ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

# Sparsity in Bayesian Inversion of Parametric Operator Equations

Cl. Schillings and Ch. Schwab

Research Report No. 2013-17
June 2013

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

- Funding ERC: FP7 Grant AdG247277

# Sparsity in Bayesian Inversion of Parametric Operator Equations

**Cl. Schillings**[1] **and Ch. Schwab**[1]

[1] Seminar for Applied Mathematics, ETH, CH-8092 Zurich, Switzerland

E-mail: `claudia.schillings@sam.math.ethz.ch`
`christoph.schwab@sam.math.ethz.ch`

**Abstract.** We establish posterior sparsity in Bayesian inversion for systems with distributed parameter uncertainty subject to noisy data. We generalize the particular case of scalar diffusion problems with random coefficients in [29] to broad classes of operator equations. For countably parametric, deterministic representations of uncertainty in the forward problem which belongs to a certain sparsity class, we quantify analytic regularity of the (countably parametric) Bayesian posterior density and prove that the parametric, deterministic density of the Bayesian posterior belongs to the same sparsity class. Generalizing [32, 29], the considered forward problems are parametric, deterministic operator equations, and computational Bayesian inversion is to evaluate expectations of quantities of interest (QoIs) under the Bayesian posterior, conditional on given data.

The sparsity results imply, on the one hand, sparsity of Legendre (generalized) polynomial chaos expansions of the Bayesian posterior and, on the other hand, convergence rates for data-adaptive Smolyak integration algorithms for computational Bayesian estimation which are independent of dimension of the parameter space. The convergence rates are, in particular, superior to Markov Chain Monte-Carlo sampling of the posterior, in terms of the number $N$ of instances of the parametric forward problem to be solved.

*Keywords*: Bayesian Inverse Problems, Parametric Operator Equations, Smolyak Quadrature, Sparsity, Uniform Prior Measures.

**Contents**

## 1. Introduction

The problem of computational and mathematical treatment of inference of responses of uncertain systems, in the presence of "large data" possibly subject to observation noise, is a key problem in engineering and the sciences. A "most likely" response of the uncertain system is offered by Bayesian Inversion, which characterizes the expected system response as an average over all realizations of system uncertainty, conditional on the given data. Computational methods for the efficient evaluation of such expectations have received considerable interest in recent years. The most widely used numerical methods for the numerical treatment of Bayesian inversion and prediction problems are based on statistical sampling from the posterior measure, and are therefore Monte-Carlo (MC) type algorithms, in particular the so-called Markov-Chain Monte-Carlo (MCMC) methods (eg. [19, 20, 26, 27]). While these methods are widely used and their theoretical foundation is well-understood, their drawbacks are slow convergence, in particular since for each draw of the Markov-Chain, one instance of the forward problem's governing equation must be solved numerically. In systems where these equations are partial differential or other operator equations, generating many samples of the Markov-Chain can be computationally costly. In [21], a stochastic Newton method is proposed with the aim of accelerating the MCMC approach by exploiting gradient and Hessian information of the posterior density. In the context of multilevel discretizations for partial differential equations, *multilevel MCMC sampling strategies* provide substantial improvements.

However, the convergence rate which can be achieved by MLMC approaches is, ultimately, limited to the order $1/2$ of convergence of MC methods. We refer to [1, 2, 18] for references and a detailed analysis.

A second challenge for computationally efficient Bayesian inversion of systems governed by PDEs and more general operator equations with random inputs is the "distributed" nature of the uncertainty: rather than expectations w.r. to a finite number of real-valued parameters, mathematical expectations over a space $X$ of uncertain coefficient functions $u$ which are distributed w.r. to a prior measure $\mu_0$ must be computed. Typical cases in point are spatially heterogeneous conductivity tensors, permeabilities in subsurface flow, dielectric tensors in electromagnetism, obstacle shapes in scattering to name but a few. Their presence mandates Bayesian inversion for uncertainty in forward problems which is described by *random fields* rather than by real-valued random parameters.

