Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

Convergence rate analysis of MCMC-FEM for Bayesian inversion of log-normal diffusion problems

V.H. Hoang and Ch. Schwab

Research Report No. 2016-19
March 2016

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

CONVERGENCE RATE ANALYSIS OF MCMC-FEM FOR BAYESIAN INVERSION OF LOG-NORMAL DIFFUSION PROBLEMS

13 MARCH 2016

VIET HA HOANG[1] AND CHRISTOPH SCHWAB[2]

ABSTRACT. Markov Chain Monte Carlo (MCMC) methods for the numerical solution of Bayesian Inverse Problems for linear second order, divergence form elliptic partial differential equations (PDEs) with lognormal random field coefficients are analyzed. The analysis of the MCMC Finite Element discretization for uniformly elliptic, random diffusion problems of [14] is extended. The complexity of MCMC sampling for the uncertain input fields from the posterior density, as well as the MCMC error due to discretization of the PDE of interest in the forward response map, are analyzed in the abstract framework of MCMC methods of Meyn and Tweedie [16]. Particular attention is given to bounds on the overall work required by the MCMC algorithms for achieving a prescribed error level $\varepsilon > 0$. We prove convergence rate estimates and bound the computational complexity of straightforward combinations of MCMC sampling strategies with Finite Element approximation of solution of the forward PDE. Due to the non-uniform ellipticity, the computational complexity analysis of the MCMC-FEM is probabilistic.

1. INTRODUCTION

We consider the inverse problem of Bayesian inference for a quantity of interest (QoI) from a linear, second order elliptic diffusion problem with unknown, lognormal random field K ie. $\log K$ is a gaussian random field in a Banach space X which we assume parametrized by a sequence of parameters u from a measurable space (U, Θ). We assume finite dimensional, noisy observation data to be given by

$$\delta = \mathcal{G}(u) + \vartheta \in \mathbb{R}^k \ . \tag{1.1}$$

Here $U \ni u \mapsto \mathcal{G}(u)$ is the forward functional (or the system response) which is assumed to be well-defined for each instance of the unknown datum u in the set U of admissible parameters. In writing (1.1), we suppose that the additive observation noise ϑ is a centred Gaussian random variable $N(0, \Sigma)$ with known, positive (co)variance $\Sigma \in \mathbb{R}^{k\times k}_{sym} > 0$.

Assuming a Gaussian prior measure γ on a measurable space U of admissible inputs, the posterior probability γ^δ given the data δ, ie. the conditional probability of the uncertain input $u \in U$ subject to given observation data $\delta \in \mathbb{R}^k$, is determined according to Bayes' theorem as (see, eg., [20])

$$\frac{d\gamma^\delta}{d\gamma} \propto \exp(-\Phi(u;\delta)), \tag{1.2}$$

where the Bayesian potential (or "misfit functional") is given by

$$\Phi(u;\delta) = \frac{1}{2}|\delta - \mathcal{G}(u)|_\Sigma^2 = \frac{1}{2}(\delta - \mathcal{G}(u))^\top \Sigma^{-1}(\delta - \mathcal{G}(u)) \,. \tag{1.3}$$

We consider the problem where $\mathcal{G}(u)$ is a linear functional of the solution of a so-called "forward" linear, elliptic diffusion problem with unknown diffusion coefficient K. Generalizing our previous analysis in [14], and in line with modelling practice in the computational geosciences, we consider in the present paper the unknown diffusion coefficient K to be a lognormal random field in the physical domain D, ie., $\log K$ is a Gaussian random field in a bounded Lipschitz domain $D \subset \mathbb{R}^d$. Gaussian random fields are determined by their mean and covariance (see, eg., [1, Sec. 3.3]). Determining the Gaussian random field $\log K$ taking values in a separable Banach space X of functions in D, given noisy observation data $\delta \in \mathbb{R}^k$ is, therefore, equivalent to determining the mean and covariance through the sequence u of normally distributed, random coordinates in a basis expansion, such as the Karhúnen-Loève expansion (in the sequel, we consider the inverse problem of predicting a Quantity of Interest (QoI) as mathematical expectation over all possible realizations of the uncertain input u, conditional on noisy measurements). As these coordinates enter, in turn, the forward problems as parameters, the key technical issue in the present paper, as compared to [14], is the analysis of Bayesian estimation in the presence of parametric forward PDE problems and their discretizations on unbounded parameter domains.

Recently there have been several attempts to numerically sample posterior probability measures of Bayesian inverse problems for inferring log-normal random fields by MCMC methods (e.g., [3, 6] and the survey [4]). However, rigorous analysis of the mathematical properties and of the combined effect of the statistical error of the MCMC approximation, and of the discretization and modelling errors of the forward problem, on the numerically realized approximations of the Bayesian posterior measure, does not seem to be available. The case of bounded coefficients was studied in detail in [14]. In this paper we develop a complete analysis on the mathematical properties, and of the discretization and statistical errors as well as the computational complexity of the MCMC process for the case of unbounded coefficients, in particular for Bayesian inference of log-normal random fields. As in [14], our analysis requires some notation for the probability of the Markov chains for sampling the exact and the approximated posterior probability measures which, in turn, depends on the discretization parameters: the truncation order J of the Karhúnen-Loève expansion and the level l of mesh refinement of the FE discretization. We denote by $\mathcal{P}_{u^{(0)}}$ the probability measure on the probability space generated by the MCMC processes starting at the initial point $u^{(0)}$ with the acceptance probability for the Metropolis-Hastings Markov chain defined as α in (2.1). Correspondingly, we denote by $\mathcal{P}^{J,l}_{u^{(0)}}$ the probability measure on the probability space generated by the MCMC process starting at $u^{(0)}$ with the acceptance probability $\alpha^{J,l}$ in (4.22) for the approximated problem with a J-term truncation of the Karhúnen-Loève expansion and with l denoting the mesh level of the finite element discretization. We denote by $\mathcal{E}_{u^{(0)}}$ and $\mathcal{E}^{J,l}_{u^{(0)}}$ the mathematical expectations with respect to the probability measures $\mathcal{P}_{u^{(0)}}$ and $\mathcal{P}^{J,l}_{u^{(0)}}$. When the initial state $u^{(0)}$ of the Markov chain is distributed according to a probability measure μ on U, we denote the expectation accordingly as $\mathcal{E}^\mu$ and $\mathcal{E}^{\mu,J,l}$.

The outline of this paper is as follows: in Section 2, we present a general framework of Bayesian inverse problems in function spaces, following [20], [16] with particular emphasis on technical aspects arising in the lognormal setting. In Section 3, we present a concrete class of elliptic PDEs where the coefficient K is an unknown lognormal stochastic diffusion coefficient. Section 4 is devoted to the convergence and complexity analysis of the MCMC for this class of problems. Appendices A and B contain the proof of Proposition 4.7 and the derivation of several estimates used in the paper. Throughout this paper, by c, C and Λ we denote various constants which do not depend on the discretization parameters, whose values can vary from one line to the next.

2. Bayesian inverse problems in measure spaces for unbounded functionals

We consider abstract Bayesian inverse problems set on a measurable space (U, Θ) where Θ is a σ-algebra such as the Borel σ-algebra on U. Let the system response for given input $u \in U$, $\mathcal{G} : (U, \Theta) \to (\mathbb{R}^k, \mathcal{B}^k)$, be strongly measurable. We assume available observation data δ of $\mathcal{G}$ which is subject to an unbiased, additive Gaussian observation noise $\vartheta \in \mathbb{R}^k$:

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

The noise ϑ is assumed centred Gaussian with law $N(0, \Sigma)$ where $\Sigma \in \mathbb{R}^{k\times k}_{sym}$ is positive definite. Assuming given a prior probability measure γ on (U, Θ), our purpose is to numerically sample from the conditional probability $\gamma^\delta = \gamma(u|\delta)$ on (U, Θ). As each draw involves solving the partial differential equation, one purpose is to assess the impact of Finite Element (FE) discretization of the data-to-observation map $\mathcal{G}(\cdot)$ on the accuracy of the samples generated by the MCMC approach, and on the overall convergence of the MCMC-FE approximation. Numerical approximations for the MCMC approach to sample the posterior measure γ^δ of this problem have been analyzed in [14] in the particular case where the forward functional $\mathcal{G}(u)$ is essentially bounded (with respect to the measure γ) for $u \in U$. In the present work we consider the case where the observation $\mathcal{G}(u)$ is possibly unbounded, as a function of the parameter sequence u. This is the case, for example, when $\mathcal{G}$ is a linear functional of the solution of an elliptic partial differential equation with a log-normal isotropic random field. The existence of the posterior distribution is proved in [3]; and the well-posedness of the posterior distribution of these problems is proved in [13] and [14], generalizing the arguments of [3]. The next result was shown in Cotter et al. [3, Theorem 2.1].

Proposition 2.1. *Assume that $\mathcal{G} : U \to \mathbb{R}^k$ is measurable. The posterior measure $\gamma^\delta(du) = \gamma(du|\delta)$ given data δ is absolutely continuous with respect to the prior measure $\gamma(du)$ and has the Radon-Nikodym derivative* (1.2) *with Φ given by* (1.3).

For the well-posedness of the posterior measures, we show that the Hellinger distance of the posterior measure corresponding to data δ is locally Lipschitz continuous with respect to δ. In order for this to hold, we need assumptions on the boundedness of the potential function $\Phi(u; \delta)$ and its Lipschitz dependence on the data δ.

Assumption 2.2. *Let γ be a probability measure on the measurable space (U, Θ). The potential $\Phi : U \times \mathbb{R}^k \to \mathbb{R}$ satisfies:*

(i) *For each $\lambda > 0$, there is a constant $\Lambda(\lambda) > 0$ such that if $|\delta| < \lambda$,*

$$\int_U \Phi(u;\delta) d\gamma(u) < \Lambda(\lambda) \; .$$

Here, and throughout the following, $|\circ|$ denotes the euclidean norm on $\mathbb{R}^k$.

(ii) *There exists a mapping $G : \mathbb{R}_{>0} \times U \mapsto \mathbb{R}$ such that for each $\lambda > 0$, $G(\lambda, \cdot) \in L^2(U,\gamma)$ and for each $\delta, \delta' \in \mathbb{R}^k$ with $|\delta|, |\delta'| \le \lambda$ holds*

$$|\Phi(u;\delta) - \Phi(u;\delta')| \le G(\lambda, u)|\delta - \delta'| \; .$$

Under Assumption 2.2, the posterior measure as defined in (1.2) depends continuously on the data δ.

Proposition 2.3. *Under Assumption 2.2, the measure γ^δ depends locally Lipschitz continuously on the data δ with respect to the Hellinger distance: for each positive constant λ there is a positive constant $C(\lambda)$ such that if $|\delta|$, $|\delta'| \le \lambda$, then*

$$d_{\text{Hell}}(\gamma^\delta, \gamma^{\delta'}) \le C(\lambda)|\delta - \delta'| \; .$$

Proposition 2.3 is obtained from Assumption 2.2 by a similar argument as for the results in [13, Theorem 2.4] and of [14, Proposition 25], using the uniform positiveness of the normalizing constant in (1.2) which is shown in the proof of Proposition 4.6 below.

To sample from the posterior measure γ^δ, we employ a Metropolis-Hastings MCMC method designed to be reversible and ergodic with respect to the posterior measure γ^δ: to this end, given data δ, we define for any $u, v \in U$

$$\alpha(u,v) = 1 \wedge \exp\big(\Phi(u;\delta) - \Phi(v;\delta)\big) \; . \tag{2.1}$$

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 γ appearing in (1.2). Let $\{w^{(k)}\}_{k\ge1}$ 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

$$u^{(k+1)} = \mathbf{1}\big(\alpha(u^{(k)}, v^{(k)}) \ge w^{(k)}\big)v^{(k)} + \Big(1 - \mathbf{1}\big(\alpha(u^{(k)}, v^{(k)}) \ge w^{(k)}\big)\Big)u^{(k)} \; . \tag{2.2}$$

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 (2.2) defines a Markov chain $\{u^{(k)}\}_{k=0}^\infty$ which is reversible with respect to γ^δ. To see this let $\nu(du, dv)$ denote the product measure $\gamma^\delta(du) \otimes \gamma(dv)$ and define $\nu^\dagger(du, dv) = \nu(dv, du)$. Note that ν describes the probability distribution of the pair $(u^{(k)}, v^{(k)})$ on $U \times U$ when $u^{(k)}$ is drawn from γ^δ, and $\nu^\dagger$ designates the same measure with the roles of u and v reversed. These two measures are equivalent (as measures on $(U \times U, \Theta \otimes \Theta)$; in case Θ is the Borel σ-algebra in a topological space U, the product σ-algebra becomes $\mathcal{B}(U \times U)$) if γ^δ and γ are equivalent. Then

$$\frac{d\nu^\dagger}{d\nu}(u,v) = \exp\big(\Phi(u;\delta) - \Phi(v;\delta)\big) \; , \quad (u,v) \in U \times U \; . \tag{2.3}$$

From Proposition 1 and Theorem 2 in [22] we deduce that (2.2) is a Metropolis-Hastings Markov chain which is γ^δ reversible, since $\alpha(u,v)$ given by (2.1) 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*. In [14], we studied ergodicity of this chain under the stronger condition that, for given data δ, the function $\Phi(u;\delta)$

is uniformly bounded from above with respect to $u \in U$. When the forward map $U \ni u \mapsto \mathcal{G}(u)$ is not uniformly bounded with respect to the parameter sequence u (as will be the case in the presently considered applications with unknown lognormal random field inputs), this does not hold and results of [14] do not apply directly. We therefore develop generalizations of [14] to the case where $\mathcal{G}(u)$ may be unbounded.

Theorem 2.4. *Let Assumption 2.2 hold. Let further $\mathcal{V}(u) \geq 1$ be a function in $L^2(U, \gamma)$.* [1] *Then γ^δ is equivalent to γ. In particular, the Markov chain* (2.2) *is well-defined and reversible with respect to γ^δ.*

Let $p(u, \cdot)$ denote the transition kernel for the Markov chain, and $p^n(u, \cdot)$ its n^{th} iterate. Then there are constants $r > 1$ and $R > 0$ so that for all $u \in U$ holds geometric ergodicity

$$\sum_{n=1}^{\infty} r^n \|p^n(u, \cdot) - \gamma^\delta\|_{\mathcal{V}} \leq R\mathcal{V}(u) \, . \tag{2.4}$$

For $g : U \to \mathbb{R}$ such that $(g(u))^2 \leq \mathcal{V}(u)$ for all $u \in U$, and $\mathcal{P}_{u^{(0)}}$ almost surely,

$$\frac{1}{M}\sum_{k=1}^{M} g(u^{(k)}) = \mathbb{E}^{\gamma^\delta}[g(u)] + c\xi_M M^{-\frac{1}{2}} \tag{2.5}$$

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 R, r and on $\mathcal{V}(\cdot)$. Furthermore, when $|g(u)| \leq \mathcal{V}(u)$ for all $u \in U$, there holds the mean square error bound

$$\left(\mathcal{E}^{\gamma^\delta}\left[\left|\mathbb{E}^{\gamma^\delta}[g(u)] - \frac{1}{M}\sum_{k=1}^{M} g(u^{(k)})\right|^2\right]\right)^{1/2} \leq CM^{-1/2} \, . \tag{2.6}$$

Proof. For $A \in \Theta$,

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

so if $\gamma(A) = 0$ then $\gamma^\delta(A) = 0$. Now, we show that if $\gamma^\delta(A) = 0$, then $\gamma(A) = 0$. To this end, we assume the contrary, ie. that $\gamma^\delta(A) = 0$ while $\gamma(A) > 0$. As $\int_A \Phi(u; \delta)d\gamma(u) < c$ we may fix a constant Λ such that $\gamma\left(\{u \in A : \Phi(u; \delta) > \Lambda\}\right) < c/\Lambda$. Choosing Λ sufficiently large, we also have $\gamma\left(\{u \in A : \Phi(u; \delta) < \Lambda\}\right) > \gamma(A) - c/\Lambda > 0$. From this, we deduce that there exists a constant $c > 0$ such that

$$\gamma^\delta(A) = Z^{-1}\int_A \exp(-\Phi(u; \delta))d\gamma(u) > c(\gamma(A) - c/\Lambda)\exp(-\Lambda) > 0 \, .$$

Thus $\gamma(A) = 0$. The measures γ and γ^δ are equivalent. We remark that instead of the preceding argument, we may refer to Remark 1 at the bottom of page 3 in the paper by Tierney [22]: his result requires the density functions of the target measure with respect to a measure ν and of the proposal density both to be positive. This holds in the presently considered case for the choice $\nu = \gamma$ with the positive density in (1.2).

From (2.1) it also follows that the proposed random draw from γ has acceptance probability at least $\exp(-\Phi(v; \delta))$ and there holds the *minorization condition*

$$\forall\, u \in U : \quad p(u, A) \geq \int_A \exp(-\Phi(v; \delta))d\gamma(v) \, . \tag{2.7}$$

[1] $\mathcal{V}(u)$ corresponds to the comparison function $V(x)$ in Sections 14 and 15 of [16].

For $\mathcal{V} \in L^2(U,\gamma)$ and $\mathcal{V}(u) \geq 1$ for all $u \in \Gamma$,

$$\begin{aligned}\int_U p(u,dv)\mathcal{V}(v) &= \int_U \alpha(u,v)\mathcal{V}(v)d\gamma(v) + (1 - \int_U \alpha(u,v)d\gamma(v))\mathcal{V}(u) \\ &\leq \int_U \mathcal{V}(v)d\gamma(v) + (1 - \int_U \exp(-\Phi(v;\delta))d\gamma(v))\mathcal{V}(u)\ .\end{aligned}$$

From Assumption 2.2, there exists a constant $\Lambda > 0$ so that

$$\gamma(\{v:\ \Phi(v;\delta) < \Lambda\}) > 0\ .$$

Thus there exists $c_0 > 0$ so that

$$1 \geq \int_U \exp(-\Phi(v;\delta))d\gamma(v) > c_0\ . \tag{2.8}$$

Therefore

$$\int_U p(u,dv)\mathcal{V}(v) \leq \int_U \mathcal{V}(v)d\gamma(v) + (1-c_0)\mathcal{V}(u).$$

Let $\hat\beta > \int_U \mathcal{V}(v)d\gamma(v)$ and $0 < \check\beta < c_0$. Then, the equation

$$\triangle\mathcal{V}(u) := \int_U p(u,dv)\mathcal{V}(v) - \mathcal{V}(u) \leq -\check\beta\mathcal{V}(u) + \hat\beta$$

is satisfied, where $\triangle\mathcal{V}(u)$ denotes the drift operator of Meyn and Tweedie [16, Eqn. (8.1)]. Therefore condition (iii) of Theorem 15.0.1 in [16] holds with the "petite" set U and we obtain (2.4).

