Sparse MCMC gpc Finite Element Methods for Bayesian Inverse Problems

Report

Author(s):
Hoang, Viet Ha; Schwab, Christoph; Stuart, Andrew M.

Sparse MCMC gpc Finite Element Methods
for Bayesian Inverse Problems

V.H. Hoang, Ch. Schwab and A.M. Stuart

Sparse MCMC gpc Finite Element Methods for Bayesian Inverse Problems

Viet Ha Hoang[1]**, Christoph Schwab**[2] **and Andrew M. Stuart**[3]

[1] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, 637371

E-mail: `vhhoang@ntu.edu.sg`
[2] Seminar for Applied Mathematics, ETH, 8092 Zurich, Switzerland

E-mail: `christoph.schwab@sam.math.ethz.ch`
[3] Mathematics Institute, Warwick University, Coventry CV4 7AL, UK.

E-mail: `a.m.stuart@warwick.ac.uk`

Abstract. Several classes of MCMC methods for the numerical solution of Bayesian Inverse Problems for partial differential equations (PDEs) with unknown random field coefficients are considered. A general framework for their numerical analysis is presented. The complexity of MCMC sampling for the unknown fields from the posterior density, as well as the convergence of the discretization error of the PDE of interest in the forward response map, is analyzed. Particular attention is given to bounds on the overall work required by the MCMC algorithms for achieving a prescribed error level ε. We show that the computational complexity of straightforward combinations of MCMC sampling strategies with standard PDE solution methods is generally excessive. Two computational strategies for substantially reducing the computational complexity of MCMC methods for Bayesian inverse problems prising in PDEs are studied: a parametric, deterministic gpc-type (generalized polynomial chaos) representation of the forward solution map of the PDE with uncertain coefficients, which has been proposed and implemented in the engineering literature (e.g. [17, 15, 16]); and a new Multi-Level Monte Carlo sampling strategy of the Markov Chain (MLMCMC) with sampling from a multilevel discretization of the posterior and a multilevel discretization of the forward PDE. We compare the computational complexity of these gpc-MCMC and MLMCMC methods to that of the plain MCMC method, and provide sufficient conditions on the regularity of the unknown coefficient for both, the gpc-MCMC and MLMCMC method, to afford substantial complexity reductions over the plain MCMC approach.

1. Introduction

Many inverse problems arising in differential equations require determination of unknown parameters u from finite dimensional data δ which we assume to be related by

$$\delta = \mathcal{G}(u) + \vartheta . \tag{1}$$

Here u, which we assume to belong to a function space, is an unknown input to a differential equation and $\mathcal{G}$ is the "forward" mapping taking one instance of the input u into a usually finite and noisy set of observations. We model these observations mathematically as continuous linear functionals on the solution of the governing partial differential equation. In (1), the parameter ϑ represents noise arising when observing and we assume that this is a single realization of a centred Gaussian $N(0, \Sigma)$. A Bayesian formulation of the inverse problem then leads to the problem of probing the probability measure ρ^δ given by

$$\frac{d\rho^\delta}{d\rho}(u) \propto \exp(-\Phi(u; \delta)) , \tag{2}$$

where

$$\Phi(u; \delta) = \frac{1}{2} |\delta - \mathcal{G}(u)|_\Sigma^2 \tag{3}$$

and $| \cdot |_\Sigma = |\Sigma^{-\frac{1}{2}} \cdot |$ with $| \cdot |$ the Euclidean norm, and where ρ is a prescribed prior probability measure.

The purpose of this paper is to analyze the computational complexity of several Markov-Chain Monte-Carlo (MCMC) approaches to probing the distribution ρ^δ. These methods will incur two principal sources of error. First, the *sampling error* arising from estimating expectations conditional on given data δ by sample averages of M realizations of the unknown u, drawn from ρ^δ. The error in doing so will decay as $M^{-\frac{1}{2}}$ as the number M of draws of u tends to ∞. Second, the *discretization error* arising from approximation of the system response for each draw of u, i.e. the error of approximating $\mathcal{G}(u)$. For expository purposes, and to cover a wide range of discretization techniques, we assume the discretization error to decay as N_{dof}^{-a}, where N_{dof} is the total number of degrees of freedom‡ and $a > 0$; and we assume that the work per step scales as N_{dof}^b as $N_{\mathrm{dof}} \to \infty$ for some power $b > 0$ so that the total work necessary for M draws in the MCMC scales as $M N_{\mathrm{dof}}^b$. If (as we show in the present paper) the constant in the mean square MCMC error bound of order $\mathcal{O}(M^{-\frac{1}{2}})$ is independent of N_{dof}, then a straightforward calculation shows that the work to obtain error ε will grow asymptotically, as $\varepsilon \to 0$, as $\varepsilon^{-2-b/a}$. The ratio b/a is thus crucial to the overall computational complexity of the algorithm.

In this paper, we develop three ideas to speed up MCMC-based algorithms for Bayesian inversion in systems governed by partial differential equations. The first idea, which underlies the preceding expository calculation concerning complexity, is that MCMC methods can be constructed whose convergence rate is independent of the number of degrees of freedom N_{dof} used in the approximation of the forward map; the

‡ logarithmic corrections also occur, and will be detailed explicitly in later sections

key idea here is to use MCMC algorithms which are defined on function spaces, as overviewed in [23] and to use Galerkin discretizations of the forward map which employ *Riesz bases* in these function spaces. The second idea is that *sparse, deterministic parametric representations of the forward map* $u \mapsto \mathcal{G}(u)$, as in [3, 21, 4], can significantly reduce b/a and allow, therefore, for reducing computational complexity, as the sparse approximation of $\mathcal{G}$ can be precomputed prior to running the Markov chain, thereby decreasing the constant b [17, 15, 16]. The third idea is that the sparse representation of the forward map can be truncated adaptively at different discretization levels of the physical system of interest. Then, we propose a multilevel Monte Carlo acceleration of the MCMC method in the spirit of [9] and prove that this allows further improvement of the computational complexity.

The paper is organized as follows. In Section 2 we discuss the Bayesian formulation of inverse problems and describe and analyze an MCMC method which is defined on metric spaces. Section 3 is devoted to the specific elliptic inverse problem which we study for illustrative purposes. In section 4 we study what we term the *standard MCMC method* where the forward map $\mathcal{G}(\cdot)$ is computed at each step of the Markov chain. Section 5 is concerned with showing that the complexity of MCMC can be reduced if we precompute a parametric representation of the forward map $\mathcal{G}$ prior to running the Markov chain, and simply evaluate it at each step of the chain, leading to the *improved MCMC method*. Finally, in Section 6, we combine this idea to further improve the efficiency of the MCMC algorithm. To this end, we employ the hierarchic nature of the gpc-Finite Element Galerkin discretizations. We combine discretizations on multiple levels $\ell = 0, 1, ..., L$ and combine these judiciously with *a level dependent sample size* M_ℓ. We show that this leads to a *multilevel MCMC-gpcFE method* that can significantly improve the overall algorithmic complexity under certain assumptions.

We will use standard notation: $\mathcal{B}^k$ denotes the sigma algebra of Borel subsets of $\mathbb{R}^k$. For a probability space $(\Omega, \mathcal{A}, \rho)$ consisting of the set Ω of elementary events, a sigma algebra $\mathcal{A}$ and a probability measure ρ, and a separable Hilbert space H with norm $\| \cdot \|_H$ and for a summability exponent $0 < p \leq \infty$ we denote by $L^p(\Omega, \rho; H)$ the Bochner space of strongly measurable mappings from Ω to H which are p-summable.

Because of the various discretizations employed, and in particular the multi-level structure of some of these, it will be helpful to the reader to have a clear overview of the notation used to describe the range of probability spaces which arise, both through the Bayesian formulation of the inverse problem, and the Markov chains used to probe it. We now give such an overview. Throughout the paper, we denote by $\mathbb{E}^\mu$ the expectation with respect to a probability measure μ on the subspace U containing the unknown function u. In the following we will finite-dimensionalize both the subspace U, in which the unknown function u lies, and the space containing the response of the forward model. The parameter J denotes the truncation level of the coefficient expansion (15) used for the unknown function, and the parameter l denotes the spatial finite element discretization level introduced in section 4. The parameters N and $\mathcal{L}$ denote the cardinality of the set of the chosen active gpc coefficients and the set of

finite element discretization levels for these coefficients in section 5. The measure μ will variously be chosen as the prior ρ, the posterior ρ^δ, and various approximations of the posterior such as $\rho^{J,l,\delta}$.

We denote by $\mathcal{P}_{u^{(0)}}$, $\mathcal{P}_{u^{(0)}}^{J,l}$ and $\mathcal{P}_{u^{(0)}}^{N,\mathcal{L}}$ probability measures on the probability space generated by the MCMC processes detailed in the following, when conditioned on the initial point $u^{(0)}$. The acceptance probability for the Metropolis-Hastings Markov chain is defined as α in (6), $\alpha^{J,l}$ in (31), and $\alpha^{N,\mathcal{L}}$ in (44) for the problems on the full infinite dimensional space and its truncations. We then denote by $\mathcal{E}_{u^{(0)}}$, $\mathcal{E}_{u^{(0)}}^{J,l}$ and $\mathcal{E}_{u^{(0)}}^{N,\mathcal{L}}$ expectation with respect to $\mathcal{P}_{u^{(0)}}$, $\mathcal{P}_{u^{(0)}}^{J,l}$ and $\mathcal{P}_{u^{(0)}}^{N,\mathcal{L}}$ respectively.

If the initial point $u^{(0)}$ of these Markov chains is distributed with respect to an initial probability measure μ on U, then we denote the probability measure on the space that describes these Markov chains by $\mathcal{P}^\mu$, $\mathcal{P}^{\mu,J,l}$ and $\mathcal{P}^{\mu,N,\mathcal{L}}$, and the corresponding expectation accordingly by $\mathcal{E}^\mu$, $\mathcal{E}^{\mu,J,l}$ and $\mathcal{E}^{\mu,N,\mathcal{L}}$.

Finally, in Section 6, we will work with the probability measure $\mathfrak{P}_L$, on the space that generates a sequence of uncorrelated Markov chains created by the multilevel-MCMC procedure, and with $\mathfrak{E}_L$, the expectation with respect to this probability measure. The definition of these measures will be given at the beginning of Section 6.

2. Bayesian inverse problems in measure spaces

On a measurable space (U, Θ) where Θ is a σ-algebra consider a measurable map $\mathcal{G} : U \to (\mathbb{R}^k, \mathcal{B}^k)$. The data δ is assumed to be an observation of $\mathcal{G}$ subject to an unbiased observation noise ϑ:

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

We assume that ϑ is a centred Gaussian with law $N(0, \Sigma)$. Let ρ be a prior probability measure on (U, Θ). Our purpose is to determine the conditional probability $\mathbb{P}(u|\delta)$ on (U, Θ). The following result holds.

Proposition 1 *Assume that $\mathcal{G} : U \to \mathbb{R}^k$ is measurable. The posterior measure $\rho^\delta(du) = \mathbb{P}(du|\delta)$ given data δ is absolutely continuous with respect to the prior measure $\rho(du)$ and has the Radon-Nikodym derivative (2) with Φ given by (3).*

This result is established in Cotter et al.[8] and Stuart [23]. Though the setting in [8] and [23] is in a Banach space X, the proofs of Theorem 2.1 in [8] and Theorem 6.31 of [23] hold for any measurable spaces as long as the mapping $\mathcal{G}$ is measurable.

To study the well-posedness of the posterior measures, that is continuity with respect to changes in the observed data, we use the Hellinger distance, as in Cotter et al. [8]; see below for the definition. In that paper it is proved that when U is a Banach space, if the prior measure ρ is Gaussian, and under the conditions that Φ grows polynomially with respect to u, and is locally Lipschitz with respect to u fixing y and with respect to y fixing u, in the second case with Lipschitz constant which also

grows polynomially in u, then the posterior measure given the data δ, i.e. ρ^δ, is locally Lipschitz in the Hellinger distance d_{Hell}:

$$d_{\text{Hell}}(\rho^\delta, \rho^{\delta'}) \leq c|\delta - \delta'| \;,$$

where (recall) $|\cdot|$ denotes the Euclidean distance in $\mathbb{R}^k$. The Fernique theorem plays an essential role in the proofs, exploiting the fact that polynomially growing functions are integrable with respect to Gaussians. In this section, we extend this result to measurable spaces under more general conditions than those in Assumption 2.4 of Cotter et al. [8]; in particular we do not assume a Gaussian prior. We make the following assumption concerning the local boundedness, and locally Lipschitz dependence of Φ on δ.

Assumption 2 *Let ρ be a probability measure on the measure space (U, Θ). The function $\Phi : U \times \mathbb{R}^k \to \mathbb{R}$ satisfies:*

(i) for each $r > 0$, there is a constant $M(r) > 0$ and a set $U(r) \subset U$ of positive ρ measure such that for all $u \in U(r)$ and for all δ such that $|\delta| \leq r$

$$0 \leq \Phi(u; \delta) < M(r);$$

(ii) there is a mapping $G : \mathbb{R} \times U \mapsto \mathbb{R}$ such that for each $r > 0$, $G(r, \cdot) \in L^2(U, \rho)$; and for every $|\delta|, |\delta'| \leq r$ holds

$$|\Phi(u; \delta) - \Phi(u; \delta')| \leq G(r, u)|\delta - \delta'| \;.$$

Under Assumption 2, the definition (2) of the posterior measure ρ^δ is meaningful as we now demonstrate.