The design and the numerical analysis of efficient, deterministic algorithms for computational Bayesian inversion of PDE problems with distributed parameter uncertainty is the purpose of the present paper. A computational framework for the treatment of distributed uncertainties based on linearization of the infinite dimensional inverse problem is proposed in [4]. The linearization about a nominal state in combination with low-rank approximations of the covariance of the posterior density allows to derive dimension independent convergent rates for the linearized inverse problem. The adaptive, infinite-dimensional quadrature approach in [29] and the present work does not rely on linearization, is (through the posterior-density) data-adaptive, and quantifies uncertainty over all scenarios (not just those which are close to nominal). The use of polynomial chaos expansions in the Bayesian posterior to accelerate computational Bayesian inversion has been pioneered in [24, 23, 22] and further analyzed in [18]. Here, as in our previous work [32, 29], we reformulate the Bayesian inversion problem as a deterministic, infinite-dimensional quadrature problem with respect to the posterior measure, given noisy observation data $\delta$ of a QoI $\phi$, and analyze the regularity of the deterministic posterior in terms of the parameters in the parametrization of the uncertainty in the forward problem.

This infinite-dimensional, deterministic quadrature problem is subsequently treated numerically by either a dimension-adaptive Smolyak quadrature algorithm as proposed in [29] or by a Quasi Monte-Carlo quadrature rule such as a polynomial lattice rule.

Affine parametric dependence of the forward problems results, for example, from expansions of the uncertainty in terms of principal components as afforded by Karhunen-Loève expansions. Then, under certain regularity assumptions on the covariance spectrum of the unknown system parameters, our results imply that these dimension adaptive integration algorithms can converge at higher rates than the rate $1/2$ (in terms of the number of solutions of the forward problem for $N$ instances of the uncertain input $u$) which is best possible for the Markov-Chain MC algorithm. This program has been implemented recently in [32, 29] for a class of scalar, isotropic diffusion problems with uncertain diffusion coefficient. Here, we generalize this approach to systems governed by an abstract class of parametric operator equations; while the technicalities of the analysis, in particular the proof of sparsity of the posterior, as well as the convergence analysis of the Smolyak quadrature, are analogous to [32] and to [29], respectively, the increase in scope afforded by the presently considered abstraction is as follows: the approach is equally applicable for definite or indefinite elliptic and for parabolic evolution problems, with scalar or tensorial unknowns (such as arise, for example, in models of anisotropic media) with single or multiple scales (as, eg., in [1, 16]), and also to Bayesian inversion subject to uncertainty in coefficients, in loadings and in domains. Also, the Smolyak quadrature convergence result given in [29] is generalized herein: whereas in [29], the integrand functions were required to allow for analytic extensions into polydiscs, here this condition is weakened to analyticity in polyellipses of possibly large eccentricities, thereby allowing poles in the analytic continuations of integrand functions which are situated arbitrarily close to the domain of integration; in [10], this was shown in certain cases to allow global analytic continuation of parametric solutions, also for large data.

The outline of this paper is as follows: in Section 2, we present the Bayesian approach to inverse problems for PDEs set in function spaces. We consider, in particular, an abstract class of operator equations which depend on a sequence $\boldsymbol{y} = (y_j)_{j\geq 1}$ of parameters which will be the forward model in the ensuing analysis, and examples for conrete instance of such equations. Section 4 presents new results on sparsity of the posterior density, generalizing the results in [32]. These results will be used in Section 5 presents the sparse Smolyak quadratur algorithm and shows that this algorithm can realize the (dimension-independent) convergence rates afforded by the sparsity of the Bayesian posterior. Section 6 presents detailed numerical experiments for parabolic evolution problems with distributed uncertainty which corroborate the theoretical results.

Finally, in Section 7 we summarize the principal conclusions and indicate the application to new quadrature algorithms as well as to sparse tensor discretizations.

## 2. Bayesian Inversion of Parametric Operator Equations

We define a class of operator equations which depend on an uncertain datum $u$ taking values in a separable Banach space $X$ via a possibly countably infinite sequence $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$ of parameters.

### 2.1. Linear Operator Equations

We denote by $\mathcal{X}$ and $\mathcal{Y}$ two reflexive Banach spaces over $\mathbb{R}$ (for some of the technical arguments which follow, we shall require also extensions of these spaces to Banach spaces over the coefficient field $\mathbb{C}$; we shall use these without distinguishing these extensions notationally) with (topological) duals $\mathcal{X}'$ and $\mathcal{Y}'$, respectively. By $\mathcal{L}(\mathcal{X},\mathcal{Y}')$, we denote the set of bounded linear operators $A : \mathcal{X} \to \mathcal{Y}'$. Via the Riesz representation theorem, we associate to each $A \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ in a one-to-one correspondence a bilinear form $\mathcal{G}(\cdot,\cdot) : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ via (with ${}_{\mathcal{Y}}\langle\cdot,\cdot\rangle_{\mathcal{Y}'}$ denoting the $\mathcal{Y}\times\mathcal{Y}'$-duality pairing)

$$\mathfrak{a}(v,w) :=_{\mathcal{Y}} \langle w, Av\rangle_{\mathcal{Y}'} \quad \text{for all } v\in\mathcal{X}, w\in\mathcal{Y}\ . \tag{1}$$

A key role will be played by bounded invertibility of differentials of operator equations at so-called "nominal" parameter values (being either a mathematical expectation, or an otherwise fixed reference for the uncertain input). Consider therefore the linear operator equation $Au = f$ where $A \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ is boundedly invertible; necessary and sufficient criteria for this are the so-called "inf-sup" conditions; we state these conditions here for reference (see, eg., [3]).