As $g^2 \leq \mathcal{V}$, the second result follows from [16], Theorem 17.0.1. To see that the constant c in (2.5) is bounded by a constant that depends only on r, R and $\mathcal{V}(u)$ we note that it is given by

$$c^2 = \mathcal{E}^{\gamma^\delta}|\bar g(u^{(0)})|^2 + 2\sum_{n=1}^\infty \mathcal{E}^{\gamma^\delta}[\bar g(u^{(0)})\bar g(u^{(n)})] \tag{2.9}$$

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

$$\begin{aligned}2\sum_{n=1}^\infty \mathcal{E}^{\gamma^\delta}[\bar g(u^{(0)})\bar g(u^{(n)})] &\leq 2\mathbb{E}^{\gamma^\delta}\left[|\bar g(u^{(0)})|\sum_{n=1}^\infty |\mathcal{E}_{u^{(0)}}[\bar g(u^{(n)})]|\right] \\ &\leq 2\mathbb{E}^{\gamma^\delta}\left[|\bar g(u^{(0)})|\sum_{n=1}^\infty |\mathcal{E}_{u^{(0)}}[g(u^{(n)})] - \mathbb{E}^{\gamma^\delta}[g]|\right] \\ &\leq 2R\mathbb{E}^{\gamma^\delta}\left[|\bar g(u^{(0)})|\mathcal{V}(u^{(0)})\right]\sum_{n=0}^\infty r^{-n}\ .\end{aligned}$$

As $g^2(u) \leq \mathcal{V}(u)$, and $\mathcal{V}(u) \geq 1$, we have that $g(u) \leq \mathcal{V}(u)$. This expression is finite as $\mathcal{V} \in L^2(U,\gamma)$. For the mean square approximation, using the stationarity of the

Markov chain conditioned to start in $U \ni u^{(0)} \sim \gamma^\delta$, we have

$$\begin{aligned}
\frac{1}{M}\mathcal{E}^{\gamma^\delta}\Big[\Big|\sum_{k=1}^{M}\bar{g}(u^{(k)})\Big|^2\Big] &= \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=1}^{M}\sum_{j=k+1}^{M}\mathcal{E}^{\gamma^\delta}[\bar{g}(u^{(k)})\bar{g}(u^{(j)})] \\
&= \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathcal{E}^{\gamma^\delta}[\bar{g}(u^{(0)})\bar{g}(u^{(j)})] \\
&= \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})\mathcal{E}_{u^{(0)}}[\bar{g}(u^{(j)})]] \\
&\le \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] \\
&\quad + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\gamma^\delta}[|\bar{g}(u^{(0)})||\mathcal{E}_{u^{(0)}}[g(u^{(j)})] - \mathbb{E}^{\gamma^\delta}[g]|] \\
&\le \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] + 2R\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k} r^{-j}\mathbb{E}^{\gamma^\delta}[|\bar{g}(u^{(0)})|\mathcal{V}(u^{(0)})] \\
&\le \mathbb{E}^{\gamma^\delta}[\bar{g}(u^{(0)})^2] + 2R\sum_{j=1}^{\infty} r^{-j}\mathbb{E}^{\gamma^\delta}[|\bar{g}(u^{(0)})|\mathcal{V}(u^{(0)})],
\end{aligned}$$

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

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

□

Remark 2.5. *The minorization condition* (2.7) *is essentially the condition that* $w^* < \infty$ *in Roberts and Rosenthal* [18] *resp. condition (19) of Mengersen and Tweedie* [15]*. However, for independence samplers considered in this paper, Roberts and Rosenthal* [18] *and Mengersen and Tweedie* [15] *only consider geometric ergodicity for the case where the function* $\mathcal{V}$ *in* (2.4) *equals 1. For the more general* $\mathcal{V}$ *under consideration here, Mengersen and Tweedie* [15] *consider the case where the transition kernel of the proposals is symmetric which is not our interest here.*

Remark 2.6. *From the case* $m = 1$ *in* [16, Theorem 16.2.4]*, with the minorization condition* (2.7)*, we deduce that given data* δ*, there exists a constant* $0 < \rho < 1$ *such that, for every* n*,*

$$\forall u \in U: \ \ \|p^n(u,\cdot) - \gamma^\delta\|_{\mathrm{TV}} \le 2\rho^n\ . \tag{2.10}$$

Using this, the convergence results in Theorem 2.4 will hold for bounded functions g *without using the geometric ergodicity property* (2.4)*. The main purpose of this paper is to prove convergence rates and accuracy versus work estimates for MCMC sampling with Finite Element discretization, for the computation of the expectation with respect to the posterior measure of linear functionals of the solution to partial differential equations with log-normal coefficients, which are unbounded with respect to the i.i.d. sequence of standard normal random variables* $u = (u_1, u_2, \ldots)$ *in the expansion of the coefficient in* (3.1) *ahead, and to verify* (2.4) *with* $r > 1$ *and* $R > 0$ *independent of the discretization parameters.*

3. Elliptic equations with log-normal random coefficients

We present the countably-parametric, deterministic forward map $\mathcal{G}(\cdot)$ and recapitulate basic results on its measurability under the Gaussian prior γ.

3.1. Elliptic problems with log-normal coefficients.

We present the setting up of the problem with log-normal coefficients. Let $D \subset \mathbb{R}^d$ be an open domain. In this section, by $\mathbb{R}^{\mathbb{N}}$ we denote the set of all infinite sequences $(u_1, u_2, \ldots)$ of real numbers. We consider the parametric, deterministic coefficient $K : D \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ which is *formally* (ie., ignoring for now questions of convergence) given, for $u = (u_1, u_2, \ldots) \in \mathbb{R}^{\mathbb{N}}$, by

$$K(\cdot, u) = K_*(\cdot) + \exp\left(\bar{K}(\cdot) + \sum_{j=1}^{\infty} \psi_j(\cdot) u_j\right) . \tag{3.1}$$

In (3.1), the sequence $\{\psi_j\}_{j\geq 1}$ is assumed to be a Schauder basis of a suitable separable (subspace of a) Banach space $X \subset L^\infty(D)$ of uncertain input data on the domain D. Examples for the series expansions in (3.1) are furnished, in particular, by Karhúnen-Loève expansions of gaussian random fields $\log(K(\cdot, u) - K_*(\cdot, u))$ in the bounded domain D. Any representation of (3.1) is ambiguous as long as no assumption on the scaling of the $\|\psi_j\|_X$ is specified. We model the uncertainty in the coefficient $K(\cdot, u)$ by placing a Gaussian measure on X. To this end, we assume the coordinates u_j to be independently, identically standard normally distributed random variables, i.e., $u_j \sim \mathcal{N}(0, 1)$. With γ_1 denoting the standard Gaussian measure in $\mathbb{R}^1$, we define the product measure

$$\gamma = \bigotimes_{j=1}^{\infty} \gamma_1 \tag{3.2}$$

as *Gaussian prior probability measure* on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. We shall work under the following assumption on the coefficients K_*, $\bar{K}$ and ψ_i.

Assumption 3.1. *The functions $\bar{K}$, K_* and ψ_j in* (3.1) *are in $L^\infty(D)$ and* $0 \leq \operatorname{essinf} K_*(x) \leq \operatorname{esssup} K_*(x) < \infty$. *Moreover,* $\boldsymbol{b} := (\|\psi_j\|_{L^\infty(D)})_{j\geq 1} \in \ell^1(\mathbb{N})$.

With $b_j := \|\psi_j\|_{L^\infty(D)}$, Assumption 3.1 implies $\boldsymbol{b} \in \ell^2(\mathbb{N})$. Then, the set

$$\Gamma_{\boldsymbol{b}} := \{u \in \mathbb{R}^{\mathbb{N}}, \ \sum_{j=1}^{\infty} b_j |u_j| < \infty\} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) \tag{3.3}$$

has full Gaussian measure, i.e. $\gamma(\Gamma_{\boldsymbol{b}}) = 1$ (see, e.g., [24, p. 153] or [19, Lemma 2.28]).

Let $\mathcal{A}_{\boldsymbol{b}} \subset \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ denote the sub σ-algebra which is obtained as restriction of $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ to $\Gamma_{\boldsymbol{b}} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$, and let $\gamma_{\boldsymbol{b}}$ denote the restriction of the gaussian measure γ to $(\mathcal{A}_{\boldsymbol{b}}, \Gamma_{\boldsymbol{b}})$:

$$\gamma_{\boldsymbol{b}} := \gamma\,|_{(\mathcal{A}_{\boldsymbol{b}}, \Gamma_{\boldsymbol{b}})} \ . \tag{3.4}$$

For proposals $u \in \Gamma_{\boldsymbol{b}}$ of the parameter sequence, we consider the *parametric, deterministic PDE*

$$-\nabla \cdot (K(\cdot, u) \nabla P(\cdot, u)) = f \quad \text{in} \quad D\,, \qquad K(\cdot, u)|_{\partial D} = 0\,, \tag{3.5}$$

where $P(\cdot,u) \in V := H_0^1(D)$ and $f \in V^*$. To address the well-posedness of (3.5), we observe that for $u \in \Gamma_{\boldsymbol{b}}$, the following quantities are $(\Gamma_{\boldsymbol{b}}, \mathcal{A}_{\boldsymbol{b}})$ measurable:

$$\hat{K}(u) = \operatorname{esssup}_{x\in D} K_*(x) + \exp(\|\bar{K}\|_{L^\infty(D)} + \sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}|u_j|), \tag{3.6}$$

and

$$\check{K}(u) = \operatorname{essinf}_{x\in D} K_*(x) + \exp(\operatorname{essinf}_{x\in D} \bar{K}(x) - \sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}|u_j|) \ . \tag{3.7}$$

Assumption 3.1 implies $\sum_{j=1}^{\infty} \|\psi_j\|_{L^\infty(D)}|u_j| < \infty$ for $u \in \Gamma_{\boldsymbol{b}}$, so that

$$\check{K}(u) > 0 \quad \text{for} \quad u \in \Gamma_{\boldsymbol{b}} \ . \tag{3.8}$$

By the Lax-Milgram Lemma, then, for every $u \in \Gamma_{\boldsymbol{b}}$ exists a unique solution $P(\cdot,u) \in V$ of (3.5) which satisfies the estimate

$$\|P(\cdot,u)\|_V \leq \frac{1}{\check{K}(u)}\|f\|_{V^*} \ , \quad u \in \Gamma_{\boldsymbol{b}} \ . \tag{3.9}$$

The apriori bound (3.9) is not uniform with respect to $u \in \Gamma_{\boldsymbol{b}}$. Still, (3.9) and the separability of V imply the following result on the strong measurability of P as a map from $(\Gamma_{\boldsymbol{b}}, \mathcal{A}_{\boldsymbol{b}})$ to $(V, \mathcal{B}(V))$ (see also, for example, [8], [7, 2]).

Proposition 3.2. *The solution map $\Gamma_{\boldsymbol{b}} \ni u \mapsto P(\cdot,u) \in V$ of the parametric problem* (3.5) *is a strongly measurable map from $(\Gamma_{\boldsymbol{b}}, \mathcal{A}_{\boldsymbol{b}})$ to $(V, \mathcal{B}(V))$.*

We also remark that $\check{K}$, $\hat{K} \in L^p(U;\gamma_{\boldsymbol{b}})$ for every $0 < p < \infty$ [2, Prop.2.3].

3.2. **Bayesian inverse problem.** Following [14, 20], we assume that the data vector $\delta \in \mathbb{R}^k$ is generated by $k \in \mathbb{N}$ independent, continuous linear observation functionals $\mathcal{O}_i \in V^*$ for $i = 1, \dots, k$. We denote the forward map $\mathcal{G}(u) : \Gamma_{\boldsymbol{b}} \to \mathbb{R}^k$ as

$$\mathcal{G}(u) = (\mathcal{O}_1(P(\cdot,u)), \mathcal{O}_2(P(\cdot,u)), \dots, \mathcal{O}_k(P(\cdot,u))) \ . \tag{3.10}$$

Let $\vartheta \sim N(0,\Sigma)$ denote the unbiased, Gaussian observation noise in (1.1) with positive definite covariance $\Sigma > 0$. We consider the Bayesian inverse problem of determining the conditional probability γ^δ on $U = \Gamma_{\boldsymbol{b}}$ with the Gaussian prior measure $\gamma_{\boldsymbol{b}}$. The conditional posterior measure γ^δ is determined by (1.2).

Proposition 3.3. *Under Assumption 3.1, the posterior measure $\gamma^\delta(du) = \gamma(du|\delta)$ satisfies*

$$\frac{d\gamma^\delta}{d\gamma_{\boldsymbol{b}}} \propto \exp(-\Phi(u;\delta)), \qquad u \in \Gamma_{\boldsymbol{b}} \tag{3.11}$$

with the Bayesian potential $\Phi(u;\delta) = \frac{1}{2}|\mathcal{G}(u) - \delta|_\Sigma^2$ as in (1.3)*. Moreover, there holds continuous dependence on the data in the following sense: for $\delta, \delta' \in \mathbb{R}^k$ such that $|\delta|, |\delta'| < \lambda$,*

$$d_{Hell}(\gamma^\delta, \gamma^{\delta'}) \leq C(\lambda)|\delta - \delta'|$$

for a constant $C(\lambda)$ depending only on λ.

Proof. From Proposition 3.2, we deduce that $\mathbb{R}^{\mathbb{N}} \supset \Gamma_{\boldsymbol{b}} \ni u \mapsto \mathcal{O}_i(P(u))$, viewed as a map from $U = \Gamma_{\boldsymbol{b}}$ to $\mathbb{R}$, is $\gamma_{\boldsymbol{b}}$ measurable. This implies the $\gamma_{\boldsymbol{b}}$-measurability of the forward map $\mathcal{G}(u)$. The $\gamma_{\boldsymbol{b}}$-measurability of $\mathcal{G}(u)$ in turn implies (3.11).

To show the continuous dependence of γ^δ with respect to δ, we verify that the conditions of Assumption 2.2 hold. To this end, we note that there exists $c > 0$ such that for all δ and for all $u \in \Gamma_{\boldsymbol{b}}$

$$\Phi(u;\delta) \le c(|\delta| + |\mathcal{G}(u)|)^2 \; .$$

From (3.10), we find that $|\mathcal{G}(u)| \le c\|P(\cdot,u)\|_V$ for every $u \in \Gamma_{\boldsymbol{b}}$, for some constant $c > 0$ independent of $u \in \Gamma_{\boldsymbol{b}}$ (but depending on the observation functionals $\mathcal{O}_i(\cdot)$). From (3.9), we deduce that for $|\delta| \le \lambda$ holds

$$\forall u \in \Gamma_{\boldsymbol{b}} : \quad \Phi(u;\delta) \le c(\lambda + \frac{1}{\check{K}(u)})^2 \; .$$

Thus $\int_U \Phi(u;\delta) d\gamma(u) \le \Lambda(\lambda)$ for some $\Lambda(\lambda) > 0$ depending only on λ.

To verify the local Lipschitz condition (ii) of Assumption 2.2, we note that for every $u \in \Gamma_{\boldsymbol{b}}$

$$|\Phi(u;\delta) - \Phi(u;\delta')| \le \frac{1}{2}|\langle \Sigma^{-1/2}(\delta + \delta' - 2\mathcal{G}(u)), \Sigma^{-1/2}(\delta - \delta')\rangle|$$
$$\le \frac{1}{2}\|\Sigma^{-1/2}\|^2_{L(\mathbb{R}^k,\mathbb{R}^k)}(|\delta| + |\delta'| + 2|\mathcal{G}(u)|)|\delta - \delta'| \le c(\lambda + \frac{1}{\check{K}(u)})|\delta - \delta'| \; .$$

With the choice $G(\lambda;u) = c(\lambda + 1/\check{K}(u))$ for $u \in \Gamma_{\boldsymbol{b}}$, the assertion follows. □

4. MCMC Finite Element Method

4.1. **Dimension truncation of the parametric forward problems.** First we consider the solution of problem (3.5) by truncating the Karhúnen-Loève expansion of the coefficient K. Denoting the coordinate vector $u \in \Gamma_{\boldsymbol{b}}$ truncated to finite dimension $J \in \mathbb{N}$ by $u^J = (u_1, u_2, \ldots, u_J, 0, 0, \ldots)$. Let $K^J(\cdot,u) = K(\cdot,u^J)$. For $u \in \Gamma_{\boldsymbol{b}}$, we consider the partial differential equation corresponding to the J-term approximation of the coefficient K:

$$-\nabla\cdot(K^J(\cdot,u)\nabla P^J(\cdot,u)) = f \quad \text{in} \quad D\,, \quad P^J(\cdot,u)|_{\partial D} = 0 \tag{4.1}$$

where $P^J(\cdot,u) \in V$. For every $u \in \Gamma_{\boldsymbol{b}}$ the dimensionally truncated, parametric deterministic problem (4.1) admits a unique solution $P^J(\cdot,u) \in V$ and from (3.9) we obtain the apriori bound

$$\forall u \in \Gamma_{\boldsymbol{b}} : \qquad \|P^J(\cdot,u)\|_V \le \frac{\|f\|_{V^*}}{\check{K}(u^J)} \; . \tag{4.2}$$

Remark 4.1. *In* (4.1)*, the parametric solution* $P^J(\cdot,u)$ *depends on parameters* $u^J \in \mathbb{R}^J$ *which is of full gaussian measure with respect to the product gaussian measure* γ^J *on* $\mathbb{R}^J$ *equipped with the* σ*-algebra of Borel sets.*

To estimate the error incurred by the dimension truncation of the parametric coefficient $K(\cdot,u)$ in (3.1), from (3.5) and (4.1) there holds for any $u \in \Gamma_{\boldsymbol{b}}$

$$-\nabla\cdot(K(\cdot,u)\nabla(P(\cdot,u) - P^J(\cdot,u))) = -\nabla\cdot((K(\cdot,u) - K^J(\cdot,u))\nabla P^J(\cdot,u)) \; .$$

We therefore have for every $u \in U = \Gamma_{\boldsymbol{b}}$ the bound

$$\begin{aligned}\|P(\cdot,u) - P^J(\cdot,u)\|_V &\le \frac{1}{\check{K}(u)}\|K(\cdot,u) - K^J(\cdot,u)\|_{L^\infty(D)}\|P^J(\cdot,u)\|_V \\ &\le \frac{1}{\check{K}(u)\check{K}(u^J)}\|K(\cdot,u) - K^J(\cdot,u)\|_{L^\infty(D)}\|f\|_{V^*}\ .\end{aligned}$$

Using the elementary inequality $|e^x - e^y| \le |x-y|(e^x + e^y)$ for $x, y \in \mathbb{R}$, we infer

$$\|K(\cdot,u) - K^J(\cdot,u)\|_{L^\infty(D)} \le 2\exp\left(b_0 + \sum_{j=1}^{\infty} b_j|u_j|\right)\sum_{j=J+1}^{\infty} b_j|u_j|\ .$$

This implies that there exists a constant $C > 0$ such that for every $u \in U = \Gamma_{\boldsymbol{b}}$

$$\|P(\cdot,u) - P^J(\cdot,u)\|_V \le C\exp\left(3\sum_{j=1}^{\infty} b_j|u_j|\right)\Big(\sum_{j=J+1}^{\infty} b_j|u_j|\Big)\ . \tag{4.3}$$

Remark 4.2. *Approximating the posterior probability measure by truncating the Karhúnen-Loève expansion is studied in Dashti and Stuart in* [5]*. However, the assumption of Theorem 2.4 in* [5] *requires imposing a Gaussian prior on a Banach space of functions that possess sufficient smoothness. Dashti and Stuart* [5] *verify that this assumption holds when the prior probability is defined on the Hölder space $C^t(D)$ for $t > 0$. In the present paper, we achieve a rate of convergence for approximating the Bayesian posterior measure by truncating the Karhúnen-Loève expansion of the random diffusion coefficient by assuming only a rate of decay for the $L^\infty(D)$ norm of the coefficients ψ_j in the expansion* (3.1)*. Indeed, define the observation functional for the dimensionally truncated forward model as*

$$\mathcal{G}^J(u) = (\mathcal{O}_1(P^J(u)), \ldots, \mathcal{O}_k(P^J(u)))\ .$$

We define the approximated potential as

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

and the approximated posterior measure $\gamma^{J,\delta}$ on $(\Gamma_{\boldsymbol{b}}, \mathcal{A}_{\boldsymbol{b}})$ by

$$\frac{d\gamma^{J,\delta}}{d\gamma_{\boldsymbol{b}}} \propto \exp(-\Phi^J(u;\delta)).$$

Then we obtain from the proof of Proposition 4.6 below the estimate

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

Indeed, this is the result in Proposition 4.6 ahead upon formally letting the meshwidth in the FE approximation in D tend to 0. The rate J^{-q} due to dimensionally truncating the uncertainty parametrization in the forward problem is therefore obtained by solely imposing a decay rate for $\|\psi_j\|_{L^\infty(D)}$.