Proposition 3 *Under Assumption 2, the measure ρ^δ depends locally Lipschitz continuously on the data δ with respect to the Hellinger metric: for each positive constant r there is a positive constant $C(r)$ such that if $|\delta|, |\delta'| \leq r$, then*

$$d_{\text{Hell}}(\rho^\delta, \rho^{\delta'}) \leq C(r)|\delta - \delta'| \;.$$

Proof Throughout this proof $K(r)$ denotes a constant depending on r, possibly changing from instance to instance. The normalization constant in (2) is

$$Z(\delta) = \int_U \exp(-\Phi(u; \delta))d\rho(u) \;. \tag{4}$$

We first show that for each $r > 0$, there is a positive constant $K(r)$ such that $Z(\delta) \geq K(r)$ when $|\delta| \leq r$. From (4) and Assumption 2(i) it follows that, that when $|\delta| \leq r$,

$$Z(\delta) \geq \rho(U(r)) \exp(-M(r)) > 0 \;. \tag{5}$$

Using the inequality $|\exp(-x) - \exp(-y)| \leq |x - y|$ which holds for $x, y \geq 0$ we find

$$|Z(\delta) - Z(\delta')| \leq \int_U |\Phi(u; \delta) - \Phi(u; \delta')|d\rho(u) \;.$$

From Assumption 2(ii),

$$|\Phi(u; \delta) - \Phi(u; \delta')| \leq G(r, u)|\delta - \delta'| \;.$$

As $G(r, u)$ is ρ-integrable, there is $K(r)$ such that

$$|Z(\delta) - Z(\delta')| \leq K(r)|\delta - \delta'| .$$

The Hellinger distance satisfies

$$2d_{\text{Hell}}(\rho^\delta, \rho^{\delta'})^2 = \int_U \left(Z(\delta)^{-1/2} \exp(-\frac{1}{2}\Phi(u; \delta)) - Z(\delta')^{-1/2} \exp-\frac{1}{2}\Phi(u; \delta') \right)^2 d\rho(u)$$
$$\leq I_1 + I_2,$$

where

$$I_1 = \frac{2}{Z(\delta)} \int_U \left(\exp(-\frac{1}{2}\Phi(u; \delta)) - \exp(-\frac{1}{2}\Phi(u; \delta')) \right)^2 d\rho(u),$$

and

$$I_2 = 2|Z(\delta)^{-1/2} - Z(\delta')^{-1/2}|^2 \int_U \exp(-\Phi(u; \delta'))d\rho(u).$$

Using again $|\exp(-x) - \exp(-y)| \leq |x - y|$, we have, for constant $K(r) > 0$,

$$I_1 \leq K(r) \int_X |\Phi(u; \delta) - \Phi(u; \delta')|^2 d\rho(u)$$
$$\leq \int_X (G(r, u))^2 d\rho(u)|\delta - \delta'|^2 \leq K(r)|\delta - \delta'|^2 .$$

Furthermore,

$$|Z(\delta)^{-1/2} - Z(\delta')^{-1/2}|^2 \leq K(r)|Z(\delta) - Z(\delta')|^2 \leq K(r)|\delta - \delta'|^2 .$$

The conclusion follows. □

We now introduce a Metropolis-Hastings MCMC method designed to be reversible and ergodic with respect to the posterior measure ρ^δ: to this end, given data δ, we define for any $u, v \in U$

$$\alpha(u, v) = 1 \wedge \exp(\Phi(u, \delta) - \Phi(v, \delta)) . \tag{6}$$

The Markov chain $\{u^{(k)}\}_{k=1}^\infty \subset U$ is then constructed as follows: given the current state $u^{(k)}$, we draw a proposal $v^{(k)}$ independently of $u^{(k)}$ from the prior measure ρ appearing in (2). Let $\{w^{(k)}\}_{k \geq 1}$ denote an i.i.d sequence with $w^{(1)} \sim \mathcal{U}[0, 1]$ and with $w^{(k)}$ independent of both $u^{(k)}$ and $v^{(k)}$. We then determine the next state $u^{(k+1)}$ via the formula

$$u^{(k+1)} = \mathbf{1}(\alpha(u^{(k)}, v^{(k)}) \geq w^{(k)})v^{(k)} + \left(1 - \mathbf{1}(\alpha(u^{(k)}, v^{(k)}) \geq w^{(k)}) \right)u^{(k)} . \tag{7}$$

Thus we choose to move from $u^{(k)}$ to $v^{(k)}$ with probability $\alpha(u^{(k)}, v^{(k)})$, and to remain at $u^{(k)}$ with probability $1 - \alpha(u^{(k)}, v^{(k)})$. We claim that (7) defines a Markov chain $\{u^{(k)}\}_{k=0}^\infty$ which is reversible with respect to ρ^δ. To see this let $\nu(du, dv)$ denote the product measure $\rho^\delta(du) \otimes \rho(dv)$ and $\nu^\dagger(du, dv) = \nu(dv, du)$. Note that ν describes the probability distribution of the pair $(u^{(k)}, v^{(k)})$ on $U \times U$ when $u^{(k)}$ is drawn from ρ^δ, and $\nu^\dagger$ designates the same measure with the roles of u and v reversed. These two measures are equivalent (as measures) if ρ^δ and ρ are equivalent, and then

$$\frac{d\nu^\dagger}{d\nu}(u, v) = \exp(\Phi(u; \delta) - \Phi(v; \delta)) , \quad (u, v) \in U \times U . \tag{8}$$

From Proposition 1 and Theorem 2 in [24] we deduce that (7) is a Metropolis-Hastings Markov chain which is ρ^δ reversible, since $\alpha(u, v)$ given by (6) is equal to $\min\{1, \frac{d\nu^\dagger}{d\nu}(u, v)\}$. Since $v^{(k)}$ is chosen independently of the current state $u^{(k)}$ the Markov chain is, in fact, an *independence sampler*. We now give sufficient conditions which render the Markov chain ergodic.

Theorem 4 *Let Assumption 2 hold with $U(r) = U$. Then ρ^δ is equivalent to ρ so that the Markov chain (7) is well-defined and reversible with respect to ρ^δ. If $p(u, A)$ denotes the transition kernel for the Markov chain, and $p^n(u, A)$ its nth iterate, then for all $n \in \mathbb{N}$,*

$$\|p^n(u, A) - \rho^\delta\|_{\mathrm{TV}} \leq 2(1 - \exp(-M(r)))^n .$$

For every bounded, continuous function $g : U \to \mathbb{R}$, there holds, $\mathcal{P}_{u^{(0)}}$ almost surely,

$$\frac{1}{M} \sum_{k=1}^M g(u^{(k)}) = \mathbb{E}^{\rho^\delta}[g(u)] + c\xi_M M^{-\frac{1}{2}} \tag{9}$$

where ξ_M is a sequence of random variables which converges weakly as $M \to \infty$ to $\xi \sim N(0, 1)$ and where c is a positive constant which depends only on $M(r)$ and on $\sup_{u \in U} |g(u)|$. Furthermore, we have the mean square error bound

$$\left(\mathcal{E}^\rho\left[\left|\mathbb{E}^{\rho^\delta}[g(u)] - \frac{1}{M} \sum_{k=1}^M g(u^{(k)})\right|^2\right]\right)^{1/2} \leq CM^{-1/2} .$$

Proof Equivalence of ρ^δ and ρ follows since the negative of the log-density is bounded from above and below, uniformly on U:

$$0 \leq \Phi(u) \leq M(r) \quad \forall u \in U . \tag{10}$$

Using these bounds and (6) it also follows that the proposed random draw from ρ has probability greater than $\exp(-M(r))$ of being accepted. Thus

$$p(u, A) \geq \exp(-M(r))\rho(A) \quad \forall u \in U.$$

The first result follows from [18], Theorem 16.2.4 with $X = U$. The second result follows from [18], Theorem 17.0.1. To see that c in (9) is bounded by a constant that depends only on $M(r)$ and $\sup_{u \in U} |g(u)|$, we note that it is given by

$$c^2 = \mathcal{E}^{\rho^\delta} |\bar{g}(u^{(0)})|^2 + 2 \sum_{n=1}^\infty \mathcal{E}^{\rho^\delta}[\bar{g}(u^{(0)})\bar{g}(u^{(n)})] \tag{11}$$

where the function $\bar{g}$ is defined as $\bar{g} = g - \mathbb{E}^{\rho^\delta}(g)$. Now

$$2 \sum_{n=0}^\infty \mathcal{E}^{\rho^\delta}[\bar{g}(u^{(0)})\bar{g}(u^{(n)})] \leq 2 \sup_u |\bar{g}(u)| \mathbb{E}^{\rho^\delta} \sum_{n=0}^\infty |\mathcal{E}_{u^{(0)}}[\bar{g}(u^{(n)})]|$$

$$\leq 2 \sup_u |\bar{g}(u)| \mathbb{E}^{\rho^\delta} \sum_{n=0}^\infty |\mathcal{E}_{u^{(0)}}[g(u^{(n)})] - \mathbb{E}^{\rho^\delta}[g]|$$

$$\leq 4 \sup_u |\bar{g}(u)|^2 \sum_{n=0}^\infty (1 - \exp(-M(r)))^n .$$

For the mean square approximation, using the stationarity of the Markov chain conditioned to start in $U \ni u^{(0)} \sim \rho^\delta$, we have

$$\frac{1}{M}\mathcal{E}^{\rho^\delta}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] = \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=1}^{M}\sum_{j=k+1}^{M}\mathcal{E}^{\rho^\delta}[\bar{g}(u^{(k)})\bar{g}(u^j)]$$

$$= \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathcal{E}^{\rho^\delta}[\bar{g}(u^{(0)})\bar{g}(u^{(j)})]$$

$$= \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})\mathcal{E}_{u^{(0)}}[\bar{g}(u^{(j)})]]$$

$$\leq \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2]$$
$$+ 2\frac{1}{M}\sum_{k=0}^{M-1}\sup_u|\bar{g}(u)|\sum_{j=1}^{M-k}\mathbb{E}^{\rho^\delta}[|\mathcal{E}_{u^{(0)}}[g(u^{(j)})] - \mathbb{E}^{\rho^\delta}[g]|]$$

$$\leq \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2] + 4\frac{1}{M}\sum_{k=0}^{M-1}\sup_u|\bar{g}(u)|^2\sum_{j=1}^{M-k}(1 - \exp(-M(r)))^j$$

$$\leq \mathbb{E}^{\rho^\delta}[\bar{g}(u^{(0)})^2] + 4\sup_u|\bar{g}(u)|^2\sum_{j=1}^{\infty}(1 - \exp(-M(r)))^j,$$

which is clearly bounded uniformly with respect to M. Thus we have shown that there exists $C > 0$ such that

$$\mathcal{E}^{\rho^\delta}\Big[\Big|\frac{1}{M}\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] \leq \frac{C}{M} \ .$$

It remains to show that the expectation with respect to the unknown posterior ρ^δ can be replaced by an expectation with respect to the prior measure ρ.

To this end we note that

$$\mathcal{E}^{\rho}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] = \int_U\mathcal{E}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]d\rho(u^{(0)})$$

$$= \int_U\mathcal{E}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]\frac{d\rho}{d\rho^\delta}(u^{(0)})d\rho^\delta(u^{(0)})$$

$$\leq \mathcal{E}^{\rho^\delta}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]Z(\delta)\sup_{u\in U}\Big[\exp(\Phi(u;\delta))\Big] \ .$$

As $Z(\delta) \leq 1$ and $\Phi(\cdot;\delta)$ is assumed to be bounded uniformly, we deduce that

$$\mathcal{E}^{\rho}\Big[\Big|\frac{1}{M}\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] \leq \frac{C}{M},$$

for a constant C independent of M. The conclusion then follows. $\qquad\square$

Remark 5 *A key observation in the previous theorem is that all constants depend only on $M(r)$ and on the supremum of g. Hence, if we can show for finite element and*

Karhúnen-Loève approximations of our forward map $\mathcal{G}$ in the physical domain D that we obtain uniform upper bounds on our approximation of Φ, then the conclusions of the preceding theorem will hold with constants that are uniformly bounded with respect to the approximation parameters.

3. Model Elliptic Inverse Problem

In the remainder of this paper, we develop the previously outlined abstract ideas for a specific class of model elliptic inverse problems where the unknown parameter is the diffusion coefficient and where each realization of the observed data comprises finitely many continuous, linear functionals of the forward solution. However we hasten to add that a similar analysis is possible for other PDE inverse problems. For example, for linear parabolic or second order hyperbolic PDEs, the analysis of the parametric forward problems, central to the approach developed herein, is available in [13, 14].

We start by discussing the forward problem of elliptic PDEs with random coefficients, enabling the construction of a prior measure on an infinite dimensional space of unknown coefficients. We then show how this prior may be combined with properties of the forward solution map to obtain a well-defined inverse problem.

3.1. A Class of Elliptic Problems With Random Coefficients

Let D be a bounded Lipschitz domain in $\mathbb{R}^d$. For $f \in L^2(D)$, we consider the elliptic problem

$$-\nabla \cdot (K(x,\omega)\nabla P(x,\omega)) = f(x) \text{ in } D, \quad P = 0 \text{ on } \partial D . \qquad (12)$$

Throughout we assume that the domain D is a convex polyhedron with plane sides. The coefficient $K(x,\omega)$ is a random field from the probability space $(\Omega, \Xi, \mathbb{P})$ to $L^\infty(D)$. We assume that the random coefficient $K(x,\omega)$ can be represented by a sequence of pairwise uncorrelated independent random variables $u_j : \Omega \to [-1, 1]$ in the series expansion

$$K(x,\omega) = \bar{K}(x) + \sum_{j \geq 1} u_j(\omega)\psi_j(x) . \qquad (13)$$

Here, the sum is either finite or infinite. To render (13) meaningful, we impose the following assumptions on $\bar{K}$ and ψ_j.

Assumption 6 *The functions $\bar{K}$ and ψ_j in (13) are in $L^\infty(D)$ and there exists a positive constant κ such that*

$$\sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)} \leq \frac{\kappa}{1 + \kappa} \bar{K}_{\min},$$

where $\bar{K}_{\min} = \text{essinf}_x \bar{K}(x) > 0$.

It follows from this assumption that there exist finite positive constants $K_{\min}$ and $K_{\max}$ such that for all $\omega \in \Omega$

$$K_{\min} \leq K(x,\omega) \leq K_{\max}, \quad \forall \, x \in D. \qquad (14)$$

In particular we may choose $K_{\min} = \bar{K}_{\min}/(1+\kappa)$ and $K_{\max} = \mathrm{esssup}_x \bar{K}(x) + \kappa \bar{K}_{\min}/(1+\kappa)$.