**Proposition 2.1.** *A bounded, linear operator $A \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ is boundedly invertible* if and only if *its bilinear form satisfies* inf-sup conditions*: there exists $\mu > 0$ such that*

$$\inf_{0\neq v\in\mathcal{X}} \sup_{0\neq w\in\mathcal{Y}} \frac{\mathfrak{a}(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu\ , \quad \inf_{0\neq w\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}} \frac{\mathfrak{a}(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu\ . \tag{2}$$

*If* (2) *holds then for every $f \in \mathcal{Y}'$ the operator equation*

$$\text{find } q\in\mathcal{X}: \quad \mathfrak{a}(q,v) = \langle f, v\rangle_{\mathcal{Y}'\times\mathcal{Y}} \quad \forall v\in\mathcal{Y} \tag{3}$$

*admits a unique response* $q \in \mathcal{X}$ *and there holds* $\|q\|_{\mathcal{X}} = \|A^{-1}f\|_{\mathcal{X}} \leq \mu^{-1}\|f\|_{\mathcal{Y}'}$.

### 2.2. Bayesian Inversion

By $G : X \to R$ we denote a "forward" response map from some separable Banach space $X$ of unknown parameters $u$ in the operator $A$ into a Banach space $R$ of responses which contains the Quantity of Interest (QoI) in the Bayesian inversion. We equip $X$ and $\mathcal{X}$ with norms $\|\cdot\|_X$ and with $\|\cdot\|_{\mathcal{X}}$, respectively. We have in mind an abstract (possibly nonlinear) operator equation of the form

$$\text{Given } u \in X, \text{ find } q \in \mathcal{X} : \quad A(u;q) = f \tag{4}$$

where $f \in \mathcal{Y}'$ is assumed to be known, and where the uncertain operator $A(u;\cdot) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ is assumed to be boundedly invertible, at least locally for the uncertain input $u$ sufficiently close to a nominal input $u_0 \in X$, ie. for $\|u - u_0\|_X$ small enough so that, for such $u$, the response of the forward problem (4) is

$$q(u) = G(u;f) \in \mathcal{X} \ .$$

In our notation, we omit the dependence of the response on $f$ and simply write $q = G(u)$. We also assume given an observation functional $\mathcal{O}(\cdot) : \mathcal{X} \to \mathbb{R}^K$ denoting a *bounded linear observation operator* on the space $\mathcal{X}$ of system responses, i.e. $\mathcal{O} \in (\mathcal{X}^*)^K$, the dual space of the space $\mathcal{X}$ of system responses. We assume that the number of observations is finite so that $K < \infty$, and equip $\mathbb{R}^K$ with the Euclidean norm, denoted by $|\cdot|$.

In this setting, we wish to predict *computationally* an expected (under the Bayesian posterior) system response of the QoI, conditional on given, noisy measurement data $\delta$. Specifically, we assume the data $\delta$ to consist of observations of QoI system responses corrupted by additive noise, ie.

$$\delta = \mathcal{O}(G(u)) + \eta \in \mathbb{R}^K \tag{5}$$

where $\eta \in \mathbb{R}^K$ represents gaussian noise in the $K$-vector of observation funcationals $\mathcal{O}(\cdot) = (o_k(\cdot))_{k=1}^K$. In the present paper, we assume that the noise process $\eta$ is Gaussian, i.e. a random vector $\eta \sim \mathcal{N}(0,\Gamma)$, for a positive definite covariance operator $\Gamma$ on $\mathbb{R}^K$ (ie., a symmetric, positive definite $K \times K$ covariance matrix $\Gamma$) which we assume to be known. We define the *uncertainty-to-observation map* of the system by $\mathcal{G} : X \to \mathbb{R}^K$ by $\mathcal{G} = \mathcal{O} \circ G$, so that