In the same vein, for uncertain inputs $K^J(\cdot,u)$ in (4.1) with a given, fixed number J parameters, *and for circulant embedding, collocation based access of the gaussian random field inputs in Finite Element discretizations, as described in Remark 4.15 ahead, no dimension truncation error needs to be accounted for.* The geometric ergodicity result and all FE error bounds which follow remain valid verbatim in this case, by formally dropping the terms J^{-q} from the error bounds.

4.2. **Finite Element discretization of the forward problem.** We now approximate the truncated, parametric problem (4.1) by discretizing it with continuous, piecewise linear Finite Elements in D. As D is a bounded polyhedron with plane sides, we consider in D a nested sequence $\{\mathcal{T}^l\}_{l=1}^\infty$ of regular, simplicial triangulations of D which are defined inductively in the usual way: $\{\mathcal{T}^l\}$ is obtained from $\{\mathcal{T}^{l-1}\}$ by dividing each simplex in $\{\mathcal{T}^{l-1}\}$ into 2^d subsimplices. On the sequence $\{\mathcal{T}^l\}_{l\geq 1}$ of regular, simplicial triangulations of D thus obtained, we define a nested sequence $\{V^l\}_{l\geq 1}$ of spaces of continuous, piecewise linear functions on $\mathcal{T}^l$ as

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

where $\mathbb{P}_1(T)$ is the set of linear polynomials in the simplex T. Let $h_l = O(2^{-l})$ denote the maximum diameter of the simplices in $\mathcal{T}^l$. Then, for $0 < t \leq 1$ and for $P \in H^{1+t}(D) \cap H^1_0(D)$, there holds the *approximation property*

$$\inf_{Q\in V^l} \|P - Q\|_V \leq C h_l^t \|P\|_{H^{1+t}(D)} \; . \tag{4.4}$$

If, in particular, $P \in W := H^2(D) \cap H^1_0(D)$, there holds

$$\inf_{Q\in V^l} \|P - Q\|_V \leq C h_l \|P\|_{H^2(D)} \; . \tag{4.5}$$

Here, the constant $C > 0$ is independent of l and of P, and depends only on the shape of $\mathcal{T}^0$.

We consider the finite element approximation of the dimensionally truncated, parametric deterministic problem (4.1) as: given $u \in \Gamma_{\boldsymbol{b}}$, find $P^{J,l}(u) \in V^l$ such that for every $\phi \in V^l$

$$\int_D K^J(x,u) \nabla P^{J,l}(x,u) \cdot \nabla \phi(x) dx = \int_D f(x)\phi(x) dx \; . \tag{4.6}$$

From Cea's lemma, we deduce that for every $u \in \Gamma_{\boldsymbol{b}}$

$$\|P^J(\cdot,u) - P^{J,l}(\cdot,u)\|_V \leq \frac{\hat{K}(u^J)}{\check{K}(u^J)} \inf_{Q\in V^l} \|P^J(\cdot,u) - Q\|_V \; . \tag{4.7}$$

In this section, we analyze how the convergence rate of the Finite Element approximation (4.6) influences the convergence rate of the MCMC algorithm to sample the posterior measure. For simplicity, we consider the case where $P \in W$ almost surely. We therefore impose sufficient conditions to ensure that the solution $P^J(\cdot,u)$ of (4.1) belongs to W for $\gamma_{\boldsymbol{b}}$-almost all $u \in \Gamma_{\boldsymbol{b}}$.

Assumption 4.3. *The domain $D \subset \mathbb{R}^d$ is convex, $f \in L^2(D)$ and the functions K_*, $\bar{K}$ and ψ_j in the expansion* (3.1) *of the parametric coefficient $K(\cdot,u)$ belong to $W^{1,\infty}(D)$, and $\bar{\boldsymbol{b}} := (\|\psi_j\|_{W^{1,\infty}(D)})_{j\geq 1} \in \ell^1(\mathbb{N})$.*

With Assumption 4.3, we define the set

$$\Gamma_{\bar{\boldsymbol{b}}} = \{u \in \mathbb{R}^{\mathbb{N}} : \; \sum_{j=1}^\infty |u_j| \bar{b}_j < \infty\} \; . \tag{4.8}$$

Assumption 4.3 implies that $\bar{\boldsymbol{b}} \in \ell^2(\mathbb{N})$. Then, as the set $\Gamma_{\boldsymbol{b}}$ in (3.3), also the set $\Gamma_{\bar{\boldsymbol{b}}}$ in (4.8) has Gaussian measure one. For finite J, for each $u \in \mathbb{R}^{\mathbb{N}}$, $\nabla K^J(\cdot,u) \in$

$L^\infty(D)^d$ is well defined and admits the explicit representation (with convergence in $L^\infty(D)^d$ as $J \to \infty$ for $u \in \Gamma_{\bar{\boldsymbol{b}}}$ in (4.8))

$$\nabla K^J(\cdot,u) = \nabla K_*(\cdot) + \exp\left(\bar{K}(\cdot) + \sum_{j=1}^J \psi_j(x) u_j\right)\left(\nabla\bar{K}(\cdot) + \sum_{j=1}^J \nabla\psi_j(\cdot) u_j\right) .$$

Moreover, there exists $C > 0$ such that for all J and for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\|\nabla K^J(\cdot,u)\|_{L^\infty(D)} \le C\left(1 + \exp\left(\sum_{j=1}^\infty b_j|u_j|\right)(1 + \sum_{j=1}^J \bar{b}_j|u_j|)\right) .$$

From (4.1), for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$-\Delta P^J(\cdot,u) = \frac{1}{K^J(\cdot,u)}(\nabla K^J(\cdot,u)\cdot\nabla P^J(\cdot,u) + f) \quad \text{in} \quad L^2(D) .$$

Under Assumption 4.3, the classical $H^2(D)$ regularity result for the Dirichlet Laplacian in convex domains D implies the (uniform w.r. to J for every $u \in \Gamma_{\boldsymbol{b}}$) bound

$$\begin{aligned}
\|P^J(\cdot,u)\|_{H^2(D)} &\le c\frac{1}{\check{K}(u^J)}(\|\nabla K^J(\cdot,u)\|_{L^\infty(D)}\|P^J(\cdot,u)\|_V + \|f\|_{L^2(D)}) \\
&\le c\frac{1}{\check{K}(u^J)}(\|\nabla K^J(\cdot,u)\|_{L^\infty(D)}\frac{\|f\|_{V^*}}{\check{K}(u^J)} + \|f\|_{L^2(D)}) \\
&\le C\exp\left(3\sum_{j=1}^\infty b_j|u_j|\right)\left(1 + \sum_{j=1}^J \bar{b}_j|u_j|\right) . \qquad (4.9)
\end{aligned}$$

Here, the constant $C > 0$ is independent of $J \in \mathbb{N}$ and of $u \in \Gamma_{\boldsymbol{b}}$ in (4.8), but depends on $\|f\|_{L^2(D)}$. From this and from equation (4.7) and (4.3), we obtain

Lemma 4.4. *Under Assumption 4.3, for every proposal $u \in \Gamma_{\boldsymbol{b}}$, the solution $P^J(\cdot,u)$ of the dimensionally truncated, parametric problem* (4.1) *and its Finite Element approximation $P^{J,l}(\cdot,u)$ in* (4.6) *satisfy the a-priori estimate*

$$\|P^J(\cdot,u) - P^{J,l}(\cdot,u)\|_V \le C\exp\left(5\sum_{j=1}^\infty b_j|u_j|\right)\left(1 + \sum_{j=1}^J \bar{b}_j|u_j|\right)2^{-l} \qquad (4.10)$$

for some constant $C > 0$ which is independent of J and l, and of $u \in \Gamma_{\boldsymbol{b}}$.

For every proposal $u \in \Gamma_{\boldsymbol{b}}$ and for every $J, l \in \mathbb{N}$, the solutions $P(\cdot,u)$ of (3.5) *and $P^{J,l}(\cdot,u)$ of* (4.6) *satisfy the error bound*

(4.11)

$$\|P(\cdot,u) - P^{J,l}(\cdot,u)\|_V \le C\exp\left(5\sum_{j=1}^\infty b_j|u_j|\right)\left(2^{-l}(1 + \sum_{j=1}^J \bar{b}_j|u_j|) + \sum_{j=J+1}^\infty b_j|u_j|\right)$$

where, again, the constant $C > 0$ is independent of $u \in \Gamma_{\boldsymbol{b}}$, and of $J, l \in \mathbb{N}$.

4.3. **Finite element approximation of the posterior probability measure.** For $u \in \Gamma_{\boldsymbol{b}}$, we consider the *vector of observables with dimension truncation and finite element approximation*, defined by

$$\mathcal{G}^{J,l}(u) = (\mathcal{O}_1(P^{J,l}(u)), \dots, \mathcal{O}_k(P^{J,l}(u))) : \Gamma_{\boldsymbol{b}} \mapsto \mathbb{R}^k .$$

We define the corresponding approximate potential function

$$\Phi^{J,l}(u;\delta) = \frac{1}{2}|\delta - \mathcal{G}^{J,l}(u)|^2_\Sigma\ , \quad u \in \Gamma_{\boldsymbol{b}}\ . \tag{4.12}$$

We denote the corresponding approximate posterior measure by $\gamma^{J,l,\delta}$. It is defined as

$$\frac{d\gamma^{J,l,\delta}}{d\gamma_{\boldsymbol{b}}}(u) \propto \exp(-\Phi^{J,l}(u;\delta))\ , \quad u \in \Gamma_{\boldsymbol{b}}\ . \tag{4.13}$$

As we justify below, the positive normalizing constant in (4.13) is uniformly (w.r. to $J, l \in \mathbb{N}$) bounded from below away from 0. Therefore, $\gamma^{J,l,\delta}$ is indeed the Bayesian posterior probability measure for the inverse problem with the approximated observation functional $\mathcal{G}^{J,l}$. With respect to the Hellinger distance, the measure $\gamma^{J,l,\delta}$ is an approximation of the 'true' (i.e., not involving J-term dimension-truncation of (3.1) or Galerkin FE approximation (4.6)) Bayesian posterior γ^δ. To prove a rate of convergence estimate for approximation (4.3), we impose an assumption on the decay rate of $\|\psi_j\|_{L^\infty(D)}$.

Assumption 4.5. *There exist constants $c > 0$ and $s > 1$ such that for every $j \geq 1$ holds $\|\psi_j\|_{L^\infty(D)} \leq c/j^s$.*

Evidently, Assumption 4.5 implies $\boldsymbol{b} \in \ell^1(\mathbb{N})$. We define

$$q = s - 1 > 0\ . \tag{4.14}$$

Proposition 4.6. *Under Assumptions 4.3 and 4.5, there exists a positive constant c (depending only on the data δ) such that, with q as defined in (4.14) holds*

$$\forall J, l \in \mathbb{N}: \qquad d_{\mathrm{Hell}}(\gamma^\delta, \gamma^{J,l,\delta}) \leq c(\delta)(J^{-q} + 2^{-l})\ .$$

Proof The proof of this proposition is an extension of the argument in [14] to the case where the forward functional $\mathcal{G}(u)$ and its dimension truncated Finite Element discretization $\mathcal{G}^{J,l}(u)$ may not be uniformly bounded with respect to $u \in \Gamma_{\boldsymbol{b}}$. We denote the exact and approximate normalizing constants, respectively, as

$$Z(\delta) = \int_{\Gamma_{\boldsymbol{b}}} \exp(-\Phi(u;\delta))d\gamma_{\boldsymbol{b}}(u), \quad Z^{J,l}(\delta) = \int_{\Gamma_{\boldsymbol{b}}} \exp(-\Phi^{J,l}(u;\delta))d\gamma_{\boldsymbol{b}}(u)\ .$$

From (3.9), we have that $\Gamma_{\boldsymbol{b}} \ni u \mapsto \|P(\cdot,u)\|_V \in L^1(\Gamma_{\boldsymbol{b}}, \gamma)$. Therefore, there exists $c > 0$ such that, for each positive constant Λ, $\gamma_{\boldsymbol{b}}(\{u \in \Gamma_{\boldsymbol{b}} : \|P(\cdot,u)\|_V > \Lambda\}) < c/\Lambda$. Thus, for $\Lambda > 0$ sufficiently large, the measure of the set of all $u \in \Gamma_{\boldsymbol{b}}$ such that $\|P(\cdot,u)\|_V \leq \Lambda$ is larger than $1 - c/\Lambda$. There exists a constant $c > 1$ such that for $\lambda > 0$ given by $|\delta|^2_\Sigma = \delta^\top \Sigma \delta < \lambda^2$ holds the bound

$$\forall u \in \Gamma_{\boldsymbol{b}}: \quad \Phi(u;\delta) \leq |\delta|^2_\Sigma + |\mathcal{G}(u)|^2_\Sigma \leq c(\lambda + \Lambda)^2\ .$$

Thus we obtain that

$$Z(\delta) > (1 - c/\Lambda)\exp(-c(\lambda + \Lambda))^2 =: c(\lambda) > 0\ . \tag{4.15}$$

We then estimate

$$\begin{aligned} & d_{\mathrm{Hell}}(\gamma^\delta, \gamma^{J,l,\delta})^2 \\ = \ & \frac{1}{2}\int_{\Gamma_{\boldsymbol{b}}} \Big(Z(\delta)^{-1/2} \exp\Big(-\frac{1}{2}\Phi(u;\delta)\Big) - (Z^{J,l}(\delta))^{-1/2} \exp\Big(-\frac{1}{2}\Phi^{J,l}(u;\delta)\Big)\Big)^2 d\gamma_{\boldsymbol{b}}(u) \\ \leq \ & I_1 + I_2, \end{aligned}$$

where I_1 and I_2 are given by

$$I_1 := \frac{1}{Z(\delta)} \int_{\Gamma_{\boldsymbol{b}}} \Big(\exp\Big(-\frac{1}{2}\Phi(u;\delta)\Big) - \exp\Big(-\frac{1}{2}\Phi^{J,l}(u;\delta)\Big)\Big)^2 d\gamma_{\boldsymbol{b}}(u),$$

$$I_2 := |Z(\delta)^{-1/2} - Z^{J,l}(\delta)^{-1/2}|^2 \int_{\Gamma_{\boldsymbol{b}}} \exp(-\Phi^{J,l}(u;\delta)) d\gamma_{\boldsymbol{b}}(u) .$$

To bound I_1, we estimate for $u \in \Gamma_{\boldsymbol{b}}$,

$$\begin{aligned} \Big| \exp\Big(-\frac{1}{2}\Phi(u;\delta)\Big) - \exp\Big(-\frac{1}{2}\Phi^{J,l}(u;\delta)\Big)\Big| &\le \frac{1}{2}|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| \\ &\le c(|\delta| + |\mathcal{G}(u)| + |\mathcal{G}^{J,l}(u)|)|\mathcal{G}(u) - \mathcal{G}^{J,l}(u)| . \end{aligned} \tag{4.16}$$

Under Assumption 4.3, there exists a constant $C > 0$ independent of J and of l such that for each $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\begin{aligned} |\mathcal{G}(u) - \mathcal{G}^{J,l}(u)| &\le C \max\{\|\mathcal{O}_i\|_{V^*}\} \|P(\cdot,u) - P^{J,l}(\cdot,u)\|_V \\ &\le C \exp\left(5 \sum_{j=1}^{\infty} b_j |u_j| \right) \left(2^{-l}(1 + \sum_{j=1}^{J} \bar{b}_j |u_j|) + \sum_{j=J+1}^{\infty} b_j |u_j| \right) . \end{aligned}$$

There exists a constant $C > 0$ such that, for each $u \in \Gamma_{\boldsymbol{b}}$, holds the bound

$$|\mathcal{G}(u)| \le C\|P(\cdot,u)\|_V \le C \frac{1}{\check{K}(u)} \le C \exp\left(\sum_{j=1}^{\infty} b_j |u_j| \right) .$$

From (4.6), we deduce that for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\|P^{J,l}(\cdot,u)\|_V \le \frac{\|f\|_{V^*}}{\check{K}(u^J)} \le C \exp\left(\sum_{j=1}^{\infty} b_j |u_j| \right) \tag{4.17}$$

so that, for $u \in \Gamma_{\boldsymbol{b}}$, there holds the bound

$$|\mathcal{G}^{J,l}(u)| \le C \exp\left(\sum_{j=1}^{\infty} b_j |u_j| \right) . \tag{4.18}$$

Therefore, with $|\delta| < \lambda$, there exists a constant $c(\lambda) > 0$ (cf. (4.15)) that depends only on λ so that for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\begin{aligned} &\Big| \exp\Big(-\frac{1}{2}\Phi(u;\delta)\Big) - \exp\Big(-\frac{1}{2}\Phi^{J,l}(u;\delta)\Big)\Big| \\ \le\ & c(\lambda) \exp\left(6 \sum_{j=1}^{\infty} b_j |u_j| \right) \left(2^{-l}(1 + \sum_{j=1}^{J} \bar{b}_j |u_j|) + \sum_{j=J+1}^{\infty} b_j |u_j| \right) . \end{aligned}$$