We denote by $U = [-1,1]^{\mathbb{N}}$ the set of all sequences $u = (u_j)_{j\geq 1}$ of coordinates u_j taking values in $[-1,1]$ and note that this is the unit ball in $\ell^\infty(\mathbb{N})$. We equip the parameter domain U with the product sigma algebra $\Theta = \bigotimes_{j=1}^\infty \mathcal{B}([-1,1])$. On the measurable space (U, Θ) thus obtained, we define the countable product probability measure

$$\rho = \bigotimes_{j\geq 1} \frac{du_j}{2},$$

where du_j is the Lebesgue measure on $[-1,1]$. As u_j are uniformly distributed on $[-1,1]$, the measure ρ is the law of the random vector $u = (u_1, u_2, \ldots)$ in U. As random variables $u_j(\omega)$ in the sequence u were assumed independent, the probability measure on realizations of random vectors $u \in U$ is a product measure: for $S = \prod_{j\geq 1} S_j$,

$$\rho(S) = \prod_{j\geq 1} \mathbb{P}(\{\omega : u_j \in S_j\}) .$$

For each $u \in U$, we define the parametric, deterministic coefficient function

$$K(x, u) = \bar{K}(x) + \sum_{j\geq 1} u_j \psi_j(x) . \tag{15}$$

Due to Assumption 6, for any $u \in U$ the series (15) converges in $L^\infty(D)$. Therefore, for each $u \in U$, we consider the model *parametric, deterministic diffusion problem in D*

$$- \nabla \cdot (K(x, u) \nabla P(x, u)) = f(x) \text{ in } D, \quad P = 0 \text{ on } \partial D . \tag{16}$$

We let $V = H_0^1(D)$, whilst V^* denotes its' dual space. We equip V with the norm $\|P\|_V = \|\nabla P\|_{L^2(D)}$. By (13), $K(x, u)$ is bounded below uniformly with respect to $(x, u) \in D \times U$ and we infer for every $u \in U$

$$
\begin{aligned}
K_{min} \|P(\cdot, u)\|_V^2 &= K_{min}(\nabla P(\cdot, u), \nabla P(\cdot, u)) \leq (K(\cdot, u)\nabla P(\cdot, u), \nabla P(\cdot, u)) \\
&= (f, P(\cdot, u)) \leq \frac{\|f\|_{V^*}}{K_{min}} \|P(\cdot, u)\|_V .
\end{aligned}
$$

It follows that

$$\sup_{u \in U} \|P(\cdot, u)\|_V \leq \frac{\|f\|_{V^*}}{K_{min}} . \tag{17}$$

The solution $P(\cdot, u)$ of (16) is the law of the solution $P(\cdot, \omega)$ of the equation (12) as the following result shows.

Proposition 7 *Under Assumption 6, the solution $P : U \mapsto V = H_0^1(D)$ is Lipschitz when viewed as a mapping from the unit ball in $\ell^\infty(\mathbb{N})$ to V. It is in particular measurable, as a mapping from the measurable space (U, Θ) to $(V, \mathcal{B}(V))$.*

Proof From (16), we have for every $\phi \in V$

$$\int_D K(x, u)(\nabla P(x, u) - \nabla P(x, u')) \cdot \nabla \phi(x) dx$$

$$= \int_D (K(x, u') - K(x, u)) \nabla P(x, u') \cdot \nabla \phi(x) dx . \tag{18}$$

Again using (14), i.e. that $K(x, u)$ is bounded below uniformly with respect to $(x, u) \in D \times U$, it follows that there exists $C > 0$ such that for all $u \in U$

$$\|P(\cdot, u) - P(\cdot, u')\|_V \leq C\|P(\cdot, u)\|_V \|K(\cdot, u') - K(\cdot, u)\|_{L^\infty(D)} . \tag{19}$$

Due to (17), it follows from (19) that there exists a constant $C > 0$ such that

$$\forall u \in U : \quad \|P(\cdot, u) - P(\cdot, u')\|_V \leq C\|K(\cdot, u') - K(\cdot, u)\|_{L^\infty(D)} . \tag{20}$$

From (13) and Assumption 6 it follows with $C > 0$ as in (20) that

$$\|P(\cdot, u) - P(\cdot, u')\|_V \leq C \sum_{\geq 1} |u_j - u'_j| \|\psi_j\|_{L^\infty(D)}$$

$$\leq C\|u - u'\|_{\ell^\infty(\mathbb{N})} \sum_{j \geq 1} \|\psi_j\|_{L^\infty(D)}$$

$$\leq C \frac{\kappa}{1 + \kappa} \bar{K}_{\min} \|u - u'\|_{\ell^\infty(\mathbb{N})} .$$

This establishes the desired Lipschitz continuity, which implies the asserted measurability. $\qquad \square$

3.2. Bayesian Elliptic Inverse Problem

We now define the Bayesian inverse problem. For $\mathcal{O}_i \in V^*$, $i = 1, \ldots, k$, which denote k continuous, linear "observation" functionals on V, we define a map $\mathcal{G} : U \to \mathbb{R}^k$ as

$$U \ni u \mapsto \mathcal{G}(u) := (\mathcal{O}_1(P(\cdot, u)), \mathcal{O}_2(P(\cdot, u)), \ldots, \mathcal{O}_k(P(\cdot, u))) \in \mathbb{R}^k .$$

By ϑ we denote an unbiased noise which follows a Gaussian distribution $N(0, \Sigma)$ in $\mathbb{R}^k$. We consider the observed data δ for $\mathcal{G}(u)$ subject to the noise ϑ, i.e.

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

and define Φ as in (3). We take as prior on u the measure ρ defined in the preceding subsection. The posterior measure on u given δ can be explicitly written.

Proposition 8 *The conditional probability measure $\rho^\delta(du) = \mathbb{P}(du|\delta)$ on U satisfies*

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

Furthermore, for $\delta, \delta' \in \mathbb{R}^k$ such that $|\delta|, |\delta'| \leq r$ there exists $C = C(r) > 0$ such that

$$d_{\text{Hell}}(\rho^\delta, \rho^{\delta'}) \leq C(r)|\delta - \delta'| .$$

Proof We have

$$\forall u, u' \in U : \quad |\mathcal{G}(u) - \mathcal{G}(u')| \leq c \max_i \{\|\mathcal{O}_i\|_{V^*}\} \|P(\cdot, u) - P(\cdot, u')\|_V .$$

From (20) there exists a constant $c > 0$ such that

$$\forall u, u' \in U : \quad |\mathcal{G}(u) - \mathcal{G}(u')| \leq c\|K(\cdot, u) - K(\cdot, u')\|_{L^\infty(D)} .$$

Proceeding as in the proof of Proposition 7, we deduce that $\mathcal{G}$ as map from $U \subset \ell^\infty(\mathbb{N})$ to $\mathbb{R}^k$ is Lipschitz and, hence, ρ-measurable. We then apply Proposition 1 to deduce the existence of ρ^δ and the formula for its Radon-Nikodym derivative with respect to ρ.

For the well-posedness of ρ^δ, we verify Assumption 2. For the function $\mathcal{G}(u)$ we have

$$|\mathcal{G}(u)| \leq \max_i\{\|\mathcal{O}_i\|_{V^*}\}\|P(\cdot, u)\|_V .$$

From (17) $\sup\{|\mathcal{G}(u)| : u \in U\} < \infty$. We note that for given data δ, there holds

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

and hence, since $|\mathcal{G}(u)|$ is uniformly bounded in U, the set $U(r)$ in Assumption 2(i) can be chosen as U for all r. We have, for every $u \in U$,

$$|\Phi(u; \delta) - \Phi(u; \delta')| \leq \frac{1}{2}|\langle \Sigma^{-1/2}(\delta + \delta' - 2\mathcal{G}(u)), \Sigma^{-1/2}(\delta - \delta')\rangle|$$

$$\leq \frac{1}{2}\|\Sigma^{-1/2}\|^2_{L(\mathbb{R}^k, \mathbb{R}^k)}(|\delta| + |\delta'| + 2|\mathcal{G}(u)|)|\delta - \delta'| .$$

Choosing the function $G(r, u)$ in Assumption 2(ii) as

$$G(r, u) = \frac{1}{2}\|\Sigma^{-1/2}\|^2_{L(\mathbb{R}^k, \mathbb{R}^k)}(2r + c),$$

for a sufficiently large constant $c > 0$, we have shown that Assumption 2(ii) holds. From Proposition 3, we get the conclusion. $\square$

Remark 9 *In the the preceding proof we have shown that $|\mathcal{G}(u)|$ is uniformly bounded for u in U. As a consequence there exists $M(r) > 0$ which is a uniform bound on $\Phi(u; \delta)$ for all $|\delta| \leq r$ and for all $u \in U$. This bound is also uniform with respect to truncation of the infinite series (13) for K, since this corresponds to a particular choice of some of the coefficients of $u \in U$, and with respect to finite element approximation of the solution of (16), since the uniform upper bound on $|\mathcal{G}(u)|$ will hold in finite element subspaces.*

4. Standard MCMC

We study computational complexity of the MCMC method defined by (7) to sample the conditional probability measure ρ^δ determined in the previous section. The complexity results will be obtained here for the model scalar, elliptic inverse problem (12). We mention again that analogous results (with identical proofs) hold for inverse problems for general second order elliptic problems. While the complexity analysis of the MCMC algorithm is of independent interest, the results in the present section will be the foundation for several accelerations of the basic MCMC algorithm presented in Sections 5 and 6 ahead. In order to obtain a constructive version of the MCMC algorithm, we will approximate solutions of the forward problem (12) by applying the Finite Element Method in the physical domain D to its parametric version (16) and by truncation of the polynomial expansion of the diffusion coefficient K given by (13). To obtain error bounds for the Finite Element discretization of the parametric forward problem (16) in the domain D, we require differentiability of the coefficient functions $\psi_j(x)$ with respect to the spatial variable x.

Assumption 10 *The functions $\bar{K}$ and ψ_j in (13) are in $W^{1,\infty}(D)$ and*

$$\sum_{j \geq 1} \|\psi_j\|_{W^{1,\infty}(D)} < \infty .$$

Furthermore, we assume that there exist positive constants C and q such that for all $J \in \mathbb{N}$, the sequence $\{\psi_j\}$ in (13) satisfies

$$\sum_{j > J} \|\psi_j\|_{L^\infty(D)} < C J^{-q} .$$

Assumption 10 shall be imposed throughout what follows. From Assumption 10 and from (15), we deduce that $K(\cdot, u) \in W^{1,\infty}(D)$ for all $u \in U$.

4.1. FE Approximation of the Forward Problem

We describe an approximation of the forward problem based on finite element representation of the solution P of (16), together with truncation of the series (15). We start by discussing the finite element approximation. Recalling that the domain D is a bounded Lipschitz polyhedron with plane sides, we denote by $\{\mathcal{T}^l\}_{l=1}^\infty$ the nested sequence of simplices which is defined inductively as follows: first we divide D into a regular family $\mathcal{T}^0$ of simplices; the set of regular simplices $\mathcal{T}^l$ is determined by dividing each simplex in $\mathcal{T}^{l-1}$ into 2^d subsimplices. Based on these triangulations, we define a nested multilevel family of spaces of continuous, piecewise linear functions on $\mathcal{T}^l$ as

$$V^l = \{u \in V : u|_T \in \mathcal{P}_1(T) \ \forall T \in \mathcal{T}^l\},$$

where $\mathcal{P}_1(T)$ denotes the set of linear polynomials in the simplex $T \in \mathcal{T}^l$. Approximating the solution of the parametric, deterministic problem (16) from the finite element spaces V^l introduces a *discretization error* which is well-known to be bounded by the *approximation property* of the V^l: there exists a positive constant $C > 0$ which is independent of l such that for all $P \in W = (H^2 \cap H_0^1)(D)$ and for every $0 < h_l \leq 1$ holds

$$\inf_{Q \in V^l} \|P - Q\|_V \leq C h_l \|P\|_{H^2(D)}, \tag{21}$$

where $h_l = O(2^{-l}) = \max\{\text{diam}(T) : T \in \mathcal{T}^l\}$ is the mesh width of triangulation $\mathcal{T}^l$ and where the constant $C > 0$ depends only on the shape regularity of $\mathcal{T}^l$.

To bound the cost of the Finite Element discretization, we assume in the following that *the union of all Finite Element basis functions w_j^l of the spaces $V^l = \text{span}\{w_j^l : j = 1, ..., N_l\}$, $l = 0, 1, 2, ...$, constitutes a Riesz basis in V*. We remark that such bases are available in two and three dimensional polyhedral domains (see, e.g., [20]) (the following assumption of availability of V-stable Riesz bases is made for convenience, and may also be replaced by the availability of a linear complexity, optimal preconditioning, such as a BPX preconditioner).

Assumption 11 *(Riesz Basis Property in V) For each $l \in \mathbb{N}_0$ there exists a set of indices $I^l \subset \mathbb{N}^d$ of cardinality $N_l = O(2^{ld})$ and a family of basis functions $w_k^l \in H_0^1(D)$*

indexed by a multi-index $k \in I^l$ *such that* $V^l = \text{span}\{w_k^l : k \in I^l\}$, *and there exist constants* c_1 *and* c_2 *which are independent of the discretization level* l *such that if* $w \in V^l$ *is written as* $w = \sum_{k \in I^l} c_k^l w_k^l \in V^l$, *then*

$$c_1 \sum_{k \in I^l} |c_k^l|^2 \le \|w\|_V^2 \le c_2 \sum_{k \in I^l} |c_k^l|^2 .$$

Multiscale Finite Element bases entail, in general, larger supports than the standard, single scale basis functions which are commonly used in the Finite Element Method, which implies that the stiffness matrices in these bases have additional nonzero entries, as compared to $O(\dim V^l) = O(2^{dl})$ many nonzero entries of the stiffness matrices that result when one-scale bases, such as the hat functions, are used.

To bound the number of nonzero entries, we shall work under

Assumption 12 *(Support overlap) For all* $l \in \mathbb{N}_0$ *and for every* $k \in I^l$, *for every* $l' \in \mathbb{N}_0$ *the support intersection* $\text{supp}(w_k^l) \cap \text{supp}(w_{k'}^{l'})$ *has positive measure for at most* $O(\max(1, 2^{l'-l}))$ *values of* k'.

We now discuss the effect of dimensional truncation, ie. of truncating the infinite series for the coefficient K of problem (16) after J terms, as

$$K^J(x, u) = \bar{K}(x) + \sum_{j=1}^{J} u_j \psi_j(x) . \tag{22}$$

To this end, we consider the approximating diffusion problem

$$-\nabla \cdot (K^J(x, u) \nabla P^J(x, u)) = f(x), \quad P^J = 0 \text{ on } \partial D . \tag{23}$$

From (19), there exists a constant $C > 0$ such that $J \in \mathbb{N}$ and all $u \in U$

$$\begin{aligned} \sup_{u \in U} \|P(\cdot, u) - P^J(\cdot, u)\|_V &\le C\|P(\cdot, u)\|_V \|K(\cdot, u) - K^J(\cdot, u)\|_{L^\infty(D)} \\ &\le C\|P(\cdot, u)\|_V J^{-q} \le \frac{C}{K_{min}} J^{-q} \|f\|_{V^*} . \end{aligned} \tag{24}$$

To bound the error incurred by Finite Element discretization, we require regularity of $P(\cdot, u)$. Assumption 10 implies the following regularity results.