$$\delta = \mathcal{G}(u) + \eta = (\mathcal{O} \circ G)(u) + \eta : X \mapsto L^2_\Gamma(\mathbb{R}^K)$$

where $L^2_\Gamma(\mathbb{R}^K)$ denotes random vectors taking values in $\mathbb{R}^K$ which are square integrable with respect to the Gaussian measure on $\mathbb{R}^K$. In view of Bayes' formula, we define the least squares functional (also referred to as "potential" in what follows) $\Phi : X \times \mathbb{R}^K \to \mathbb{R}$ by $\Phi(u;\delta) = \frac{1}{2}|\delta - \mathcal{G}(u)|^2_\Gamma$ where $|\cdot|_\Gamma = |\Gamma^{-\frac{1}{2}} \cdot|$. Then, the Bayesian potential takes the form

$$\Phi_\Gamma(u;\delta) = \frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1}(\delta - \mathcal{G}(u))\right) \ . \tag{6}$$

In [34] an infinite-dimensional version of the Bayes rule is shown to hold in the present setting. It states that, under appropriate continuity conditions on the uncertainty-to-observation map $\mathcal{G} = (\mathcal{O} \circ G)(\cdot)$ and the prior measure on $u$, the posterior distribution $\mu^\delta$ of $u$ given data $\delta$ is absolutely continuous with respect to the prior measure $\mu_0$. In particular, then, the Radon-Nikodym derivative of the Bayesian posterior w.r.t. to the prior measure admits a bounded density $\Theta$ w.r.t. to the prior $\mu_0$.

*2.3. Uncertainty Parametrization*

We parametrize the uncertain datum $u$ in the forward equation (4). In parametric statistical estimation, $u$ is a (low-dimensional) vector containing a few unknown parameters $(y_j)_{j\in\mathbb{J}}$, for a finite index set $\mathbb{J} = \{1, 2, ..., J\}$ with small cardinality $J$ so that $X \simeq \mathbb{R}^J$. In the context of PDEs, often the case where $u \in X$, a separable Banach space is of interest. We assume that there exists a countable, unconditional base $(\psi_j)_{j\in\mathbb{J}}$ of $X$ such that, for some "nominal" value $\langle u\rangle \in X$ of the uncertain datum $u$, and for some coefficient sequence $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$ (depending on $u - \langle u\rangle \in X$) the uncertainty $u$ is parametrized by $\boldsymbol{y}$ in the sense that there holds

$$u = u(\boldsymbol{y}) := \langle u\rangle + \sum_{j\in\mathbb{J}} y_j\psi_j \in X \tag{7}$$

with unconditional convergence. We refer to $u - \langle u\rangle$ as "fluctuation" of $u$ about the nominal value $\langle u\rangle \in X$.

Many choices for the functions $\psi_j$ in (7) are conceivable; among them are standard spline bases, but also Karhunen-Loève eigenfunctions. If the uncertain datum is an $X$-valued random field $u$ in arbitary domains $D$, and for general covariance kernels, the $\psi_j$ must be obtained numerically, eg. by Fast Multipole Methods together with a Krylov subspace iteration, cp. [33]. With $y_j$ denoting the coordinate variables, the parametrization (7) is deterministic. In order to place (5), (7) into the (probabilistic) Bayesian setting of [34], we introduce (after possibly rescaling the fluctuations) a "reference" parameter domain $U = [-1,1]^{\mathbb{J}} = \prod_{j\in\mathbb{J}}[-1,1]$, and equip this countable cartesian product of sets with the product sigma-algebra $\mathcal{B} = \bigotimes_{j\in\mathbb{J}} \mathcal{B}^1$, with $\mathcal{B}^1$ the sigma-algebra of Borel sets on $[-1,1]$. On the measurable space $(U, \mathcal{B})$ we introduce a probability measure $\mu_0$ (which will serve a Bayesian prior in what follows), and which we shall choose as $\mu_0 = \bigotimes_{j\in\mathbb{J}} \frac{1}{2}\lambda^1$ with $\lambda^1$ denoting the Lebesgue measure on $[-1,1]$. Then $(U, \mathcal{B}, \mu_0)$ becomes (as countable product of probability spaces) a probability space on the set $U$ of all sequences of coefficient vectors $\boldsymbol{y}$. Then the uncertain datum $u$ in (7) becomes a random field, with $\mu_0$ charging the possible realizations of $u$. As indicated in [12, 32, 29], *analyticity of uncertainty parametrization* (7) *with respect to the parameter sequence* $\boldsymbol{y}$ can be used to derive sparsity results for this posterior.