As $\Gamma_{\bar{\boldsymbol{b}}}$, $\Gamma_{\boldsymbol{b}} \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ are sets of full measure under the gaussian prior γ in (3.2), we have

$$\begin{aligned} I_1 \le c(\lambda) \bigg(& 2^{-2l} \int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12 \sum_{j=1}^{\infty} b_j |u_j| \Big) (1 + \sum_{j=1}^{J} \bar{b}_j |u_j|)^2 d\gamma_{\boldsymbol{b}} \\ & + \int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12 \sum_{j=1}^{\infty} b_j |u_j| \Big) \Big(\sum_{j=J+1}^{\infty} b_j |u_j| \Big)^2 d\gamma_{\boldsymbol{b}} \bigg) . \end{aligned}$$

We note that

$$\int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12\sum_{j=1}^{\infty} b_j|u_j|\Big)(1+\sum_{j=1}^{J}\bar{b}_j|u_j|)^2 d\gamma_{\boldsymbol{b}} \leq \int_{\Gamma_{\boldsymbol{b}}} \exp\left(\sum_{j=1}^{\infty} 12 b_j|u_j| + \sum_{j=1}^{J} 2\bar{b}_j|u_j|\right) d\gamma_{\boldsymbol{b}}$$
$$\leq \prod_{j=1}^{\infty}\int_{\mathbb{R}} \exp\big((12b_j+2\bar{b}_j)|u_j|\big)\, d\gamma_1(u_j)\ .$$

Employing the inequality

$$\int_{-\infty}^{\infty} \exp(-z^2/2+|z|t)\frac{dz}{\sqrt{2\pi}} \leq \exp(t^2/2+t\sqrt{2/\pi}), \tag{4.19}$$

which we prove in Appendix B, we deduce that,

$$\int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12\sum_{j=1}^{\infty} b_j|u_j|\Big)(1+\sum_{j=1}^{J}\bar{b}_j|u_j|)^2 d\gamma_{\boldsymbol{b}}$$
$$\leq \quad \exp\left(\sum_{j=1}^{\infty}(12b_j+2\bar{b}_j)^2/2+\sum_{j=1}^{\infty}(12b_j+2\bar{b}_j)\sqrt{2/\pi}\right) < \infty$$

due to our assumption that the sequences $\boldsymbol{b}=(b_j)_{j\geq 1}$, $\bar{\boldsymbol{b}}:=(\bar{b}_j)_{j\geq 1}\in \ell^1(\mathbb{N})$. We now estimate

$$\int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12\sum_{j=1}^{\infty} b_j|u_j|\Big)\Big(\sum_{j=J+1}^{\infty} b_j|u_j|\Big)^2 d\gamma_{\boldsymbol{b}}$$
$$= \int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12\sum_{j=1}^{\infty} b_j|u_j|\Big)\Big(\sum_{i,j=J+1}^{\infty} b_ib_j|u_i||u_j|\Big)d\gamma_{\boldsymbol{b}}$$
$$\leq \sum_{i=J+1}^{\infty} b_i^2\int_{-\infty}^{\infty}\exp(12b_i|u_i|)u_i^2 d\gamma_1(u_i)\prod_{\substack{k=1\\k\neq i}}^{\infty}\int_{-\infty}^{\infty}\exp(12b_k|u_k|)d\gamma_1(u_k)$$
$$+\sum_{\substack{i,j=J+1\\i\neq j}}^{\infty} b_ib_j\int_{-\infty}^{\infty}\exp(12b_i|u_i|)|u_i|d\gamma_1(u_i)\cdot\int_{-\infty}^{\infty}\exp(12b_j|u_j|)|u_j|d\gamma_1(u_j)$$
$$\cdot\prod_{\substack{k=1\\k\neq i,j}}^{\infty}\int_{-\infty}^{\infty}\exp(12b_k|u_k|)d\gamma_1(u_k)\ .$$

In Appendix B we show that there exists $c>0$ (independent of b_j) such that

$$\int_{-\infty}^{\infty} z^2\exp(-z^2/2+|z|t)\frac{dz}{\sqrt{2\pi}} \leq c\exp(t^2/2)(1+t^2), \tag{4.20}$$

and

$$\int_{-\infty}^{\infty} |z|\exp(-z^2/2+|z|t)\frac{dz}{\sqrt{2\pi}} \leq c\exp(t^2/2)(1+t)\ . \tag{4.21}$$

We deduce that there exists $C > 0$ such that for all $J \geq 1$

$$\begin{aligned}
&\int_{\Gamma_{\boldsymbol{b}}} \exp\Big(12\sum_{j=1}^{\infty} b_j|u_j|\Big)\Big(\sum_{j=J+1}^{\infty} b_j|u_j|\Big)^2 d\gamma_{\boldsymbol{b}} \\
\leq\ & C\sum_{i=J+1}^{\infty} b_i^2(1+b_i^2)\exp\left(\sum_{k=1}^{\infty} 12^2 b_k^2/2 + 12 b_k\sqrt{2/\pi}\right) \\
&+C\sum_{\substack{i,j=J+1\\ i\neq j}}^{\infty} b_i b_j(1+b_i)(1+b_j)\exp\left(\sum_{k=1}^{\infty} 12^2 b_k^2/2 + 12 b_k\sqrt{2/\pi}\right) \\
\leq\ & C\left(\sum_{j=J+1}^{\infty} b_i\right)^2 \leq CJ^{-2q}
\end{aligned}$$

by (4.14) and Assumption 4.5.

Hence, there exists a constant $C > 0$ which is independent of $l, J \in \mathbb{N}$ such that

$$I_1 \leq C(2^{-2l} + (\sum_{j=J+1}^{\infty} b_j)^2) \leq C(2^{-2l} + J^{-2q}) .$$

To estimate term I_2, we observe that there exists 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

$$\begin{aligned}
|Z(\delta) - Z^{J,l}(\delta)| &\leq \int_{\Gamma_{\boldsymbol{b}}} |\exp(-\Phi(u;\delta)) - \exp(-\Phi^{J,l}(u;\delta))| d\gamma_{\boldsymbol{b}}(u) \\
&\leq \int_{\Gamma_{\boldsymbol{b}}} |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| d\gamma_{\boldsymbol{b}}(u) \\
&\leq \Big(\int_{\Gamma_{\boldsymbol{b}}} |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)|^2 d\gamma_{\boldsymbol{b}}(u)\Big)^{1/2} .
\end{aligned}$$

The proof for the uniform boundedness of $Z^{J,l}(\delta)$ from below is similar to that for $Z(\delta)$, using (4.17). Therefore, as $Z(\delta)$ and $Z^{J,l}(\delta)$ are uniformly (with respect to $J, l \in \mathbb{N}$) bounded away from zero for all observation data δ, an analysis similar to that for I_1 shows that

$$I_2 \leq c(2^{-2l} + J^{-2q}) .$$

Thus, there exists a constant $c(\delta) > 0$ which is independent of J and l such that

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

which completes the proof of Proposition 4.6. □

4.4. **Convergence of the MCMC-FE algorithm.** For computational approximation of integrals of functions $g : \Gamma_{\boldsymbol{b}} \to \mathbb{R}$ with respect to the posterior measure γ^{δ}, we perform the Markov-Chain Monte-Carlo method (2.2) with the acceptance probability $\alpha(u, v)$ in (2.1) replaced by

$$\alpha^{J,l}(u,v) = 1 \wedge \exp(\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta)) , \quad (u,v) \in \Gamma_{\boldsymbol{b}} \times \Gamma_{\boldsymbol{b}} . \tag{4.22}$$

The integral of g over $\Gamma_{\boldsymbol{b}}$ with respect to $\gamma^{J,l,\delta}$ is then approximated by

$$E_M^{\gamma^{J,l,\delta}}[g] := \frac{1}{M}\sum_{k=1}^{M} g(u^{(k)}). \tag{4.23}$$

To estimate the MCMC sampling error incurred by Finite Element approximation,

$$\mathbb{E}^{\gamma^\delta}[g] - E_M^{\gamma^{J,l,\delta}}[g] \ , \tag{4.24}$$

we develop an asymptotic, probabilistic error bound for $\mathcal{P}^{J,l}_{u^{(0)}}$-a.e. realization of the Markov chain and a mean square bound when $M \to \infty$.

By $p^{J,l}(u,\cdot)$ we denote the transition kernel of the Markov chain generated by the MCMC process with acceptance probability $\alpha^{J,l}$ defined in (4.22). Then we have

Proposition 4.7. *For sufficiently large constants $a > 0$ and $B > 0$, there exist $J_0 \in \mathbb{N}$, $l_0 \in \mathbb{N}$, $r > 1$ and $R > 0$ (generally depending on a, B and on δ) such that for every $J > J_0$ and every $l > l_0$, the function $\mathcal{V} : U \to \mathbb{R}$ defined by*

(4.25)

$$\mathcal{V}(u) = \begin{cases} \exp\left(a\sum_{j=1}^\infty (b_j + \bar{b}_j)|u_j| + \frac{1}{\varepsilon}\sum_{j>J} b_j|u_j|\right) & \text{if} \quad u \in \Gamma_{\bar{\boldsymbol{b}}} \\ \exp\left((a\sum_{j=1}^\infty b_j + a\sum_{j=1}^J \bar{b}_j)|u_j| + \frac{1}{\varepsilon}\sum_{j>J} b_j|u_j|\right) & \text{if} \quad u \in \Gamma_{\boldsymbol{b}} \setminus \Gamma_{\bar{\boldsymbol{b}}}, \end{cases}$$

with $\varepsilon = B\sum_{j>J} b_j$ satisfies, for every $n \in \mathbb{N}$ and for every $u \in U = \Gamma_{\boldsymbol{b}}$,

$$\|(p^{J,l})^n(u,\cdot) - \gamma^{J,l,\delta}\|_{\mathcal{V}} \le Rr^{-n}\mathcal{V}(u) \ .$$

We prove this proposition in Appendix A.

Remark 4.8. *From Proposition 4.7 and the proof of Theorem 2.4, the Markov chain generated by the MCMC process on the approximate forward problem* (4.6) *with acceptance probability* (4.22) *satisfies* (2.5) *and* (2.6)*, with the probability measure $\gamma^{J,l,\delta}$. The constants c in* (2.5) *and C in* (2.6) *depend on r, R and $\mathbb{E}^{\gamma^{J,l,\delta}}[\mathcal{V}(u)^2]$ in Proposition 4.7. From Lemma 6.2, $\|\mathcal{V}\|_{L^2(U,\gamma_{\boldsymbol{b}})}$ is uniformly bounded above with respect to J and l. As shown in the proof of Proposition 4.6, the normalizing constant $Z^{J,l}$ in* (4.13) *is bounded below from* 0 *uniformly with respect to J and l. Thus $\mathbb{E}^{\gamma^{J,l,\delta}}[\mathcal{V}(u)^2]$ is uniformly bounded with respect to J and l. Therefore, the constants c and C in* (2.5) *and* (2.6) *for the Markov chain generated by the MCMC process with the acceptance probability* (4.22) *are bounded above uniformly with respect to J and l.*

From (4.18), we deduce that there are positive constants c_1 and c_2 (depending only on r) so that, for $|\delta| < r$ and for every $J, l \in \mathbb{N}$ holds

$$\forall u \in \Gamma_{\boldsymbol{b}} : \qquad |\Phi^{J,l}(u;\delta)| \le c_1 + c_2 \exp\left(2\sum_{j=1}^\infty b_j|u_j|\right) . \tag{4.26}$$

To estimate effect of discretization and dimension truncation in the forward map on the Bayesian estimate, we require the following result.

Lemma 4.9. *For $U = \Gamma_{\boldsymbol{b}}$, the integral*

$$\kappa = \int_U \exp\left(-c_2 \exp(2\sum_{j=1}^\infty b_j|u_j|)\right) d\gamma_{\boldsymbol{b}}(u)$$

is finite and strictly positive. Here, c_2 is as in (4.26).

Proof First we note from Appendix B that

$$\int_{\mathbb{R}^{\mathbb{N}}} \exp(2\sum_{j=1}^{\infty} b_j|u_j|)d\gamma(u) = \lim_{J\to\infty}\prod_{j=1}^{J}\int_{-\infty}^{\infty}\exp(2b_j|u_j| - u_j^2/2)\frac{du_j}{\sqrt{2\pi}}$$

$$\leq \prod_{j=1}^{\infty}\exp(2b_j^2 + 2b_j\sqrt{2/\pi}) = \exp(2\sum_{j=1}^{\infty} b_j^2 + 2\sqrt{2/\pi}\sum_{j=1}^{\infty} b_j)) < \infty.$$

For every fixed $c > 0$ there exists a constant $\Lambda > 0$ so that

$$\gamma_{\boldsymbol{b}}\left(\left\{u \in \Gamma_{\boldsymbol{b}} :\ \exp(2\sum_{j=1}^{\infty} b_j|u_j|) \geq \Lambda\right\}\right) < c/\Lambda < 1\ ,$$

and we obtain

$$\gamma_{\boldsymbol{b}}\left(\left\{u \in \Gamma_{\boldsymbol{b}} :\ \exp(2\sum_{j=1}^{\infty} b_j|u_j|) < \Lambda\right\}\right) > 1 - c/\Lambda > 0\ .$$

Thus

$$\kappa > \exp(-c_2\Lambda)(1 - c/\Lambda) > 0\ .$$

The lemma is thus proved. □

Using Lemma 4.9, we may introduce a probability measure $\bar{\gamma}$ on $\Gamma_{\boldsymbol{b}} \subset U$ via

$$d\bar{\gamma}(u) = \frac{1}{\kappa}\exp(-c_2\exp(2\sum_{j=1}^{\infty} b_j|u_j|))d\gamma_{\boldsymbol{b}}(u)\ . \tag{4.27}$$

Let $\mathcal{V} : U \to \mathbb{R}$ be the function defined in (4.25). Then, there holds

Proposition 4.10. *Let $g : \Gamma_{\boldsymbol{b}} \to \mathbb{R}$ be a function such that $|g(u)|^2 \leq \mathcal{V}(u)$ for $u \in \Gamma_{\boldsymbol{b}}$. Then, for every initial condition $u^{(0)}$ and for $\mathcal{P}^{J,l}_{u(0)}$-almost every realization of the Markov chain holds the error bound*

$$\left|\mathbb{E}^{\gamma^\delta}[g(u)] - E_M^{\gamma^{J,l,\delta}}[g(u)]\right| \leq C_1 M^{-\frac{1}{2}} + C_2\big(J^{-q} + 2^{-l}\big) \tag{4.28}$$

where C_1 is a random variable which satisfies a.s. $C_1 \leq c_3|\xi_M|$ for random variables ξ_M on the probability space generating the randomness in the Markov chain which converge weakly as $M \to \infty$ to $\xi \sim N(0,1)$. In (4.28), C_2 is a constant which is independent of M, J and l.

Moreover, there exists a constant $c_4 > 0$ (which depends only on the data δ, and which is, in particular, independent of J,l and M) such that

$$\left(\mathcal{E}^{\bar{\gamma},J,l}\left[\left|\mathbb{E}^{\gamma^\delta}[g] - E_M^{\gamma^{J,l,\delta}}[g]\right|^2\right]\right)^{1/2} \leq c_4(M^{-1/2} + J^{-q} + 2^{-l})\ . \tag{4.29}$$

Here, the constant $q > 0$ is defined in (4.14).

Proof We observe that $|g(u)|^2 < \mathcal{V}(u)$ for $u \in \Gamma_{\boldsymbol{b}}$ implies that $g \in L^2(\Gamma_{\boldsymbol{b}}, \gamma_{\boldsymbol{b}})$ and $g \in L^2(\Gamma_{\boldsymbol{b}}, \gamma^\delta)$. We have from Proposition 4.6 and properties of the Hellinger metric (specifically, from (2.7) in [3])

$$|\mathbb{E}^{\gamma^\delta}[g] - \mathbb{E}^{\gamma^{J,l,\delta}}[g]| \leq \bar{c}(g)d_{\rm Hell}(\gamma^\delta, \gamma^{J,l,\delta}) \leq \bar{c}(g)c(\delta)(J^{-q} + 2^{-l})\ . \tag{4.30}$$

Here, $c(\delta)$ is as in Proposition 4.6 and $\bar{c}(g)$ depends on $\|g\|_{L^2(\Gamma_{\mathfrak{b}},\gamma)}$, but is independent of $J, l \in \mathbb{N}$. From Theorem 17.0.1 of [16] we deduce the existence of a constant $C > 0$, such that

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

where the sequence ξ_M of random variables converges weakly as $M \to \infty$ to $\xi \sim N(0,1)$. The constant C is determined by

$$C^2 = \mathcal{E}^{\gamma^{J,l,\delta}}[\bar{g}(u^{(0)})]^2 + 2\sum_{n=1}^{\infty} \mathcal{E}^{\gamma^{J,l,\delta}}[\bar{g}(u^{(0)})\bar{g}(u^{(n)})]$$

where $\bar{g}(u) = g(u) - \mathbb{E}^{\gamma^{J,l,\delta}}[g]$. As the normalizing constant $Z^{J,l}$ is uniformly bounded away from zero for sufficiently large J and l (cp. the proof of Proposition 4.6), there exists a constant $c > 0$ such that for all $J, l \in \mathbb{N}$

$$|\mathbb{E}^{\gamma^{J,l,\delta}}[g]| \le c\mathbb{E}^{\gamma}[|g|] \le c\mathbb{E}^{\gamma}[\mathcal{V}] \ .$$

Thus we have the uniform w.r. to J and l bound

$$\mathcal{E}^{\gamma^{J,l,\delta}}[\bar{g}(u^{(0)})]^2 \le c\mathbb{E}^{\gamma}(\mathcal{V})^2 \ .$$

We also have

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathcal{E}^{\gamma^{J,l,\delta}}[\bar{g}(u^{(0)})\bar{g}(u^{(n)})] &\le \mathbb{E}^{\gamma^{J,l,\delta}}\left[|\bar{g}(u^{(0)})| \sum_{n=1}^{\infty} |\mathcal{E}^{J,l}_{u^{(0)}}[\bar{g}(u^{(n)})]|\right] \\
&\le \mathbb{E}^{\gamma^{J,l,\delta}}\left[|\bar{g}(u^{(0)})| \sum_{n=1}^{\infty} |\mathcal{E}^{J,l}_{u^{(0)}}[g(u^{(n)})] - \mathbb{E}^{\gamma^{J,l,\delta}}[g]|\right] \\
&\le \mathbb{E}^{\gamma^{J,l,\delta}}[|\bar{g}(u)|\mathcal{V}(u)] \sum_{n=1}^{\infty} Rr^{-n}
\end{aligned}$$

which is uniformly bounded above for all J and l. Combining this with (4.30) shows the first assertion.