Proposition 13 *If* D *is convex and* $f \in L^2(D)$, *and if Assumption 10 holds, then, for every* $u \in U$, *the solution* $P^J(\cdot, u)$ *of (23) belongs to the space* $W := H^2(D) \cap H_0^1(D)$ *and there exists a positive constant* $C > 0$ *such that*

$$\sup_{J \in \mathbb{N}} \sup_{u \in U} \|P^J(\cdot, u)\|_W \le C\|f\|_{L^2(D)} .$$

Proof By (14), $K^J(x, u) \ge K_{min} > 0$ and we may rewrite the PDE in (23) as

$$-\Delta P^J(x, u) = \frac{1}{K^J(x, u)} (f(x) + \nabla K^J(x, u) \cdot \nabla P^J(x, u)).$$

By our assumptions, the right hand side is uniformly bounded with respect to $u \in U$ in the space $L^2(D)$. As the domain D is convex, we deduce that P^J are uniformly bounded with respect to J and $u \in U$ in the space W: it holds

$$\begin{aligned} \sup_{u \in U} \|\Delta P^J(\cdot, u)\|_{L^2(D)} &\le \frac{1}{K_{min}} \sup_{u \in U} \sup_{J \ge 1} \left[\|f\|_{L^2(D)} + \|K^J(\cdot, u)\|_{W^{1,\infty}(D)} \|P^J(\cdot, u)\|_V \right] \\ &\le C < \infty . \end{aligned}$$

The desired, uniform (w.r. to J and l) bound in the W norm then follows from the $W^{1,\infty}(D)$-summability of the ψ_j implied by Assumption 10, the L^2 bound on Δu and (17). $\qquad\square$

Finally, we consider the finite element approximation of the truncated problem (23): given $J, l \in \mathbb{N}$, find $P^{J,l}(\cdot, u) \in V^l$ such that

$$\int_D K^J(x, u)\nabla P^{J,l}(x, u) \cdot \nabla\phi(x)dx = \int_D f(x)\phi(x)dx, \quad \forall\, \phi \in V^l . \tag{25}$$

By the uniform positivity (14) of $K^J(x, u)$, the following error estimate holds:

$$\|P^J(\cdot, u) - P^{J,l}(\cdot, u)\|_V \le C2^{-l}\|P^J(\cdot, u)\|_W . \tag{26}$$

Therefore combining (24) and (26) and repeating the argument in the proof of (17), we obtain:

Proposition 14 *Consider the approximation of the elliptic problem (12) via the gpc finite element solution of the truncated problem (23), under Assumptions 6, 10 and 11. Then there exists a constant $C > 0$ such that for every $J, l \in \mathbb{N}$ and for every $u \in U$ it holds that*

$$\sup_{u \in U} \|P(\cdot, u) - P^{J,l}(\cdot, u)\|_V \le C(2^{-l}\|P^J(\cdot, u)\|_W + J^{-q}\|P(\cdot, u)\|_V) . \tag{27}$$

Moreover, the Finite Element solutions $P^{J,l}(\cdot, u)$ are V-stable in the sense that

$$\sup_{J,l \in \mathbb{N}} \sup_{u \in U} \|P^{J,l}(\cdot, u)\|_V \le \frac{C}{K_{min}}\|f\|_{V^*} . \tag{28}$$

4.2. FE Approximation of the Posterior Measure

We denote the vector of observables from the discretized parametric system's forward solution map by

$$\mathcal{G}^{J,l}(u) = (\mathcal{O}_1(P^{J,l}(u)), \ldots, \mathcal{O}_k(P^{J,l}(u))) : U \mapsto \mathbb{R}^k$$

and define the function

$$\Phi^{J,l}(u; \delta) = \frac{1}{2}|\delta - \mathcal{G}^{J,l}(u)|_\Sigma^2 . \tag{29}$$

We define an approximate conditional posterior probability measure $\rho^{J,l,\delta}$ on the measurable space (U, Θ) as

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

The measure $\rho^{J,l,\delta}$ is an approximation of ρ^δ, with error in the Hellinger metric which scales with J and l as the forward error in Proposition 14. We show this in the following Proposition whose proof generalizes the method introduced in [8].

Proposition 15 *If the domain D is convex and if $f \in L^2(D)$, there exists a positive constant c depending only on the data δ such that*

$$d_{\text{Hell}}(\rho^\delta, \rho^{J,l,\delta}) \le c(\delta)(J^{-q} + 2^{-l})\|f\|_{L^2(D)} .$$

Proof We denote the normalizing constants as

$$Z(\delta) = \int_U \exp(-\Phi(u;\delta))d\rho(u), \quad Z^{J,l}(\delta) = \int_U \exp(-\Phi^{J,l}(u;\delta))d\rho(u) .$$

We then estimate

$$2d_{\text{Hell}}(\rho^\delta, \rho^{J,l,\delta})^2$$
$$= \int_U \left(Z(\delta)^{-1/2} \exp\left(-\frac{1}{2}\Phi(u;\delta)\right) - (Z^{J,l}(\delta))^{-1/2} \exp\left(-\frac{1}{2}\Phi^{J,l}(u;\delta)\right) \right)^2 d\rho(u)$$
$$\leq I_1 + I_2,$$

where we defined

$$I_1 := \frac{2}{Z(\delta)} \int_U \left(\exp\left(-\frac{1}{2}\Phi(u;\delta)\right) - \exp\left(-\frac{1}{2}\Phi^{J,l}(u;\delta)\right) \right)^2 d\rho(u),$$

$$I_2 := 2|Z(\delta)^{-1/2} - Z^{J,l}(\delta)^{-1/2}|^2 \int_U \exp(-\Phi^{J,l}(u;\delta))d\rho(u) .$$

We estimate I_1 and I_2. To bound I_1, given data δ, for every $u \in U$ holds

$$\left| \exp\left(-\frac{1}{2}\Phi(u;\delta)\right) - \exp\left(-\frac{1}{2}\Phi^{J,l}(u;\delta)\right) \right| \leq \frac{1}{2}|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)|$$
$$\leq c(2|\delta| + |\mathcal{G}(u)| + |\mathcal{G}^{J,l}(u)|)|\mathcal{G}(u) - \mathcal{G}^{J,l}(u)| . \tag{30}$$

Moreover, by Proposition (14), there exists a constant $C > 0$ independent of J and of l such that, for all $u \in U$, there holds

$$|\mathcal{G}(u) - \mathcal{G}^{J,l}(u)| \leq C \max\{\|\mathcal{O}_i\|_{V^*}\}\|P(\cdot,u) - P^{J,l}(\cdot,u)\|_V$$
$$\leq C(2^{-l}\|P^J(\cdot,u)\|_W + J^{-q}\|P(\cdot,u)\|_V) .$$

By (17) and Proposition 13, $\|P(\cdot,u)\|_V$ and $\|P^J(\cdot,u)\|_W$ are uniformly bounded with respect to $u \in U$, so that

$$I_1 \leq c(\delta)\mathbb{E}^\rho(2^{-l}\|P^J(\cdot,u)\|_W + J^{-q}\|P(\cdot,u)\|_V)^2$$
$$\leq c(\delta)(J^{-2q}\|f\|_{V^*}^2 + 2^{-2l}\|f\|_{L^2(D)}^2) .$$

To estimate term I_2, we observe that there is a positive constant $c > 0$ such that for every $J, l \in \mathbb{N}$ holds

$$|Z(\delta)^{-1/2} - Z^{J,l}(\delta)^{-1/2}|^2 \leq c(Z(\delta)^{-3} \vee Z^{J,l}(\delta)^{-3})|Z(\delta) - Z^{J,l}(\delta)|^2 .$$

We note that

$$|Z(\delta) - Z^{J,l}(\delta)| \leq \int_U |\exp(-\Phi(u;\delta)) - \exp(-\Phi^{J,l}(u;\delta))|d\rho(u)$$
$$\leq \int_U |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)|d\rho(u) .$$

Therefore, as $Z(\delta)$ and $Z^{J,l}(\delta)$ are uniformly bounded below for all δ, analysis similar to that for I_1 shows that

$$I_2 \leq c(2^{-2l} + J^{-2q})\|f\|_{L^2(D)}^2.$$

Thus

$$d_{\text{Hell}}(\rho^\delta, \rho^{J,l,\delta}) \leq c(\delta)(2^{-l} + J^{-q})\|f\|_{L^2(D)} .$$

$\square$

4.3. Computational Complexity of Standard MCMC

Given $J, l \in \mathbb{N}$, and data δ, we use the MCMC method (7) to sample the probability measure $\rho^{J,l,\delta}$. In so doing we create a method for approximating integrals of functions $g : U \to \mathbb{R}$ with respect to ρ^{δ}. We use the following notation for the empirical measure generated by the Markov chain designed to sample $\rho^{J,l,\delta}$:

$$E_M^{\rho^{J,l,\delta}}[g] := \frac{1}{M} \sum_{k=1}^{M} g(u^{(k)}),$$

where the Markov chain $(u^{(k)})_{k \in \mathbb{N}}$ is generated from the process (7) with the acceptance probability being replaced by

$$\alpha^{J,l}(u, v) = 1 \wedge \exp(\Phi^{J,l}(u; \delta) - \Phi^{J,l}(v; \delta)), \quad (u, v) \in U \times U. \tag{31}$$

Given $M \in \mathbb{N}$ we wish to estimate the MC sampling error

$$\mathbb{E}^{\rho^{\delta}}[g] - E_M^{\rho^{J,l,\delta}}[g]. \tag{32}$$

We develop in the following two types of error bounds as $M \to \infty$ for (32): a probabilistic error bound for $\mathcal{P}_{u^{(0)}}^{J,l}$ almost every realization of the Markov chain and a mean square bound.

Proposition 16 *Let $g : U \to \mathbb{R}$ be a bounded continuous function on U with respect to the supremum norm. Then, for every initial condition $u^{(0)}$ and for $\mathcal{P}_{u^{(0)}}^{J,l}$-almost every realization of the Markov chain holds the error bound*

$$\left| \mathbb{E}^{\rho^{\delta}} g(u) - E_M^{\rho^{J,l,\delta}}[g] \right| \leq c_1 M^{-\frac{1}{2}} + c_2(J^{-q} + 2^{-l})$$

where $c_1 \leq c_3|\xi_M|$, ξ_M is a random variable (on the probability space generating the randomness within the Markov chain) which converges weakly as $M \to \infty$ to $\xi \sim N(0, 1)$ and c_2 is a non-random constant independent of M, J and l.

Moreover, there exists a constant c_4 (which is deterministic and depends only on the data δ, and which is, in particular, independent of J, l) such that

$$\left(\mathcal{E}^{\rho, J, l} \left[\left| \mathbb{E}^{\rho^{\delta}}[g] - E_M^{\rho^{J,l,\delta}}[g] \right|^2 \right] \right)^{1/2} \leq c_4(M^{-1/2} + J^{-q} + 2^{-l}) \tag{33}$$

Here, the constant $q > 0$ is as in Assumption 10.

Proof As g is bounded, we have from Proposition 15 and properties of the Hellinger metric (specifically, from (2.7) in [8]) for every $u \in U$ that

$$|\mathbb{E}^{\rho^{\delta}} g(u) - \mathbb{E}^{\rho^{J,l,\delta}} g(u)| \leq \bar{c}(g) d_{\text{Hell}}(\rho^{\delta}, \rho^{J,l,\delta}) \leq \bar{c}(g) c(\delta)(J^{-q} + 2^{-l}). \tag{34}$$

Here, $c(\delta)$ is as in Proposition 15 and $\bar{c}(g)$ depends on the supremum of $g(u)$ over U, but is independent of J, l. By Theorem 4 (and Remarks 5 and 9) we deduce the existence of a constant $C > 0$, independent of M, J and l, such that

$$|\mathbb{E}^{\rho^{J,l,\delta}} g - \frac{1}{M} \sum_{k=1}^{M} g(u^{(k)})| \leq C|\xi_M| M^{-1/2}$$

where ξ_M converges weakly as $M \to \infty$ to $\xi \sim N(0,1)$. Combining this with (34) gives the first assertion.

