

where Γ is the domain boundary and $f \in [H^{-1}(D)]^d$. We consider the Bayesian inverse problem with observation on the stress tensor, so it is convenient to consider the Hellinger-Reissner mixed form. We note that the compliance tensor $\mathcal{C}^{-1}$ can be expressed in terms of μ and λ as

$$\mathcal{C}^{-1}_{ijkl} = \frac{1}{4\mu}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) + \left(\frac{1}{d^2\lambda + 2d\mu} - \frac{1}{2d\mu}\right)\delta_{ij}\delta_{kl}. \tag{2.4}$$

With the given primal form of the linear elasticity equation (2.3), we have the following equivalent Hellinger-Reissner mixed form: Find $\sigma \in [L^2(D)]^{d\times d}$ and $u \in [L^2(D)]^d$ such that

$$\begin{cases} (\mathcal{C}^{-1}\sigma, \tau) - (\tau, \nabla^{(s)}u) = 0 & \forall\tau \in [L^2(D)]^{d\times d}, \\ -(\sigma, \nabla^{(s)}v) = -(f, v) & \forall v \in [H^1_0(D)]^d, \end{cases} \tag{2.5}$$

where $(\cdot,\cdot)$ denotes the inner product in $[L^2(D)]^{d\times d}$ and in $[L^2(D)]^d$, extended to the duality pairing between $[H^{-1}_0(D)]^d$ and $[H^1_0(D)]^d$.

We assume that the Lamé parameters can be represented in the parametric form as

$$\mu(x,z) = \mu_*(x) + \exp\left(\bar{\mu}(x) + \sum_{j\geq 1} \mu_j(x) z_j\right),$$
$$\lambda(x,z) = \lambda_*(x) + \exp\left(\bar{\lambda}(x) + \sum_{j\geq 1} \lambda_j(x) z_j\right), \tag{2.6}$$

where $x \in D$ and $z = (z_1, z_2, ...) \in \mathbb{R}^{\mathbb{N}}$. We assume the functions μ_*, λ_* are non-negative. Hence μ and λ are both positive and can be arbitrary close to 0. We specify a prior probability measure on the coefficient space by assuming that the coordinates z_j are independently, identically distributed according to the standard Gaussian measure, i.e. $z_j \sim N(0,1)$. We denote by γ_1 the standard Gaussian measure in $\mathbb{R}^1$. We equip $\mathbb{R}^{\mathbb{N}}$ with the product σ-algebra $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) = \bigotimes_{j=1}^{\infty} \mathcal{B}(\mathbb{R})$ where $\mathcal{B}$ is the Borel σ-algebra. The Gaussian probability measure γ on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ is the product measure,

$$\gamma = \bigotimes_{i=1}^{\infty} \gamma_1.$$

In section 2.2, we consider a Bayesian inverse problem for inferring the Lamé parameters μ, λ or a quantity of interest depending on the Lamé parameters from a set of noisy observations consisting of linear functionals of the stress tensor σ. As σ, u depend on z, we denote them as $\sigma(x,z)$ and $u(x,z)$. We impose the following assumption on the functions μ_*, $\bar{\mu}$, μ_j and λ_*, $\bar{\lambda}$, λ_j.

Assumption 2.1. The functions μ_*, $\bar{\mu}$, μ_j and λ_*, $\bar{\lambda}$, λ_j are in $L^\infty(D)$ and $\operatorname{ess\,inf} \mu_*(x) \geq 0$, $\operatorname{ess\,inf} \lambda_*(x) \geq 0$. There are $c > 0$ and $s > 1$ such that $\|\mu_j\|_{L^\infty(D)} \leq cj^{-s}$, $\|\lambda_j\|_{L^\infty(D)} \leq cj^{-s}$.

We emphasize that in assumption 2.1, $\mu_* = 0$ and $\lambda_* = 0$ are admissible. We denote by $q = s - 1$ and $\mathbf{b}_\mu := \{\|\mu_j\|_{L^\infty(D)}\} \in \ell^1(\mathbb{N})$, $\mathbf{b}_\lambda := \{\|\lambda_j\|_{L^\infty(D)}\} \in \ell^1(\mathbb{N})$. Let $b_\mu^j := \|\mu_j\|_{L^\infty(D)}$ and $b_\lambda^j := \|\lambda_j\|_{L^\infty(D)}$. Assumption 2.1 implies that the set

$$U := \{z \in \mathbb{R}^{\mathbb{N}}, \sum_{j=1}^{\infty} b_\mu^j |z_j| < \infty, \sum_{j=1}^{\infty} b_\lambda^j |z_j| < \infty\} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$$

has full Gaussian measure, i.e. $\gamma(U) = 1$, (see [4, 27]). For every $z \in U$, the Lamé parameters (2.6) are well defined functions of $L^\infty(D)$. Let Θ denote the restriction of the product σ-algebra $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ to U. The prior probability is the restriction of the Gaussian measure γ to U; we denote it also as γ. For $z \in U$, we define

$$\mu_{\max}(z) = \operatorname*{ess\,sup}_{x\in D} \mu_*(x) + \exp(\|\bar{\mu}\|_{L^\infty(D)} + \sum_{j=1}^{\infty} \|\mu_j\|_{L^\infty(D)} |z_j|),$$
$$\lambda_{\max}(z) = \operatorname*{ess\,sup}_{x\in D} \lambda_*(x) + \exp(\|\bar{\lambda}\|_{L^\infty(D)} + \sum_{j=1}^{\infty} \|\lambda_j\|_{L^\infty(D)} |z_j|), \tag{2.7}$$

and

$$\mu_{\min}(z) = \operatorname*{ess\,inf}_{x\in D} \mu_*(x) + \exp(\operatorname*{ess\,inf}_{x\in D}(\bar{\mu}) - \sum_{j=1}^{\infty} \|\mu_j\|_{L^\infty(D)}|z_j|),$$
$$\lambda_{\min}(z) = \operatorname*{ess\,inf}_{x\in D} \lambda_*(x) + \exp(\operatorname*{ess\,inf}_{x\in D}(\bar{\lambda}) - \sum_{j=1}^{\infty} \|\lambda_j\|_{L^\infty(D)}|z_j|). \tag{2.8}$$

For $z \in U$ and for $x \in D$, we note $0 < \mu_{\min}(z) \leq \mu(x,z) \leq \mu_{\max}(z) < \infty$ and $0 < \lambda_{\min}(z) \leq \lambda(x,z) \leq \lambda_{\max}(z) < \infty$. We observe that $\mu_{\max}(z)$, $\lambda_{\max}(z)$ and $\mu_{\min}(z)$, $\lambda_{\min}(z)$ are (U,Θ) measurable. To formulate the mixed problem (2.5) as an abstract saddle point problem, we introduce the following two spaces

$$X = [L^2(D)]^{d\times d}_{\text{sym}}, \quad M = [H^1_0(D)]^d.$$

We define the following two bilinear forms $a \in \mathcal{L}(X \times X, \mathbb{R})$ and $b \in \mathcal{L}(X \times M, \mathbb{R})$,

$$\begin{aligned} a(z;\sigma,\tau) &= (\mathcal{C}^{-1}(z)\sigma,\tau) \\ &= \left(\frac{1}{2\mu(z)}\sigma,\tau\right) + \left(\left(\frac{1}{d^2\lambda(z)+2d\mu(z)} - \frac{1}{2d\mu(z)}\right) tr(\sigma), tr(\tau)\right), \\ b(\tau,v) &= -(\tau,\nabla^{(s)}v), \end{aligned}$$

where $(\cdot,\cdot)$ denotes the inner product in $[L^2(D)]^{d\times d}$, as well as in $[L^2(D)]^d$. Then the Hellinger-Reissner mixed problem can be written as the following saddle point problem

$$\begin{cases} \text{Seek } \sigma \in X \text{and } u \in M \text{ such that} \\ a(z;\sigma,\tau) + b(u,\tau) = 0 & \forall \tau \in X, \\ b(\sigma,v) = -(f,v) & \forall v \in M. \end{cases} \tag{2.9}$$

The well-posedness of the mixed problem is well known. We note the following inf-sup conditions.

Proposition 2.2. *The bilinear operators $a(z;\sigma,\tau)$ and $b(\tau,v)$ satisfy the following inf-sup conditions.*

(1)

$$\exists \alpha(z) > 0 \textit{ such that } \inf_{w\in X} \sup_{\tau\in X} \frac{a(z;w,\tau)}{\|w\|_X\|\tau\|_X} \geq \alpha(z),$$

(2)

$$\exists \beta > 0 \textit{ such that } \inf_{v\in M} \sup_{\tau\in X} \frac{b(\tau,v)}{\|\tau\|_X\|v\|_M} \geq \beta.$$

Proof. We observe that $\sum_{i,j=1}^{d}(\tau_{ii},\tau_{jj}) \leq d\sum_{i,j=1}^{d}(\tau_{ii},\tau_{ii}) \leq d\sum_{i,j=1}^{d}(\tau_{ij},\tau_{ij})$, for $d = 2,3$. Hence, the coercivity is due to

(2.10)
$$\begin{aligned} a(z;\tau,\tau) &= \sum_{i,j=1}^{d}\left(\frac{1}{2\mu(z)}\tau_{ij},\tau_{ij}\right) + \sum_{i=1}^{d}\left(\left(\frac{1}{d^2\lambda(z)+2d\mu(z)} - \frac{1}{2d\mu(z)}\right)\tau_{ii},\tau_{jj}\right) \\ &\geq \sum_{i,j=1}^{d}\left(\frac{1}{2\mu(z)}\tau_{ij},\tau_{ij}\right) + \sum_{i,j=1}^{d}\left(\left(\frac{1}{d\lambda(z)+2\mu(z)} - \frac{1}{2\mu(z)}\right)\tau_{ij},\tau_{ij}\right) \\ &\geq \frac{1}{d\lambda_{\max}(z)+2\mu_{\max}(z)}\|\tau\|_X^2. \end{aligned}$$

Hence the first inf-sup condition follows. For the second inf-sup condition, we refer to the proof of lemma 3.6 in Braess [5]. For $v \in H_0^1(D)^d$, let $\tau = -\nabla^{(s)}v$,

$$\frac{-(\tau,\nabla^{(s)}v)}{\|\tau\|} = \frac{\|\nabla^{(s)}v\|^2}{\|\nabla^{(s)}v\|} \geq c\|v\|_1$$

due to the Korn's inequality. This establishes the inf-sup condition. ☐

Given both the inf-sup conditions, the problem is well-posed by theorem 2.34 in [15]. Furthermore, we have the following estimates,

$$\|\sigma\|_X \leq c_1\|f\|_{M'}, \quad \|v\|_M \leq c_2\|f\|_{M'}, \tag{2.11}$$

with $c_1 = \frac{1}{\beta}(1+\frac{\|a\|}{\alpha(z)})$ and $c_2 = \frac{\|a\|}{\beta^2}(1+\frac{\|a\|}{\alpha(z)})$. We emphasize that $\alpha(z) = 1/(d\lambda_{\max}(z)+2\mu_{\max}(z))$ is a constant depending on z. We denote by $\mathbf{b}$ the sequence with components $b_j = \max\{b_\mu^j, b_\lambda^j\}$, where b_μ^j, b_λ^j are the components of $\mathbf{b}_\mu$, $\mathbf{b}_\lambda$. It follows that,

$$\begin{aligned} \|\sigma(z)\|_X &\leq C\left(1+\frac{d\lambda_{\max}(z)+2\mu_{\max}(z)}{2\mu_{\min}(z)}\right)\|f\|_{M'} \\ &\leq C\left(1+\exp\left(2\sum_{j=1}^{\infty}b_j|z_j|\right)\right). \end{aligned} \tag{2.12}$$

2.2. **Existence and wellposedness of the Bayesian inverse problem.** We present the problem setting for the Bayesian inverse problem. We consider general observations which are linear functions of the solution of the forward problem. Let $\mathcal{O}_i \in X', i = 1,...,k$ be k continuous and bounded linear observation functionals on X. We define the forward observation map $\mathcal{G}: U \to \mathbb{R}^k$ for all $z \in U$ as

$$\mathcal{G}(z) := (\mathcal{O}_1(\sigma(\cdot,z)), \mathcal{O}_2(\sigma(\cdot,z)), ..., \mathcal{O}_k(\sigma(\cdot,z))). \tag{2.13}$$

Let ϑ be the observation noise. It is assumed Gaussian and independent of the parameter z. Thus the random variable ϑ has value in $\mathbb{R}^k$ and follows the normal distribution $N(0,\Sigma)$, where Σ is a known $k\times k$ symmetric positive definite covariance matrix. The noisy observation data δ is

$$\delta = \mathcal{G}(z) + \vartheta.$$

We define the mismatch function (also known as potential function)

$$\Phi(z;\delta) = \frac{1}{2}|\delta - \mathcal{G}(z)|_\Sigma^2, \tag{2.14}$$

where $|\cdot|_\Sigma^2 = \langle \Sigma^{-1/2}, \Sigma^{-1/2} \rangle$ with $\langle \cdot,\cdot \rangle$ being the inner product in $\mathbb{R}^k$. We have the following proposition.