*2.4. $(p,\varepsilon)$-Analyticity*

Analytic dependence of responses on the components $y_j$ of the parameter $\boldsymbol{y} \in U$ plays an important role for polynomial approximation results, as well as for the sparsity of the Bayesian posterior. To state it, we recall the notion of Bernstein-ellipse which denotes the closed ellipse $\mathcal{E}_r \subset \mathbb{C}$ with foci at $z = \pm 1$ and with semiaxis sum $r > 1$, ie. $\mathcal{E}_r = \{(w + 1/w)/2 : 1 \leq |w| \leq r\}$. Note that $\operatorname{dist}(\partial\mathcal{E}_r, [-1,1]) = r - 1$ and that in the limit $r \downarrow 1$, $\mathcal{E}_r$ degenerates to $[-1,1]$.

**Definition 2.2.** *Given a summability exponent $0 < p < 1$ and a real number $\varepsilon > 0$, we say that the parametric family $\{q(\boldsymbol{y}) : \boldsymbol{y} \in U\} \subset \mathcal{X}$ is $(p,\varepsilon)$-analytic if*

$(p,\varepsilon) : 1$ *(well-posedness of the forward problem)*
*for each $\boldsymbol{y} \in U$, there exists a unique realization $u(\boldsymbol{y}) \in X$ of the uncertainty and a unique solution $q(\boldsymbol{y}) \in \mathcal{X}$ of the forward problem $(4)$. This solution satisfies the a-priori estimate*

$$\forall \boldsymbol{y} \in U : \quad \|q(\boldsymbol{y})\|_{\mathcal{X}} \leq C_0(\boldsymbol{y}) \tag{8}$$

*where $U \ni \boldsymbol{y} \mapsto C_0(\boldsymbol{y}) \in L^1(U;\mu_0)$; we say that $(4)$ is* uniformly well-posed *if in $(8)$ the bound $C_0$ does not depend on $\boldsymbol{y}$.*

$(p,\varepsilon) : 2$ *(analyticity)*
*There exists $0 \leq p \leq 1$ and a sequence $b = (b_j)_{j\in\mathbb{J}} \in \ell^p(\mathbb{J})$ such that for every $0 < \varepsilon \leq 1$, there exists $C_\varepsilon > 0$ and a sequence $\rho = (\rho_j)_{j\in\mathbb{J}}$ of poly-radii $\rho_j > 1$ such that*

$$\sum_{j\in\mathbb{J}} \rho_j b_j \leq 1 - \varepsilon\,, \tag{9}$$

*and such that solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y}) \in \mathcal{X}$ admits an analytic continuation to the open polyellipse $\mathcal{E}_\rho := \prod_{j\in\mathbb{J}} \mathcal{E}_{\rho_j} \subset \mathbb{C}^{\mathbb{J}}$ and satisfies the bound*

$$\forall \boldsymbol{z} \in \mathcal{E}_\rho : \quad \|q(\boldsymbol{z})\|_{\mathcal{X}} \leq C_\varepsilon(\boldsymbol{y}) \tag{10}$$

*where $\boldsymbol{y} := \Re(\boldsymbol{z}) \in \bigotimes_{j\in\mathbb{J}}[-\rho_j,\rho_j] \subset \mathbb{R}^{\mathbb{J}}$ and where $C_\varepsilon(\boldsymbol{y}) \in L^1(U;\mu_0)$. If in $(10)$ the bound $C_\varepsilon$ does not depend on $\boldsymbol{y}$, we say that $q(\boldsymbol{y})$ is uniformly $(p,\varepsilon)$-sparse.*

In the ensuing arguments, we shall occasionally also consider the stronger $(p,\varepsilon)$-analyticity in *open polydiscs* centered at the origin $z_j = 0 \in \mathbb{C}$ which have radii $\rho_j > 0$. We denote these polydiscs by

$$\mathcal{U}_\rho = \bigotimes_{j\in\mathbb{J}} \{z_j \in \mathbb{C} | |z_j| < \rho_j\} \subset \mathbb{C}^{\mathbb{J}}\,, \quad \rho = (\rho_j)_{j\in\mathbb{J}} \tag{11}$$

and we refer to $\rho$ as *poly-radius*. In the case when $\rho_j = 1$ for all $j \in \mathbb{J}$, we simply write $\mathcal{U}$ in place of $\mathcal{U}_\rho$ to denote the unit disc in $\mathbb{C}^{\mathbb{J}}$. Observe that $U \subset \mathcal{E}_\rho \subset \mathcal{U}_\rho$.