To prove the mean square error bound (33), we define

$$\bar{g}(u) := g(u) - \mathbb{E}^{\rho^{J,l,\delta}}[g] \ .$$

Due to the invariance of the stationary measures $\rho^{J,l,\delta}$, we may write

$$\frac{1}{M}\mathcal{E}^{\rho^{J,l,\delta},J,l}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] = \mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=1}^{M}\sum_{j=k+1}^{M}\mathcal{E}^{\rho^{J,l,\delta},J,l}[\bar{g}(u^{(k)})\bar{g}(u^j)]$$

$$= \mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathcal{E}^{\rho^{J,l,\delta},J,l}[\bar{g}(u^{(0)})\bar{g}(u^{(j)})]$$

$$= \mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})\mathcal{E}^{J,l}_{u^{(0)}}[\bar{g}(u^{(j)})]]$$

$$\leq \mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2]$$

$$+ 2\frac{1}{M}\sum_{k=0}^{M-1}\sup|\bar{g}|\sum_{j=1}^{M-k}\mathbb{E}^{\rho^{J,l,\delta}}[|\mathcal{E}^{J,l}_{u^{(0)}}g(u^{(j)}) - \mathbb{E}^{\rho^{J,l,\delta}}[g]|]$$

$$\leq \mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2] + 4\frac{1}{M}\sum_{k=0}^{M-1}\sup|\bar{g}|^2\sum_{j=1}^{M-k}(1-R)^j,$$

where, as in Remark 5, due to $\sup_{u\in U}\|P^{J,l}(u)\|_V$ being bounded uniformly with respect to the (discretization) parameters J and l, the constant $0 < R < 1$ is independent of the parameters J and l. Since $\sup_{J,l}\mathbb{E}^{\rho^{J,l,\delta}}[\bar{g}(u^{(0)})^2]$ is bounded independently of J and of l, we deduce that

$$\sup_{J,l,M\in\mathbb{N}} M\,\mathcal{E}^{\rho^{J,l,\delta},J,l}\Big[\Big|\frac{1}{M}\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] < \infty \ .$$

As

$$\mathcal{E}^{\rho,J,l}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] = \int_U \mathcal{E}^{J,l}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]d\rho(u^{(0)})$$

$$= \int_U \mathcal{E}^{J,l}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]\frac{d\rho}{d\rho^{J,l,\delta}}(u^{(0)})d\rho^{J,l,\delta}(u^{(0)})$$

$$\leq \mathcal{E}^{\rho^{J,l,\delta},J,l}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big]Z^{J,l}(\delta)\sup_{u\in U}\exp(\Phi^{J,l}(u;\delta)) \ .$$

As $Z^{J,l}(\delta) \leq 1$ and as $\sup_{u\in U}|\Phi^{J,l}(u;\delta)|$ is bounded uniformly with respect to J and l, we get the conclusion after using the bound from Proposition 15 on the Hellinger distance between ρ^δ and $\rho^{J,l,\delta}$. $\qquad\square$

We consider the case where $g(u) = \ell(P(u))$ with ℓ being a bounded linear functional in V^*. As

$$|\mathbb{E}^{\rho^\delta}[\ell(P(u))] - \mathbb{E}^{\rho^\delta}[\ell(P^{J,l}(u))]| \leq c(J^{-q} + 2^{-l})$$

and

$$|\mathbb{E}^{\rho^\delta}[\ell(P^{J,l}(u))] - \mathbb{E}^{\rho^{J,l,\delta}}[\ell(P^{J,l}(u))]| \leq c(J^{-q} + 2^{-l}),$$

we have

$$|\mathbb{E}^{\rho^\delta}[\ell(P(u))] - \mathbb{E}^{\rho^{J,l,\delta}}[\ell(P^{J,l}(u))]| \leq c(J^{-q} + 2^{-l}).$$

We therefore perform an MCMC algorithm to approximate $\mathbb{E}^{\rho^{J,l,\delta}}[\ell(P^{J,l}(u))]$. As $\ell(P^{J,l}(u))$ and $\Phi^{J,l}(u;\delta)$ depend only on the finite set of coordinates $\{u_1, \ldots, u_J\}$ in expansion (15), we perform the Metropolis-Hastings MCMC method on this set with proposals being drawn from the restriction of the prior measure ρ to this finite set.

Proposition 17 *Let $g(u) = \ell(P(u))$ where ℓ is a bounded linear functional in V^*. The approximate evaluation of the sample average $\frac{1}{M}\sum_{k=1}^{M}\ell(P^{J,l}(u^{(k)}))$ by the Markov Chain Monte Carlo Finite Element Method (MCMC-FEM for short) with M realizations of the chain, with Finite Element discretization in the domain D at mesh level l as described above, and with J-term truncated coefficient representation (22), requires $\mathcal{O}(l^{d-1}2^{dl}MJ)$ floating point operations.*

Proof From Assumption 12, we deduce that the total number of non-zero entries of the stiffness matrix for solving the Finite Element equation (25) is $O(l^{d-1}2^{dl})$. To compute each of these entries, we require $O(J)$ operations for computing the coefficients K^J at the quadrature points. Therefore the cost of constructing the stiffness matrix is $O(l^{d-1}2^{dl}J)$. From Assumption 11, the Riesz basis property of the Finite Element basis implies that the condition number of the stiffness matrix is bounded uniformly for all l. Hence, the conjugate gradient method for the approximate solution of the linear system resulting from the Finite Element discretization with an accuracy comparable to the order of the discretization error requires $O(l^{d-1}2^{dl})$ float point operations. The total cost for solving the approximated forward problem at each step of the Markov chain requires at most $O(l^{d-1}2^{dl}J)$ floating point operations. Computing $\ell(P^{J,l}(u^{(k)}))$ requires $O(2^{dl})$ floating point operations. Since we draw M samples of the chain, the assertion follows. $\square$

We have the following result.

Theorem 18 *Let Assumptions 6, 10 and 11 hold. For $g(u) = \ell(P(u))$ where ℓ is a bounded linear functional in V^*, with probability $p_{N_{\mathrm{dof}}}(t)$ the conditional expectation $\mathbb{E}^{\rho^\delta}g(u)$ can be approximated using N_{dof} degrees of freedom to approximate the forward PDE and $t^2 N_{\mathrm{dof}}^{2/d}$ MCMC steps (with a total of $t^2 N_{\mathrm{dof}}^{1+2/d}$ degrees of freedom), incurring an error of $O(N_{\mathrm{dof}}^{-1/d})$, and using not more than*

$$ct^2 \log(N_{\mathrm{dof}})^{d-1} N_{\mathrm{dof}}^{1+(2+1/q)/d}$$

floating point operations, where

$$\lim_{N_{\mathrm{dof}}\to\infty} p_{N_{\mathrm{dof}}}(t) \to \int_{-c't}^{c't} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx \,,$$

for some positive constants c, c' independent of N_{dof}.

In mean square with respect to the measure $\mathcal{P}^{\rho,J,l}$, $\mathbb{E}^{\rho^\delta}[g(u)]$ can be approximated with an error $O(N_{dof}^{-1/d})$, using not more than $N_{\text{dof}}^{1+2/d}$ number of degrees of freedom in total, and not more than $O(\log(N_{\text{dof}})^{d-1}N_{\text{dof}}^{1+(2+1/q)/d}$ floating point operations. Here, the constant $q > 0$ is as in Assumption 10.

Proof We invoke the error estimate in Proposition 16, and choose the parameters M, J and l so as to balance the bounds $M^{-1/2}$, J^{-q} and 2^{-l}, taking into account the fact that the coefficient of $M^{-\frac{1}{2}}$ is only known through its asymptotic normality. We select $J = 2^{l/q}$ and $M = t^2 N_{dof}^{2/d}$ where $t = c_3|\xi_M|$, with N_{dof} denoting the number of degrees of freedom at each step being $N_{dof} = O(2^{dl})$; the constant c_3 and the random variable ξ_M is as in Proposition 16. Then the total number of floating point operations required as $l \to \infty$ is not larger than $O(t^2 l^{d-1} 2^{(d+2+1/q)l})$.

We then arrive at the conclusion.   □

5. Sparse gpc-MCMC

We again study computational complexity of the MCMC method defined by (7) to sample the conditional probability measure ρ^δ determined in the previous section. However we study a computational method which effects a reduction in computational cost by precomputing the parametric dependence of the forward model, which enters the likelihood. This method is introduced, and used in practice, in the series of papers [17, 15, 16]. The major cost in MCMC methods is the repeated solution of the forward equations, with varying coefficients from the MCMC sampler of ρ^δ. The complexity of these repeated forward solves can be drastically reduced by precomputing an approximate, deterministic parametric representation of the system's response which is valid for *all* possible realizations of $u \in U$. Specifically, we precompute a sparse tensor finite element approximation of the parametric, deterministic forward problem, by an approximate polynomial chaos representation of the solution's dependence on u and by discretization of the forward solutions' spatial dependence from a multilevel hierarchy of Finite Element spaces in D. As we shall show, this strategy is particularly effective, if only linear functionals $\ell(\cdot)$ of the system's response are of interest: in this case, only scalar coefficients of the gpc expansion need to be stored and evaluated. We use this to reduce the cost per step of the MCMC method. We again work under Assumption 10 and, furthermore, we make the following assumptions throughout this section:

5.1. Sparse Tensor gpc-Finite Element Approximation of the Parametric Forward Problem

5.1.1. Best N term parametric approximation By (17), the solution $P(\cdot, u)$ of problem (16) is uniformly bounded in V by $\|f\|_{V^*}/K_{\min}$. Therefore, from Proposition 7, we deduce that $P(\cdot, \cdot) \in L^2(U, \rho; V)$. Therefore, the parametric solution admits a polynomial chaos type representation in $L^2(U, \rho; V)$. To define it, we denote by $L_n(u_n)$

the Legendre polynomial of degree n, normalized such that

$$\frac{1}{2} \int_{-1}^{1} |L_n(\xi)|^2 d\xi = 1 .$$

By $\mathcal{F}$ we denote the set of all sequences $\nu = (\nu_1, \nu_2, \ldots)$ of nonnegative integers such that ν_j are integers for all j and only a finite numbers of them are non-zero. We define

$$L_\nu(u) = \prod_{j \geq 1} L_{\nu_j}(u_j) . \tag{35}$$

Since $L_0 \equiv 1$, for each $\nu \in \mathcal{F}$ the products contain only finitely many nontrivial factors. The set $\{L_\nu : \nu \in \mathcal{F}\}$ forms an orthonormal basis for $L^2(U, \rho)$. We can therefore expand $P(\cdot, u)$ into the Legendre expansion

$$P(\cdot, u) = \sum_{\nu \in \mathcal{F}} P_\nu(\cdot) L_\nu(u),$$

where $P_\nu := \int_U P(\cdot, u) L_\nu(u) d\rho(u) \in V$. By the $L^2(U, \rho)$ orthonormality of the set $\{L_\nu : \nu \in \mathcal{F}\}$, Parseval's equation in the Bochner space $L^2(U, \rho; V)$ takes the form

$$\forall P \in L^2(U, \rho; V) : \quad \|P\|^2_{L^2(U,\rho;V)} = \sum_{\nu \in \mathcal{F}} \|P_\nu\|^2_V .$$

For the ensuing analysis, we shall impose the following assumption on the summability of the gpc expansion of P:

Assumption 19 *There exists a constant $0 < p < 1$ such that the coefficients P_ν of the gpc expansion of P satisfy $(\|P_\nu\|_V)_\nu \in \ell^p(\mathcal{F})$.*

This assumption is valid under the provision of suitable decay of the coefficient functions ψ_j such as Assumption 10. We refer to [5, 6] for details. By a classical argument ("Stechkin's Lemma"), this implies the following, so-called "best N-term approximation property".

Proposition 20 *Under Assumption 19, there exists a nondecreasing sequence $\{\Lambda_N\}_{N \in \mathbb{N}} \subset \mathcal{F}$ of subsets Λ_N whose cardinality does not exceed N, such that*

$$\left\| P - \sum_{\nu \in \Lambda_N} P_\nu(u) L_\nu \right\|^2_{L^2(U,\rho;V)} = \sum_{\nu \in \mathcal{F} \setminus \Lambda_N} \|P_\nu\|^2_V \leq C N^{-2r}, \tag{36}$$

where the convergence rate $r = 1/p - 1/2 > 1/2$ and where the constant $C = \|(\|P_\nu\|_V)_{\nu \in \mathcal{F}}\|^2_{\ell^p(\mathcal{F})}$ is independent of N.

5.1.2. Finite element approximation The best N-term approximations

$$u_{\Lambda_N} := \sum_{\nu \in \Lambda_N} P_\nu(u) L_\nu \tag{37}$$

in Proposition 20 indicate that sampling the parametric forward map with evaluation of N solutions $P_\nu(u)$, $\nu \in \Lambda_N$ of the parametric, elliptic problem with accuracy N^{-r} is possible; since $r > 1/2$, this is superior to what can be expected from N MC samples.

There are, however, two obstacles which obstruct the practicality of this idea: first, the proof of Proposition 20 is nonconstructive, and does not provide concrete choices for the sets Λ_N of "active" gpc coefficients which realize (36) and, second, even if Λ_N were available, the "coefficients" $P_\nu \in V$ can not be obtained exactly, in general, but must be approximated for example by a Finite Element discretization in D.

As $u \in L^2(U, \rho; V)$, we consider the variational form "in the mean" of (16) as

$$\int_U \int_D K(x,u)\nabla P(x,u) \cdot \nabla Q(x,u)dxd\rho(u) = \int_U \int_D f(x)Q(x,u)dxd\rho(u), \tag{38}$$

for all $Q \in L^2(U, \rho; V)$. For each set $\Lambda_N \subset \mathcal{F}$ of cardinality not more than N that satisfies Proposition 20, and each vector $\mathcal{L} = (l_\nu)_{\nu \in \Lambda_N}$ of nonnegative integers, we define finite dimensional approximation spaces as

$$X_{N,\mathcal{L}} = \{P_{N,\mathcal{L}} = \sum_{\nu \in \Lambda_N} P_{\nu,\mathcal{L}}(x)L_\nu(u); \; P_{\nu,\mathcal{L}} \in V^{l_\nu}\} . \tag{39}$$

Evidently, $X_{N,\mathcal{L}} \subset L^2(U, \rho; V)$ is a finite-dimensional (hence closed) subspace for any N and any selection $\mathcal{L}$ of the discretization levels.

The total number of degrees of freedom, $N_{dof} = \dim(X_{N,\mathcal{L}})$, necessary for the sparse representation of the parametric forward map is given by

$$N_{\text{dof}} = O\left(\sum_{\nu \in \Lambda_N} 2^{dl_\nu}\right) \quad \text{as} \quad N, \; l_\nu \to \infty . \tag{40}$$

The *stochastic, sparse tensor Galerkin approximation* of the parametric forward problem (16), based on the index sets $\Lambda_N \subset \mathcal{F}$, and $\mathcal{L} = \{l_\nu : \nu \in \Lambda_N\}$, reads: find $P_{N,\mathcal{L}} \in X_{N,\mathcal{L}}$ such that for all $Q_{N,\mathcal{L}} \in X_{N,\mathcal{L}}$ holds

$$\begin{aligned} b(P_{N,\mathcal{L}}, Q_{N,\mathcal{L}}) &:= \int_U \int_D K(x,u)\nabla P_{N,\mathcal{L}} \cdot \nabla Q_{N,\mathcal{L}}dxd\rho(u) \\ &= \int_U \int_D f(x)Q_{N,\mathcal{L}}(x,u)dxd\rho(u) . \end{aligned} \tag{41}$$

The coercivity of the bilinear form $b(\cdot, \cdot)$ ensures the existence and uniqueness of $P_{N,\mathcal{L}}$ as well as their quasioptimality in $L^2(U, \rho; V)$: by Cea's lemma, for a constant $C > 0$ which is independent of Λ and of $\mathcal{L}$,

$$\|P - P_{N,\mathcal{L}}\|_{L^2(U,\rho;V)} \le C \inf_{Q_{\nu,\mathcal{L}} \in V^{l_\nu}} \|P - \sum_{\nu \in \Lambda} Q_{\nu,\mathcal{L}}L_\nu\|_{L^2(U,\rho;V)} .$$

We obtain the following error bound which consists of the error in the best N-term truncation for the gpc expansion and of the Finite Element approximation error for the "active" gpc coefficients.

$$\|P - P_{N,\mathcal{L}}\|^2_{L^2(U,\rho;V)} \le C(N^{-2r} + \sum_{\nu \in \Lambda_N} \inf_{Q_{\nu,\mathcal{L}} \in V^{l_\nu}} \|P_\nu - Q_{\nu,\mathcal{L}}\|^2_V) . \tag{42}$$

The following assumption is, therefore, a stronger requirement than the mere p-summability of the gpc coefficient sequence $\{\|P_\nu\|_V\}_{\nu \in \mathcal{F}}$.