Proposition 2.3. *The forward map $\mathcal{G} : U \to \mathbb{R}^k$ of the parametric problem* (2.9) *is a strongly measurable map from (U, Θ) to $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$.*

Proof. From equation (2.5), we have, for every $\tau \in X, z, z' \in U$,

$$\begin{cases} (\mathcal{C}^{-1}(z')\sigma(z') - \mathcal{C}^{-1}(z)\sigma(z), \tau) - (\tau, \nabla^{(s)}u(z') - \nabla^{(s)}u(z)) = 0, \\ -(\sigma(z') - \sigma(z), \nabla^{(s)}v) = 0. \end{cases}$$

Rearranging the equation, we have

$$\begin{cases} (\mathcal{C}^{-1}(z')(\sigma(z') - \sigma(z)), \tau) - (\tau, \nabla^{(s)}(u(z') - u(z))) = ((\mathcal{C}^{-1}(z) - \mathcal{C}^{-1}(z'))\sigma(z), \tau), \\ -(\sigma(z') - \sigma(z), \nabla^{(s)}v) = 0. \end{cases} \tag{2.15}$$

With the a priori estimate for saddle point problems from theorem 2.34 in [15], we have

$$\begin{aligned} \|\sigma(z') - \sigma(z)\|_X &\le C\left(1 + \frac{d\lambda_{\max}(z') + 2\mu_{\max}(z')}{2\mu_{\min}(z')}\right)\|(\mathcal{C}^{-1}(z) - \mathcal{C}^{-1}(z'))\sigma(z)\|_X \\ &\le C\left(1 + \frac{d\lambda_{\max}(z') + 2\mu_{\max}(z')}{2\mu_{\min}(z')}\right)\|\mathcal{C}^{-1}(z') - \mathcal{C}^{-1}(z)\|_{L^\infty(D)}\|\sigma(z)\|_X, \end{aligned}$$

with a constant C independent of z. Next we estimate $\|\mathcal{C}^{-1}(z') - \mathcal{C}^{-1}(z)\|_{L^\infty(D)}$. From the formula of the compliance tensor $\mathcal{C}^{-1}$ in (2.4), we have

$$\begin{aligned} \|\mathcal{C}^{-1}(z') - \mathcal{C}^{-1}(z)\|_{L^\infty(D)} &\le C\left\|\frac{1}{2\mu(z')} - \frac{1}{2\mu(z)}\right\|_{L^\infty(D)} \\ &\quad + C\left\|\frac{1}{d^2\lambda(z') + 2d\mu(z')} - \frac{1}{d^2\lambda(z) + 2d\mu(z)}\right\|_{L^\infty(D)}. \end{aligned}$$

We have

$$\begin{aligned} &\left\|\frac{1}{2\mu(z')} - \frac{1}{2\mu(z)}\right\|_{L^\infty(D)} \le C\left(\frac{\|\mu(z) - \mu(z')\|_{L^\infty(D)}}{\mu_{\min}(z)\mu_{\min}(z')}\right) \\ &\le C\left(\frac{1}{\mu_{\min}(z)\mu_{\min}(z')}\right)\Bigg\|\exp(\bar{\mu} + \sum_{j=1}^\infty \mu_j z_j) \\ &\qquad + \exp(\bar{\mu}' + \sum_{j=1}^\infty \mu_j z_j')\Bigg\|_{L^\infty(D)}\left\|\sum_{j=1}^\infty \mu_j(z_j - z_j')\right\|_{L^\infty(D)} \\ &\le C\exp\left(c\sum_{j=1}^\infty b_j(|z_j| + |z_j'|)\right)\sum_{j=1}^\infty |z_j - z_j'|. \end{aligned}$$

Similarly,

$$\begin{aligned}\Big\| \frac{1}{d^2\lambda(z')+2d\mu(z')} - \frac{1}{d^2\lambda(z)+2d\mu(z)} \Big\|_{L^\infty(D)} \\ \le C \left(\frac{\|\lambda(z)-\lambda(z')\|_{L^\infty(D)} + \|\mu(z)-\mu(z')\|_{L^\infty(D)}}{(d^2\lambda_{\min}(z)+2d\mu_{\min}(z))(d^2\lambda_{\min}(z')+2d\mu_{\min}(z'))} \right) \\ \le C \exp\left(c\sum_{j=1}^{\infty} b_j(|z_j|+|z_j'|) \right) \sum_{j=1}^{\infty} |z_j - z_j'|.\end{aligned}$$

Hence we have the following bound

$$\begin{aligned}\|\sigma(z)-\sigma(z')\|_X \\ \le C\left(1+\frac{d\lambda_{\max}(z')+2\mu_{\max}(z')}{2\mu_{\min}(z')}\right)\|\mathcal{C}^{-1}(z')-\mathcal{C}^{-1}(z)\|_{L^\infty(D)}\|\sigma(z)\|_X \\ \le \exp\left(c\sum_{j=1}^{\infty} b_j(|z_j|+|z_j'|)\right)\sum_{j\ge 1}|z_j-z_j'|.\end{aligned} \tag{2.16}$$

Let $J\in\mathbb{N}$. For $z\in U$ let $z^J=(z_1,z_2,...,z_J,0,0,...)\in U$. We define $\sigma^J(z)=\sigma(z^J)$ and $\mathcal{G}^J(z)=\mathcal{G}(z^J)$. Regarding σ^J as a map from $\mathbb{R}^J \ni (z_1,...,z_J)\mapsto \sigma^J(z)\in [L^2(D)]^{d\times d}_{sym}$, from equation (2.16), it is continuous. Thus $\mathcal{G}^J$ regarded as a map from $\mathbb{R}^J\ni(z_1,...,z_J)\mapsto \mathcal{G}^J(z)\in\mathbb{R}^k$ is continuous. Therefore, for each $E\in\mathcal{B}(\mathbb{R}^k)$, there exists a set $E^{-1}\in\mathcal{B}(\mathbb{R}^J)$ such that the preimage $(\mathcal{G}^J)^{-1}(E)$ is the set of $z\in U$ such that $(z_1,...,z_J)\in E^{-1}$. This set is in the σ-algebra Θ. Thus $\mathcal{G}^J$ is a measurable function from U to $\mathbb{R}^k$. From equation (2.16) with $z'=z^J$, we deduce that for $z\in U$,

$$\lim_{J\to\infty}\|\sigma^J(z)-\sigma(z)\|_X=0,$$

and

$$\lim_{J\to\infty}|\mathcal{G}^J(z)-\mathcal{G}(z)|_{\mathbb{R}^k}=0.$$

As $\mathcal{G}(z)$ is the pointwise limit of a sequence of measurable maps, $\mathcal{G}: U\ni z\mapsto \mathcal{G}(z)\in\mathbb{R}^N$ is measurable. ☐

Proposition 2.4. *The posterior γ^δ is absolutely continuous with respect to the prior γ. The Radon-Nikodym derivative is given by*

$$\frac{d\gamma^\delta}{d\gamma}\propto\exp(-\Phi(z;\delta)). \tag{2.17}$$

Proof. As the forward map $\mathcal{G}$ is measurable, from theorem (2.1) in [10], γ^δ is absolutely continuous with respect to the prior γ. ☐

Next we show the continuity in the Hellinger distance of the posterior measure with respect to the observation data, which will imply the well-posedness of the posterior measure.

Proposition 2.5. *The measure γ^δ depends locally Lipschitz continuously on the data δ with respect to the Hellinger metric: for every $r > 0$ and every $\delta, \delta' \in \mathbb{R}^d$ such that for $|\delta|_\Sigma, |\delta'|_\Sigma \le r$, there exists $C = C(r) > 0$ such that*

$$d_{Hell}(\gamma^\delta, \gamma^{\delta'}) \le C(r)|\delta - \delta'|_\Sigma.$$

We include the proof in appendix A.

2.3. **Posterior approximation by finitely truncating the Lamé parameters.** Next, we consider the approximation of the forward equation by truncating the series expansion (2.6) for the Lamé parameters after J terms. We consider

$$\begin{aligned} \mu^J(x,z) &= \mu_*(x) + \exp\left(\bar{\mu}(x) + \sum_{j=1}^{J} \mu_j(x) z_j\right), \\ \lambda^J(x,z) &= \lambda_*(x) + \exp\left(\bar{\lambda}(x) + \sum_{j=1}^{J} \lambda_j(x) z_j\right), \end{aligned} \tag{2.18}$$

with $x \in D$, $z \in U$. We consider the truncated problem,

$$\begin{cases} \text{Find } (\sigma^J, u^J) \in X \times M \text{such that} \\ (\mathcal{C}_J^{-1}(z)\sigma^J, \tau) - (\tau, \nabla^{(s)} u^J) = 0 & \forall \tau \in X \\ -(\sigma^J, \nabla^{(s)} v) = -(f, v) & \forall v \in M, \end{cases} \tag{2.19}$$

where the truncated compliance tensor $\mathcal{C}_J^{-1}$ is,

$$(\mathcal{C}_J^{-1})_{ijkl} = \frac{1}{4\mu^J}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) + \left(\frac{1}{d^2\lambda^J + 2d\mu^J} - \frac{1}{2d\mu^J}\right)\delta_{ij}\delta_{kl}. \tag{2.20}$$

Proposition 2.6. *Under assumption 2.1, there are positive constants C and c such that the solution (σ^J, u^J) to problem (2.19) satisfies the following estimates,*

$$\begin{aligned} \|\sigma - \sigma^J\|_X &\le C\left(\sum_{j>J} b_j|z_j|\right) \exp\left(c\sum_{j=1}^{\infty} b_j|z_j|\right), \\ \|u - u^J\|_M &\le C\left(\sum_{j>J} b_j|z_j|\right) \exp\left(c\sum_{j=1}^{\infty} b_j|z_j|\right). \end{aligned}$$

Proof. By subtracting equation (2.5) from equation (2.19), we find that $(\sigma^J - \sigma, u^J - u)$ is the solution of the saddle point problem

$$\begin{cases} (\mathcal{C}_J^{-1}(z)(\sigma^J - \sigma), \tau) - (\tau, \nabla^{(s)}(u^J - u)) = ((\mathcal{C}^{-1}(z) - \mathcal{C}_J^{-1}(z))\sigma, \tau) & \forall \tau \in X, \\ -(\sigma^J - \sigma, \nabla^{(s)} v) = 0 & \forall v \in M. \end{cases} \tag{2.21}$$

With theorem 2.34 from [15], we have the following estimate,

$$\begin{aligned}\|\sigma(z)-\sigma^J(z)\|_X &\le C\left(1+\frac{\|a\|}{\alpha(z)}\right)\|(\mathcal{C}^{-1}(z)-\mathcal{C}_J^{-1}(z))\sigma(z)\|_X\\ &\le C\left(1+\exp\left(2\sum_{j=1}^{\infty}b_j|z_j|\right)\right)\|\mathcal{C}^{-1}(z)-\mathcal{C}_J^{-1}(z)\|_{L^\infty(D)}\|\sigma(z)\|_X.\end{aligned}$$

Similarly to the proof of Proposition 2.3 we have

$$\begin{aligned}&\|\mathcal{C}^{-1}(z)-\mathcal{C}_J^{-1}(z)\|_{L^\infty(D)}\\ &\qquad\le C\left(\|\lambda^J(z)-\lambda(z)\|_{L^\infty(D)}+\|\mu^J(z)-\mu(z)\|_{L^\infty(D)}\right)\exp\left(c\sum_{j=1}^{\infty}b_j|z_j|\right).\end{aligned}$$

Next we establish bounds for the truncated error of the Lamé parameters. We have

$$\begin{aligned}\|\mu-\mu^J\|_{L^\infty(D)} &= \left\|\exp\left(\bar{\mu}+\sum_{j=1}^{\infty}z_j\mu_j\right)-\exp\left(\bar{\mu}+\sum_{j=1}^{J}z_j\mu_j\right)\right\|_{L^\infty(D)}\\ &\le\left(\sum_{j>J}b_j|z_j|\right)\left\|\exp\left(\bar{\mu}+\sum_{j=1}^{\infty}z_j\mu_j\right)+\exp\left(\bar{\mu}+\sum_{j=1}^{J}z_j\mu_j\right)\right\|_{L^\infty(D)}\\ &\le C\left(\sum_{j>J}b_j|z_j|\right)\exp\left(\sum_{j=1}^{\infty}b_j|z_j|\right).\end{aligned}$$

Similarly we have

$$\|\lambda-\lambda^J\|_{L^\infty(D)}\le C\left(\sum_{j>J}b_j|z_j|\right)\exp\left(\sum_{j=1}^{\infty}b_j|z_j|\right).$$

From equation (2.12), we have the following bound

$$\begin{aligned}\|\sigma(z)-\sigma^J(z)\|_X &\le C\left(1+\exp\left(2\sum_{j=1}^{\infty}b_j|z_j|\right)\right)\|\mathcal{C}^{-1}-\mathcal{C}_J^{-1}\|_{L^\infty(D)}\|\sigma(z)\|_X\\ &\le C\left(\sum_{j>J}b_j|z_j|\right)\exp\left(c\sum_{j=1}^{\infty}b_j|z_j|\right).\end{aligned}$$

Similarly, we also have

$$\begin{aligned}\|u(z)-u^J(z)\|_M &\le C\left(1+\exp\left(2\sum_{j=1}^{\infty}b_j|z_j|\right)\right)\|(\mathcal{C}^{-1}-\mathcal{C}_J^{-1})\sigma(z)\|_{L^\infty(D)}\|\sigma(z)\|_X\\ &\le C\left(\sum_{j>J}b_j|z_j|\right)\exp\left(c\sum_{j=1}^{\infty}b_j|z_j|\right).\end{aligned}$$

□

We define the approximated forward map by

$$\mathcal{G}^J(z) = (\mathcal{O}_1(\sigma^J(z)), ..., \mathcal{O}_k(\sigma^J(z))). \tag{2.22}$$

The approximated posterior measure $\gamma^{J,\delta}$ is defined as,

$$\frac{d\gamma^{J,\delta}}{d\gamma} \propto \exp(-\Phi^J(z;\delta)),$$

where $\Phi^J(z;\delta)$ is the potential function,

$$\Phi^J(z;\delta) = \frac{1}{2}|\delta - \mathcal{G}^J(z)|^2_\Sigma. \tag{2.23}$$

The measure $\gamma^{J,\delta}$ is an approximation of the Bayesian posterior. Next we estimate the error of this posterior approximation by the solution of the truncated equation in the Hellinger metric.

Proposition 2.7. *Under assumption 2.1, there is a constant $C(\delta) > 0$ such that for every J,*

$$d_{Hell}(\gamma^\delta, \gamma^{J,\delta}) \le C(\delta)J^{-q}. \tag{2.24}$$

We present the proof in appendix A.