We now prove the mean square bound (4.29). Due to the invariance of the stationary measures $\gamma^{J,l,\delta}$, we may write

$$\begin{aligned}
\frac{1}{M}\mathcal{E}^{\gamma^{J,l,\delta},J,l}\Big[\Big|\sum_{k=1}^{M}\bar g(u^{(k)})\Big|^2\Big] &= \mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=1}^{M}\sum_{j=k+1}^{M}\mathcal{E}^{\gamma^{J,l,\delta},J,l}[\bar g(u^{(k)})\bar g(u^{j})]\\
&= \mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathcal{E}^{\gamma^{J,l,\delta},J,l}[\bar g(u^{(0)})\bar g(u^{(j)})]\\
&= \mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2] + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})\mathcal{E}^{J,l}_{u^{(0)}}[\bar g(u^{(j)})]]\\
&\le \mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2]\\
&\quad + 2\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k}\mathbb{E}^{\gamma^{J,l,\delta}}[|\bar g(u^{(0)})||\mathcal{E}^{J,l}_{u^{(0)}}g(u^{(j)}) - \mathbb{E}^{\gamma^{J,l,\delta}}[g]|]\\
&\le \mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2] + 2R\frac{1}{M}\sum_{k=0}^{M-1}\sum_{j=1}^{M-k} r^{-j}\mathbb{E}^{\gamma^{J,l,\delta}}[|\bar g(u^{(0)})|\mathcal{V}(u^{(0)})|],
\end{aligned}$$

where, as in Proposition 4.7 the constant $r > 1$ is independent of the parameters J and l. Since $\sup_{J,l}\mathbb{E}^{\gamma^{J,l,\delta}}[\bar g(u^{(0)})^2]$ is bounded independently of J and of l, we deduce that

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

We now show that this estimate also holds for the expectation with respect to the Markov chain where the initial state $u^{(0)}$ is distributed according to $\bar\gamma$ defined in (4.27). We have

$$\begin{aligned}
&\mathcal{E}^{\bar\gamma,J,l}\Big[\Big|\sum_{k=1}^{M}\bar g(u^{(k)})\Big|^2\Big] = \int_{\Gamma_{\boldsymbol b}}\mathcal{E}^{J,l}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar g(u^{(k)})\Big|^2\Big]d\bar\gamma(u^{(0)})\\
&= \int_{\Gamma_{\boldsymbol b}}\mathcal{E}^{J,l}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar g(u^{(k)})\Big|^2\Big]\frac{d\bar\gamma}{d\gamma_{\boldsymbol b}}(u^{(0)})\frac{d\gamma_{\boldsymbol b}}{\gamma^{J,l,\delta}}(u^{(0)})d\gamma^{J,l,\delta}(u^{(0)})\\
&= \int_{\Gamma_{\boldsymbol b}}\mathcal{E}^{J,l}_{u^{(0)}}\Big[\Big|\sum_{k=1}^{M}\bar g(u^{(k)})\Big|^2\Big]\frac{1}{\kappa}\exp\left(-c_2\exp(2\sum_{j=1}^{\infty}b_j|u_j|)\right)Z^{J,l}(\delta)\exp(\Phi^{J,l}(u;\delta))d\gamma^{J,l,\delta}(u^{(0)})\,.
\end{aligned}$$

From (4.26), we deduce that

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

□

For $u \in \Gamma_{\boldsymbol b}$ and for $\ell \in V^*$, define $g(u) := \ell(P(\cdot,u))$ and, for $J,l \in \mathbb{N}$, $g^{J,l}(u) := \ell(P^{J,l}(\cdot,u))$. From (3.11) and (4.11), there exists a constant $c > 0$ which is independent of $J,l \in \mathbb{N}$ such that

$$\mathbb{E}^{\gamma^{\delta}}[|g(\cdot) - g^{J,l}(\cdot)|] \le c(J^{-q} + 2^{-l})\,.$$

We estimate $\mathbb{E}^{\gamma^{\delta}}[g^{J,l}(\cdot)]$ by performing the MCMC-FE algorithm with the acceptance probability (4.22). Then, Proposition 4.10 holds for $g^{J,l}$ since there exists

$c > 0$ such that $|g^{J,l}(u)| \le c\hat{K}(u)/\check{K}(u)$ for every $u \in \Gamma_{\boldsymbol{b}}$. Thus, the function $\mathcal{V}(u)$ can be chosen so as to majorize $c(\hat{K}(u)/\check{K}(u))^2$ which is in $L^2(U, \gamma_{\boldsymbol{b}})$.

Proposition 4.11. *For $u \in \Gamma_{\boldsymbol{b}}$ and for $\ell \in V^*$, define $g(u) := \ell(P(\cdot, u))$. Then, for every initial condition $u^{(0)}$ and for $\mathcal{P}^{J,l}_{u^{(0)}}$-almost every realization of the Markov chain, there holds the error bound*

$$\left|\mathbb{E}^{\gamma^\delta} g(\cdot) - E_M^{\gamma^{J,l,\delta}}[g^{J,l}(\cdot)]\right| \le c_1 M^{-\frac{1}{2}} + c_2\big(J^{-q} + 2^{-l}\big)\ . \tag{4.31}$$

There also holds the mean square error bound

$$\left(\mathcal{E}^{\bar{\gamma},J,l}\left[\left|\mathbb{E}^{\gamma^\delta}[g(\cdot)] - E_M^{\gamma^{J,l,\delta}}[g^{J,l}(\cdot)]\right|^2\right]\right)^{1/2} \le c_4(M^{-1/2} + J^{-q} + 2^{-l})\ . \tag{4.32}$$

Here, the constants c_i are as in Proposition 4.10.

4.5. **Complexity analysis for the MCMC FEM.** In the previous section, we assumed that the exact solution P^J to the dimensionally truncated problem (4.1) belongs to $H^2(D)$, which implied the FE convergence rate (4.5). For this to hold, we imposed Assumption 4.3 and required that the domain D is convex and $f \in L^2(D)$. However, Assumption 4.3 may be violated: for general exponential covariances or in nonconvex, polyhedral domains it does not hold. Instead, the coefficient functions of the Karhúnen-Loève expansion of the random coefficient K only belong to some Hölder classes. The solution P^J of (4.1) is thus at best in $H^{1+t}(D)$ for some $0 < t < 1$. This is also the case where the domain D is a non-convex polygon or if the function f does not belong to $L^2(D)$. For these cases, the convergence rate of the FE approximation reduces to $O(h^t) = O(N^{-t/d})$ where $N = \dim(V^l)$ denotes the number of degrees of freedom and h denotes the meshwidth, cp. (4.4). Then, the error bound on the Hellinger distance in Proposition 4.6 becomes

$$d_{Hell}(\gamma^\delta, \gamma^{J,l,\delta}) \le c(J^{-q} + 2^{-tl})\ . \tag{4.33}$$

Therefore, for this general case, in Proposition 4.10, the error estimates (4.28) and (4.29) become the $\mathcal{P}^{J,l}_{u^{(0)}}$-almost sure bound

$$\left|\mathbb{E}^{\gamma^\delta}[\ell(P(\cdot,u))] - E_M^{\gamma^{J,l,\delta}}[\ell(P^{J,l}(\cdot,u))]\right| \le c_1 M^{-1/2} + c_2(J^{-q} + 2^{-tl}) \tag{4.34}$$

where $c_1 \le c_3|\xi_M|$ with an asymptotically $N(0,1)$ random variable ξ_M, and the mean square bound

$$\left(\mathcal{E}^{\bar{\gamma},J,l}\left[\left|\mathbb{E}^{\gamma^\delta}[\ell(P(\cdot,u))] - E_M^{\gamma^{J,l,\delta}}[\ell(P^{J,l}(\cdot,u))]\right|^2\right]\right)^{1/2} \le c_4(M^{-1/2} + J^{-q} + 2^{-tl}), \tag{4.35}$$

respectively.

We next develop a complexity analysis for the MCMC FE procedure proposed above for estimating the expectation with respect to the posterior probability measure of a function $\Gamma_{\boldsymbol{b}} \ni u \mapsto g(u) = \ell(P(u))$ for $\ell \in V^*$, ie. for a bounded linear functional of the parametric solution. As the present, lognormal parametric dependence (3.1) does not allow for uniform (w.r. to $u \in \Gamma_{\boldsymbol{b}}$) work bounds, we opt for error vs. work bounds that are *probabilistic*, unlike the results in [14]. To this end, we address the cost for the numerical solution of the FE equations (4.6), which amounts to generating, for each instance of $u \in \Gamma_{\boldsymbol{b}}$ proposed by the Markov Chain, the "stiffness matrix"

$$\mathbf{A}^{J,l}(u) = \left(\int_D K^J(x,u)\nabla w_k^l(x) \cdot \nabla w_{k'}^l(x)dx : k, k' \in I^l\right)\ , \quad u \in \Gamma_{\boldsymbol{b}}\ . \tag{4.36}$$

From (3.8), $\mathbf{A}^{J,l}$ is a symmetric positive definite matrix, for all $u \in \Gamma_{\boldsymbol{b}}$, J and l in $\mathbb{N}$. In each numerical realization of (4.6), for each proposal $u \in U$ of the sampler, one realization of the "stiffness" matrix $\mathbf{A}^{J,l}(u)$ must be computed and inverted. Assuming exact evaluation of matrix entries (see, however, Remark 4.15), the linear system with matrix $\mathbf{A}^{J,l}(u)$ will not be inverted exactly, but rather solved iteratively except in space dimension $d = 1$ (when $\mathbf{A}^{J,l}(u)$ is tridiagonal, symmetric and, for every realization $u \in \Gamma_{\boldsymbol{b}}$, positive definite). The iteration is stopped when the iteration error will match the discretization errors. Under Assumption 3.1 we have (3.8), i.e. $\mathbf{A}^{J,l}(u)$ is symmetric, positive definite for all $u \in \Gamma_{\boldsymbol{b}}$. We opt for iterative solution by preconditioned conjugate gradient (CG) iteration. For preconditioning and in order to relate euclidean norm of the residual vector in the iteration to the V-norm of the corresponding Finite Element error, we assume available a Riesz basis for $V = H^1_0(D)$ spanning the finite element spaces V^l.

Assumption 4.12. *(Riesz Basis Property in V) For each $L \in \mathbb{N}_0$ there exists a set of indices $\mathcal{I}^L \subset \mathbb{N}^d$ of cardinality $N_L = O(2^{Ld})$ and a hierarchic family of basis functions $w^L_k \in H^1_0(D)$ indexed by a multi-index $k \in \mathcal{I}^L$ such that $V^L = \mathrm{span}\{w^L_k : k \in \mathcal{I}^L\}$. The collection $\bigcup_{l\geq 0}\{w^l_k : k \in \mathcal{I}^l\}$ constitutes a Riesz-basis for $L^2(D)$ which, when rescaled according to $\{2^l w^l_k : k \in \mathcal{J}^l := \mathcal{I}^l \setminus \mathcal{I}^{l-1}\}_{l\geq 0}$, becomes a Riesz-basis for the space V, i.e. there exist constants $0 < \check{c}_{Riesz} \leq \hat{c}_{Riesz} < \infty$ which are independent of $L \in \mathbb{N}$ such that*

$$\check{c}_{Riesz} \sum_{k\in\mathcal{I}^L} |c^L_k|^2 \leq \|w^L\|^2_{L^2(D)} \leq \hat{c}_{Riesz} \sum_{k\in\mathcal{I}^L} |c^L_k|^2 \,, \tag{4.37}$$

and

$$\check{c}_{Riesz} \sum_{l=0}^{L} 2^{2l} \sum_{k\in\mathcal{J}^l} |c^l_k|^2 \leq \|w^L\|^2_V \leq \hat{c}_{Riesz} \sum_{l=0}^{L} 2^{2l} \sum_{k\in\mathcal{J}^l} |c^l_k|^2 \,. \tag{4.38}$$

Concrete constructions of such bases, in polygonal and polyhedral domains D can be found in [17] and [23] and the references there. To bound the number of nonzero entries in the stiffness matrix (and thus the number of float point operations necessary for one matrix-vector multiplication), we impose further assumptions on the overlap of the support of the basis functions.

Assumption 4.13. *(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 $\mathrm{supp}(w^l_k) \cap \mathrm{supp}(w^{l'}_{k'})$ has positive measure for at most $O(\max(1, 2^{l'-l}))$ values of k'.*

We emphasize that the support overlap and the Riesz basis assumptions 4.12 and 4.13 hold independently of the realization of $u \in \Gamma_{\boldsymbol{b}}$. In the ensuing discussion of the complexity of the MCMC-FEM, we denote by $N_{dof} = \#(I^l) = O(h^{-d})$ the number of degrees of freedom (or the number of unknowns) which are to be determined in the Galerkin approximation (4.6). Assumption 4.13 implies that a) the number of nonvanishing entries of the matrix $A(u)$ in (4.36) is of $O(N_{dof} \log N_{dof})$, for any $J \in \mathbb{N}$ and for any $u \in \Gamma_{\boldsymbol{b}}$, and b) the number of float point operations of a matrix vector multiplication is likewise bounded by $O(N_{dof} \log N_{dof})$ operations, with the constant implied in $O(\cdot)$ being independent of $u \in \Gamma_{\boldsymbol{b}}$. Denote by $x \in \mathbb{R}^{N_{dof}}$ the solution vector corresponding to the FE equation (4.6), ie. the solution of $\mathbf{A}^{J,l}(u)\boldsymbol{x} = \mathbf{f}$ for some proposal $u \in \Gamma_{\boldsymbol{b}}$ of the Markov chain and for the load vector $\mathbf{f} = \{(f, w^l_k) : k \in I^l\}$. Due to the Riesz basis property (4.37), we may insert a diagonal preconditioner $\mathbf{D}$ and apply the CG algorithm [9, Alg. 10.2.1] to the *diagonally preconditioned* linear system $\mathbf{D}^{-1/2}\mathbf{A}^{J,l}(u)\mathbf{D}^{-1/2}\mathbf{D}^{1/2}\boldsymbol{x} = \mathbf{D}^{-1/2}\mathbf{f}$. For every $u \in \Gamma_{\boldsymbol{b}}$, CG iterations for this linear system generate, starting from an arbitrary initial choice $\boldsymbol{x}_0 \in \mathbb{R}^{N_{dof}}$, a sequence $\{x_j\}_{j\geq 0}$ which converges to $\boldsymbol{x}$. For every $u \in \Gamma_{\boldsymbol{b}}$, there holds (see, eg. [9, Thm. 10.2.6]) for iteration error at step j, $e_j := x - x_j \in \mathbb{R}^{N_{dof}}$,

$$\|e_j\|_{\mathbf{A}^{J,l}(u)} \leq 2\left(\frac{\sqrt{\kappa(u)}-1}{\sqrt{\kappa(u)}+1}\right)^j \|e_0\|_{\mathbf{A}^{J,l}(u)}\;, \quad j = 1, 2, \ldots\;. \tag{4.39}$$

In (4.39), we denote for a SPD matrix $\mathbf{B}$, the "energy" norm $\|x\|_{\mathbf{B}} := \sqrt{x^\top \mathbf{B} x}$ and, for $u \in \Gamma_{\boldsymbol{b}}$, $\kappa = \kappa(u) \geq 1$ denotes the condition number of the matrix $\mathbf{D}^{-1/2}\mathbf{A}^{J,l}(u)\mathbf{D}^{-1/2}$ arising in (4.6). This condition number is independent of J and of l by the Riesz basis assumption (4.38), but depends on the realization $u \in \Gamma_{\boldsymbol{b}}$ of the uncertain input. Specifically, denote by $\mathbf{D}$ the block diagonal matrix with entries 2^l on diagonal elements with indices $k \in J^l$. Then, under Assumption 4.12, there exists a constant $C > 0$ such that for all $J, l \in \mathbb{N}$ and for all $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\kappa(u) = \mathrm{cond}_2(\mathbf{D}^{-1/2}\mathbf{A}^{J,l}(u)\mathbf{D}^{-1/2}) = \frac{\lambda_{max}(\mathbf{D}^{-1/2}\mathbf{A}^{J,l}(u)\mathbf{D}^{-1/2})}{\lambda_{min}(\mathbf{D}^{-1/2}\mathbf{A}^{J,l}(u)\mathbf{D}^{-1/2})} \leq C\frac{\hat{K}(u)}{\check{K}(u)}\;.$$

This means that with diagonal preconditioning, starting from any initial vector $\boldsymbol{x}_0$ the CG iteration produces a sequence $\{\boldsymbol{x}_j\}_{j\geq 0}$ which converges in the norm $\|\circ\|_2$ to the solution $\boldsymbol{x}$ of the linear system, *at a rate which is independent of the discretization level* but *which will in general depend on the proposal of the chain.* This is different from the situation encountered in [14] and will only allow probabilistic work estimates for the MCMC-FEM.

We equilibrate the terms in the error bounds (4.34) and (4.35), to obtain for a FE discretization at level l of mesh refinement, with spatial solution regularity $H^{1+t}(D)$ in (4.4). This leads to the choices

$$J = O(2^{tl/q})\;, \quad M = O(2^{2tl})\;, \qquad l \in \mathbb{N}\;. \tag{4.40}$$

Starting the CG iteration with $\boldsymbol{x}_0 = 0$, and denoting by $P_j^{J,l}(u) \in V^l$ the FE solution corresponding to iteration vector $\boldsymbol{x}_j$, from (4.38) follows for any fixed $u \in \Gamma_{\boldsymbol{b}}$

$$\begin{aligned}\|P^{J,l}(u) - P_j^{J,l}(u)\|_V^2 &\simeq \|e_j\|^2_{\mathbf{A}^{J,l}(u)}\\ &\lesssim \left(\frac{\sqrt{\kappa(u)}-1}{\sqrt{\kappa(u)}+1}\right)^{2j} \|e_0\|^2_{\mathbf{A}^{J,l}(u)}\\ &\simeq \left(\frac{\sqrt{\kappa(u)}-1}{\sqrt{\kappa(u)}+1}\right)^{2j} \|P^{J,l}(u)\|_V^2\\ &\lesssim \left(\frac{\sqrt{\kappa(u)}-1}{\sqrt{\kappa(u)}+1}\right)^{2j} \frac{1}{\check{K}(u)^2}\|f\|^2_{L^2(D)}\;.\end{aligned}$$

Here, the constants implied in $\simeq$ and $\lesssim$ are independent of J, l and of $u \in \Gamma_{\boldsymbol{b}}$. To ensure that the error due to stopping the preconditioned conjugate gradient iteration after j^* many steps be smaller than the error (4.35), we find that a sufficient condition (based on the bound (4.39)) is

$$(4.41) \qquad j^*(u) \geq C(|\log h| + |\log \check{K}(u)|)\sqrt{\frac{\hat{K}(u)}{\check{K}(u)}}\,, \quad \text{for given } u \in \Gamma_{\boldsymbol{b}}\,.$$

Here $C > 0$ is independent of h, l and of $u \in \Gamma_{\boldsymbol{b}}$ (but depends on t, $\check{c}_{Riesz}$ and $\hat{c}_{Riesz}$ in (4.37). Since $\hat{K}, \check{K}^{-1} \in L^p(U; \gamma_{\boldsymbol{b}})$ for every $0 < p < \infty$ [2, Prop.2.3], the lower bound in (4.41) is strongly $\gamma_{\boldsymbol{b}}$-measurable.