Assumption 21 *There are positive constants τ, α and β such that with a total budget of N_{dof} degrees of freedom, and with at most $N = N_{dof}^{\tau/r}$ active gpc modes ν, an active set of gpc modes Λ_N such that $|\nu| = O(\log N) \ \forall \ \nu \in \Lambda_N$, and a combined gpc-Finite Element approximation $P_{N,\mathcal{L}} \in X_{N,\mathcal{L}}$ with rate of convergence*

$$\|P - P_{N,\mathcal{L}}\|_{L^2(U,\rho;V)} \leq C N_{dof}^{-\tau},$$

can be found in $O(N_{dof}^{\alpha}(\log N_{dof})^{\beta})$ float point operations.

Note that the constant $C > 0$ in Assumption 21 could, in general be considerably larger that the best possible constant in the N-term approximation result Proposition 20.

Let us indicate sufficient conditions that ensure Assumptions 19, 21. The first condition is quantitative decay rate of the coefficient functions ψ_j in the parametric representation (13) of the random input.

Assumption 22 *The coefficients ψ_j are arranged in decreasing order of magnitude of $\|\psi_j\|_{L^\infty(D)}$ and there is a constant $s > 1$ and $C > 0$ such that*

$$\forall j \in \mathbb{N}: \quad \|\psi_j\|_{L^\infty(D)} \leq C j^{-s} \ .$$

To obtain convergence rates for the FE-discretization in the domain D, i.e. of the term $\|P_\nu - Q_{\nu,\mathcal{L}}\|_V$ in (42), it is also necessary to ensure spatial regularity of the solution $P(x,u)$ of the parametric problem (16). To this end, we require

Assumption 23 *For all $j \in \mathbb{N}$, $\psi_j \in W^{1,\infty}(D)$ and there exists a constant $C > 0$ such that*

$$\forall j \in \mathbb{N}: \quad \|\nabla\psi_j\|_{W^{1,\infty}(D)} \leq C j^{-s'} \quad \text{for some} \quad 1 < s' \leq s \ .$$

We remark that Assumptions 22 and 23 imply Assumption 10 with $q = s - 1 > 0$. Under these assumptions, the following proposition holds.

Proposition 24 *Under Assumptions 22, 23, if, moreover, the domain D is convex and $f \in L^2(D)$, the solution $P(\cdot, u)$ of the parametric, deterministic problem (16) belongs to the space $L^2(U, \rho; W)$.*

From estimate (42), we get with Proposition 24 and standard approximation properties of continuous, piecewise linear FEM the error bound

$$\|P - P_{N,\mathcal{L}}\|_{L^2(U,\rho;V)}^2 \leq C(N^{-2r} + \sum_{\nu \in \Lambda_N} 2^{-2l_\nu}\|P_\nu\|_{H^2(D)}^2) \ . \tag{43}$$

In order to obtain an error bound in terms of N_{dof} defined in (40) which is uniform in terms of N, we select, for $\nu \in \Lambda_N$ the discretization levels l_ν of the active gpc coefficient P_ν so that both terms in the upper bound (43) are of equal order of magnitude. This constrained optimization problem was solved, for example, in [5], under the assumption that $(\|P_\nu\|_{H^2(D)})_\nu \in \ell^p(\mathcal{F})$.

In recent years, several algorithms have appeared or are under current development which satisfy Assumption 21 with various exponents $\alpha \geq 1$ and $\beta \geq 0$. We mention only the references [3, 21, 11, 2, 4]

5.2. Approximation of the Posterior Measure

For the solution $P_{N,\mathcal{L}}$ in Assumption 21, we define

$$\mathcal{G}^{N,\mathcal{L}}(u) = (\mathcal{O}_1(P_{N,\mathcal{L}}(u)), \ldots, \mathcal{O}_d(P_{N,\mathcal{L}}(u))),$$

and the function

$$\Phi^{N,\mathcal{L}}(u;\delta) = \frac{1}{2}|\delta - \mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma^2 .$$

The conditional measure $\rho^{N,\mathcal{L},\delta}$ on the measurable space (U, Θ) is defined as

$$\frac{d\rho^{N,\mathcal{L},\delta}}{d\rho} \propto \exp(-\Phi^{N,\mathcal{L}}(u;\delta)) .$$

We then have the following approximation result.

Proposition 25 *Let Assumptions 19, 21 hold. Then there is a constant $c = c(\delta)$ which only depends on the data δ such that*

$$d_{\mathrm{Hell}}(\rho^\delta, \rho^{N,\mathcal{L},\delta}) \le c(\delta)N_{\mathrm{dof}}^{-\tau} .$$

Proof The proof for this proposition is similar to that for Proposition 15, differing only in a few details; hence we highlight only the differences. These are due to estimates on the forward error from Assumptions 21 being valid only in the mean square sense whilst Proposition 14 holds pointwise for $u \in U$. Nonetheless, at the point in the estimation of I_1 and I_2 where the forward error estimate is used, it is possible to use a mean square forward error estimate instead of a pointwise forward error estimate. From Assumption 21, we deduce that there is a positive constant c such that:

$$\rho\{u : \ |\mathcal{G}(u) - \mathcal{G}^{N,\mathcal{L}}(u)| > 1\} \le cN_{dof}^{-2\tau} .$$

As $\|P(u)\|_V$ is uniformly bounded for all u, there is a constant $c_1(\delta) > 0$ such that $|\delta - \mathcal{G}(u)|_\Sigma < c_1(\delta)$. Choose a constant $c_2(\delta) > 0$ sufficiently large. If $|\delta - \mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma > c_2(\delta)$, then

$$|\mathcal{G}^{N,\mathcal{L}}(u) - \mathcal{G}(u)|_\Sigma \ge |\delta - \mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma - |\delta - \mathcal{G}(u)|_\Sigma > c_2(\delta) - c_1(\delta) > 1 .$$

Let $U_1 \subset U$ be the set of $u \in U$ such that $|\delta - \mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma > c_2(\delta)$. We have that $\rho(U_1) \le cN_{dof}^{-2\tau}$. Thus,

$$\frac{1}{Z(\delta)} \int_{U_1} \left| \exp\left(-\frac{1}{2}\Phi(u;\delta)\right) - \exp\left(-\frac{1}{2}\Phi^{N,\mathcal{L}}(u;\delta)\right) \right| d\rho(u) \le c(\delta)N_{dof}^{-2\tau} .$$

When $u \notin U_1$, $|\delta - \mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma \le c_2(\delta)$ so there is a constant $c_3(\delta)$ so that $|\mathcal{G}^{N,\mathcal{L}}(u)| \le c_3(\delta)$. An argument similar to that for (30) shows that

$$\left| \exp\left(-\frac{1}{2}\Phi(u;\delta)\right) - \exp\left(-\frac{1}{2}\Phi^{N,\mathcal{L}}(u;\delta)\right) \right| \le$$
$$c(2|\delta| + |\mathcal{G}(u)| + |\mathcal{G}^{N,\mathcal{L}}(u)|)|\mathcal{G}(u) - \mathcal{G}^{N,\mathcal{L}}(u)|.$$

Therefore

$$I_1 = \frac{1}{Z(\delta)} \int_U \Big| \exp\Big(-\frac{1}{2}\Phi(u;\delta)\Big) - \exp\Big(-\frac{1}{2}\Phi^{N,\mathcal{L}}(u;\delta)\Big)\Big|^2 d\rho(u)$$
$$\leq c(\delta) N_{dof}^{-2\tau} +$$
$$c \int_U (2|\delta| + |\mathcal{G}(u)| + c_3(\delta))^2 |\mathcal{G}(u) - \mathcal{G}^{N,\mathcal{L}}(u)|^2 d\rho(u)$$
$$\leq c(\delta) N_{dof}^{-2\tau} + c(\delta) \int_U \|P(\cdot, u) - P_{N,\mathcal{L}}(\cdot, u)\|_V^2 d\rho(u)$$
$$\leq c(\delta) N_{dof}^{-2\tau} .$$

To show that $I_2 < c(\delta) N_{\mathrm{dof}}^{-2\tau}$ we still need to verify that

$$Z^{N,\mathcal{L}}(\delta) = \int_U \exp(-\Phi^{N,\mathcal{L}}(u;\delta)) d\rho(u)$$

is uniformly bounded from below by a positive bound for all N and $\mathcal{L}$. As $P_{N,\mathcal{L}}$ is uniformly bounded in $L^2(U, \rho; V)$,

$$\int_U |\mathcal{G}^{N,\mathcal{L}}(u)| d\rho(u) \leq c \int_U \|P_{N,\mathcal{L}}(u)\|_V d\rho(u) \leq c.$$

Fixing $r > 0$ sufficiently large, the ρ measure of the set $u \in U$ such that $|\mathcal{G}^{N,\mathcal{L}}(u)| > r$ is bounded by c/r. Therefore the measure of the set of $u \in U$ such that $|\mathcal{G}^{N,\mathcal{L}}(u)| \leq r$ is bounded from below by $1 - c/r$. Thus we have proved

$$Z^{N,\mathcal{L}}(\delta) \geq \int_U \exp(-\frac{1}{2}(|\delta|_\Sigma + |\mathcal{G}^{N,\mathcal{L}}(u)|_\Sigma)^2) d\rho(u) > c(\delta) > 0 .$$

$$\square$$

Let $(u^{(k)})_k$ be the Markov chain generated by the sampling process (7) with the acceptance probability being replaced by

$$\alpha^{N,\mathcal{L}}(u, v) = 1 \wedge \exp(\Phi^{N,\mathcal{L}}(u, \delta) - \Phi^{N,\mathcal{L}}(v, \delta)) . \tag{44}$$

We denote by

$$E_M^{\rho^{N,\mathcal{L},\delta}}[g] = \frac{1}{M} \sum_{k=1}^M g(u^{(k)}) .$$

We then have:

Proposition 26 *Let g be a bounded continuous function from U to $\mathbb{R}$. Then*

$$|\mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g]| \leq c_6 M^{-1/2} + c_7 N_{\mathrm{dof}}^{-\tau}, \tag{45}$$

$\mathcal{P}\rho^{N,\mathcal{L},\delta,N,\mathcal{L}}$ almost surely, where $c_6 \leq c_8 |\xi_M|$ where ξ_M is a random variable which converges weakly as $M \to \infty$ to $\xi \sim N(0, 1)$; the constants c_7 and c_8 are deterministic and do not depend on M, N and N_{dof}.

There exists a deterministic positive constant c_9 such that the gpc-MCMC converges in the mean square with the same rate of convergence

$$\Big(\mathcal{E}^{\rho,N,\mathcal{L}}\Big[\Big|\mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g]\Big|^2\Big]\Big)^{1/2} \leq c_9(M^{-1/2} + N_{\mathrm{dof}}^{-\tau}) .$$

Proof Using (44), the probability that a random draw from ρ has probability larger than $\exp(-\Phi^{N,\mathcal{L}}(v;\delta))$ of being accepted. Therefore the transition kernel of the Markov chain generated by (7) with the acceptance probability (44) satisfies

$$p(u, A) \geq \int_A \exp(-\Phi^{N,\mathcal{L}}(v;\delta)) d\rho(v).$$

Using Theorem 16.2.4 of [18], we deduce that the nth iteration of the transition kernel satisfies

$$\|p^n(u, \cdot) - \rho^{N,\mathcal{L},\delta}\|_{\mathrm{TV}} \leq 2 \left(1 - \int_U \exp(-\Phi^{N,\mathcal{L}}(v;\delta)) d\rho(v) \right)^n .$$

From the proof of Proposition 25, we have

$$\int_U \exp(-\Phi^{N,\mathcal{L}}(v;\delta)) d\rho(v) \geq \exp(-c_2(\delta)^2/2) + cN_{dof}^{-2\tau} .$$

Thus, we can choose a constant $R < 1$ independent of the approximating parameters N and $\mathcal{L}$ so that for all $n \in \mathbb{N}$ holds

$$\|p^n(u, \cdot) - \rho^{N,\mathcal{L},\delta}\|_{\mathrm{TV}} \leq 2(1 - R)^n .$$

In a similar manner as for Proposition 16, we deduce the probabilistic bound. For the mean square bound, similar to the proof of Proposition 16, we have

$$\mathcal{E}^{\rho^{N,\mathcal{L}},N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] \leq C(M^{-1/2} + N_{dof}^{-\tau})^2 .$$

Let $U' := \{u \in U : |\mathcal{G}^{N,\mathcal{L}}(u) - \mathcal{G}(u)| > 1\}$. We deduce that there exists a constant $c > 0$ independent of $\mathcal{L}$, N_{dof}, N such that $\rho(U') \leq cN_{dof}^{-2\tau}$ and such that we may estimate

$$\begin{aligned}
& \mathcal{E}^{\rho,N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] \\
=\ & \int_{U'} \mathcal{E}_{u^{(0)}}^{N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] d\rho(u^{(0)}) \\
& + \int_{U \setminus U'} \mathcal{E}_{u^{(0)}}^{N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] d\rho(u^{(0)}) \\
\leq\ & CN_{dof}^{-2\tau} + \int_{U \setminus U'} \mathcal{E}_{u^{(0)}}^{N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] d\rho(u^{(0)}) \\
\leq\ & CN_{dof}^{-2\tau} + \\
& \int_{U \setminus U'} \mathcal{E}_{u^{(0)}}^{N,\mathcal{L}} \left[\left| \mathbb{E}^{\rho^\delta}[g] - E_M^{\rho^{N,\mathcal{L},\delta}}[g] \right|^2 \right] Z^{N,\mathcal{L}}(\delta) \exp(\Phi^{N,\mathcal{L}}(u;\delta)) d\rho^{N,\mathcal{L},\delta}(u^{(0)}) .
\end{aligned}$$