2.4. **FE approximation of the truncated problem.** We describe the FE approximation of the solution (σ^J, u^J) of (2.19) with the truncated coefficients in (2.18). In the Lipschitz polyhedron domain D with plane sides, we define the following nested family $\{\mathcal{T}^l\}_{l=0}^\infty$ of regular simplicial partition of D. The nested sequence of simplices is defined recursively as: the domain D is first subdivided into a regular family $\mathcal{T}^0$ of simplices T; then for $l \ge 1$, each simplex in $\mathcal{T}^l$ is obtained by dividing each simplex in $\mathcal{T}^{l-1}$ into 2^d congruent subsimplices. Hence the mesh size $h_l = \max\{\text{diam}(T) : T \in \mathcal{T}^l\}$ of $\mathcal{T}^l$ is $h_l = 2^{-l}h_0$. We define the following nested multi-level family of spaces of piecewise constant and continuous piecewise linear functions on $\mathcal{T}^l$ as

$$X^l = \{\sigma \in X : \sigma|_T \in [\mathcal{P}_0(T)]^{2\times 2}_{sym} \quad \forall T \in \mathcal{T}^l\},$$
$$M^l = \{u \in M : u|_T \in [\mathcal{P}_1(T)]^2 \quad \forall T \in \mathcal{T}^l\},$$

where $\mathcal{P}_0(T)$ denotes the set of constant functions in simplex $T \in \mathcal{T}^l$ and $\mathcal{P}_1(T)$ denotes the set of linear polynomials. With the FE approximation space defined, we consider the FE approximation of the truncated problem in spaces X^l and M^l,

$$\begin{cases} \text{Seek } \sigma^{J,l} \in X^l \text{ and } u^{J,l} \in M^l \text{ such that} & \\ a^J(z;\sigma^{J,l},\tau^l) + b(u^{J,l},\tau^l) = 0 & \forall \tau^l \in X^l, \\ b(\sigma^{J,l}, v^l) = -(f, v^l) & \forall v^l \in M^l, \end{cases} \tag{2.25}$$

where $a^J(z;\sigma^{J,l},\tau^l) = (\mathcal{C}_J^{-1}(z)\sigma^{J,l},\tau^l)$. With the coercivity of the bilinear operator a^J, we have the same inf-sup condition in proposition 2.2 for the bilinear form a^J. Hence problem (2.25) is well-posed. Consequently we have the following estimate (lemma 2.44 of [15]).

Lemma 2.8. *Under the inf-sup condition in proposition 2.2, the solution* $(\sigma^{J,l}, u^{J,l})$ *to problem (2.25) satisfies the following estimates*

$$\|\sigma^J - \sigma^{J,l}\|_X \le c_1 \inf_{\tau^l \in X^l} \|\sigma^J - \tau^l\|_X + c_2 \inf_{v^l \in M^l} \|u^J - v^l\|_M$$

$$\|u^J - u^{J,l}\|_M \le c_3 \inf_{\tau^l \in X^l} \|\sigma^J - \tau^l\|_X + c_4 \inf_{v^l \in M^l} \|u^J - v^l\|_M$$

with $c_1 = (1 + \frac{\|a\|}{\alpha(z)})(1 + \frac{\|b\|}{\beta})$, $c_2 = \frac{\|b\|}{\alpha(z)}$, $c_3 = c_1 \frac{\|a\|}{\beta}$, $c_4 = 1 + \frac{\|b\|}{\beta} + c_2 \frac{\|a\|}{\beta}$, *where* $\|a\| = \|a\|_{X,X}$ *and* $\|b\| = \|b\|_{X,M}$.

We emphasize that $\alpha(z) = 1/(d\lambda_{\max}(z) + 2\mu_{\max}(z))$ is a constant dependent on z, but independent of J and l. We make the following assumption on the regularity of the solution.

Assumption 2.9. We assume $\mu_*, \bar{\mu}, \mu_j, \lambda_*, \bar{\lambda}, \lambda_j \in W^{1,\infty}(D)$. Let $b_j^* = \max\{\|\mu_j\|_{W^{1,\infty}(D)}, \|\lambda_j\|_{W^{1,\infty}(D)}\}$. We further assume $\sum_{j\ge 1} b_j^* < \infty$. For $J \in \mathbb{N}$, solution $\sigma^J(\cdot, z)$ is in $W_1 := [H^1(D)]^{d\times d}_{\text{sym}}$ and $u^J(\cdot, z)$ is in $W_2 := [H^2(D)]^d \cap [H^1_0(D)]^d$ for all $z \in U$. We further assume the following bound from the classical regularity theory for elasticity problem [25, 31],

$$\|\sigma^J\|_{W_1} \le C \exp\left(c \sum_{j=1}^J b_j |z_j|\right) \left(1 + \sum_{j=1}^J b_j^* |z_j|\right)$$

$$\|u^J\|_{W_2} \le C \exp\left(c \sum_{j=1}^J b_j |z_j|\right) \left(1 + \sum_{j=1}^J b_j^* |z_j|\right).$$

Remark 2.10. We consider $f \in L^2(D)$. For a polygonal domain, the solution to the elasticity equation may not possess the H^2 regularity. It may have singularity behaviour at the corners. Finite element spaces with a graded mesh at the singularities achieve the finite element convergence rate $O(N^{-1/d})$, where N is the total number of degrees of freedom (see, e.g., Nistor and Schwab [26]). For a uniform mesh as considered above, when the solution only belongs to $H^{1+t}(D)$ for $t < 1$, the FE convergence rate is $O(h_l^{-t})$, i.e. $O(N^{-t/d})$ in terms of the total number of degrees of freedom N. The MLMCMC method achieves the convergence rate $O(N^{-t/d})$ for the total number of degrees of freedom N (apart from a possible logarithmic multiplying factor), which is optimal (see remark 4.1).

The regularity assumption holds when D is a smooth domain. When D is convex, to solve the problem, we can construct a convex polygonal domain D^l that is inscribed in D. The affine parts of the boundary of D^l are of the same size as the mesh size h_l which is $\mathcal{O}(2^{-l})$. To construct our family of simplices $\mathcal{T}^l$, we first divide D into simplices each with a curvilinear boundary segment which is a subset of ∂D. Then we refine the simplices by dividing each of them into 2^d subsimplices as described at the beginning of this section. However for each simplex at the boundary with a curvilinear boundary segment, we divide them into 2^d subsimplices where 2^{d-1} are with a curvilinear segment which is a subset of ∂D. We denote by D^l the convex polygonal domain whose vertices are the ones of $\mathcal{T}^l$ on ∂D. We then solve problem (2.25) with D being replaced by D^l. For a domain with a sufficiently smooth boundary, the solution of the elasticity equation u belongs to $H^2(D)^d$. Examining the $H^2(D)$ regularity proof of u in [25], $\|u\|_{H^2(D)^d}$ is bounded by a polynomial of $\|\lambda(z)\|_{W^{1,\infty}(D)}$, $\|\mu(z)\|_{W^{1,\infty}(D)}$, $|\lambda_{\max}(z)|$,$|\mu_{\max}(z)|$, $|\frac{1}{\lambda_{\min}(z)}|$, $|\frac{1}{\mu_{\min}(z)}|$. Hence assumption 2.9 holds for smooth domains.

From this assumption, we have the following discretization error which is a consequence of the approximation property of X^l and M^l: There exists a positive constant $C > 0$ which is independent of l such that for every $l \in \mathbb{N}$ holds

$$\inf_{\tau^l \in X^l} \|\sigma^J - \tau^l\|_X \leq C h_l \|\sigma^J\|_{W_1}$$
$$\inf_{v^l \in M^l} \|u^J - v^l\|_M \leq C h_l \|u^J\|_{W_2},$$

where $h_l = O(2^{-l})$ and the constant C depends only on $\mathcal{T}^0$. Further, we have the following error bound.

Proposition 2.11. *Consider the FE approximation of the truncated mixed problem, under assumption 2.9. There is a constant $C > 0$ such that for every $J, l \in \mathbb{N}$ and for every $z \in U$, the following error bound holds*

$$\|\sigma(\cdot, z) - \sigma^{J,l}(\cdot, z)\|_X \leq C \exp\left(c \sum_{j=1}^{\infty} b_j |z_j|\right) \left(\sum_{j>J} b_j |z_j| + \left(1 + \sum_{j=1}^{J} b_j^* |z_j|\right) 2^{-l}\right),$$

$$\|u(\cdot, z) - u^{J,l}(\cdot, z)\|_M \leq C \exp\left(c \sum_{j=1}^{\infty} b_j |z_j|\right) \left(\sum_{j>J} b_j |z_j| + \left(1 + \sum_{j=1}^{J} b_j^* |z_j|\right) 2^{-l}\right).$$

The proposition is a direct consequence of proposition 2.6 and the FE error estimate. Next we consider the FE approximation of the posterior measure. We denote the vector of observable from the FE approximation of the forward map by

$$\mathcal{G}^{J,l}(z) = (\mathcal{O}_1(\sigma^{J,l}(z)), ..., \mathcal{O}_k(\sigma^{J,l}(z))) : U \to \mathbb{R}^k. \tag{2.26}$$

The approximated potential function is defined as

$$\Phi^{J,l}(z; \delta) = \frac{1}{2} |\delta - \mathcal{G}^{J,l}(z)|_\Sigma^2. \tag{2.27}$$

Now we define the approximated posterior probability measure $\gamma^{J,l,\delta}$ on the measurable space (U, Θ) as

$$\frac{d\gamma^{J,l,\delta}}{d\gamma} \propto \exp(-\Phi^{J,l}(z; \delta)).$$

In the Hellinger metric, the error of this approximation of the posterior measure is as in the following theorem.

Theorem 2.12. *Under assumption 2.1 and assumption 2.9, there exists a positive constant $C(\delta)$ depending only on the data δ such that for every J and l*

$$d_{Hell}(\gamma^\delta, \gamma^{J,l,\delta}) \leq C(\delta)(J^{-q} + 2^{-l}).$$

The proof is similar to that of proposition 2.7.

3. Linear elasticity problem in the Hellinger-Reissner dual mixed form

3.1. **Parametric forward problem.** We consider the elasticity problem (2.3) in section 2.1. Assuming that the material is isotropic, we consider the stiffness matrix $\mathcal{C}$ in equation (2.1) and the constitutive equations (2.2). In this section, we consider the Hellinger-Reissner dual mixed form with the Dirichlet boundary condition. We introduce the following space,

$$H(\mathrm{div}, D) := \{\tau \in [L^2(D)]^{d\times d}; \quad \mathrm{div}\ \tau \in [L^2(D)]^d\},$$

with the norm $\|\tau\|_{H(\mathrm{div},D)} := (\|\tau\|^2_{[L^2(D)]^{d\times d}} + \|\mathrm{div}\tau\|^2_{[L^2(D)]^d})^{1/2}$. We denote $H(\mathrm{div}, D, \mathbb{S})$ the closure of $C^\infty(D, \mathbb{S})$ with respect to the norm $\|\tau\|_{H(\mathrm{div},D)}$, where $\mathbb{S} = \mathbb{R}^{d\times d}_{\mathrm{sym}}$. With this space, we consider the following Hellinger-Reissner dual mixed problem,

$$\begin{cases} \text{Find } \sigma \in H(\mathrm{div}, D, \mathbb{S}), u \in [L^2(D)]^d \text{ such that} & \\ (\mathcal{C}^{-1}\sigma, \tau) + (\mathrm{div}\tau, u) = 0 & \forall \tau \in H(\mathrm{div}, D, \mathbb{S}), \\ (\mathrm{div}\sigma, v) = -(f, v) & \forall v \in [L^2(D)]^d, \end{cases} \tag{3.1}$$

where $f \in [L^2(D)]^d$. In section 3.2, we consider a Bayesian inverse problem for inferring the Lamé parameters μ, λ or a quantity of interest depending on the Lamé parameters from a set of noisy observations consisting of linear functionals of the stress tensor σ. The well-posedness of mixed problem (3.1) is well known (see [5]). However, ensuring the symmetry of the stress tensor σ is often a major challenge in numerically solving problem (3.1). To avoid the difficulties arising from the symmetry of the stress tensor, the following weakly symmetric weak form is often solved instead of the original Hellinger-Reissner dual mixed form,

$$\begin{cases} \text{Find } \sigma \in H(\mathrm{div}, D), u \in [L^2(D)]^d \text{ and } \eta \in L^2(D), & \text{such that} \\ (\mathcal{C}^{-1}\sigma, \tau) + (\mathrm{div}\tau, u) + (as(\tau), \eta) = 0 & \forall \tau \in H(\mathrm{div}, D), \\ (\mathrm{div}\sigma, v) = -(f, v) & \forall v \in [L^2(D)]^d, \\ (as(\sigma), \psi) = 0. & \forall \psi \in L^2(D), \end{cases} \tag{3.2}$$

where $as(\tau) := \tau - \tau^\top \in [L^2(D)]^{d\times d}$ is the anti-symmetric part. For $d = 2$, we have $(as(\tau), \psi) := \int_D (\tau_{12} - \tau_{21})\psi dx$. With lemma 4.1 of chapter 6 in [5], problem (3.2) is equivalent to problem (3.1). We introduce the kernel space $W = \{\tau \in H(\mathrm{div}, D); (\mathrm{div}\tau, v) = 0 \text{ for } v \in [L^2(D)]^d, (as(\tau), \psi) = 0 \text{ for } \psi \in L^2(D)\}$, and the following operators a and b',

$$\begin{aligned} a(z; \sigma, \tau) &= (\mathcal{C}^{-1}(z)\sigma, \tau) \\ &= \left(\frac{1}{2\mu(z)}\sigma, \tau\right) + \left(\left(\frac{1}{d^2\lambda(z) + 2d\mu(z)} - \frac{1}{2d\mu(z)}\right) tr(\sigma), tr(\tau)\right), \\ b'(\tau, (v, \psi)) &= (\mathrm{div}\tau, v) + (as(\tau), \psi). \end{aligned} \tag{3.3}$$

We denote the space for v and ψ by $A = [L^2(D)]^d$, $B = L^2(D)$ and $N = A \times B$. Let $X = H(\mathrm{div}, D)$. We have the saddle point problem,

$$\begin{cases} \text{Seek } \sigma \in X \text{ and } (u,\eta) \in N \text{ such that} & \\ a(z;\sigma,\tau) + b'(\tau,(u,\eta)) = 0 & \forall \tau \in X \\ b'(\sigma,(v,\psi)) = -(f,v) & \forall (v,\psi) \in N. \end{cases} \tag{3.4}$$

Let $\|(v,\psi)\|_N = \|v\|_A + \|\psi\|_B$. We show that the inf-sup conditions are satisfied in the following proposition.