We estimate the expected (under the gaussian measure γ) work involved in performing j^* steps of the CG iteration. For any $u \in \Gamma_{\boldsymbol{b}}$, the stiffness matrix $\mathbf{A}^{J,l}(u)$ in (4.36) has $O(l^{d-1}2^{dl})$ non zero entries (with $O(\cdot)$ uniform w.r. to u). Assuming a fixed number of quadrature points per matrix entry, the computation of each of these entries requires $O(J)$ many floating point operations, due to the necessary evaluation of the J-term truncation of the parametric coefficient $K(u)$. Therefore the number of floating point operations for forming the stiffness matrix $\mathbf{A}^{J,l}(u)$ is bounded by $O(Jl^{d-1}2^{dl})$, with $O(\cdot)$ uniform w.r. to $u \in \Gamma_{\boldsymbol{b}}$. Then, the number of floating point operations at each CG step is bounded by $O(l^{d-1}2^{dl})$. This leads to a *bound for the total work for the approximate solution of the linear system* corresponding to the Galerkin equation (4.6), for *any single, given proposal* $v \in \Gamma_{\boldsymbol{b}}$ *of the Markov chain* (with the constants implied in $\lesssim$ being independent of $J, l \in \mathbb{N}$ and of $v \in \Gamma_{\boldsymbol{b}}$)

(4.42)

$$\begin{aligned} W_{PCG}(v) &\sim Jl^{d-1}2^{dl} + (\hat{K}(v))^{1/2}(\check{K}(v))^{-1/2}(l + |\log \check{K}(v)|)l^{d-1}2^{dl} \\ &\lesssim Jl^{d-1}2^{dl} + (\hat{K}(v))^{1/2}(\check{K}(v))^{-1/2}(1 + |\log \check{K}(v)|)l^{d}2^{dl} \\ &\sim l^{d-1}2^{tl/q+ld} + (\hat{K}(v))^{1/2}(\check{K}(v))^{-1/2}(1 + |\log \check{K}(v)|)l^{d}2^{ld}\,. \end{aligned}$$

Multiplying this with the (deterministic) bound (4.40) on the number M of steps in the chain that is sufficient in order for the chain to attain (in mean square) the FE discretization error, we find bounds on the expectation (with respect to the probability space generating the randomness of the samples $v^{(k)}$ which are distributed independently, identically according to the Gaussian prior γ) of the total work required for running the chain to convergence in mean square.

Solving approximately the Galerkin discretized forward equation for each proposal $u \in \Gamma_{\boldsymbol{b}}$ produced by the chain with $j^*(u)$ CG iterations as in (4.41), we realize numerically the approximate posterior measure $\gamma_*^{J,l,\delta}$ defined by

$$\frac{d\gamma_*^{J,l,\delta}}{d\gamma} \propto \exp(-\Phi_*^{J,l}(u;\delta)) \quad \text{where} \quad \Phi_*^{J,l}(u;\delta) = \frac{1}{2}|\delta - \mathcal{G}_*^{J,l}(u;\delta)|$$

and

$$\mathcal{G}_*^{J,l}(u;\delta) = (\mathcal{O}_1(P^{J,l}_{j^*(u)}(\cdot,u)), \ldots, \mathcal{O}_k(P^{J,l}_{j^*(u)}(\cdot,u)))\,.$$

Arguing as in the proof of Proposition 4.6, we obtain a bound on the Hellinger distance between the true posterior γ^δ and the approximate posterior measure $\gamma_*^{J,l,\delta}$ obtained numerically by dimension truncation, Galerkin-discretization and incomplete iterative solution of the linear systems,

$$d_{\text{Hell}}(\gamma^\delta, \gamma_*^{J,l,\delta}) \leq c(J^{-q} + 2^{-tl})\,.$$

We therefore deduce that for $\ell \in V^*$

$$|\mathbb{E}^{\gamma^\delta}[\ell(P(\cdot,u))] - \mathbb{E}^{\gamma_*^{J,l,\delta}}[\ell(P^{J,l}_{j^*(u)}(\cdot,u))]| \leq c(J^{-q} + 2^{-tl})\,.$$

Therefore if for each proposal $v^{(k)} \in \Gamma_{\boldsymbol{b}}$ of the MCMC process, $j^*(v^{(k)})$ steps where $j^*()$ is assumed to be the lower bound in (4.41) of the CG iteration are performed, we get (4.34) and (4.35) with $E_M^{\gamma^{J,l,\delta}}[\ell(P^{J,l}(\cdot,u))]$ being replaced by $E_M^{\gamma_*^{J,l,\delta}}[\ell(P^{J,l}_{j^*(u)}(\cdot,u))]$. This, in turn, leads to a bound for the *expected* (over all proposals $u \in \Gamma_{\boldsymbol{b}}$ and with respect to the Bayesian prior $\gamma_{\boldsymbol{b}}$ in (3.4)) work for performing one step of the MCMC FE process given by

$$\mathbb{E}_{\gamma_{\boldsymbol{b}}}[W_{PCG}(\cdot)] \le Cl^{d-1}2^{2tl+dl+tl/q} \,.$$

As the proposals are chosen independently, the expected work under the probability measure of the space of sequences of M independently generated proposals $\{v^{(k)}\}$ for performing M steps of the MCMC process on the discretized forward problem is bounded by $CMl^{d-1}2^{2tl+dl+tl/q}$. For the probabilistic convergence estimate of (4.34), we balance the error contributions in (4.28), and choose $J = \lceil 2^{tl/q} \rceil$ and $M = \lceil s^2 N_{dof}^{2t/d} \rceil$ where $s = c_3|\xi_M|$. To attain the mean square convergence estimate (4.29), we choose $J = \lceil 2^{tl/q} \rceil$ and $M = \lceil 2^{2tl} \rceil$. With these choices, the *expected number of floating point operations under the gaussian prior* $\gamma_{\boldsymbol{b}}$ on $(\Gamma_{\boldsymbol{b}}, \mathcal{B}(\Gamma_{\boldsymbol{b}}))$ is bounded by $Cs^2l^{d-1}2^{tl+dl+tl/q}$ to attain the almost sure convergence bound (4.34) and by $Cl^{d-1}2^{tl+dl+tl/q}$ to attain the mean square convergence bound (4.35).

Theorem 4.14. *Under Assumptions 4.12 and 4.13, for* $g(u) = \ell(P(u))$ *where* $\ell \in V^*$*, given data* $\delta \in \mathbb{R}^k$*, for any* $s > 0$ *with probability* $p_{N_{dof}}(s)$*, the conditional expectation* $\mathbb{E}^{\gamma^\delta}[g(u)]$ *can be approximated by solving discretized forward problems* (4.6) *with* N_{dof} *degrees of freedom per step of the MCMC algorithm, with* $s^2 N_{dof}^{2t/d}$ *MCMC steps (with a total of* $s^2 N_{dof}^{1+2t/d}$ *degrees of freedom), incurring an error of* $O(N_{dof}^{-t/d})$ *and the expectation of the total number of floating point operations with respect to the sequence of independent proposals* $v^{(k)}$ *is not larger than*

$$cs^2(\log N_{dof})^{d-1}N_{dof}^{1+t(2+1/q)/d},$$

where

$$\lim_{N_{dof}\to\infty} p_{N_{dof}}(s) \to \int_{-c's}^{c's} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx,$$

for a positive constant c' *independent of* N_{dof} *and of* s.

In the mean square with respect to the measure $\mathcal{P}^{\tilde{\gamma},J,l}$, $\mathbb{E}^{\gamma^\delta}[g(u)]$ *can be approximated with an error* $O(N_{dof}^{-t/d})$*, using not more than* $O(N_{dof}^{1+2t/d})$ *number of degrees of freedom in total, and the expectation of the total number of floating point operations with respect to the sequence of random proposals* $v^{(k)}$ *is bounded by* $O((\log N_{dof})^{d-1}N_{dof}^{1+t(2+1/q)/d})$.

Remark 4.15. The preceding error vs. work analysis was performed under the assumption of exact numerical integration, with $O(1)$ cost per integral. In practice, however, the entries $A_{ij}^{J,l}(u)$ of the stiffness matrix in (4.36) can not be evaluated exactly and numerical integration is used. Under Assumption 4.3, the parametric coefficient $K^J(\cdot,u)$ is, for $J \in \mathbb{N} \cup \{\infty\}$ and for every proposal $u \in \Gamma_{\bar{\boldsymbol{b}}}$, Lipschitz w.r.t. to $x \in D$. The matrix entries $A_{ij}^{J,l}(u)$ in (4.36) may be replaced by numerically integrated approximations $\tilde{A}_{ij}^{J,l}(u)$ which are obtained, for example, by the midpoint rule (the $d+1$ point quadrature rule averaging the values of $K^J(\cdot,u)$ at the vertices of $T \in \mathcal{T}^l$ results in the same error bound). Denoting by $T \subset D$ a simplex in $\mathcal{T}^l$, we set

$$\tilde{A}^{J,l}_{ij}(u) := \sum_{|T\cap S(i,j)|>0} |T \cap S(i,j)| K^J(u, x_T) \nabla w^l_i|_T \cdot \nabla w^l_j|_T \ , \tag{4.43}$$

where $S(i,j)$ denotes the intersection of the supports of the one-scale basis functions w^l_i and w^l_j and $x_T = |T|^{-1} \int_T x$ denotes the barycenter of T. Note that the gradients $\nabla w^l_i|_T$ and $\nabla w^l_j|_T$ are constant in T. With approximation (4.43), we estimate for every $u \in \Gamma_{\bar{b}}$ its impact on the parametric FE approximation for proposal u by the Strang Lemma. To this end, we denote by $P^{J,l} \in V^l$ the $H^1_0(D)$-projection of P^J onto V^l and define $g^J(x,u) := \log(K^J(x,u) - K_*(x,u))$. By $a(u;\cdot,\cdot)$ we denote the parametric bilinear form corresponding to (3.5) with exact coefficient K in (3.1), and with $a^J(u;\cdot,\cdot)$ its approximation with coefficient K^J in (4.1), and $a^{J,l}(u;\cdot,\cdot)$ its approximation with coefficient $K^J(\cdot,u)$ approximated by the stepfunction of its values $K^J(u,x_T)$ in the barycenter x_T for $T \in \mathcal{T}^l$. Then, for abitrary $w^l \in V^l$, we estimate

$$\begin{aligned}
&\left|a^J(u;P^{J,l},w^l) - a^{J,l}(u;P^{J,l},w^l)\right| \\
&\le \sum_{T\in\mathcal{T}^l} \left| \int_T (\exp(g^J(x;u)) - \exp(g^J(x_T;u))) \nabla P^{J,l} \cdot \nabla w^l dx \right| \\
&\lesssim \sum_{T\in\mathcal{T}^l} \|\exp(g^J(\cdot;u)) - \exp(g^J(x_T;u))\|_{L^\infty(T)} \left| \int_T \nabla P^{J,l} \cdot \nabla w^l dx \right| \ .
\end{aligned}$$

Here, and in the remainder of this remark, the constant implied in $\lesssim$ is independent of J, L and of $u \in \Gamma_{\boldsymbol{b}}$. For every $u \in \Gamma_{\bar{\boldsymbol{b}}}$ and for every $J \in \mathbb{N} \cup \{\infty\}$ the Lipschitz regularity of $x \mapsto g^J(x;u)$ implies for every $T \in \mathcal{T}^l$

$$\|\exp(g^J(\cdot;u)) - \exp(g^J(x_T;u))\|_{L^\infty(T)} \lesssim h_T \|\exp(g^J(\cdot;u))\|_{L^\infty(T)} \|\nabla g^J(\cdot;u)\|_{L^\infty(T)} \ .$$

The shape regularity of $\mathcal{T}^l$ implies that $\max\{h_T : T \in \mathcal{T}^l\} \lesssim 2^{-l}$, with the constant implied in $\lesssim$ depending only on the shape regularity parameter of the family $\{\mathcal{T}^l\}_{l\ge 0}$. We estimate for every $T \in \mathcal{T}^l$ and for every $u \in \Gamma_{\bar{\boldsymbol{b}}}$

$$\begin{aligned}
& \|\exp(g^J(\cdot;u)) - \exp(g^J(x_T;u))\|_{L^\infty(D)} \\
= \ & \max_{T\in\mathcal{T}^l} \|\exp(g^J(\cdot;u)) - \exp(g^J(x_T;u))\|_{L^\infty(T)} \\
\lesssim \ & 2^{-l} \exp\left(\sum_{j=1}^J u_j(b_j + \bar{b}_j) \right) \ .
\end{aligned}$$

Inserting into the above bound, and using Cauchy-Schwarz, we obtain for every $u \in \Gamma_{\bar{\boldsymbol{b}}}$ and for every $J \in \mathbb{N} \cup \{\infty\}$ the Strang type consistency error bound

$$\sup_{0\ne w^l\in V^l} \frac{\left|a^J(u;P^{J,l},w^l) - a^{J,l}(u;P^{J,l},w^l)\right|}{\|w^l\|_V} \lesssim 2^{-l} \exp\left(\sum_{j=1}^J u_j(b_j + \bar{b}_j) \right) \|P^{J,l}(\cdot,u)\|_V \ . \tag{4.44}$$

The projection $P^{J,l}(\cdot,u)$ also satisfies the a-priori bound (4.2) with $\check{K}(u)$ as in (3.7). The consistency bound (4.44) is of the same type as the FE error bounds in Section 4.2, indicating that the present convergence analysis remains valid under the same hypotheses also in the presence of the one-point quadrature approximation (4.43).

5. Concluding Remarks

We close the paper by a summary of the principal conclusions, and we indicate straightforward corollaries from the present results.

We presented a convergence and complexity analysis of Finite Element Markov Chain Monte Carlo methods for Bayesian inversion of second order, elliptic diffusion problems with isotropic, log-normal gaussian models for the uncertain diffusion coefficient in a bounded domain $D \subset \mathbb{R}^d$. We considered continuous, piecewise linear Finite Elements for the primal discretization of the forward map (relating realizations of the random diffusion coefficient to functionals of the solution), and we assumed finite-dimensional, additive centered gaussian observation noise and gaussian prior γ. The present analysis extends directly to (frequently employed in practice) *mixed* FE discretizations of the forward problem.

We approximated the log-normal diffusion coefficient $\log K$ in (3.1) by its J-term truncated Karhúnen-Loève expansion $\log K^J$ in (4.1). We proved geometric ergodicity of the Markov chains running on the Finite Element discretization of the dimension truncated forward problem (4.6).

We assumed γ-a.s. $W^{1,\infty}(D)$ regularity of the uncertain diffusion coefficient. This assumption was made only to avoid unnecessary technicalities. This $W^{1,\infty}(D)$ regularity is necessary to achieve full, first order convergence of the Finite Element discretization. It could be weakened, leading to corresponding reduced rates of convergence. If higher regularity were available γ-a.s. (e.g. for gaussian fields $\log K$ with smooth covariance), higher order Finite Element discretizations would allow for an analogous convergence analysis, with improved convergence rates.

The FE discretized forward problems were solved iteratively, by multi level preconditioning (using a wavelet Finite Element basis). Due to the unbounded parameter ranges which occur in the gaussian proposals, complexity estimates for discretization and iterative solution are proposal-dependent, mandating a *probabilistic complexity analysis* of the Finite Element solver which we provided in Section 4.5. Analogous probabilistic complexity analyses will apply also to other numerical methods for the solution of lognormal diffusion problems, such as multilevel Monte-Carlo and Quasi Monte-Carlo methods developed in [21], [11] and the references there. Error bounds in mean square and a.s. with respect to the probability of the space that describes the randomness of the Markov chain were obtained. All constants implied in our error and complexity bounds are independent of the discretization parameters in the forward models and of the proposals $u^{(k)} \in \Gamma_{\boldsymbol{b}}$ generated by the chain. They depend, however, implicitly on the observation noise covariance $\Sigma > 0$ in (1.3) and in general degenerate for $\Sigma \to 0$. The presently considered approximation error bounds accounted for a J-term truncation $K^J(x,u)$ of the Karhúnen-Loève expansion (3.1) of the gaussian random field $\log(K - K_*)(u)$. On uniform regular simplicial partitions of axiparallel rectangular domains D into simplices, and with numerical quadrature (4.43), *circulant embedding* methods can be used to sample the random coefficient $K(\cdot,u)$ in each element barycenter x_T, $T \in \mathcal{T}^l$, *without dimensionally truncating the Karhúnen-Loève representation* and at *cost proportional to* $O(l2^{ld})$, ie. log-linear in $\#(\mathcal{T}^l)$, see [10, Sec. 5.1,5.2] and the references there for details. The quadrature error analysis in Remark 4.15 indicates that the preceding error bounds also hold for the resulting "circulant embedding" MCMC method. As is readily seen, for the "circulant embedding" MCMC, all error versus work bounds in Section 4.5 remain valid, however *with q formally set* *to infinity*: the present analysis also applies to this case, giving the same bounds, without the dimension truncation term J^{-q}.

Isotropy of the uncertain coefficient K in (3.1) and convexity of the physical domain D are not essential in our MCMC convergence analysis: analogous results hold true for second order, divergence form problems with matrix coefficient K, and for nonconvex domains, provided that elliptic regularity is quantified in terms of weighted Sobolev spaces in D, and that Finite Element discretizations with local mesh refinements are used to resolve corner and edge singularities.

The present paper did not include numerical experiments. Such experiments, as well as an extension of the present approach to a multi-level MCMC setting will be presented in [12].

6. Appendix A

We prove Proposition 4.7 in this appendix.

Lemma 6.1. *Let $U = \Gamma_{\boldsymbol{b}}$. For each $n \in \mathbb{N}$, there is a constant $c(n)$ such that for all functions $g \in L^2(U; \gamma_{\boldsymbol{b}})$, $g \geq 0$,*

$$|p^n(u,\cdot)(g)| \leq c(n) \int_U g(v) d\gamma_{\boldsymbol{b}}(v) + (1-c_0)^n g(u) \tag{6.1}$$

where c_0 is the constant in the proof of Theorem 2.4 .