On $U \setminus U'$, $\sup_{u \in U} |\mathcal{G}^{N,\mathcal{L}}(u)|$ is uniformly bounded with respect to all N and $\mathcal{L}$. From this, we get the conclusion. $\qquad\square$

Remark 27 *In Proposition 26, g is assumed to be a bounded continuous function from U to $\mathbb{R}$. The sparse-MCMC is of particular interest in the case where g is given by $\ell \circ P$ where ℓ is a linear functional on V, i.e. $\ell \in V^*$. From Assumption 21 and the fact that*

$$\frac{d\rho^\delta}{d\rho}(u) = \frac{1}{Z(\delta)} \exp(-\Phi(u;\delta)) ,$$

we deduce that

$$|\mathbb{E}^{\rho^\delta}[\ell(P(u))] - \mathbb{E}^{\rho^\delta}[\ell(P_{N,\mathcal{L}}(u))]| \leq c(\delta, \ell) N_{dof}^{-\tau} .$$

On the other hand, from Proposition 25, we have (cf. [8, Eq. (2.7)])

$$|\mathbb{E}^{\rho^{\delta}}[\ell(P_{N,\mathcal{L}}(u))] - \mathbb{E}^{\rho^{N,\mathcal{L},\delta}}[\ell(P_{N,\mathcal{L}}(u))]| \leq c(\delta, \ell) N_{dof}^{-\tau} .$$

Therefore, by the triangle inequality,

$$|\mathbb{E}^{\rho^{\delta}}[\ell(P(u))] - \mathbb{E}^{\rho^{N,\mathcal{L},\delta}}[\ell(P_{N,\mathcal{L}}(u))]| \leq c(\delta, \ell) N_{dof}^{-\tau} .$$

We wish to approximate $\mathbb{E}^{\rho^{N,\mathcal{L},\delta}}[\ell(P_{N,\mathcal{L}}(u))]$ with a Markov Chain-Monte Carlo algorithm. In doing so, the following difficulty may arise: although $\ell(P(u))$ is uniformly bounded with respect to $u \in U$, $\sup_{u \in U} \ell(P_{N,\mathcal{L}}(u))$ may not be uniformly bounded with respect to N and $\mathcal{L}$. However, we can still apply Proposition 26 by using a cut-off argument: to this end, we define the continuous bounded function $\tilde{g}(u) : U \to \mathbb{R}$ by truncation, i.e.

$$\tilde{g}(u) := \begin{cases} \ell(P_{N,\mathcal{L}}(u)) & \text{if} \quad |\ell(P_{N,\mathcal{L}}(u))| \leq \sup_{u \in U} |\ell(P(u))| + 1 , \\ \sup_{u \in U} |\ell(P(u))| + 1 & \text{if} \quad \ell(P_{N,\mathcal{L}}(u)) > \sup_{u \in U} |\ell(P(u))| + 1 , \\ -\sup_{u \in U} |\ell(P(u))| - 1 & \text{if} \quad \ell(P_{N,\mathcal{L}}(u)) < -\sup_{u \in U} |\ell(P(u))| - 1 . \end{cases}$$

Define $U' := \{u \in U : |\ell(P(u)) - \ell(P_{N,\mathcal{L}}(u))| > 1\}$. From Assumption 21, we find that $\rho(U') < c(\ell, \delta) N_{dof}^{-2\tau}$. It follows then that there exists a constant $c > 0$ depending on the data δ, but independent of N and of $\mathcal{L}$ such that

$$|\mathbb{E}^{\rho^{N,\mathcal{L},\delta}}[\ell(P_{N,\mathcal{L}}(u)) - \tilde{g}(u)]| \leq \int_{U'} |\ell(P_{N,\mathcal{L}}(u)) - \tilde{g}(u)| d\rho^{N,\mathcal{L},\delta}(y)$$

$$\leq c(\delta) \int_{U} I_{U'}(u)(|\ell(P_{N,\mathcal{L}}(u))| + c) d\rho(y)$$

$$\leq c(\delta) \rho(U')^{1/2}(\|\ell(P_{N,\mathcal{L}}(u)\|_{L^2(U,\rho;\mathbb{R})} + c) \leq c(\delta) N_{dof}^{-\tau} .$$

Therefore, we may run the MCMC algorithm on $\mathbb{E}^{\rho^{N,\mathcal{L},\delta}}[\tilde{g}(u)]$.

At each step of the MCMC algorithm, we need to compute $\ell(P_{N,\mathcal{L}}(u^{(k)}))$ which, *for linear functionals* $\ell(\cdot)$, *is equal to* $\sum_{\nu \in \Lambda} \ell(P_{\nu,\mathcal{L}}) L_{\nu}(u^{(k)})$. Because the parametric solution of the elliptic problem can be precomputed before the MCMC is run, and then needs only to be *evaluated* at each state of the MCMC method, significant savings can be obtained. We illustrate this, using the ideas of the previous Remark 27, to guide the choice of test functions.

Proposition 28 *Let $g(u) = \ell(P(u))$ where ℓ is a bounded linear functional in V^*. Under Assumption 21, the total number of floating point operations required for performing M steps in the Metropolis-Hastings method as $N, M \to \infty$ is bounded by $O(N_{\text{dof}}^{\alpha}(\log N_{\text{dof}})^{\beta} + MN \log N)$.*

Proof Under Assumption 21, the cost of solving one instance of problem (41) is bounded by $O(N_{\text{dof}}^{\alpha}(\log N_{\text{dof}})^{\beta})$. At each MCMC step, we need to evaluate the observation functionals

$$\mathcal{O}_i(P_{N,\mathcal{L}}(u^{(k)})) = \sum_{\nu \in \Lambda_N} \mathcal{O}_i(P_{\nu,\mathcal{L}}) L_{\nu}(u^{(k)}) . \tag{46}$$

We note in passing that the storage of the parametric gpc-type representation of the forward map (46) requires only one real per gpc mode, *provided that only functionals of the forward solution are of interest.* We now estimate the complexity of computing one draw of the forward map (46). For $\nu \in \mathcal{F}$, each multivariate Legendre polynomial $L_\nu(u^k)$ can be evaluated with $O(|\nu|)$ float point operations. As $|\nu| = O(\log N)$, computing the observation functionals $\mathcal{O}_i(P_{N,\mathcal{L}})$ requires $O(N \log N)$ floating point operations. Thus we need $O(N_{\mathrm{dof}}^\alpha (\log N_{\mathrm{dof}})^\beta + MN \log N)$ floating point operations to perform M steps of the Metropolis-Hastings method with sampling of the surrogate, sparse gpc-Finite Element representation of the forward map.

$\square$

Theorem 29 *For $g(u) = \ell(P(u))$ where ℓ is a bounded linear functional $\ell \in V^*$, under Assumptions 6, 11 and 22, with probability $p_{N_{\mathrm{dof}}}(t)$ the conditional expectation $\mathbb{E}^{\rho^\delta}[g(u)]$ can be approximated with N_{dof} degrees of freedom, incurring an error of $O(N_{\mathrm{dof}}^{-\tau})$ using not more than*

$$cN_{\mathrm{dof}}^\alpha (\log N_{\mathrm{dof}})^\beta + ct^2 N_{\mathrm{dof}}^{2\tau+\tau/r} \log(N_{\mathrm{dof}})$$

many floating point operations, where

$$\lim_{N_{\mathrm{dof}} \to \infty} p_{N_{\mathrm{dof}}}(t) \to \int_{-c't}^{c't} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx,$$

for some constants c, c' independent of N_{dof}.

In the mean square with respect to the measure $\mathcal{P}^{\rho,N,\mathcal{L}}$, $\mathbb{E}^{\rho^\delta}[g(u)]$ can be approximated with N_{dof} degrees of freedom, with an error $N_{dof}^{-\tau}$ using not more than

$$O(N_{\mathrm{dof}}^\alpha (\log N_{\mathrm{dof}})^\beta + N_{\mathrm{dof}}^{2\tau+\tau/r} \log(N_{\mathrm{dof}}))$$

floating point operations.

Proof We relate the number of MCMC realizations M with the total number of degrees of freedom N_{dof} by equating the terms in the error bound (45). To this end, we choose $M = t^2 N_{\mathrm{dof}}^{2\tau}$ where $t = c_8 |\xi_M|$; the constant c_8 and the random variable ξ_M are as in Proposition 26 . With $N = N_{\mathrm{dof}}^{\tau/r}$, the number of floating point operations required in Proposition 28 is bounded by

$$cN_{\mathrm{dof}}^\alpha (\log N_{\mathrm{dof}})^\beta + ct^2 N_{\mathrm{dof}}^{2\tau+\tau/r} \log N_{\mathrm{dof}} \ .$$

As ξ_M converges weakly to the normal Gaussian variable, we deduce the limit for the probability density $p_{N_{\mathrm{dof}}}(t)$ of the random variable t. The proof for the mean square approximation is similar. $\square$

Remark 30 *In the previous section, the parametric PDE (16) is to be solved once at every step of the MCMC process, using N_{dof} degrees of freedom, with $O(N_{\mathrm{dof}}^{2/d})$ steps required (the multiplying constant depends on a random variable when we consider the realization-wise error). Ignoring $\log$ factors, the resulting error can be expressed in terms of the total number of floating points operations N_{fp} as $O(N_{fp}^{-1/(d+2+1/q)})$. Here,*

the forward PDE is solved for every realization before running the MCMC process. The rate of convergence of the MCMC process in terms of the total number of floating point operations used is $O(N_{fp}^{-\min(\tau/\alpha,1/(2+1/r))})$. This can be significantly smaller than the rate of convergence in Theorem 18 when α is close to 1. For example, with the decay rate of $\|\psi_j\|_\infty$ in Assumption 22, the summability constant p in Assumption 19 can be any constant that is greater than $1/s$. Therefore the constant r in Proposition 20 can be any psitive constant smaller than $s - 1/2$. On the other hand, the constant q in Assumption 10 is bounded by $s - 1$. As

$$2 + \frac{1}{s - 1/2} < d + 2 + \frac{1}{s - 1},$$

we therefore can choose r so that

$$\frac{1}{2 + 1/r} > \frac{1}{d + 2 + 1/q}.$$

As shown in [5], when $(\|P_\nu\|_{H^2(D)})_\nu \in \ell^p(\mathcal{F})$, τ can be chosen as $1/d$. Thus, when α is sufficiently close to 1, the complexity of the sparse gpc-MCMC approach is superior to that of the plain MCMC approach in the previous section.

6. Multilevel MCMC

We showed that substantial complexity reduction is possible in the "plain" MCMC FE sampling of the posterior measure ρ^δ introduced in Section 4 provided that all samples are computed from *one* precomputed sparse tensor gpc-representation of the forward map of the parametric, deterministic problem (16). We proved, in particular, that the forward map $\mathcal{G}(u)$ is obtained from *continuous, linear functionals $\mathcal{O}_i(\cdot)$ on the forward solution $U \ni u \mapsto P(\cdot, u) \in V$* allowing for a sparse approximate representation of gpc-type. Lower efficiency results if, for example, the rate of convergence of the procedure for computing the solution of the sparse tensor finite element solution in (41) is slow with respect to the total number of degrees of freedom, and/or if the complexity grows superlinearly with respect to the number of degrees of freedom.

Although an increasing number of algorithms for the efficient computation of approximate responses of the forward problem on the entire parameter space U are available (e.g. [3, 21, 11, 2, 4]) and therefore the gpc-MCMC is feasible, many systems of engineering interest do not readily admit gpc-based representations of the parametric forwards maps. Finding other non-gpc based methods for reducing complexity of 'direct' MCMC sampling under ρ^δ is therefore of interest. In this section, we give sufficient conditions on the data and on the ψ_j such that complexity reduction is possible by performing a multilevel sampling procedure where a number of samples depending on the discretization parameters are used for problem (12).

6.1. Derivation of the MLMCMC

Consider $\ell \in V^*$, i.e. a bounded linear functional on V. We aim at estimating $\mathbb{E}^{\rho^\delta}[\ell(P(u))]$ where P is the solution of problem (12). For each level $l = \lceil L/2 \rceil, \lceil L/2 \rceil +$

$1, \ldots, L$, we assume that problem (12) is discretized with the truncation of the Karhúnen-Loève expansion after J terms with $J = J_l$ as defined in (22) and with a finite element discretization mesh of width h_l. The multilevel FE-discretization of the forward problem (12) and the truncation (22) induces a corresponding hierarchy of approximations $\rho^{J,l,\delta}$ of the posterior measure ρ^{δ}.

Following [9, 1, 19, 10] the MLMCMC will be based on sampling a telescopic expansion of the discretization error with a level-dependent sample size.

We continue to work under Assumptions 6, 10, 11. We recall the sequence of discretization levels in the FE discretizations in D in Assumption 11, and the gpc input's truncation dimension J in Assumption 10. We then derive the Multilevel MCMC-FEM as follows. First, we note that there exists $C > 0$ independent of L such that

$$|\mathbb{E}^{\rho^\delta}[\ell(P(u))] - \mathbb{E}^{\rho^\delta}[\ell(P^{J_L,L}(u))]| \leq C \sup_{u \in U} \|P(u) - P^{J_L,L}(u)\|_V \leq C2^{-L} . \tag{47}$$

We then write

$$\mathbb{E}^{\rho^\delta}[\ell(P^{J_L,L})] = \mathbb{E}^{\rho^\delta}[\ell(P^{J_{\lceil L/2 \rceil},\lceil L/2 \rceil})] + \sum_{l=\lceil L/2 \rceil+1}^{L} \mathbb{E}^{\rho^\delta}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})] . \tag{48}$$

For $l = \lceil L/2 \rceil, \lceil L/2 \rceil + 1, \ldots, L$, let

$$\mathcal{G}^{J_l,l}(u) = \{\mathcal{O}_1(P^{J_l,l}(u)), \mathcal{O}_2(P^{J_l,l}(u)), \ldots, \mathcal{O}_d(P^{J_l,l}(u))\}$$

and

$$\Phi^{J_l,l} = \frac{1}{2}|\delta - \mathcal{G}^{J_l,l}(u)|^2.$$

We introduce, for each $l \in \mathbb{N}$, the Markov chains $\mathcal{C}_l = (u^{(k)})_k$ which are generated by (7) with the acceptance probability $\alpha(u,v)$ in (6) being replaced by

$$\alpha^{J_l,l}(u,v) = 1 \wedge \exp(\Phi^{J_l,l}(u;\delta) - \Phi^{J_l,l}(v;\delta)) , \quad (u,v) \in U \times U . \tag{49}$$

Then the chains $\mathcal{C}_l$ are pairwise uncorrelated.