Proposition 3.1. *The bilinear operators $a(z;\sigma,\tau)$ and $b'(\tau,(v,\psi))$ satisfy the following inf-sup conditions:*

(1)

$$\exists\, \alpha(z) > 0 \text{ such that } \inf_{w\in W} \sup_{\tau \in W} \frac{a(z;w,\tau)}{\|w\|_X \|\tau\|_X} \geq \alpha(z),$$

(2)

$$\exists\, \beta > 0 \text{ such that } \inf_{(v,\psi)\in N} \sup_{\tau \in X} \frac{b'(\tau,(v,\psi))}{\|\tau\|_X \|(v,\psi)\|_N} \geq \beta.$$

Proof. These inf-sup conditions are standard. Since $\mathrm{div}\tau = 0, \forall \tau \in W$, we derive the following coercivity and boundedness of a in the kernel space W as in the proof of proposition 2.2,

$$\begin{aligned} a(z;\tau,\tau) &\geq \frac{1}{d\lambda_{\max}(z) + 2\mu_{\max}(z)} (\|\tau\|^2_{L^2(D)} + \|\mathrm{div}\tau\|^2_{L^2(D)}) \\ &= \frac{1}{d\lambda_{\max}(z) + 2\mu_{\max}(z)} \|\tau\|^2_{H(\mathrm{div},D)}. \end{aligned} \tag{3.5}$$

The inf-sup condition of a then follows. The inf-sup condition of b' is proved in page 320 of [5]. □

With these inf-sup conditions, problem (3.4) is well-posed by theorem 2.34 in [15]. Furthermore, we have the estimates,

$$\|\sigma\|_X \leq c_1 \|f\|_{A'}, \quad \|(u,\eta)\|_N \leq c_2 \|f\|_{A'} \tag{3.6}$$

with $c_1 = \frac{1}{\beta}\left(1 + \frac{\|a\|}{\alpha(z)}\right)$ and $c_2 = \frac{\|a\|}{\beta^2}\left(1 + \frac{\|a\|}{\alpha(z)}\right)$. Thus

$$\|\sigma(z)\|_X \leq C\left(1 + \frac{d\lambda_{\max}(z) + 2\mu_{\max}(z)}{2\mu_{\min}(z)}\right)\|f\|_{A'} \leq C\left(1 + \exp\left(2\sum_{j=1}^{\infty} b_j |z_j|\right)\right). \tag{3.7}$$

3.2. **Existence and wellposedness of the Bayesian inverse problem.** We present the setting for the Bayesian inverse problem for the Hellinger-Reissner dual mixed forward problem with Gaussian prior. Let $X = H(\mathrm{div}, D)$, we recall the definitions in section 2.2 for the linear functionals $\mathcal{O}_i \in X', i = 1, ..., k$, the forward observation map $\mathcal{G}$, observation noise ϑ, the noisy observation data δ and mismatch function $\Phi(z;\delta)$. We have the following proposition.

Proposition 3.2. *The parametric forward map $\mathcal{G} : U \to \mathbb{R}^k$ is strongly measurable from (U,Θ) to $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$.*

The proof is similar to that of proposition 2.3. With theorem 2.1 in [10], we have the following proposition.

Proposition 3.3. *The posterior γ^δ is absolutely continuous with respect to the prior γ. The Radon-Nikodym derivative is given by*

$$\frac{d\gamma^\delta}{d\gamma} \propto \exp(-\Phi(z;\delta)).$$

Next we show the continuity of the posterior in the Hellinger distance with respect to the observation data, which implies the well-posedness of the posterior measure.

Proposition 3.4. *The measure γ^δ depends locally Lipschitz continuously on the data δ with respect to the Hellinger metric: for every $r > 0$ and $\delta, \delta' \in \mathbb{R}^d$ such that for $|\delta|_\Sigma, |\delta'|_\Sigma \le r$, there exists $C = C(r) > 0$ such that*

$$d_{Hell}(\gamma^\delta, \gamma^{\delta'}) \le C(r)|\delta - \delta'|_\Sigma.$$

The proof is similar to that for proposition 2.5.

3.3. **Posterior approximation by finitely truncating the Lamé parameters.** Next, we consider the approximation of the forward equation by truncating the series expansion (2.6) for the Lamé parameters after J terms, with the truncated expansion (2.18) in section 2.3. Under assumption 2.1, we consider the truncated problem,

$$\begin{cases} \text{Find } (\sigma^J, u^J, \eta^J) \in X \times A \times B \text{ such that} \\ (\mathcal{C}_J^{-1}(z)\sigma^J, \tau) + (\nabla\cdot\tau, u^J) + (as(\tau), \eta^J) = 0 & \tau \in X \\ (\nabla\cdot\sigma^J, v) = (f, v) & v \in A, \\ (as(\sigma^J), \psi) = 0, & \psi \in B. \end{cases} \tag{3.8}$$

Proposition 3.5. *Under assumption 2.1, there is a positive constant C such that the solution (σ^J, u^J, η^J) to problem* (3.8) *satisfies the following estimates,*

$$\|\sigma(z) - \sigma^J(z)\|_X \le C\left(\sum_{j>J} b_j|z_j|\right)\exp\left(c\sum_{j=1}^{\infty} b_j|z_j|\right),$$

$$\|u(z) - u^J(z)\|_A + \|\eta(z) - \eta^J(z)\|_B \le C\left(\sum_{j>J} b_j|z_j|\right)\exp\left(c\sum_{j=1}^{\infty} b_j|z_j|\right).$$

The proof is similar to that of proposition 2.6. We recall the definition in section 2.3 for the approximated forward map $\mathcal{G}^J(z)$ in (2.22), the approximated potential function $\Phi^J(z;\delta)$ in (2.23) and the approximated posterior measure $\gamma^{J,\delta}$. We have the error estimate for this posterior approximation by the solution of the truncated equation in the Hellinger metric.

Proposition 3.6. *Under assumption 2.1, there is a constant $C(\delta) > 0$ such that for every J, it holds*

$$d_{Hell}(\gamma^\delta, \gamma^{J,\delta}) \le C(\delta)J^{-q}. \tag{3.9}$$

The proof is similar to that of proposition 2.7.

3.4. **FE approximation of the truncated problem.** In this section we describe the FE approximation of the solution (σ^J, u^J, η^J) in (3.8). As mentioned earlier, it is not trivial to find a proper finite element space to solve problem (3.1). The challenge arises from the symmetry of the stress tensor σ. Arnold and Winther [2] establish a FE space for the stress tensor σ with full symmetry. However there are 24 local degrees of freedom for a two dimensional problem and to our best knowledge there are no better FE spaces in $H(\mathrm{div}, D)$ which are fully symmetric but with less degrees of freedom. An alternative way to avoid difficulty from the symmetry of the stress tensor is to consider the problem with weak symmetry. To solve this problem, the PEERS element is developed by Arnold, Brezzi and Douglas [1] and the BDM element is developed by Brezzi, Douglas and Marini [6]. More detailed analysis of mixed FE for the elasticity problem can also be found in [29]. In this paper, we use the PEERS element. We work with a two dimensional domain D but the three dimensional case can be considered similarly.

In the two dimensional Lipschitz polyhedron domain D with plane sides, we define the nested family $\{\mathcal{T}^l\}_{l=0}^{\infty}$ of regular simplicial partition of D as in section 2.4. We introduce the following spaces for two dimensional PEERS element method,

(3.10)

$$\mathcal{M}^{k,l} := \{v \in L^2(D); \quad v|_T \in \mathcal{P}_k(T) \quad \forall T \in \mathcal{T}^l\},$$

$$\mathcal{M}_0^{k,l} := \mathcal{M}^{k,l} \cap H^1(D), \quad \mathcal{M}_{0,0}^{k,l} := \mathcal{M}^{k,l} \cap H_0^1(D),$$

$$RT_k^l := \{v \in (\mathcal{M}^{k+1,l})^2 \cap H(\mathrm{div}, D); v|_T = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + p_3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, p_1, p_2, p_3 \in \mathcal{P}_k(T)\},$$

$$B_3^l := \{v \in \mathcal{M}_0^{3,l}; v(x) = 0 \text{on every edge of the triangulation}\}.$$

where $\mathcal{P}_k(T)$ denotes the set of polynomials of degree k in simplex $T \in \mathcal{T}^l$. We define the following nested multi-level family of spaces of the simplest PEERS element on $\mathcal{T}^l$ as

$$\begin{aligned} X^l &:= [RT_0^l]^2 \oplus [\mathrm{curl}(B_3^l)]^2, \\ A^l &:= [\mathcal{M}^{0,l}]^2, \\ B^l &:= \mathcal{M}_0^{1,l}. \end{aligned} \tag{3.11}$$

Remark 3.7. The PEERS element is initially developed for two dimensional plane elasticity problem in [1]. In this section, we present our theory in two dimensions. However, the PEERS element could be extended to three dimensions easily with

$$X^l := [RT_0^l]^3 \oplus [\mathrm{curl}(B_4^l)]^3, A^l := [\mathcal{M}^{0,l}]^3, B^l := [\mathcal{M}_0^{1,l}]^3,$$

where

$$RT_k^l := \{v \in (\mathcal{M}^{k+1,l})^3 \cap H(\mathrm{div}, D); v|_T = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + p_4 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, p_1, p_2, p_3, p_4 \in \mathcal{P}_k(T)\}.$$

The analysis in this section is valid for the three dimensional version of the PEERS element. One can refer to example 1 in [3] for more details of the three dimensional version of the PEERS element.

With the FE approximation space defined, we consider the FE approximation of the truncated problem in spaces X^l, A^l and B^l, which are subspaces of X, A and B:

$$\begin{cases} \text{Seek } \sigma^{J,l} \in X^l, u^{J,l} \in A^l \text{ and } \eta^{J,l} \in B^l \text{ such that} \\ a^J(z;\sigma^{J,l},\tau^l) + b'(\tau^l,(u^{J,l},\eta^{J,l})) = 0 & \forall \tau^l \in X^l, \\ b'(\sigma^{J,l},(v^l,\psi^l)) = -(f,v^l) & \forall v^l \in A^l, \forall \psi^l \in B^l, \end{cases} \tag{3.12}$$

where $a^J(z;\sigma^{J,l},\tau^l) = (\mathcal{C}_J^{-1}(z)\sigma^{J,l},\tau^l)$. We denote by $W^l = \{\tau^l \in X^l : \mathrm{div}\tau^l = 0, (as(\tau^l,\psi^l) = 0\ \forall \psi^l \in B^l\}$ the kernel space. With the coercivity of the bilinear operator a^J on the kernel, we have the following inf-sup conditions in the approximation space.

Proposition 3.8. *The bilinear operators $a^J(z;w^l,\tau^l)$ and $b'(\tau^l,(v^l,\psi^l))$ satisfy the following inf-sup conditions.*

(1)

$$\exists \alpha(z) > 0, \quad \inf_{w^l \in W^l} \sup_{\tau^l \in W^l} \frac{a^J(z;w^l,\tau^l)}{\|w^l\|_{X^l}\|\tau^l\|_{X^l}} \geq \alpha(z),$$

(2)

$$\exists \beta > 0, \quad \inf_{\substack{v^l \in A^l \\ \psi^l \in B^l}} \sup_{\tau^l \in X^l} \frac{b'(\tau^l,(v^l,\psi^l))}{\|\tau^l\|_{X^l}(\|v^l\|_{A^l} + \|\psi^l\|_{B^l})} \geq \beta.$$

The constant $\alpha(z)$ is as in proposition 2.2. Proof of the second inf-sup condition can be found in [1]. Applying lemma 2.44 from [15], we have the following estimate.

Lemma 3.9. *Under the inf-sup condition from proposition 3.8, the solution $(\sigma^{J,l},u^{J,l},\eta^{J,l})$ to problem (3.12) satisfies the following estimates*

$$\begin{aligned} &\|\sigma^J - \sigma^{J,l}\|_X \\ &\quad \leq c_1 \inf_{\tau^l \in X^l} \|\sigma^J - \tau^l\|_X + c_2(\inf_{v^l \in A^l} \|u^J - v^l\|_A + \inf_{\psi^l \in B^l} \|\eta^J - \psi^l\|_B), \\ &\|u^J - u^{J,l}\|_A + \|\eta^J - \eta^{J,l}\|_B \\ &\quad \leq c_3 \inf_{\tau^l \in X^l} \|\sigma^J - \tau^l\|_X + c_4(\inf_{v^l \in A^l} \|u^J - v^l\|_A + \inf_{\psi^l \in B^l} \|\eta^J - \psi^l\|_B), \end{aligned}$$

with $c_1 = (1 + \frac{\|a^J\|}{\alpha(z)})(1 + \frac{\|b\|}{\beta})$, $c_2 = \frac{\|b\|}{\alpha(z)}$, $c_3 = c_1\frac{\|a^J\|}{\beta}$, $c_4 = 1 + \frac{\|b\|}{\beta} + c_2\frac{\|a^J\|}{\beta}$ where $\|a^J\| = \|a^J\|_{X,X}$ and $\|b\| = \|b\|_{X,(A,B)}$.

With assumption 2.9 and remark 2.10, we have

$$\begin{aligned}
&\|\sigma^J-\sigma^{J,l}\|_X\\
&\le C\left(1+\exp\left(c\sum_{j=1}^J b_j|z_j|\right)\right)\left(h_l\|\sigma^J\|_{H^1}+h_l\|f\|_{H^1}+h_l\|u^J\|_{H^1}+h_l\|\eta^J\|_{H^1}\right)\\
&\le C\exp\left(c\sum_{j=1}^J b_j|z_j|\right)\left(1+\sum_{j=1}^J b_j^*|z_j|\right)2^{-l},\\
&\|u^J-u^{J,l}\|_A+\|\eta^J-\eta^{J,l}\|_B\\
&\le C\left(1+\exp\left(c\sum_{j=1}^J b_j|z_j|\right)\right)\left(h_l\|\sigma^J\|_{H^1}+h_l\|f\|_{H^1}+h_l\|u^J\|_{H^1}+h_l\|\eta^J\|_{H^1}\right)\\
&\le C\exp\left(c\sum_{j=1}^J b_j|z_j|\right)\left(1+\sum_{j=1}^J b_j^*|z_j|\right)2^{-l}.
\end{aligned}$$

Hence we have the following proposition by the triangle inequality.

Proposition 3.10. *Consider the FE approximation of the truncated mixed problem, under assumption 2.9. There is a constant $C>0$ such that for $z\in U$ the following error bound holds*

$$\|\sigma(\cdot,z)-\sigma^{J,l}(\cdot,z)\|_X\le C\exp\left(c\sum_{j=1}^\infty b_j|z_j|\right)\left(\sum_{j>J}b_j|z_j|+\left(1+\sum_{j=1}^J b_j^*|z_j|\right)2^{-l}\right),$$

$$\|u(\cdot,z)-u^{J,l}(\cdot,z)\|_A+\|\eta(\cdot,z)-\eta^{J,l}(\cdot,z)\|_B$$

$$\le C\exp\left(c\sum_{j=1}^\infty b_j|z_j|\right)\left(\sum_{j>J}b_j|z_j|+\left(1+\sum_{j=1}^J b_j^*|z_j|\right)2^{-l}\right).$$

We recall the definitions of $\mathcal{G}^{J,l}(z)$, $\Phi^{J,l}(z;\delta)$ and $\gamma^{J,l,\delta}$ in section 2.4. We have the following result.