Proof We prove this Lemma by induction. When $n = 1$:

$$\begin{aligned} |p(u,\cdot)(g)| = \int_U \alpha(u,v) g(v) d\gamma_{\boldsymbol{b}}(v) + (1 - \int_U \alpha(u,v) d\gamma_{\boldsymbol{b}}(v)) g(u) \\ \leq \int_U g(v) d\gamma_{\boldsymbol{b}}(v) + (1-c_0) g(u) \ . \end{aligned}$$

Assuming that (6.1) holds for $n-1$, then

$$\begin{aligned} &|p^n(u,\cdot)(g)| = \int_U \alpha(u,v) p^{n-1}(v,\cdot)(g) d\gamma_{\boldsymbol{b}}(v) + (1 - \int_U \alpha(u,v) d\gamma_{\boldsymbol{b}}(v)) p^{n-1}(u,\cdot)(g) \\ &\leq \int_U \Big(c(n-1) \int_U g(w) d\gamma_{\boldsymbol{b}}(w) + (1-c_0)^{n-1} g(v) \Big) d\gamma_{\boldsymbol{b}}(v) \\ &\qquad + (1-c_0) \Big(c(n-1) \int_U g(v) d\gamma_{\boldsymbol{b}}(v) + (1-c_0)^{n-1} g(u) \Big) \\ &\leq c(n) \int_U g(v) d\gamma_{\boldsymbol{b}}(v) + (1-c_0)^n g(u) \end{aligned}$$

where $c(n) = c(n-1) + (1-c_0)^{n-1} + (1-c_0)c(n-1)$ which is independent of g. This proves (6.1). □

In the remainder of the paper, $c(n)$ is as in (6.1).

Let $a > 0$ be an arbitrary constant, to be fixed in the proof of Lemma 6.6 ahead. Define the function $\mathcal{V}_0 : \Gamma_{\boldsymbol{b}} \to \mathbb{R}$ by

$$\mathcal{V}_0(u) = \begin{cases} \exp\Big(a \sum_{j=1}^\infty (b_j + \bar{b}_j) |u_j| \Big) & \text{if} \quad u \in \Gamma_{\bar{\boldsymbol{b}}} \\ \exp\Big(a \sum_{j=1}^\infty b_j |u_j| \Big) & \text{if} \quad u \in \Gamma_{\boldsymbol{b}} \setminus \Gamma_{\bar{\boldsymbol{b}}}, \end{cases} \tag{6.2}$$

and, for J fixed independent of $u \in \Gamma_{\boldsymbol{b}}$, and for $\varepsilon = B\sum_{j>J} b_j$ with a constant $B > 1$ to be selected the function

(6.3)
$$\mathcal{V}(u) = \begin{cases} \exp\left(a\sum_{j=1}^\infty (b_j + \bar{b}_j)|u_j| + \frac{1}{\varepsilon}\sum_{j>J} b_j|u_j|\right) & \text{if} \quad u \in \Gamma_{\bar{\boldsymbol{b}}} \\ \exp\left((a\sum_{j=1}^\infty b_j + a\sum_{j=1}^J \bar{b}_j)|u_j| + \frac{1}{\varepsilon}\sum_{j>J} b_j|u_j|\right) & \text{if} \quad u \in \Gamma_{\boldsymbol{b}} \setminus \Gamma_{\bar{\boldsymbol{b}}}, \end{cases} .$$

We have

Lemma 6.2. *For every $B > 1$, the function $\mathcal{V}(u)$ defined in* (6.3) *satisfies $\|\mathcal{V}\|_{L^2(U,\gamma_{\boldsymbol{b}})} < C(a)$ where $C(a)$ is a constant that only depends on a.*

Proof From (4.19), with J as in (6.3),

$$\int_U \mathcal{V}(u)^2 d\gamma_{\boldsymbol{b}}(u) \le \prod_{j=1}^J \int_{\mathbb{R}} \exp((2ab_j + 2a\bar{b}_j)|u_j|)d\gamma_1 \prod_{j>J}\int_{\mathbb{R}} \exp((2ab_j + 2a\bar{b}_j + 2b_j/\varepsilon)|u_j|)d\gamma_1$$
$$\le \exp\left(\sum_{j=1}^J (2ab_j + 2a\bar{b}_j)^2/2 + (2ab_j + 2a\bar{b}_j)\sqrt{2/\pi}\right)$$
$$\times \exp\left(\sum_{j>J}(2ab_j + 2a\bar{b}_j + 2b_j/\varepsilon)^2/2 + (2ab_j + 2a\bar{b}_j + 2b_j/\varepsilon)\sqrt{2/\pi}\right)$$
$$\le \exp\left(6\sum_{j=1}^\infty a^2(b_j^2 + \bar{b}_j^2) + 6\sum_{j>J} b_j^2/\varepsilon^2 + 2a\sqrt{2/\pi}\sum_{j=1}^\infty (b_j + \bar{b}_j) + 2\sqrt{2/\pi}\sum_{j>J} b_j/\varepsilon\right) .$$

As $\boldsymbol{b} \in \ell^1$, for $\varepsilon = B\sum_{j>J} b_j$ with $B > 1$ to be fixed ahead, there holds

$$\sum_{j>J} b_j^2/\varepsilon^2 < \sum_{j>J} b_j/\varepsilon < 1 . \tag{6.4}$$

Thus $\|\mathcal{V}\|_{L^2(U;\gamma_{\boldsymbol{b}})} < C(a)$ where $C(a)$ depends only on a. □

Lemma 6.3. *There are constants $0 < \tau' < 1$ and $n_1 = n_1(\tau') \in \mathbb{N}$ such that for all $n \ge n_1$ there is a constant $B_1(n, a) > 1$ such that for $B > B_1(n, a)$ there holds*

$$\|p^n(u, \cdot) - \gamma^\delta\|_{\mathcal{V}} \le \tau'\mathcal{V}(u) \quad \forall u \in \Gamma_{\boldsymbol{b}}, \tag{6.5}$$

where p is the transition density of the Markov chain obtained from the MCMC procedure with the acceptance probability α in (2.1).

Proof From the geometric ergodicity property (2.4) of the Markov chain with the transition kernel p, with respect to the function $\mathcal{V}_0$, we have

$$\lim_{n\to\infty} \sup_{u \in \Gamma_{\boldsymbol{b}}} \|p^n(u, \cdot) - \gamma^\delta\|_{\mathcal{V}_0}/\mathcal{V}_0(u) = 0 .$$

Thus for every $0 < \tau < 1$ there exists $n_0(\tau) \in \mathbb{N}$ so that

$$\sup_{u\in\Gamma_{\boldsymbol{b}}} \frac{\|p^n(u, \cdot) - \gamma^\delta\|_{\mathcal{V}_0}}{\mathcal{V}_0(u)} \le \tau < 1 \quad \forall\, n > n_0(\tau) . \tag{6.6}$$

Since $\mathcal{V}_0$ is independent of J and l, the constants τ and n_0 can be chosen independent of J and l. As $\gamma(\Gamma_{\bar{\boldsymbol{b}}}) = 1$, we note that

$$\int_U (\mathcal{V}(v) - \mathcal{V}_0(v))d\gamma_{\boldsymbol{b}}(v) = \int_{\Gamma_{\bar{\boldsymbol{b}}}} (\mathcal{V}(v) - \mathcal{V}_0(v))d\gamma_{\boldsymbol{b}}(v) \le \int_{\Gamma_{\bar{\boldsymbol{b}}}} \mathcal{V}(v)(1 - \exp(-\frac{1}{\varepsilon}\sum_{j>J}(b_j|v_j|)))d\gamma_{\boldsymbol{b}}(v) .$$

Using $1-\exp(-x/\varepsilon)\le x/\varepsilon$ for all $x>0$ and $\varepsilon>0$, we have

$$\begin{aligned}
&\int_U (\mathcal{V}(v)-\mathcal{V}_0(v))d\gamma_{\boldsymbol{b}}(v) \le \frac{1}{\varepsilon}\int_{\Gamma_{\bar{\boldsymbol{b}}}} \mathcal{V}(v)\Big(\sum_{j>J} b_j|v_j|\Big)d\gamma_{\boldsymbol{b}}(v)\\
&\le \frac{1}{\varepsilon}\sum_{j>J} b_j \prod_{k=1}^{J}\int_{\mathbb{R}} \exp(a(b_k+\bar{b}_k)|v_k|)d\gamma_1(v_k)\\
&\times \prod_{k>J,k\neq j}\int_{\mathbb{R}}\exp((a(b_k+\bar{b}_k)+\frac{1}{\varepsilon}b_k)|v_k|)d\gamma_1(v_k)\\
&\qquad\times\int_{\mathbb{R}}|v_j|\exp((a(b_j+\bar{b}_j)+\frac{1}{\varepsilon}b_j)|v_j|)d\gamma_1(v_j)\ .
\end{aligned}$$

Using (4.19) and (4.21), to estimate the last expression, we get (with the absolute constant $c>0$ of (4.21) and with $B>1$ in (6.3))

$$\begin{aligned}
&\int_U (\mathcal{V}(v)-\mathcal{V}_0(v))d\gamma_{\boldsymbol{b}}(v)\\
\le\ & \frac{c}{\varepsilon}\sum_{j>J} b_j \exp\big(((a+1/\varepsilon)b_j+a\bar{b}_j)^2/2\big)\,(1+(a+1/\varepsilon)b_j+a\bar{b}_j)\\
\times\ & \prod_{k=1}^{J}\exp\Big((ab_k+a\bar{b}_k)^2/2+(ab_k+a\bar{b}_k)\sqrt{2/\pi}\Big)\\
\times\ & \prod_{\substack{k>J\\ k\neq j}}\exp\Big(((a+1/\varepsilon)b_k+a\bar{b}_k)^2/2+((a+1/\varepsilon)b_k+a\bar{b}_k)\sqrt{2/\pi}\Big)\\
&\le \frac{c}{\varepsilon}\exp\left(c_1(a)\sum_{k=1}^{\infty}(b_k+b_k^2+\bar{b}_k+\bar{b}_k^2)+c_1(a)\sum_{k>J}(\frac{b_k}{\varepsilon}+\frac{b_k^2}{\varepsilon^2})\right)\sum_{j>J}b_j\\
\le\ & \frac{c_2(a)}{\varepsilon}\sum_{j>J}b_j=\frac{c_2(a)}{B}
\end{aligned}$$

with the constant $B>1$ as in (6.3), and with constants $c_1(a)$ and $c_2(a)$ depending only on a. With the constant $c(n)$ in Lemma 6.1, for every $u\in\Gamma_{\boldsymbol{b}}$

$$p^n(u,\cdot)(\mathcal{V}(\cdot)-\mathcal{V}_0(\cdot))\le c(n)\frac{c_2(a)}{B}+(1-c_0)^n(\mathcal{V}(u)-\mathcal{V}_0(u))\ . \tag{6.7}$$

Let $g:U\to\mathbb{R}$ be a function such that $|g(u)|\le\mathcal{V}(u)\ \forall u\in\Gamma_{\boldsymbol{b}}$. Define further q

$$g_0(u)=\begin{cases} g(u)\exp(-\frac{1}{\varepsilon}\sum_{j>J}b_j|u_j|) & \text{if}\quad u\in\Gamma_{\bar{\boldsymbol{b}}}\\ g(u)\exp(-\frac{1}{\varepsilon}\sum_{j>J}b_j|u_j|-a\sum_{j=1}^{J}\bar{b}_j|u_j|) & \text{if}\quad u\in\Gamma_{\boldsymbol{b}}\setminus\Gamma_{\bar{\boldsymbol{b}}}\ .\end{cases}$$

Then $|g_0(u)|\le\mathcal{V}_0(u)$ for all $u\in\Gamma_{\boldsymbol{b}}$. We also have that

$$|g(u)-g_0(u)|=|g(u)|\left(1-\exp(-\frac{1}{\varepsilon}\sum_{j>J}b_j|u_j|)\right)$$

when $u\in\Gamma_{\bar{\boldsymbol{b}}}$, and

$$|g(u)-g_0(u)|=|g(u)|\left(1-\exp(-\frac{1}{\varepsilon}\sum_{j>J}b_j|u_j|-a\sum_{j=1}^{J}\bar{b}_j|u_j|)\right)$$

when $u \in \Gamma_{\boldsymbol{b}} \setminus \Gamma_{\bar{\boldsymbol{b}}}$. Thus $|g(u) - g_0(u)| \le \mathcal{V}(u) - \mathcal{V}_0(u)$ for all $u \in \Gamma_{\boldsymbol{b}}$. From (6.6), we have for every $u \in \Gamma_{\boldsymbol{b}}$ and every $n > n_0(\tau)$

$$\begin{aligned} |p^n(u,\cdot)(g) - \gamma^\delta(g)| &\le |p^n(u,\cdot)(g-g_0)| + |p^n(u,\cdot)(g_0) - \gamma^\delta(g_0)| + |\gamma^\delta(g-g_0)| \\ &\le |p^n(u,\cdot)(\mathcal{V} - \mathcal{V}_0)| + |\gamma^\delta(\mathcal{V} - \mathcal{V}_0)| + \tau \mathcal{V}_0(u) \ . \end{aligned}$$

Since

$$\gamma^\delta(\mathcal{V} - \mathcal{V}_0) \le c\gamma_{\boldsymbol{b}}(\mathcal{V} - \mathcal{V}_0),$$

where c depends only on the normalizing constant in the Radon-Nikodym derivative of γ^δ with respect to $\gamma_{\boldsymbol{b}}$, we have from (6.7) with the constant $c_0 \in]0,1[$ in (2.8)

$$\begin{aligned} |p^n(u,\cdot)(g) - \gamma^\delta(g)| &\le c(n)\frac{c_2(a)}{B} + (1-c_0)^n(\mathcal{V}(u) - \mathcal{V}_0(u)) + c\frac{c_2(a)}{B} + \tau\mathcal{V}_0(u) \\ &\le c(n)\frac{c_2(a)}{B} + (1-c_0)^n\mathcal{V}(u) + \tau\mathcal{V}(u) + \frac{cc_2(a)}{B} \end{aligned}$$

for every $u \in \Gamma_{\boldsymbol{b}}$ and for every $n \ge n_0(\tau)$ as in (6.6).

Fixing $\tau' < 1$ such that $\tau' > \tau$, we can choose $\mathbb{N} \ni n_1 > n_0$ (depending on c_0, τ and τ') such that

$$(1-c_0)^n < \frac{1}{2}(\tau' - \tau), \quad \forall\, n > n_1 \ . \tag{6.8}$$

For every $n > n_1(c_0,\tau,\tau')$, there exists a constant $B_1(n,a)$ so that

$$(c(n)+c)\frac{c_2(a)}{B} < \frac{1}{2}(\tau' - \tau) \ \ \forall B > B_1(n,a) \ . \tag{6.9}$$

Thus, for every $n > n_1$ and for $B > B_1(n,a)$ in (6.3), we have

$$|p^n(u,\cdot)(g) - \gamma^\delta(g)| \le \tau'\mathcal{V}(u) \ \ \forall u \in \Gamma_{\boldsymbol{b}} \ .$$

This is (6.5). □

We next bound the error incurred in the acceptance function $\alpha(u,v)$ in (2.1) due to dimension truncation and Galerkin discretization.

Lemma 6.4. *We have the estimate*

$$|\alpha(u,v) - \alpha^{J,l}(u,v)| \le |\Phi(u;\delta) - \Phi^{J,l}(u,\delta)| + |\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| \ . \tag{6.10}$$

Proof We first consider the case where $\alpha(u,v) \ge \alpha^{J,l}(u,v)$. If $\Phi(u;\delta) - \Phi(v;\delta) \ge 0$ then $\alpha(u,v) = 1$. If $\alpha^{J,l}(u,v) = 1$ the conclusion then follows. If $\alpha^{J,l}(u,v) < 1$ we have $\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta) < 0$. Using $1 - \exp(-x) < x$ for $x \ge 0$, we have

$$\begin{aligned} \alpha(u,v) - \alpha^{J,l}(u,v) &\le -(\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta)) \\ &\le (\Phi(u;\delta) - \Phi(v;\delta)) - (\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta)) \\ &\le |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + |\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| \ . \end{aligned}$$

If $\Phi(u;\delta) - \Phi(v;\delta) < 0$ then $\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta) < 0$. Using the inequality $|\exp(-x) - \exp(-y)| \le |x-y|$ for all $x, y > 0$, we have

$$\begin{aligned} \alpha(u,v) - \alpha^{J,l}(u,v) &= \exp(\Phi(u;\delta) - \Phi(v;\delta)) - \exp(\Phi^{J,l}(u;\delta) - \Phi^{J,l}(v;\delta) \\ &\le |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + |\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| \ . \end{aligned}$$

The proof for the case $\alpha(u,v) < \alpha^{J,l}(u,v)$ is similar. □

Using Lemma 6.4, we obtain a bound on the error in the transition kernel, due to the dimension truncation and discretization error in the forward model.

Lemma 6.5. *For each $n \in \mathbb{N}$, there is a constant $d(n,a)$ which depends on $\mathcal{V}_0$ such that for every $B > 1$ in (6.3)*

$$\|(p^{J,l})^n(u,\cdot) - p^n(u,\cdot)\|_{\mathcal{V}} \le d(n,a)\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + 2^{-l} + J^{-q}\Big)$$

$$+ n \min\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q}), (1-c_0)^n\Big)\mathcal{V}(u) \tag{6.11}$$

where c is a constant independent of J, l and n.