For $M_l \in \mathbb{N}$, $\ell \in V^*$ and for a function $Q : U \to V$, we define the sample average with respect to the multilevel approximation of the Markov chain thus defined by

$$E_{M_l}^{\rho^{J_l,l,\delta}} \ell(Q) = \frac{1}{M_l} \sum_{k=1}^{M_l} \ell(Q(u^{(k)})) .$$

We denote by $\mathfrak{C}_L = \{\mathcal{C}_{\lceil L/2 \rceil}, \mathcal{C}_{\lceil L/2 \rceil+1}, \ldots, \mathcal{C}_L\}$; and denote by $\mathfrak{P}_L$ the product probability measure on the probability space that describes the law of $\mathfrak{C}_L$:

$$\mathfrak{P}_L := \mathcal{P}^{\rho,J_{\lceil L/2 \rceil},\lceil L/2 \rceil} \otimes \mathcal{P}^{\rho,J_{\lceil L/2 \rceil+1},\lceil L/2 \rceil+1} \otimes \ldots \otimes \mathcal{P}^{\rho,J_L,L} .$$

Let $\mathfrak{E}_L$ be the expectation with respect to $\mathfrak{C}_L$ under the probability $\mathfrak{P}_L$. We then approximate the right hand side of (48) by

$$T := E_M^{\rho^{J_{\lceil L/2 \rceil},\lceil L/2 \rceil},\delta}[\ell(P^{J_{\lceil L/2 \rceil},\lceil L/2 \rceil})] + \sum_{l=\lceil L/2 \rceil+1}^{L} E_{M_l}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l} - P^{J_{l-1},l-1})] , \tag{50}$$

where the number of samples M_l and M are to be determined.

Remark 31 *As shown below, our bound on the error of the approximation of* $\mathbb{E}^{\rho^\delta}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})]$ *by* $\mathbb{E}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})]$ *involves the term* 2^{-2l}. *Thus, to achieve an approximation error of* $O(2^{-L})$ *(ie. of the order of the discretization error in one instance of the forward problem) in the estimated expectation, we only perform the telescoping process from* $\lceil L/2 \rceil$.

6.2. Complexity Analysis of the MLMCMC-FEM

From Proposition 15 we deduce that there exists $C(\delta) > 0$ such that for all $l \in \mathbb{N}$

$$
\begin{aligned}
&\left| \mathbb{E}^{\rho^\delta}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})] - \mathbb{E}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})] \right| \\
&\leq C(\delta)(J_l^{-q} + 2^{-l}) \sup_{u \in U} |\ell(P^{J_l,l}(u)) - \ell(P^{J_{l-1},l-1}(u))| \\
&\leq C(\delta)(J_l^{-q} + 2^{-l})2^{-l} .
\end{aligned}
\tag{51}
$$

Following the procedure in the proof of Proposition 16, we have

$$
\begin{aligned}
&\mathcal{E}^{\rho,J_l,l}\left[\left| \mathbb{E}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})] - E_{M_l}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})] \right|^2\right] \\
&\leq M_l^{-1} \sup_u |\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})|^2 \\
&\leq C M_l^{-1} 2^{-2l} .
\end{aligned}
\tag{52}
$$

Similarly, we have that

$$
\left| \mathbb{E}^{\rho^\delta}[\ell(P^{J_{\lceil L/2\rceil},\lceil L/2\rceil})] - \mathbb{E}^{\rho^{J_{\lceil L/2\rceil},\lceil L/2\rceil,\delta}}[\ell(P^{J_{\lceil L/2\rceil},\lceil L/2\rceil})] \right| \leq C(\delta)(J_{\lceil L/2\rceil}^{-q} + 2^{-L/2})2^{-L/2}, \tag{53}
$$

and

$$
\begin{aligned}
&\mathcal{E}^{\rho,J_{\lceil L/2\rceil},\lceil L/2\rceil}\left[\left| \mathbb{E}^{\rho^{J_{\lceil L/2\rceil},\lceil L/2\rceil,\delta}}[\ell(P^{J_{\lceil L/2\rceil},\lceil L/2\rceil})] - E_M^{\rho^{J_{\lceil L/2\rceil},\lceil L/2\rceil,\delta}}[\ell(P^{J_{\lceil L/2\rceil},\lceil L/2\rceil})] \right|^2\right] \\
&\leq C M^{-1} .
\end{aligned}
\tag{54}
$$

From equations (47), (48), (51), (52), (53) and (54) with the Cauchy Schwartz inequality that there exists a constant $C > 0$ such that, for any $L \in \mathbb{N}$, and for any choice $\{J_l\}_{l=0}^L$ and $M \in \mathbb{N}$, there holds

$$
\begin{aligned}
\left(\mathfrak{E}_L[|\mathbb{E}^{\rho^\delta}[\ell(P)] - T|^2] \right)^{1/2} &\leq C2^{-L} + C(\delta)(J_{\lceil L/2\rceil}^{-q} + 2^{-L/2})2^{-L/2} + CM^{-1/2} \\
&+ CL^{1/2} \sum_{l=\lceil L/2\rceil+1}^{L} (J_l^{-q} + 2^{-l} + M_l^{-1/2})2^{-l} .
\end{aligned}
$$

Choosing here $J_l = \lceil 2^{l/q}\rceil$, $M_l = 2^{2(L-l)}$, $J_{\lceil L/2\rceil} = 2^{L/(2q)}$ and $M = 2^{2L}$, we then find

$$
\left(\mathfrak{E}_L[|\mathbb{E}^{\rho^\delta}[\ell(P)] - T|^2] \right)^{1/2} \leq C(\delta)L^{3/2}2^{-L} .
$$

As $l \to \infty$, the number of degrees of freedom used for computing $P^{J_l,l} - P^{J_{l-1},l-1}$ is $O(2^{dl})$ for a single sample of u.

The total number of degrees of freedom for computing $E_{M_l}^{\rho^{J_l,l,\delta}}[\ell(P^{J_l,l}) - \ell(P^{J_{l-1},l-1})]$ behaves, asymptotically, as $l \to \infty$, as $O(M_l 2^{dl})$. Likewise, the total number of degrees of

freedom required for computing $E_M^{\rho^{J_{\lceil L/2 \rceil}, \lceil L/2 \rceil, \delta}}[\ell(P^{J_{\lceil L/2 \rceil}, \lceil L/2 \rceil})]$ is $O(M2^{dL/2})$ as $L \to \infty$. The total number of degrees of freedom is therefore not larger than

$$\sum_{l=\lceil L/2 \rceil+1}^{L} M_l 2^{dl} + M2^{dL/2} \lesssim \sum_{l=\lceil L/2 \rceil+1}^{L} 2^{2(L-l)}2^{dl} + 2^{2L}2^{dL/2} \lesssim 2^{(2+d/2)L} . \tag{55}$$

As in Section 4, to form the stiffness matrix to compute the Finite Element approximation $P^{J_l,l}$ of P requires, for each realization of u, not more than $O(2^{dl}l^{d-1})J_l = O(l^{d-1}2^{dl+l/q})$ float point operations. The number of floating point operations required for the computation of the Finite Element approximation $P^{J_{\lceil L/2 \rceil}, \lceil L/2 \rceil}$ is not larger than $O(L^{d-1}2^{(dL/2+L/(2q))})$. The total number of floating point operations required is asymptotically, as $L \to \infty$, bounded by

$$\sum_{l=\lceil L/2 \rceil+1}^{L} 2^{2(L-l)}l^{d-1}2^{dl+l/q} + 2^{2L}L^{d-1}2^{(dL/2+L/(2q))} \simeq L^{d-1}(2^{dL+L/q} + 2^{(d/2+1/(2q)+2)L}) .$$

Setting $N = 2^{dL}$, we have thus shown the following result.

Theorem 32 *The expectation $\mathbb{E}^{\rho^\delta}[\ell(P)]$ can be approximated by Multi-Level MCMC FEM based on a continuous, piecewise linear FEM on a family of quasiuniform, shape-regular triangulations of meshwidth h in D to a mean-square error $O(h) = O(N^{-1/d}(\log N)^{3/2})$ using a total of $O(N^{1/2+2/d})$ degrees of freedom and a total of $O((\log N)^{d-1}N^{\max(1/2+1/(2dq)+2/d,1+1/(dq))})$ floating point operations.*

Remark 33 *For the gpc-MLMCMC method, when using N_{fp} floating point operations and ignoring logarithmic terms in the cost estimates, the error can be written as*

$$O(N_{fp}^{-\min\{1/(d+1/q),1/(d/2+2+1/(2q))\}}) .$$

This rate is always superior to the rate $N_{fp}^{1/(d+2+1/q)}$ for the plain MCMC method which we found in Section 4.

    Comparing to the sparse gpc-MCMC method in section 5, assuming that we have the optimal rate $N_{\text{dof}}^{-1/d}$ i.e. $\tau = 1/d$ in Assumption 21, when $\alpha > 1/2+2/d$ (assuming that $q \sim s-1$ is large), the method presented in this section is superior in terms of accuracy versus complexity.

    Although we only presented the MLMCMC approach for the particular elliptic problem (12) with a special distribution (13), it is expected that for other distributions of the coefficient K, when Assumption 21 does not hold, or hold with a constant α that is not close to 1, the MLMCMC approach is a superior alternative method than the plain MCMC to compute the expectation of a function $g(u) = \ell(P(u))$ where $\ell \in V^*$.

7. Conclusions

We note that in [22] an entirely deterministic approach to the solution of the Bayesian inverse problem is presented. However it is to be expected that the methods presented herein will be superior in practice, for some problems, because of the ability of MCMC

based methods to sample measures which concentrate on small parts of the space. A detailed comparison of the computational performance of the gpc-accelerated MLMCMC methods of this paper with the deterministic approach in [22] will provide useful information about their relative merits.

We also observe that we have concentrated on a very special MCMC method, namely the independence sampler. This will work well when the negative log likelihood Φ does not vary too much, but will be inefficient in general. More appropriate MCMC methods may be found in [7]. However for these more general methods the analysis of the Markov chain based on the methods of [18] are not appropriate and more sophisticated arguments are required, as presented in [12].

Acknowledgments

References

[1] Andrea Barth, Christoph Schwab, and N. Zollinger. Multi-Level Monte Carlo Finite Element Method for elliptic PDEs with stochastic coefficients. *Numerische Mathematik*, 119(1):123–161, 2011.

[2] Marcel Bieri. A sparse composite collocation finite element method for elliptic SPDEs. *SIAM J. Numer. Anal.*, 49(6):2277–2301, 2011.

[3] Marcel Bieri, Roman Andreev, and Christoph Schwab. Sparse tensor discretization of elliptic sPDEs. *SIAM J. Sci. Computing*, 31(6):4281–4304, 2009.

[4] Abdellah Chkifa, Albert Cohen, Ron DeVore, and Christoph Schwab. Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic pdes. Technical report.

[5] Albert Cohen, Ronald DeVore, and Christoph Schwab. Convergence rates of best n-term Galerkin approximations for a class of elliptic spdes. *Journ. Found. Comp. Math.*, 10(6):615–646, 2010.

[6] Albert Cohen, Ronald Devore, and Christoph Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Anal. Appl. (Singap.)*, 9(1):11–47, 2011.

[7] S. Cotter, G. Roberts, A. Stuart, and D. White. MCMC methods for functions: modifying old algorithms to make them faster. *Arxiv preprint arXiv:1202.0709*, 2012.

[8] S. L. Cotter, M. Dashti, J. C. Robinson, and Andrew M. Stuart. Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse problems*, 25, 2009.

[9] Michael B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[10] C.J. Gittelson, J. Könnö, Ch. Schwab, and R. Stenberg. The multi-level monte carlo finite element method for a stochastic brinkman problem. Technical Report 2011-31, Seminar for Applied Mathematics, ETH Zürich, 2011.

[11] Claude Jeffrey Gittelson. *Adaptive Galerkin Methods for Parametric and Stochastic Operator Equations*. PhD thesis, ETH Zürich, 2011. ETH Dissertation No. 19533.

[12] M. Hairer, A. Stuart, and S. Vollmer. Spectral gaps for a Metropolis-Hastings algorithm in infinite dimensions. *Arxiv preprint arXiv:1112.1392*, 2011.

[13] Viet Ha Hoang and Christoph Schwab. Sparse Tensor Galerkin discretizations for parametric and

random parabolic pdes I: Analytic regularity and gpc approximation. *Report 2010-10, Seminar for Applied Mathematics, ETH Zürich*, ((in review)), 2010.

[14] Viet Ha Hoang and Christoph Schwab. Regularity and generalized polynomial chaos approximation of parametric and random 2nd order hyperbolic partial differential equation. *Analysis and Applications*, 2012.

[15] Y. Marzouk and D. Xiu. A stochastic collocation approach to Bayesian inference in inverse problems. 2009.

[16] Y.M. Marzouk and H.N. Najm. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228(6):1862–1902, 2009.

[17] Y.M. Marzouk, H.N. Najm, and L.A. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, 224(2):560–586, 2007.

[18] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[19] Siddhartha Mishra and Christoph Schwab. Sparse tensor multi-level monte carlo finite volume methods for hyperbolic conservation laws with random intitial data. *Report 2010-24, Seminar for Applied Mathematics, ETH Zürich*, ((to appear in Mathematics of Computation (2012))), 2010.

[20] Hoang Nguyen and Rob Stevenson. Finite-Element wavelets on manifolds. *IMA J. Numer. Anal.*, 23(1):149–173, 2003.

[21] Christoph Schwab and Claude Jeffrey Gittelson. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numerica*, 20:291–467, 2011.

[22] Christoph Schwab and A. M. Stuart. Sparse deterministic approximation of Bayesian inverse problems. *Inverse Problems*, 28(4), 2012.

[23] Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 2010.

[24] L. Tierney. A note on metropolis-hastings kernels for general state spaces. *Ann. Appl. Prob.*, 8:1–9, 1998.

Research Reports