Theorem 3.11. *Under assumption 2.1 and assumption 2.9, there exists a positive constant $C(\delta)$ depending only on the data δ such that for every J and l*

$$d_{Hell}(\gamma^\delta,\gamma^{J,l,\delta})\le C(\delta)(J^{-q}+2^{-l}).$$

The proof is similar to that of proposition 2.7.

4. Multilevel MCMC for Gaussian prior

We now consider the multi-level Markov Chain Monte Carlo (MLMCMC) for sampling the posterior probability under Gaussian prior. We follow the MLMCMC method developed in [20]. We denote the solution of the forward problem as $s=(\sigma,u)$. We consider the MLMCMC approach for approximating the posterior expectation of $\ell(s(\cdot,z))$, where ℓ is a bounded linear map on $X\times M$. To balance the contribution of the truncation error and the FE error in theorem 2.12 and theorem

3.11, we choose J s.t. $J^{-q} = \mathcal{O}(2^{-l})$, i.e. $J = J_l = \lceil 2^{l/q} \rceil$. Then we denote the FE solution of the truncated problem as $s^l = (\sigma^{J_l,l}, u^{J_l,l})$. Similarly, for convenience, we denote $\Phi^{J_l,l}$ as Φ^l and $\gamma^{J_l,l,\delta}$ as γ^l. The MLMCMC sampling procedure for Gaussian prior probability measure is developed in [20]. We recapitulate the details in appendix B. The MLMCMC estimator $E_L^{MLMCMC}[\ell(s)]$ of $\mathbb{E}^{\gamma^\delta}[\ell(s)]$ is

$$\begin{aligned}
&E_L^{MLMCMC}(\ell(s)) \\
=&\sum_{l=1}^{L}\sum_{l'=1}^{L'(l)} \Big[E_{M_{ll'}}^{\gamma^l}\left[A_1^{ll'}\right] + E_{M_{ll'}}^{\gamma^{l-1}}\left[A_2^{ll'}\right] + E_{M_{ll'}}^{\gamma^l}\left[A_3^{l}\right] \cdot E_{M_{ll'}}^{\gamma^{l-1}}\left[A_4^{ll'} + A_8^{ll'}\right] \\
&+ E_{M_{ll'}}^{\gamma^{l-1}}\left[A_5^{l}\right] \cdot E_{M_{ll'}}^{\gamma^l}\left[A_6^{ll'} + A_7^{ll'}\right]\Big] \\
&+ \sum_{l=1}^{L} \Big[E_{M_{l0}}^{\gamma^l}\left[A_1^{l0}\right] + E_{M_{l0}}^{\gamma^{l-1}}\left[A_2^{l0}\right] + E_{M_{l0}}^{\gamma^l}\left[A_3^{l}\right] \cdot E_{M_{l0}}^{\gamma^{l-1}}\left[A_4^{l0} + A_8^{l0}\right] \\
&+ E_{M_{l0}}^{\gamma^{l-1}}\left[A_5^{l}\right] \cdot E_{M_{l0}}^{\gamma^l}\left[A_6^{l0} + A_7^{l0}\right]\Big] \\
&+ \sum_{l'=1}^{L'(0)} E_{M_{0l'}}^{\gamma^0}\left[\ell\left(s^{l'} - s^{l'-1}\right)\right] + E_{M_{00}}^{\gamma^0}\left[\ell\left(s^0\right)\right]
\end{aligned}$$

where

$$\begin{aligned}
A_1^{ll'} &= \left(1 - \exp\left(\Phi^l(z;\delta) - \Phi^{l-1}(z;\delta)\right)\right) Q(z) I^l(z), \\
A_2^{ll'} &= \left(\exp\left(\Phi^{l-1}(z;\delta) - \Phi^l(z;\delta)\right) - 1\right) Q(z) \left(1 - I^l(z)\right), \\
A_3^{l} &= \left(\exp\left(\Phi^l(z,\delta) - \Phi^{l-1}(z,\delta)\right) - 1\right) I^l(z), \\
A_4^{ll'} &= Q(z) I^l(z), \\
A_5^{l} &= \left(1 - \exp\left(\Phi^{l-1}(z,\delta) - \Phi^l(z,\delta)\right)\right)\left(1 - I^l(z)\right), \\
A_6^{ll'} &= \exp\left(\Phi^l(z,\delta) - \Phi^{l-1}(z,\delta)\right) Q(z) I^l(z), \\
A_7^{ll'} &= Q(z)\left(1 - I^l(z)\right), \\
A_8^{ll'} &= \exp\left(\Phi^{l-1}(z,\delta) - \Phi^l(z,\delta)\right) Q(z)\left(1 - I^l(z)\right),
\end{aligned}$$

with $Q(z) = \ell(s^{l'} - s^{l'-1})$ when $l' \geq 1$ and $Q(z) = \ell(s^0)$ when $l' = 0$. The truncation function is

$$I^l(z) = \begin{cases} 1 & \text{if} \quad \Phi^l(z;\delta) - \Phi^{l-1}(z;\delta) \leq 0, \\ 0 & \text{if} \quad \Phi^l(z;\delta) - \Phi^{l-1}(z;\delta) > 0. \end{cases}$$

Here $E_{M_{ll'}}^{\gamma^l}$ denotes the MCMC sample average of the Markov chain generated by MCMC sampling procedure with $M_{ll'}$ samples, and with the acceptance probability

$$\alpha^l(z,z') = 1 \wedge \exp(\Phi^l(z;\delta) - \Phi^l(z';\delta)), \quad z, z' \in U$$

for the independence sampler and the preconditioned Crank-Nicolson (pCN) sampler [18]. Let $\mathcal{E}^L$ denote the expectation with respect to the probability space of the

Markov chains generated in the MLMCMC sampling process. With the following sampling choices,

$$L'(l) := L - l, \quad \text{and} \quad M_{ll'} := (l+l')^a 2^{2(L-(l+l'))} \quad \text{for} \quad l \geq 1, l' \geq 1$$

we have the following estimates for the error $\mathcal{E}^L\{|\mathbb{E}^{\gamma^\delta}[\ell(s)] - E_L^{MLMCMC}[\ell(s)]|\}$ from [20]. The proof of these error estimates in the case of independence sampler

Table 1. Total MLMCMC error and total degrees of freedom with different sample size choices for Gaussian prior

a	$M_{ll'}, l, l' > 1$	$M_{l0} = M_{0l}$	M_{00}	Total error	Total DOFs
0	$2^{2(L-(l+l'))}$	$2^{2(L-l)}/L^2$	$2^{2L}/L^4$	$O\left(L^2 2^{-L}\right)$	$O(2^{dL})$
2	$(l+l')^2\, 2^{2(L-(l+l'))}$	$2^{2(L-l)}$	$2^{2L}/L^2$	$O\left(L \log L 2^{-L}\right)$	$O(L^2 2^{dL})$
3	$(l+l')^3\, 2^{2(L-(l+l'))}$	$l 2^{2(L-l)}$	$2^{2L}/L$	$O\left(L^{1/2} 2^{-L}\right)$	$O(L^3 2^{dL})$
4	$(l+l')^4\, 2^{2(L-(l+l'))}$	$l^2 2^{2(L-l)}$	$2^{2L}/\left(\log L^2\right)$	$O\left(\log L 2^{-L}\right)$	$O(L^4 2^{dL})$

follows the same lines of [20]. For the pCN sampler, if we assume that spectral gap results of Hairer et al. [18], the error estimates can be shown rigorously as in [20].

Remark 4.1. If the FE convergence rate is $O(2^{-tl})$ for the FE mesh size $O(2^{-l})$ for $0 < t < 1$ (e.g. when the solution to the forward equation only possesses the $H^{1+t}(D)$ regularity), with the MCMC sample size $M_{ll'} = (l+l')^a 2^{2t(L-l-l')}$, we get the optimal convergence rate $O(N^{-t/d})$ for the total number of degrees of freedom N (with a possible logarithmic factor).

5. Numerical experiments

5.1. **Linear elasticity problem in the Hellinger-Reissner mixed form.** We present numerical experiments that support the theoretical error estimate for the MLMCMC method developed for the Bayesian inverse problem of the forward elasticity equation in the Hellinger-Reissner mixed form. We restrict our consideration to the periodic boundary condition as this allows for a reference solution to be computed highly accurately with the spectral collocation method [19]. The reference posterior expectation is computed by the Gauss-Hermite quadrature method with 128 quadrature points. We consider the case where the Lamé parameters depend on one random variable where we can compute a highly accurate reference posterior expectation for the purpose of comparison. The MLMCMC method works equally in the general case but a highly accurate reference posterior expectation cannot be obtained by Gaussian quadratures. We first consider the Hellinger-Reissner mixed forward problem in the $(0,1) \times (0,1)$ square domain D in two dimensions with the periodic boundary condition,

$$\begin{cases} (\mathcal{C}^{-1}\sigma, \tau) - (\tau, \nabla^{(s)} u) = 0 & \forall \tau \in [L_2(D)]^{2\times 2}, \\ -(\sigma, \nabla^{(s)} v) = -(f, v) & \forall v \in [H^1(D)]^2, \end{cases} \tag{5.1}$$

where, with $x \in D$, the forcing f is given as

$$\begin{cases} f_1 = 200 \sin(2\pi x_1), \\ f_2 = 200 \sin(2\pi x_2). \end{cases}$$

The random Lamé parameters depend on a random variable $z \sim N(0,1)$ as such,

$$\begin{cases} \mu = \exp(z\cos(2\pi x_1)\sin(2\pi x_2)), \\ \lambda = \exp(z\sin(2\pi x_1)\cos(2\pi x_2)). \end{cases} \tag{5.2}$$

We consider the following observation functional,

$$\mathcal{G}(z) = \int_D (x_1^{\frac{1}{2}}\sigma_{22} + x_2^{\frac{1}{2}}\sigma_{11})dx. \tag{5.3}$$

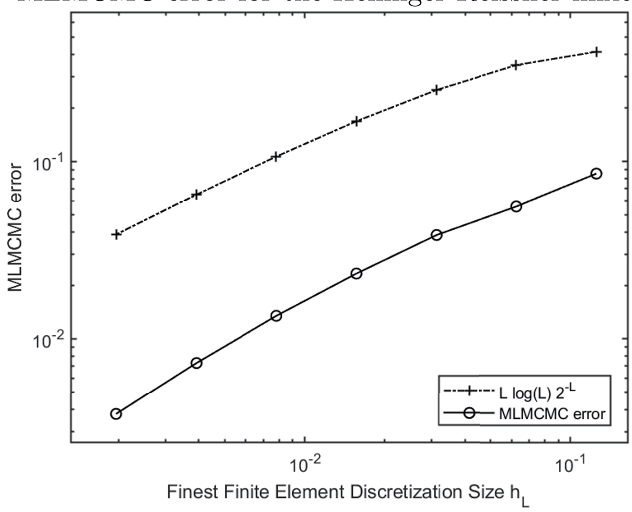
The quantity of interest is,

$$\ell(u(z)) = \int_D \left(x_1^{\frac{1}{2}} \frac{\partial u_1}{\partial x_1} + x_2^{\frac{1}{2}} \frac{\partial u_2}{\partial x_2} \right) dx. \tag{5.4}$$

We assume that the random noise is distributed according to $N(0,1)$. In the MLMCMC approximation, we solve the linear system by GMRES method with Schur complement preconditioner. A random realization of the forward equation is generated. With an additional randomly generated Gaussian noise we have the observation $\delta = -0.9676561$. We ran the MLMCMC with independent sampler for each Markov chain. The average absolute errors of 80 runs of MLMCMC is plotted in figure 1 and figure 2 against the finest mesh size h_L with the constant $a = 2, 3$ in Table 1. The theoretical convergence rate is plotted as a reference to the numerical convergence curve. We see that the error of the MLMCMC approximation behaves as numerically predicted.

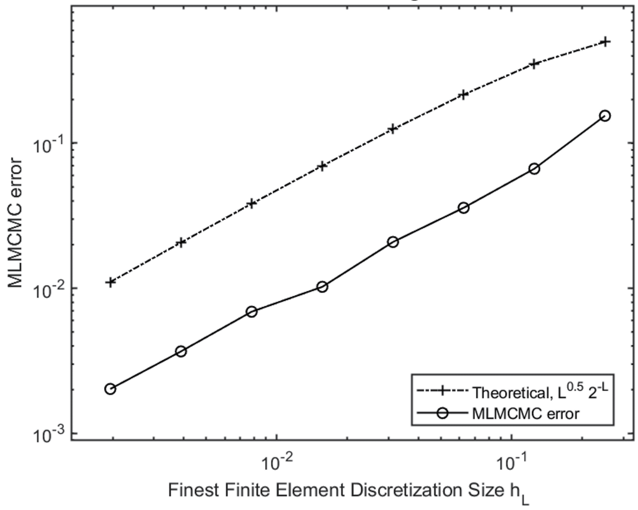
We also present the CPU time for these experiments in Table 2. We record the

FIGURE 1. MLMCMC error for the Hellinger-Reissner mixed form, a=2



CPU time for MLMCMC with sample numbers from Table 1 with $a = 2$ and $a = 3$. To demonstrate the optimal complexity of the MLMCMC sampling method, we compare the CPU time to that required by the plain MCMC method where the forward equation is solved with high resolution level $O(2^{-L})$ for every sample. With M samples, the MCMC sample error is $O(M^{-1/2})$ so to balance the MCMC sampling error and the FE $O(2^{-L})$ error, we choose $M = 2^{2L}$ in this experiment (see, [22]).

FIGURE 2. MLMCMC error for the Hellinger-Reissner mixed form, a=3



The CPU time presented in table 2 is the average CPU time of 5 independent runs. The experiments are performed on a computer with Intel Xeon E5-2620 CPU. The computational advantage of the MLMCMC method over the plain MCMC is clear. The CPU time for the plain MCMC grows so fast that it become intractable rather quickly, while the MLMCMC method's CPU time grows much slower for both cases with $a = 2$ and $a = 3$.