Proof We show (6.11) by induction with respect to n.
$n = 1$: for a function $g : \Gamma_{\boldsymbol{b}} \to \mathbb{R}$ such that $\forall u \in \Gamma_{\boldsymbol{b}} : \ |g(u)| \le \mathcal{V}(u)$

$$|p^{J,l}(u,\cdot)(g) - p(u,\cdot)(g)|$$
$$\le \int_U |\alpha^{J,l}(u,v) - \alpha(u,v)|\mathcal{V}(v)d\gamma_{\boldsymbol{b}}(v) + \Big|\int_U \alpha^{J,l}(u,v)d\gamma_{\boldsymbol{b}}(v) - \int_U \alpha(u,v)d\gamma_{\boldsymbol{b}}(v)\Big|\mathcal{V}(u)\ .$$

For $v \in \Gamma_{\boldsymbol{b}}$, we have

$$\begin{aligned}|\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| &\le c(|\delta| + |\mathcal{G}(v)| + |\mathcal{G}^{J,l}(v)|)|\mathcal{G}(v) - \mathcal{G}^{J,l}(v)| \\ &\le c(\delta)(1 + \|P(v)\|_V + \|P^{J,l}(v)\|_V)\|P(v) - P^{J,l}(v)\|_V \\ &\le c\exp(7\sum_{j=1}^\infty b_j|v_j|)\Big(2^{-l}(1 + \sum_{j=1}^J \bar{b}_j|v_j|) + \sum_{j>J} b_j|v_j|\Big) \\ &\le c\exp(7\sum_{j=1}^\infty b_j|v_j| + \sum_{j=1}^J \bar{b}_j|v_j|)\Big(2^{-l} + \sum_{j>J} b_j|v_j|\Big)\ .\end{aligned} \tag{6.12}$$

Arguing as in the proof of Proposition 4.6, using (4.19) and (4.21), we deduce that

$$\int_U |\Phi(v;\delta) - \Phi^{J,l}(v;\delta)|\mathcal{V}(v)d\gamma_{\boldsymbol{b}}(v) \le c(a)(2^{-l} + \sum_{j>J} b_j) \le c(a)(2^{-l} + J^{-q}), \tag{6.13}$$

when $B > 1$. From Lemma 6.4,

$$\int_U |\alpha^{J,l}(u,v) - \alpha(u,v)|\mathcal{V}(v)d\gamma_{\boldsymbol{b}}(v) \le |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)|\int_U \mathcal{V}(v)d\gamma_{\boldsymbol{b}}(v) + c(a)(2^{-l} + J^{-q})\ .$$

From (6.12), (4.19) and (4.21), we also have

$$\int_U |\Phi(v;\delta) - \Phi^{J,l}(v;\delta)|d\gamma_{\boldsymbol{b}}(v) \le c(2^{-l} + J^{-q})\ .$$

Therefore

$$\int_U |\alpha^{J,l}(u,v) - \alpha(u,v))|d\gamma_{\boldsymbol{b}}(v) \le |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q})\ . \tag{6.14}$$

We note that

$$\int_U \alpha^{J,l}(u,v)d\gamma_{\boldsymbol{b}}(v) \ge \int_U \exp(-\Phi^{J,l}(v;\delta))d\gamma_{\boldsymbol{b}}(v)\ .$$

From (4.26), we can choose $\Lambda > 0$ (depending on $c_1, c_2, \boldsymbol{b}$) so that

$$\gamma_{\boldsymbol{b}}(\{v : \Phi^{J,l}(v,\delta) < \Lambda\})$$

is uniformly bounded away from zero for J and l sufficiently large. Therefore, there exists $0 < c_0 < 1$ such that for J and l sufficiently large

$$\int_U \exp(-\Phi^{J,l}(v;\delta))d\gamma_{\boldsymbol{b}}(v) > c_0 > 0\ .$$

We choose c_0 as in (2.8). Then, for J and l sufficiently large,

$$\forall u \in \Gamma_{\boldsymbol{b}}: \quad \Big| \int_U (\alpha^{J,l}(u,v) - \alpha(u,v)) d\gamma_{\boldsymbol{b}}(v) \Big| \le 1 - c_0 \ .$$

We therefore deduce that for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\begin{aligned} |p^{J,l}(u,\cdot)(g) - p(u,\cdot)(g)| &\le d(1,a)(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + 2^{-l} + J^{-q}) \\ &+ \min\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q}), 1 - c_0 \Big) \mathcal{V}(u) \ . \end{aligned}$$

This proves the assertion for $n = 1$.

Assume now that (6.11) holds for $n-1$ for some $n > 1$. Then

$$|((p^{J,l})^n(u,\cdot) - p^n(u,\cdot))(g)| \le |p^{J,l}(u,\cdot)((p^{J,l})^{n-1} - p^{n-1})(g)| + |(p^{J,l} - p)(u,\cdot)p^{n-1}(g)| \ .$$

We also have from (6.1) that

$$\begin{aligned} &|(p^{J,l} - p)(u,\cdot)(p^{n-1}(\cdot,\cdot)(g))| \\ &\le \int_U |\alpha^{J,l}(u,v) - \alpha(u,v)| \Big(c(n-1) \int_U \mathcal{V}(w) d\gamma_{\boldsymbol{b}}(w) + (1-c_0)^{n-1}\mathcal{V}(v) \Big) d\gamma_{\boldsymbol{b}}(v) \\ &\qquad + \Big| \int_U (\alpha(u,v) - \alpha^{J,l}(u,v)) d\gamma_{\boldsymbol{b}}(v) \Big| \Big(c(n-1) \int_U \mathcal{V}(v) d\gamma_{\boldsymbol{b}}(v) + (1-c_0)^{n-1}\mathcal{V}(u) \Big) \ . \end{aligned}$$

From (6.10), (6.12), (6.13) and (6.14) we deduce that there exists a constant $d'(n,a)$ such that for every $u \in \Gamma_{\boldsymbol{b}}$ holds

$$\begin{aligned} |(p^{J,l} - p)(u,\cdot)(p^{n-1}(u,\cdot)(g))| &\le d'(n,a)(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + 2^{-l} + J^{-q}) \\ &+ \min\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q}), (1-c_0)^n \Big) \mathcal{V}(u) \ . \end{aligned}$$

We have further that for every $u \in \Gamma_{\boldsymbol{b}}$

$$\begin{aligned} &|p^{J,l}(u,dv)((p^{J,l})^{n-1} - p^{n-1})(v,\cdot)(g)| \\ &\le \int_U |((p^{J,l})^{n-1} - p^{n-1})(v,\cdot)(g)| d\gamma_{\boldsymbol{b}}(v) + (1-c_0)|((p^{J,l})^{n-1} - p^{n-1})(u,\cdot)(g)| \\ &\le \int_U \Big(d(n-1,a)(|\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| + 2^{-l} + J^{-q}) \\ &+ (n-1)(|\Phi(v;\delta) - \Phi^{J,l}(v;\delta)| + c(2^{-l} + J^{-q}))\mathcal{V}(v) \Big) d\gamma_{\boldsymbol{b}}(v) \\ &+ (1-c_0) d(n-1,a)(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + 2^{-l} + J^{-q}) \\ &+ (1-c_0)(n-1) \min\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q})), (1-c_0)^{n-1} \Big) \mathcal{V}(u) \\ &\le d''(n,a)(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + 2^{-l} + J^{-q}) \\ &+ (n-1) \min\Big(|\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| + c(2^{-l} + J^{-q})), (1-c_0)^n \Big) \mathcal{V}(u) \ . \end{aligned}$$

From these bounds, we obtain (6.11) for $n > 1$ and complete the proof. □

We then have the following result.

Lemma 6.6. *Given data δ, there exists a constant $M < \infty$ so that for every $B > 1$ in (6.3) and for every $J, l \in \mathbb{N}$ holds*

$$\sup_{u \in \Gamma_{\boldsymbol{b}}} \frac{\| p^{J,l}(u,\cdot) - \gamma^{J,l,\delta} \|_{\mathcal{V}}}{\mathcal{V}(u)} \le M \ . \tag{6.15}$$

Assume that a is sufficiently large. Then, there are constants $n \in \mathbb{N}$, $B_1(n,a) \in \mathbb{N}$, $J_0 \in \mathbb{N}$, $l_0 \in \mathbb{N}$, $0 < \tau'' < 1$ so that for every $B > B_1(n,a)$, for every $J > J_0$ and every $l > l_0$ holds

$$\sup_{u \in \Gamma_{\boldsymbol{b}}} \frac{\|(p^{J,l})^n(u,\cdot) - \gamma^{J,l,\delta}\|_{\mathcal{V}}}{\mathcal{V}(u)} \le \tau'' < 1 \ . \tag{6.16}$$

Proof From Proposition 4.6, $\gamma^{J,l,\delta}(\mathcal{V})$ is uniformly bounded. Further,

$$p^{J,l}(u,\cdot)(\mathcal{V}) = \int_U \alpha^{J,l}(u,v)\mathcal{V}(v)d\gamma_{\boldsymbol{b}}(v) + (1 - \int_U \alpha^{J,l}(u,v)d\gamma_{\boldsymbol{b}}(v))\mathcal{V}(u) \le c + \mathcal{V}(u) \ .$$

As the normalizing constant in (4.13) is bounded away from 0 uniformly with respect to J and l, $\int_{\Gamma_{\boldsymbol{b}}} \mathcal{V}(u)d\gamma^{J,l,\delta}(u)$ is uniformly bounded for every J and l. Thus we can choose $M > 0$ such that (6.15) holds uniformly with respect to J and l.

Using the bound $x \le \varepsilon \exp(x/\varepsilon)$, for every $x, \varepsilon > 0$, from (6.12) we deduce that for arbitrary, fixed $a \ge 7$,

$$\begin{aligned} |\Phi(u;\delta) - \Phi^{J,l}(u;\delta)| &\le c \exp\left(7\sum_{j=1}^{\infty} b_j|u_j| + \sum_{j=1}^{J} \bar{b}_j|u_j|\right) \\ &\quad \times\Big((2^{-l} + \varepsilon \exp(\frac{1}{\varepsilon}\sum_{j>J} b_j|u_j|)\Big) \\ &\le c(2^{-l} + \varepsilon)\mathcal{V}(u) \ . \end{aligned}$$

From Lemma 6.5, for $|g| \le \mathcal{V}$, for all $n \in \mathbb{N}$, and for every $J, l \in \mathbb{N}$ and every $u \in \Gamma_{\boldsymbol{b}}$

$$|(p^{J,l})^n(u,\cdot)(g) - p^n(u,\cdot)(g)| \le \Big(cd(n,a)(2^{-l} + \varepsilon + J^{-q}) + n(1-c_0)^n\Big)\mathcal{V}(u) \ .$$

For $n > n_1(\tau')$ and the constant $B > 1$ in (6.3) such that $B > B_1(n,a)$ where n_1 and $B_1(n,a)$ are the constants in Lemma 6.3, we obtain

$$\begin{aligned} &|(p^{J,l})^n(u,\dot{)}(g) - \gamma^{J,l,\delta}(g)| \\ &\le |(p^{J,l})^n(u,\cdot)(g) - p^n(u,\cdot)(g)| + |p^n(u,\cdot)(g) - \gamma^{\delta}(g)| + |\gamma^{\delta}(g) - \gamma^{J,l,\delta}(g)| \\ &\le \Big(cd(n,a)(2^{-l} + \varepsilon + J^{-q}) + n(1-c_0)^n + \tau'\Big)\mathcal{V}(u) + c(2^{-l} + J^{-q})\|\mathcal{V}\|_{L^2(U,\gamma_{\boldsymbol{b}})} \end{aligned}$$

for all $J, l \in \mathbb{N}$ and all $u \in \Gamma_{\boldsymbol{b}}$ where τ' is the constant in Lemma 6.3. Let $\tau'' < 1$ be such that $\tau'' > \tau'$. We choose $n > n_1$ so that $n(1-c_0)^n < (\tau''-\tau')/4$. From Lemma 6.2, $\|\mathcal{V}\|_{L^2(U,\gamma_{\boldsymbol{b}})} < C(a)$ where the bound $C(a)$ only depends on a. We recall that we work under Assumption 4.5. We therefore can choose $J_0 = J_0(n,a,B)$ so that with $\varepsilon = B\sum_{j>J} b_j \le cBJ^{-q}$

$$\forall J > J_0(n,a,B): \quad cd(n,a)(\varepsilon + J^{-q}) + cC(a)J^{-q} = \frac{1}{4}(\tau'' - \tau') \ .$$

There exists $l_0(n,a) \in \mathbb{N}$ such that for every $l > l_0$ holds

$$cd(n,a)2^{-l} + cC(a)2^{-l} < \frac{1}{4}(\tau'' - \tau') \ .$$

Thus, for every $l > l_0$ and for every $J > J_0$ if $|g(u)| \le \mathcal{V}(u)$ for all $u \in \Gamma_{\boldsymbol{b}}$, there holds

$$\forall u \in \Gamma_{\boldsymbol{b}}: \quad |(p^{J,l})^n(u,\cdot)(g) - \gamma^{J,l,\delta}(g)| \le \tau''\mathcal{V}(u) \ .$$

This is (6.16) and completes the proof of Lemma 6.6. □

Proof of Proposition 4.7 From the proof of Proposition 16.1.3 of Meyn and Tweedie [16], using Lemma 6.6, we find for fixed $n \in \mathbb{N}$, $u \in \Gamma_{\boldsymbol{b}}$, and for $J > J_0$, $l > l_0$ as in Lemma 6.6, that there holds

$$\forall m \in \mathbb{N}: \quad \sup_{u \in \Gamma_{\boldsymbol{b}}} \frac{\|(p^{J,l})^m(u,\cdot) - \gamma^{J,l,\delta}\|_{\mathcal{V}}}{\mathcal{V}(u)} \leq \frac{M^n}{\tau''}((\tau'')^{1/n})^m.$$

This implies Proposition 4.7.

7. Appendix B: Proof of (4.19) - (4.21)

To prove estimate (4.19) we observe that for any $t > 0$

$$\begin{aligned}
&\int_{-\infty}^{\infty} \exp(-z^2/2 + |z|t)\frac{dz}{\sqrt{2\pi}} \\
=\ & \exp(t^2/2)\int_{-\infty}^{\infty} \exp(-(|z|-t)^2/2)\frac{dz}{\sqrt{2\pi}} \\
=\ & \exp(t^2/2)\Big(\int_{-\infty}^{0} \exp(-(z+t)^2/2)\frac{dz}{\sqrt{2\pi}} + \int_{0}^{\infty} \exp(-(z-t)^2/2)\frac{dz}{\sqrt{2\pi}}\Big) \\
=\ & \exp(t^2/2)\Big(\int_{-\infty}^{t} \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}} + \int_{-t}^{\infty} \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\Big) \\
=\ & \exp(t^2/2)\left(1 + \int_{-t}^{t} \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\right) \\
\leq\ & \exp(t^2/2)(1 + t\sqrt{2/\pi}) \leq \exp(t^2/2)\exp(t\sqrt{2/\pi})\ .
\end{aligned}$$

For inequality (4.20), we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} z^2 \exp(-z^2/2 + |z|t)\frac{dz}{\sqrt{2\pi}} \\
&= \exp(t^2/2)\Big(\int_{-\infty}^{0} z^2 \exp(-(z+t)^2/2)\frac{dz}{\sqrt{2\pi}} + \int_{0}^{\infty} z^2 \exp(-(z-t)^2/2)\frac{dz}{\sqrt{2\pi}}\Big) \\
&= \exp(t^2/2)\Big(\int_{-\infty}^{t} (z-t)^2 \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}} + \int_{-t}^{\infty} (z+t)^2 \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\Big) \\
&\leq 2\exp(t^2/2)\Big(\int_{-\infty}^{t} (z^2+t^2) \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}} + \int_{-t}^{\infty} (z^2+t^2) \exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\Big) \\
&= 4\exp(t^2/2)\int_{-\infty}^{\infty} (z^2+t^2)\exp(-z^2/2)\frac{dz}{\sqrt{2\pi}} \\
&\leq c\exp(t^2/2)(1+t^2)\ .
\end{aligned}$$

Similarly, for inequality (4.21), we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} |z|\exp(-z^2/2+|z|t)\frac{dz}{\sqrt{2\pi}}\\
&=\exp(t^2/2)\Big(\int_{-\infty}^{0}|z|\exp(-(z+t)^2/2)\frac{dz}{\sqrt{2\pi}}+\int_0^{\infty}|z|\exp(-(z-t)^2/2)\frac{dz}{\sqrt{2\pi}}\Big)\\
&=\exp(t^2/2)\Big(\int_{-\infty}^{t}|z-t|\exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}+\int_{-t}^{\infty}|z+t|\exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\Big)\\
&\leq 2\exp(t^2/2)\int_{-\infty}^{\infty}(|z|+t)\exp(-z^2/2)\frac{dz}{\sqrt{2\pi}}\leq c\exp(t^2/2)(1+t)\ .
\end{aligned}$$

Acknowledgement: We acknowledge partial support under ERC AdG 247277 (to CS) and AcRF Tier 1 research grant RG69/10, the Singapore A*Star SERC grant 122-PSF-0007 and the AcRF Tier 2 grant MOE 2013-T2-1-095, ARC 44/13 (to VHH). Parts of this work are performed when VHH visited the Institute for Mathematical Research (FIM) at the ETH Zurich in 2014 and 2015, to which he is grateful for the excellent research conditions. Initial prelimiary research for this paper was done when VHH visited the Institute for Computational and Experimental Research in Mathematics (ICERM) at Brown University, USA, during the Programme "Computational Challenges in Probability"; he is grateful to ICERM for the financial support and the excellent research facilities during the visit. The authors thank Andrew Stuart for helpful discussions on related results, and for pointing out the relevance of the results in [15] for the present work.

References

[1] Robert J. Adler. *The geometry of random fields*. John Wiley & Sons Ltd., Chichester, 1981. Wiley Series in Probability and Mathematical Statistics.

[2] Julia Charrier. Strong and weak error estimates for elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.*, 50(1):216–246, 2012.

[3] 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.

[4] M. Dashti and A. M. Stuart. The bayesian approach to inverse problems. Technical Report 1302.6989v4, ArXiv.

[5] M. Dashti and A. M. Stuart. Uncertainty quantification and weak approximation of an elliptic inverse problem. *SIAM J. Numer. Anal.*, 49(6):2524–2542, 2011.

[6] T. J. Dodwell, C. Ketelsen, R. Scheichl, and A. L. Teckentrup. A Hierarchical Multilevel Markov Chain Monte Carlo Algorithm with Applications to Uncertainty Quantification in Subsurface Flow. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):1075–1108, 2015.

[7] J. Galvis and M. Sarkis. Approximating infinity-dimensional stochastic Darcy's equations without uniform ellipticity. *SIAM J. Numer. Anal.*, 47(5):3624–3651, 2009.

[8] Claude J. Gittelson. Stochastic Galerkin discretization of the log-normal isotropic diffusion problem. *Mathematical Models and Methods in Applied Sciences*, 20(2):237–263, 2010.

[9] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[10] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan. Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications. *J. Comput. Phys.*, 230(10):3668–3694, 2011.

[11] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, Ch. Schwab, and I.H. Sloan. Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numerische Mathematik*. Published online 2nd December 2014.

[12] V. H. Hoang and Ch. Schwab. In preparation. 2016.

[13] Viet Ha Hoang. Bayesian inverse problems in measure spaces with application to Burgers and Hamilton-Jacobi equations with white noise forcing. *Inverse Problems*, 28(2):025009, 29, 2012.
[14] Viet Ha Hoang, Christoph Schwab, and Andrew Stuart. Complexity analysis of accelerated MCMC methods for Bayesian inversion. *Inverse Problems*, 29(8), 2013.
[15] K. L. Mengersen and R. L. Tweedie. Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.*, 24(1):101–121, 1996.
[16] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, Cambridge, second edition, 2009.
[17] H. Nguyen. *Finite element wavelets for solving partial differential equations.* PhD thesis, University of Utrecht, 2005.
[18] Gareth O. Roberts and Jeffrey S. Rosenthal. Quantitative non-geometric convergence bounds for independence samplers. *Methodol. Comput. Appl. Probab.*, 13(2):391–403, 2011.
[19] Christoph Schwab and Claude Jeffrey Gittelson. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numerica*, 20:291–467, 2011.
[20] Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 2010.
[21] A.L. Teckentrup, R. Scheichl, M.B. Giles, and E. Ullmann. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numerische Mathematik*, 125(3):569–600, 2013.
[22] L. Tierney. A note on Metropolis-Hastings kernels for general state spaces. *Ann. Appl. Prob.*, 8:1–9, 1998.
[23] Karsten Urban. *Wavelet methods for elliptic partial differential equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2009.
[24] Y. Yamasaki. *Measures on infinite-dimensional spaces*, volume 5 of *Series in Pure Mathematics*. World Scientific Publishing Co., Singapore, 1985.

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

[2] Seminar for Applied Mathematics, ETH, 8092 Zurich, Switzerland