Next we present the results from MLMCMC with the preconditioned Crank-

TABLE 2. CPU time of MLMCMC and plain MCMC for problems in mixed form

L	1	2	3	4	5	6
MLMCMC CPU time for $a = 2$	0.022	0.089	0.489	3.26	25.3	202
MLMCMC CPU time for $a = 3$	0.022	0.131	0.949	8.83	85.4	908
plain MCMC CPU time	0.013	0.099	1.77	41.8	1180	36000

Nicolson (pCN) sampler. Independence sampler is known for low acceptance rate in practice. In the following numerical experiment, independence sampler has a low acceptance rate which is less than 10%. Hence we adopted the pCN sampler for each MLMCMC chain, where the proposal $\omega^{(k)}$ is generated by

$$\omega^{(k)} = \sqrt{1-\beta^2} z^{(k)} + \beta\xi,$$

with $\xi \sim N(0,1)$. We choose $\beta = 0.3$. The acceptance rate of all chains in MLMCMC is raised to 40-50%. We solve problem (5.1) with forcing,

$$\begin{cases} f_1 = 1000\sin(2\pi x_1), \\ f_2 = 1000\sin(2\pi x_2). \end{cases}$$

The random Lamé parameters which depend on random variable $z \sim N(0,1)$ are defined as in equation (5.2). We consider the same observation functional (5.3) and quantity of interest (5.4). We ran the MLMCMC with pCN sampler for each Markov chain. The average MLMCMC absolute errors of 80 independent runs are plotted in figure 3 and figure 4 against the finest mesh size h_L with the constant $a = 2, 3$ in Table 1. The theoretical convergence rate is plotted as a reference to the numerical convergence curve. We see that the error of the MLMCMC approximation behaves as numerically predicted.

We now present a numerical experiment for the case where the forcing f depends on

FIGURE 3. MLMCMC error for the Hellinger-Reissner mixed form by pCN sampler, a=2

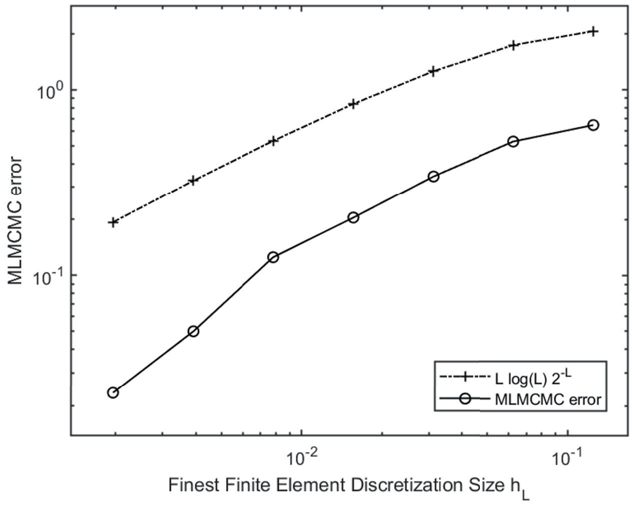
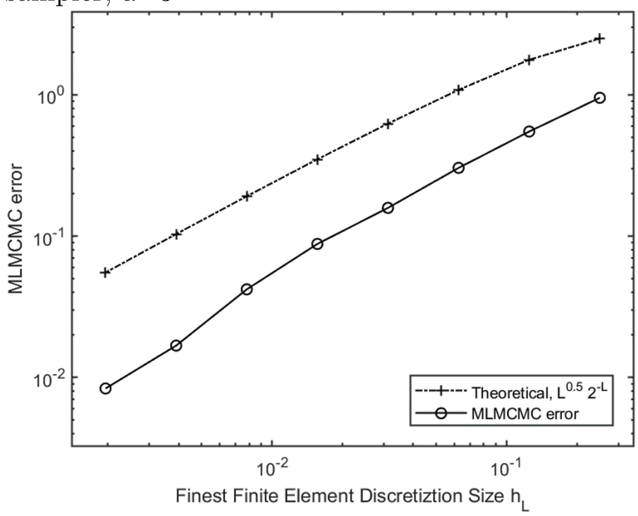


FIGURE 4. MLMCMC error for the Hellinger-Reissner mixed form by pCN sampler, a=3



25 random variables. We consider the same problem and forcing as in equation (5.1).

Let $\{\phi_n(t)\}_{n=1}^5 = \{1, \sin(2\pi t), \cos(2\pi t), \sin(4\pi t), \cos(4\pi t)\}$, we have the following expansion of μ and λ,

$$\begin{cases} \mu = \exp\left(\sum_{i=1}^5 \sum_{j=1}^5 c_{ij} z_{ij} \phi_i(x)\phi_j(y)\right), \\ \lambda = \exp\left(\sum_{i=1}^5 \sum_{j=1}^5 c_{ij} z_{ij} \phi_j(x)\phi_i(y)\right), \end{cases} \tag{5.5}$$

where c_{ij} are constants which represent the decaying rate of the power of -1.5 for the 25 coefficients as listed in table 3; and $z_{ij} \sim N(0,1)$. We consider the observation and quantities of interest as in equation (5.3) and equation (5.4). Due to the high dimensionality of the probability space, we are not able to get a highly accurate reference by quadrature rules for integrals. To obtain a reference value for the posterior expectation, we average several plain MCMC runs with a large number of samples, where the forward equation is solved with a highly accurate spectral collocation solver. The reference value is -0.957102, with a noisy observation $\delta = 0.592839$. With the reference value mentioned above, we perform MLMCMC for $L = 1$ to 6. The average absolute errors of 16 MLMCMC runs are listed in table 4. These indicate a clear trend of convergence which agrees with our theory.

TABLE 3. Constant c_{ij}

	i=1	i=2	i=3	i=4	i=5
j=1	$1^{-1.5}$	$2^{-1.5}$	$4^{-1.5}$	$7^{-1.5}$	$11^{-1.5}$
j=2	$3^{-1.5}$	$5^{-1.5}$	$8^{-1.5}$	$12^{-1.5}$	$16^{-1.5}$
j=3	$6^{-1.5}$	$9^{-1.5}$	$13^{-1.5}$	$17^{-1.5}$	$20^{-1.5}$
j=4	$10^{-1.5}$	$14^{-1.5}$	$18^{-1.5}$	$21^{-1.5}$	$23^{-1.5}$
j=5	$15^{-1.5}$	$19^{-1.5}$	$22^{-1.5}$	$24^{-1.5}$	$25^{-1.5}$

TABLE 4. MLMCMC error for the mixed forward problem with 25 random variables

L	1	2	3	4	5	6
error for $a = 2$	0.4317	0.6370	0.4503	0.2040	0.2010	0.0891
error for $a = 3$	0.4317	0.3669	0.1666	0.1198	0.0708	0.0372

5.2. **Linear elasticity problem in the Hellinger-Reissner dual mixed form.** Next we present numerical experiments that support the theoretical error estimates for Bayesian inverse problem of the Hellinger-Reissner dual mixed form of the elasticity problem with Gaussian prior. We consider the Hellinger-Reissner dual mixed forward problem in the $(0,1) \times (0,1)$ square domain D in two dimensions with the periodic boundary condition,

$$\begin{cases} (\mathcal{C}^{-1}\sigma, \tau) + (\nabla \cdot \tau, u) + (as(\tau), \eta) = 0 & \forall \tau \in H(\text{div}, D) \\ (\nabla \cdot \sigma, v) = -(f, v) & \forall v \in [L^2(D)]^2 \\ (as(\sigma), \psi) = 0 & \forall \psi \in L^2(D). \end{cases} \tag{5.6}$$

With $x = (x_1, x_2) \in D$, the forcing f is given as

$$\begin{cases} f_1 = 200\sin(2\pi x_1), \\ f_2 = 200\sin(2\pi x_2). \end{cases}$$

The random Lamé parameters depend on a random variable $z \sim N(0,1)$ as,

$$\begin{cases} \mu = \exp(z\cos(2\pi x_1)\sin(2\pi x_2)), \\ \lambda = \exp(z\sin(2\pi x_1)\cos(2\pi x_2)). \end{cases} \tag{5.7}$$

We consider the following observation functional,

$$\mathcal{G}(z) = \int_D (x_1^{\frac{1}{2}}\sigma_{22} + x_2^{\frac{1}{2}}\sigma_{11})dx. \tag{5.8}$$

The quantity of interest is,

$$\ell(u(z)) = 100\int_D \left(x_1^{\frac{1}{2}}u_1 + x_2^{\frac{1}{2}}u_2\right)dx. \tag{5.9}$$

We assume that the random noise is distributed according to $N(0,1)$. A random sample of z and a Gaussian noise is generated by random generator. We get the observation $\delta = -1.374767$. In the MLMCMC experiment, the solution of the forward problem is approximated with PEERS element. We solve the linear system with GMRES method with Schur complement preconditioner. The MLMCMC approximation experiment is ran with independent sampler for each Markov chain. The average MLMCMC errors of 80 independent runs are plotted in figure 5 and figure 6 against the finest mesh size h_L with the constant $a = 2, 3$ in table 1. The theoretical convergence rate is plotted as a reference to the numerical convergence curve. We see that the error of the MLMCMC approximation behaves as numerically predicted.

FIGURE 5. MLMCMC error for the Hellinger-Reissner dual mixed form, a=2

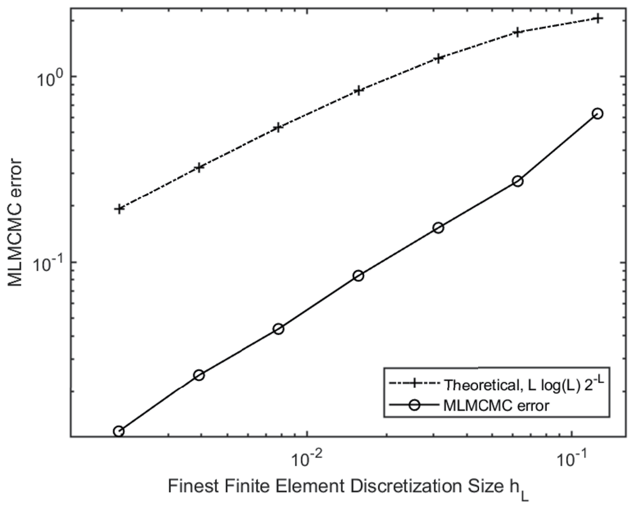
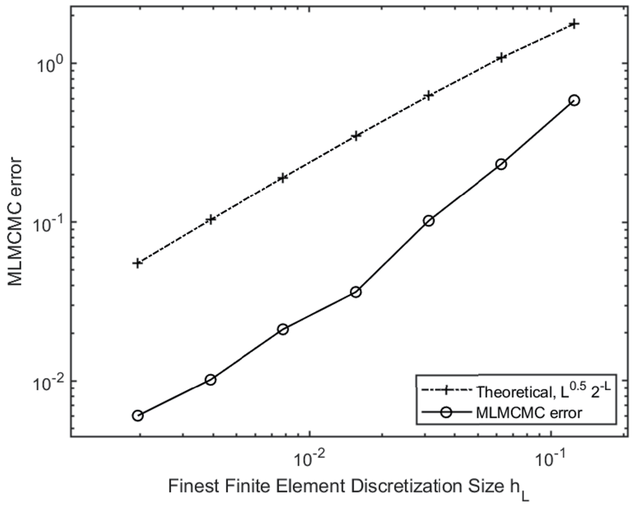


FIGURE 6. MLMCMC error for the Hellinger-Reissner dual mixed form, a=3



We present the CPU time for the numerical experiments of the dual mixed form in table 5. CPU time for MLMCMC with sample number from table 1 with $a = 2$ and $a = 3$ is presented. CPU time for plain MCMC with sample number of 2^{2L} is also included. From the table, the MLMCMC clearly outperforms plain MCMC in terms of computational efficiency.

Next, we present the results from MLMCMC approximation where each Markov

TABLE 5. CPU time of MLMCMC and plain MCMC for problems in dual mixed form

L	1	2	3	4	5	6
MLMCMC CPU time for $a = 2$	0.096	0.552	2.94	21.0	149	1110
MLMCMC CPU time for $a = 3$	0.105	0.781	5.90	49.9	494	4600
plain MCMC CPU time	0.076	0.629	10.5	237	6550	194000

chain is generated with pCN sampler. We choose $\beta = 0.5$ and the acceptance rate of all chains in MLMCMC is raised to 40-50%. With the same setting, we find that the independence sampler has acceptance rate less than 10%. We solve problem (5.6) with the forcing,

$$\begin{cases} f_1 = 500\sin(2\pi x_1), \\ f_2 = 500\cos(2\pi x_2). \end{cases}$$

The random Lamé parameters depend on a random variable $z \sim N(0,1)$ as in equation (5.7). We consider the same observation functional (5.8) and the quantity of interest (5.9). We run the MLMCMC with pCN sampler for each Markov chain with sample sizes $M_{ll'} = (l + l')^a 2^{2(L-(l+l'))}$ from table 1 for $a = 2, 3$. The average MLMCMC absolute errors of 80 independent runs are plotted in figure 7 and figure 8 against the finest mesh size h_L. The theoretical convergence rate is plotted as a reference to the numerical convergence curve. We see that the error of the MLMCMC approximation behaves as numerically predicted.

FIGURE 7. MLMCMC error for the Hellinger-Reissner dual mixed form by pCN sampler, a=2

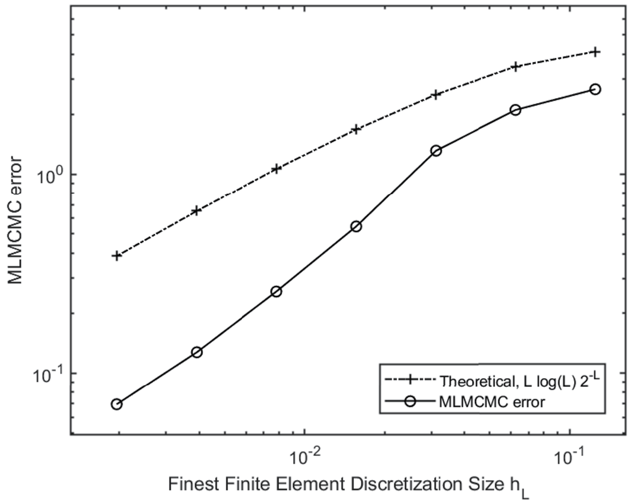
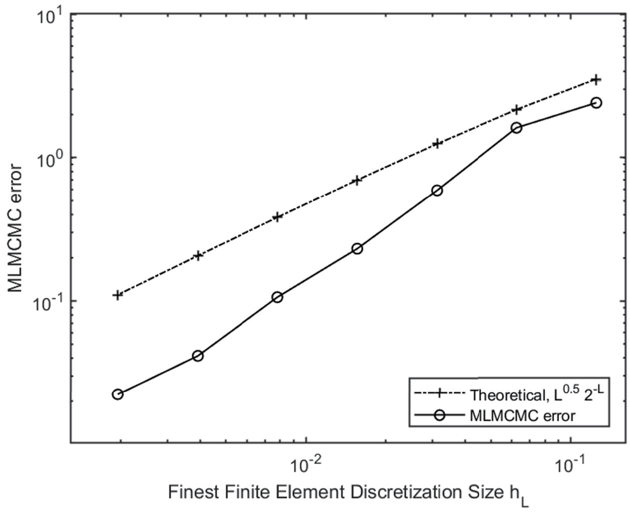


FIGURE 8. MLMCMC error for the Hellinger-Reissner dual mixed form by pCN sampler, a=3



Appendix A

Before we present the proof for proposition 2.5 and proposition 2.7, we note the following inequalities whose proof can be found in [21]. For $s > 0$, the following inequalities hold

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \exp\left(-t^2/2 + |t|s\right) \frac{dt}{\sqrt{2\pi}} \le c \exp\left(s^2/2\right) \exp(s\sqrt{2/\pi}),\\
&\int_{-\infty}^{\infty} t^2 \exp\left(-t^2/2 + |t|s\right) \frac{dt}{\sqrt{2\pi}} \le c \exp\left(s^2/2\right)\left(1+s^2\right),\\
&\int_{-\infty}^{\infty} |t| \exp\left(-t^2/2 + |t|s\right) \frac{dt}{\sqrt{2\pi}} \le c \exp\left(s^2/2\right)(1+s).
\end{aligned} \tag{5.10}
$$

Proof for proposition 2.5. The normalizing constant in (2.17) is

$$Z(\delta) = \int_U \exp(-\Phi(z;\delta)) d\gamma(z). \tag{5.11}$$

We first show that for each $r > 0$, there is a positive constant $K(r)$ such that $Z(\delta) \ge K(r)$ when $|\delta|_\Sigma \le r$. We note that for given data δ,

$$\forall z \in U: \quad |\Phi(z;\delta)| \le \frac{1}{2}(|\delta|_\Sigma + |\mathcal{G}(z)|_\Sigma)^2.$$

From (5.10), we have

$$
\begin{aligned}
&\int_U \Phi(z;\delta) d\gamma(z)\\
&\le C\left(|\delta|_\Sigma^2 + \int_U |\delta|_\Sigma \exp\left(\sum_{j=1}^{\infty} b_j|z_j|\right) d\gamma(z) + \int_U \exp\left(2\sum_{j=1}^{\infty} b_j|z_j|\right) d\gamma(z)\right)\\
&< \Lambda(r),
\end{aligned}
$$

where $\Lambda(r)$ is a constant depending on r. For each positive constant Λ, there exists $c > 0$ such that $\gamma(z \in U : \Phi(z,\delta) > c) < \Lambda/c$. Thus the measure of the set of all $z \in U$ such that $\Phi(z,\delta) \le c$ is larger than $1 - \Lambda/c$. It follows that when $|\delta|_\Sigma \le r$,

$$Z(\delta) > (1 - \Lambda/c) \exp(-c) =: K(r) > 0. \tag{5.12}$$

Using the inequality $|\exp(x) - \exp(y)| \le |x - y|(\exp(x) + \exp(y))$ for all x and y, we find

$$|Z(\delta) - Z(\delta')| \le \int_U |\Phi(z;\delta) - \Phi(z;\delta')| d\gamma(z). \tag{5.13}$$

We note that for every $z \in U$,

$$
\begin{aligned}
|\Phi(z;\delta) - \Phi(z;\delta')| &\le \frac{1}{2}\left|\left\langle \Sigma^{-1/2}(\delta + \delta' - 2\mathcal{G}(z)), \Sigma^{-1/2}(\delta - \delta')\right\rangle\right|\\
&\le \frac{1}{2}(|\delta|_\Sigma + |\delta'|_\Sigma + 2|\mathcal{G}|_\Sigma)|\delta - \delta'|_\Sigma\\
&\le c\left(r + \frac{d\lambda_{\max}(z) + 2\mu_{\max}(z)}{\mu_{\min}(z)}\right)|\delta - \delta'|_\Sigma.
\end{aligned}
$$

Let $G(r,z) = c\left(r + \frac{d\lambda_{\max}(z)+2\mu_{\max}(z)}{\mu_{\min}(z)}\right)$, we have

$$|\Phi(z;\delta) - \Phi(z;\delta)| \leq G(r,z)|\delta - \delta'|_\Sigma. \tag{5.14}$$

Next we write the Hellinger distance as

$$\begin{aligned} 2d_{Hell}(\gamma^\delta,\gamma^{\delta'})^2 &= \int_U (Z(\delta)^{-1/2}\exp(\frac{1}{2}\Phi(z;\delta)) - Z(\delta')^{-1/2}\exp(\frac{1}{2}\Phi(z;\delta')))^2 d\gamma(z) \\ &\leq I_1 + I_2, \end{aligned} \tag{5.15}$$

where

$$\begin{aligned} I_1 &= \frac{2}{Z(\delta)}\int_U (\exp(-\frac{1}{2}\Phi(z;\delta)) - \exp(-\frac{1}{2}\Phi(z;\delta')))^2 d\gamma(z) \\ I_2 &= 2|Z(\delta)^{-1/2} - Z(\delta')^{-1/2}|^2 \int_U \exp(-\Phi(z;\delta'))d\gamma(z). \end{aligned}$$

With inequalities (5.12), (5.13) and (5.14), we have

$$\begin{aligned} I_1 &\leq K(r)\int_U |\Phi(z;\delta) - \Phi(z;\delta')|^2 d\gamma(z) \\ &\leq K(r)\int_U (G(r,z))^2 d\gamma(z)|\delta - \delta'|_\Sigma^2 \leq K(r)|\delta-\delta'|_\Sigma^2, \end{aligned}$$

Similarly, we have the same bound for I_2. □

Proof for proposition 2.7. Here write the Hellinger distance as in (5.15). We first consider the bound for I_1, given data δ, for every $z \in U$

$$\begin{aligned} \left|\exp(-\frac{1}{2}\Phi(z;\delta)) - \exp(-\frac{1}{2}\Phi^J(z;\delta))\right| &\leq \frac{1}{2}|\Phi(z;\delta) - \Phi^J(z;\delta)| \\ &\leq C(2|\delta| + |\mathcal{G}(z)| + |\mathcal{G}^J(z)|)|\mathcal{G}(z) - \mathcal{G}^J(z)| \\ &\leq C(\delta)\left(\sum_{j>J} b_j|z_j|\right)\exp\left(c\sum_{j=1}^\infty b_j|z_j|\right) \end{aligned}$$

Thus with lemma 5.10, we have

$$I_1 \leq C(\delta)\int_U \exp\left(c\sum_{j=1}^\infty |z_j|b_j\right)\left(\sum_{j>J}|z_j|b_j\right)^2 d\gamma(z).$$

Consider the right hand size of the equation, we have

$$\begin{aligned}
&\int_U \exp\left(c\sum_{j=1}^{\infty}|z_j|b_j\right)\left(\sum_{j>J}|z_j|b_j\right)^2 d\gamma(z)\\
&=\int_U \exp\left(c\sum_{j=1}^{\infty}|z_j|b_j\right)\left(\sum_{i,j>J}b_ib_j|z_i||z_j|\right)d\gamma(z)\\
&\le\sum_{i>J}b_i^2\int_{-\infty}^{\infty}\exp(cb_i|z_i|)z_i^2d\gamma_1(z_i)\prod_{\substack{k=1\\k\neq i}}^{\infty}\int_{-\infty}^{\infty}\exp(cb_k|z_k|)d\gamma_1(z_k)\\
&\quad+\sum_{i,j>J}b_ib_j\int_{-\infty}^{\infty}\exp\left(cb_i\left|z_i\right|\right)\left|z_i\right|d\gamma_1\left(z_i\right)\cdot\int_{-\infty}^{\infty}\exp\left(cb_j\left|z_j\right|\right)\left|z_j\right|d\gamma_1\left(z_j\right)\\
&\qquad\cdot\prod_{\substack{k=1\\k\neq i,j}}^{\infty}\int_{-\infty}^{\infty}\exp\left(cb_k\left|z_k\right|\right)d\gamma_1\left(z_k\right).
\end{aligned}$$

With the inequalities from (5.10), we have

$$\begin{aligned}
&\int_U \exp\left(c\sum_{j=1}^{\infty}b_j\left|z_j\right|\right)\left(\sum_{j>J}b_j\left|z_j\right|\right)^2 d\gamma(z)\\
&\le C\sum_{i>J}b_i^2\left(1+b_i^2\right)\exp\left(\sum_{k=1}^{\infty}c^2b_k^2/2+cb_k\sqrt{2/\pi}\right)\\
&\quad+C\sum_{i,j>1}b_ib_j\left(1+b_i\right)\left(1+b_j\right)\exp\left(\sum_{k=1}^{\infty}c^2b_k^2/2+cb_k\sqrt{2/\pi}\right)\\
&\le C\left(\sum_{j>J}b_i\right)^2\le CJ^{-2q}.
\end{aligned}\tag{5.16}$$

Conclude the estimation for both I_1 and I_2, we have

$$d_{Hell}(\gamma^{\delta},\gamma^{J,\delta})\le C(\delta)J^{-q}.$$

□

Appendix B

For completeness, we present the derivation of MLMCMC for the Gaussian prior in [20] in this appendix. We follow the setting in section 4. There is a positive constant C independent of L such that

$$|\mathbb{E}^{\gamma^{\delta}}[\ell(s)]-\mathbb{E}^{\gamma^{\delta}}[\ell(s^L)]|\le C2^{-L}.\tag{5.17}$$

Before we derive the MLMCMC, we introduce the following indication function to deal with the unboundedness of $\exp(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta))$. We denote by

$$I^l(z)=\begin{cases}1 & \text{if} \quad \Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\le 0,\\ 0 & \text{if} \quad \Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)> 0.\end{cases} \tag{5.18}$$

Let $Q(z)$ be the quantity of interest. Let $Z^l=\int_U \exp(-\Phi^l(z;\delta))d\gamma(z)$. For $l\ge 1$, we have

$$\begin{aligned}
&\mathbb{E}^{\gamma^l}[Q(z)]-\mathbb{E}^{\gamma^{l-1}}[Q(z)]\\
&=\frac{1}{Z^l}\int_U \exp\left(-\Phi^l(z;\delta)\right)Q(z)I^l(z)d\gamma(z)\\
&\quad-\frac{1}{Z^{l-1}}\int_U \exp\left(-\Phi^{l-1}(z;\delta)\right)Q(z)I^l(z)d\gamma(z)\\
&\quad+\frac{1}{Z^l}\int_U \exp\left(-\Phi^l(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
&\quad-\frac{1}{Z^{l-1}}\int_U \exp\left(-\Phi^{l-1}(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
&=\frac{1}{Z^l}\int_U \left(\exp\left(-\Phi^l(z;\delta)\right)-\exp\left(-\Phi^{l-1}(z;\delta)\right)\right)Q(z)I^l(z)d\gamma(z)\\
&\quad+\left(\frac{1}{Z^l}-\frac{1}{Z^{l-1}}\right)\int_U \exp\left(-\Phi^{l-1}(z;\delta)\right)Q(z)I^l(z)d\gamma(z)\\
&\quad-\frac{1}{Z^{l-1}}\int_U \left(\exp\left(-\Phi^{l-1}(z;\delta)\right)-\exp\left(-\Phi^l(z;\delta)\right)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
&\quad+\left(\frac{1}{Z^l}-\frac{1}{Z^{l-1}}\right)\int_U \exp\left(-\Phi^l(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z).
\end{aligned}$$

Let $A_1^l=\left(1-\exp\left(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\right)\right)Q(z)I^l(z)$. We have

$$\begin{aligned}
&\frac{1}{Z^l}\int_U \left(\exp\left(-\Phi^l(z;\delta)\right)-\exp\left(-\Phi^{l-1}(z;\delta)\right)\right)Q(z)I^l(z)d\gamma(z)\\
&=\mathbb{E}^{\gamma^l}\left[\left(1-\exp\left(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\right)Q(z)I^l(z)\right)\right]=\mathbb{E}^{\gamma^l}\left[A_1^l\right].
\end{aligned}$$

Let $A_2^l=\left(\exp\left(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\right)-1\right)Q(z)\left(1-I^l(z)\right)$. We write

$$\begin{aligned}
&-\frac{1}{Z^{l-1}}\int_U \left(\exp\left(-\Phi^{l-1}(z;\delta)\right)-\exp\left(-\Phi^l(z;\delta)\right)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
&=\mathbb{E}^{\gamma^{l-1}}\left[\left(\exp\left(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\right)-1\right)Q(z)\left(1-I^l(z)\right)\right]=\mathbb{E}^{\gamma^{l-1}}\left[A_2^l\right].
\end{aligned}$$

We note that

$$\begin{aligned}
&\frac{1}{Z^l}-\frac{1}{Z^{l-1}}\\
&=\frac{1}{Z^lZ^{l-1}}\int_U\Big(\exp\Big(-\Phi^{l-1}(z;\delta)\Big)-\exp\Big(-\Phi^l(z;\delta)\Big)\Big)\Big(I^l(z)+1-I^l(z)\Big)d\gamma(z)\\
&=\frac{1}{Z^lZ^{l-1}}\int_U\exp\Big(-\Phi^l(z;\delta)\Big)\Big(\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)-1\Big)I^l(z)d\gamma(z)\\
&+\frac{1}{Z^lZ^{l-1}}\int_U\exp\Big(-\Phi^{l-1}(z;\delta)\Big)\Big(1-\exp\Big(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\Big)\Big)\Big(1-I^l(z)\Big)d\gamma(z)\\
&=\frac{1}{Z^{l-1}}\mathbb{E}^{\gamma^l}\Big[\Big(\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)-1\Big)I^l(z)\Big]+\\
&+\frac{1}{Z^l}\mathbb{E}^{\gamma^{l-1}}\Big[\Big(1-\exp\Big(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\Big)\Big)\Big(1-I^l(z)\Big)\Big].
\end{aligned}$$

Thus

$$\begin{aligned}
&\left(\frac{1}{Z^l}-\frac{1}{Z^{l-1}}\right)\int_U\exp\Big(-\Phi^{l-1}(z;\delta)\Big)Q(z)I^l(z)d\gamma(z)\\
=&\mathbb{E}^{\gamma^l}\Big[\Big(\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)-1\Big)I^l(z)\Big]\cdot\\
&\frac{1}{Z^{l-1}}\int_U\exp\Big(-\Phi^{l-1}(z;\delta)\Big)Q(z)I^l(z)d\gamma(z)\\
&+\mathbb{E}^{\gamma^{l-1}}\Big[\Big(1-\exp\Big(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\Big)\Big)\Big(1-I^l(z)\Big)\Big]\cdot\\
&\frac{1}{Z^l}\int_U\exp\Big(-\Phi^l(z;\delta)\Big)\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)Q(z)I^l(z)d\gamma(z)\\
=&\mathbb{E}^{\gamma^l}\Big[A_3^l\Big]\mathbb{E}^{\gamma^{l-1}}\Big[A_4^l\Big]+\mathbb{E}^{\gamma-1}\Big[A_5^l\Big]\mathbb{E}^{\gamma^l}\Big[A_6^l\Big],
\end{aligned}$$

where

$$\begin{aligned}
A_3^l&=\Big(\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)-1\Big)I^l(z)\\
A_4^l&=Q(z)I^l(z)\\
A_5^l&=\Big(1-\exp\Big(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\Big)\Big)\Big(1-I^l(z)\Big)\\
A_6^l&=\exp\Big(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\Big)Q(z)I^l(z).
\end{aligned}$$

Similarly, defining for $l\geq 1$

$$A_7^l=Q(z)\Big(1-I^l(z)\Big)\text{ and }A_8^l=\exp\Big(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\Big)Q(z)\Big(1-I^l(z)\Big),$$

we have

$$
\begin{aligned}
&\left(\frac{1}{Z^l}-\frac{1}{Z^{l-1}}\right)\int_U \exp\left(-\Phi^l(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
=&\mathbb{E}^{\gamma^{l-1}}\left[\left(1-\exp\left(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\right)\right)\left(1-I^l(z)\right)\right].\\
&\frac{1}{Z^l}\int_U \exp\left(-\Phi^l(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
&+\mathbb{E}^{\gamma^l}\left[\left(\exp\left(\Phi^l(z;\delta)-\Phi^{l-1}(z;\delta)\right)-1\right)I^l(z)\right].\\
&\frac{1}{Z^{l-1}}\int_U \exp\left(-\Phi^{l-1}(z;\delta)\right)\exp\left(\Phi^{l-1}(z;\delta)-\Phi^l(z;\delta)\right)Q(z)\left(1-I^l(z)\right)d\gamma(z)\\
=&\mathbb{E}^{\gamma^{l-1}}\left[A_5^l\right]\mathbb{E}^{\gamma^l}\left[A_7^l\right]+\mathbb{E}^{\gamma^l}\left[A_3^l\right]\mathbb{E}^{\gamma^{l-1}}\left[A_8^l\right].
\end{aligned}
$$

We conclude that, for every $l \geq 1$,

$$
\begin{aligned}
&\mathbb{E}^{\gamma^l}[Q(z)]-\mathbb{E}^{\gamma^{l-1}}[Q(z)]\\
&\qquad=\mathbb{E}^{\gamma^l}\left[A_1^l\right]+\mathbb{E}^{\gamma^{l-1}}\left[A_2^l\right]+\mathbb{E}^{\gamma^l}\left[A_3^l\right]\cdot\mathbb{E}^{\gamma^{l-1}}\left[A_4^l+A_8^l\right]\\
&\qquad\quad+\mathbb{E}^{\gamma^{l-1}}\left[A_5^l\right]\cdot\mathbb{E}^{\gamma^l}\left[A_6^l+A_7^l\right].
\end{aligned}
$$

When $Q=\ell(s^{l'}-s^{l'-1})$, we denote A_1^l as $A_1^{ll'}, A_2^l$ as $A_2^{ll'}, A_4^l$ as $A_4^{ll'}, A_6^l$ as $A_6^{ll'}, A_7^l$ as $A_7^{ll'}$ and A_8^l as $A_8^{ll'}$. In the case of $Q=\ell\left(s^0\right)$, we denote A_1^l as A_1^{l0}, A_2^l as A_2^{l0}, A_4^l as A_4^{l0}, A_6^l as A_6^{l0}, A_7^l as A_7^{l0} and A_8^l as A_8^{l0}.
We therefore approximate $\mathbb{E}^{\gamma^L}[\ell(s(z))]$ by

$$
\begin{aligned}
&\sum_{l=1}^{L}\sum_{l'=1}^{L'(l)}\left[\mathbb{E}^{\gamma^l}\left[A_1^{ll'}\right]+\mathbb{E}^{\gamma^{l-1}}\left[A_2^{ll'}\right]+\mathbb{E}^{\gamma^l}\left[A_3^l\right]\cdot\mathbb{E}^{\gamma^{l-1}}\left[A_4^{ll'}+A_8^{ll'}\right]\right.\\
&\left.+\mathbb{E}^{\gamma^{l-1}}\left[A_5^l\right]\cdot\mathbb{E}^{\gamma^l}\left[A_6^{ll'}+A_7^{ll'}\right]\right]\\
&+\sum_{l=1}^{L}\left[\mathbb{E}^{\gamma^l}\left[A_1^{l0}\right]+\mathbb{E}^{\gamma^{l-1}}\left[A_2^{l0}\right]+\mathbb{E}^{\gamma^l}\left[A_3^l\right]\cdot\mathbb{E}^{\gamma^{l-1}}\left[A_4^{l0}+A_8^{l0}\right]\right.\\
&\left.+\mathbb{E}^{\gamma^{l-1}}\left[A_5^l\right]\cdot\mathbb{E}^{\gamma^l}\left[A_6^{l0}+A_7^{l0}\right]\right]\\
&+\sum_{l'=1}^{L'(0)}\mathbb{E}^{\gamma^0}\left[\ell\left(s^{l'}-s^{l'-1}\right)\right]+\mathbb{E}^{\gamma^0}\left[\ell\left(s^0\right)\right]
\end{aligned}
$$

The Multilevel Markov Chain Monte Carlo estimator is now defined by replacing the expectations in the preceding expression by finite sample MCMC averages, i.e.

$$\begin{aligned}
&E_L^{MLMCMC}(\ell(s)) \\
=&\sum_{l=1}^{L}\sum_{l'=1}^{L'(l)}\Big[E_{M_{ll'}}^{\gamma^l}\Big[A_1^{ll'}\Big]+E_{M_{ll'}}^{\gamma^{l-1}}\Big[A_2^{ll'}\Big]+E_{M_{ll'}}^{\gamma^l}\Big[A_3^{l}\Big]\cdot E_{M_{ll'}}^{\gamma^{l-1}}\Big[A_4^{ll'}+A_8^{ll'}\Big] \\
&+E_{M_{ll'}}^{\gamma^{l-1}}\Big[A_5^{l}\Big]\cdot E_{M_{ll'}}^{\gamma^l}\Big[A_6^{ll'}+A_7^{ll'}\Big]\Big] \\
&+\sum_{l=1}^{L}\Big[E_{M_{l0}}^{\gamma^l}\Big[A_1^{l0}\Big]+E_{M_{l0}}^{\gamma^{l-1}}\Big[A_2^{l0}\Big]+E_{M_{l0}}^{\gamma^l}\Big[A_3^{l}\Big]\cdot E_{M_{l0}}^{\gamma^{l-1}}\Big[A_4^{l0}+A_8^{l0}\Big] \\
&+E_{M_{l0}}^{\gamma^{l-1}}\Big[A_5^{l}\Big]\cdot E_{M_{l0}}^{\gamma^l}\Big[A_6^{l0}+A_7^{l0}\Big]\Big] \\
&+\sum_{l'=1}^{L'(0)}E_{M_{0l'}}^{\gamma^0}\Big[\ell\Big(s^{l'}-s^{l'-1}\Big)\Big]+E_{M_{00}}^{\gamma^0}\big[\ell\left(s^0\right)\big]
\end{aligned}$$

where $E_{M_{ll'}}^{\gamma^l}$ denotes the $M_{ll'}$ sample average of the MCMC chain. As the FE error estimate of the forward equation is Gaussian integrable, we derive the same error estimate for the MLMCMC estimator as in Table 1, following the procedure in [20].

Acknowledgement

Yang Juntao would like to thank the Singapore Economics Development Board and Nvidia Singapore for the Industrial Postgraduate Program PhD scholarship. He is also grateful to Dr. Simon See and Jeff Adie from Nvidia AI technology center for supporting his pursuit in the direction of data inversion problems. Viet Ha Hoang acknowledges the Tier 2 grant MOE2017-T2-2-144 awarded by the Singapore Ministry of Education.

References

[1] D. N. Arnold, F. Brezzi and J. Douglas, Jr., *PEERS: a new mixed finite element for plane elasticity*, Japan J. Appl. Math. **1** (1984), 347–367.

[2] D. N. Arnold and R. Winther, *Mixed finite elements for elasticity*, Numer. Math. **92** (2002), 401–419.

[3] D. Boffi, F. Brezzi and M. Fortin, *Reduced symmetry elements in linear elasticity*, Commun. Pure Appl. Anal. **8** (2009), 95–121.

[4] V. I. Bogachev, *Gaussian Measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998.

[5] D. Braess, *Finite Elements*, Cambridge University Press, Cambridge, second ed., 2001. Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker.

[6] F. Brezzi, J. Douglas, Jr. and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235.

[7] A. Cohen and R. DeVore, *Approximation of high-dimensional parametric PDEs*, Acta Numer. **24** (2015), 1–159.

[8] A. Cohen, R. DeVore, and C. Schwab, *Convergence rates of best N -term Galerkin approximations for a class of elliptic sPDEs*, Found. Comput. Math. **10** (2010), 615–646.

[9] A. Cohen, R. Devore and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's*, Anal. Appl. (Singap.) **9** (2011), 1–47.

[10] S. L. Cotter, M. Dashti, J. C. Robinson and A. M. Stuart, *Bayesian inverse problems for functions and applications to fluid mechanics*, Inverse Problems **25** (2009): 115008, 43.

[11] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Computer Science and Applied Mathematics, Academic Press, Inc., Orlando, FL, second ed., 1984.

[12] J. Dick, R. N. Gantner, Q. T. Le Gia and C. Schwab, *Multilevel higher-order quasi-Monte Carlo Bayesian estimation*, Math. Models Methods Appl. Sci. **27** (2017), 953–995.

[13] T. J. Dodwell, C. Ketelsen, R. Scheichl, and A. L. Teckentrup, *A hierarchical multilevel Markov chain Monte Carlo algorithm with applications to uncertainty quantification in subsurface flow*, SIAM/ASA J. Uncertain. Quantif. **3** (2015), 1075–1108.

[14] Y. Efendiev, B. Jin, M. Presho and X. Tan, *Multilevel Markov chain Monte Carlo method for high-contrast single-phase flow problems*, Commun. Comput. Phys. **17** (2015), 259–286.

[15] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004.

[16] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numer., **24** (2015), 259–328.

[17] M. B. Giles, *An introduction to multilevel Monte Carlo methods*, in: Proceedings of the International Congress of Mathematicians–Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3571–3590.

[18] M. Hairer, A. M. Stuart and S. J. Vollmer, *Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions*, Ann. Appl. Probab., **24** (2014), 2455–2490.

[19] J. S. Hesthaven, S. Gottlieb and D. Gottlieb, *Spectral Methods for Time-Dependent Problems*, Cambridge Monographs on Applied and Computational Mathematics, vol. 21, Cambridge University Press, Cambridge, 2007. Press, Cambridge, 2007.

[20] V. H. Hoang, J. H. Quek and C. Schwab, *Analysis of a multilevel Markov chain Monte Carlo finite element method for Bayesian inversion of log-normal diffusions*, Inverse Problems **36** (2020): 035021, 46.

[21] V. H. Hoang and C. Schwab, *Convergence rate analysis of mcmc-fem for bayesian inversion of log-normal diffusion problems*, in: Research reports/Seminar for Applied Mathematics, vol. 2016, ETH Zurich, 2016.

[22] V. H. Hoang, C. Schwab and A. M. Stuart, *Complexity analysis of accelerated MCMC methods for Bayesian inversion*, Inverse Problems **29** (2013): 085010, 37.

[23] H. K. Hoel, K. J. H. Law and R. Tempone, *Multilevel ensemble Kalman filtering*, SIAM J. Numer. Anal. **54** (2016), 1813–1839.

[24] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Applied Mathematical Sciences, Springer-Verlag, New York, 2005.

[25] W. McLean, *Strongly Elliptic Systems and Boundary Integral Equations*, Cambridge University Press, Cambridge, 2000.

[26] V. Nistor and C. Schwab, *High-order Galerkin approximations for parametric second-order elliptic partial differential equations*, Math. Models Methods Appl. Sci. **23** (2013), 1729–1760.

[27] C. Schwab and C. J. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer. **20** (2011), 291–467.

[28] C. Schillings and C. Schwab, *Sparse, adaptive Smolyak quadratures for Bayesian inverse problems*, Inverse Problems **29** (2013): 065011, 28.

[29] R. Stenberg, *A family of mixed finite elements for the elasticity problem*, Numer. Math. **53** (1988), 513–538.

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

[31] B. Xia and V. H. Hoang, *High-dimensional finite element method for multiscale linear elasticity*, IMA J. Numer. Anal. **35** (2015), 1277–1314.

[32] J. Yang and V. H. Hoang, *Multilevel Markov Chain Monte Carlo for Bayesian inverse problem for Navier-Stokes equation*, Inverse Probl. Imaging **17** (2023), 106–135.

Manuscript received January 2 2022
revised September 17 2022

J. T. Yang
Nvidia AI Technology Center, 8 Temasek Blvd, Singapore 038988
Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371
E-mail address: yjuntao@nvidia.com

V. H. Hoang
Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371
E-mail address: vhhoang@ntu.edu.sg