*2.5. $(p,\varepsilon)$-Analyticity of Affine Parametric Operator Families*

At the core of the deterministic approach proposed and analyzed here is a *reformulation of the forward problem $(4)$ with uncertain distributed parameter $u \in X$ of the* *form* (7) *as parametric deterministic operator equation, on a possibly infinite-dimensional parameter space U.* We are thus interested in expressing the posterior distribution in terms of the parametric representation (7) of the uncertain parameter $u$. For Bayesian inversion, we think of the uncertain input $u \in X$ to the forward map as random field input to the operator $A$. Assuming that the uncertain random input has finite second moments under the Bayesian prior $\mu_0$ as $X$-valued random variable, it can be represented as a ($L^2(\Omega, \mu_0; X)$-convergent) Karhunen-Loève expansion with independent random coefficients. We accordingly choose the prior $\mu_0$ on the uncertainty in Bayes' theorem 2.7 ahead as countable product probability measure. The above assumption on affine parametrization of the distributed system uncertainty by the sequence $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$ of (possibly countably many) parameters results in a parametric operator equation of the form

$$A(\boldsymbol{y}) = A_0 + \sum_{j\in\mathbb{J}} y_j A_j \in \mathcal{L}(\mathcal{X}, \mathcal{Y}') \ . \tag{12}$$

Recall that either $\mathbb{J} = \{1, 2, ..., J\}$ or $\mathbb{J} = \mathbb{N}$. All assertions proved in the sequel hold in either case, and in the former case all bounds are in particular independent of $J = \#(\mathbb{J})$.

In (12), $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$ is assumed to be an i.i.d sequence of real-valued random variables $y_j \sim \mathcal{U}(-1, 1)$, $A_0$ is a "nominal operator" (representing the non-perturbed system) and the sequence $(A_j)_{j\in\mathbb{J}} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ denotes a sequence of "fluctuations" about the "nominal operator"$A_0 = A(0)$. Affine parameter dependences (12) result when the unknown $u$ in (7) is modelled as random field via its Karhunen-Loève expansion in $X$ (or in a closed subspace $X' \subset X$).

In order for the sum in (12) to converge, we impose the following assumptions on the sequence $\{A_j\}_{j\geq 0} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y}')$. In doing so, we associate with the operator $A_j$ the bilinear forms $\mathfrak{a}_j(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ via

$$\forall v \in \mathcal{X},\ w \in \mathcal{Y} : \quad \mathfrak{a}_j(v, w) = {}_{\mathcal{Y}}\langle w, A_j v\rangle_{\mathcal{Y}'} \ , \quad j = 0, 1, 2....$$

**Assumption 2.3.** *The operator family* $\{A_j\}_{j\geq 0} \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ *in* (12) *satisfies:*

*(i) The "nominal" or "mean field" operator* $A_0 \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ *is boundedly invertible, i.e. (cf. Proposition 2.1) there exists* $\mu_0 > 0$ *such that*

$$\textbf{\textit{A1}} \quad \inf_{0\neq v\in\mathcal{X}} \sup_{0\neq w\in\mathcal{Y}} \frac{\mathfrak{a}_0(v, w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0 \ , \ \inf_{0\neq w\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}} \frac{\mathfrak{a}_0(v, w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0 \ . \tag{13}$$

*(ii) The "fluctuation" operators* $\{A_j\}_{j\geq 1}$ *are small with respect to* $A_0$ *in the following sense: there exists a constant* $0 < \kappa < 1$ *such that for* $\mu_0$ *as in* (13) *holds*

$$\textbf{\textit{A2}} \quad \sum_{j\in\mathbb{J}} b_j \leq \kappa < 1 \ , \quad \textit{where} \quad b_j := \|A_0^{-1} A_j\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y}')} \ , \tag{14}$$

*(iii) (p summability) For some* $0 < p < 1$, *the operators* $B_j$ *are p-summable, in the sense that with the sequence* $b = (b_j)_{j\in\mathbb{J}}$ *as in* (14) *holds*

$$\textbf{\textit{A3}} \quad \|b\|^p_{\ell^p(\mathbb{J})} = \sum_{j\in\mathbb{J}} b_j^p < \infty \ . \tag{15}$$

Condition (14) (and, hence, Assumption 2.3) is sufficient for the bounded invertibility of $A(\boldsymbol{y})$, uniformly with respect to the parameter sequence $\boldsymbol{y} \in U = [-1,1]^{\mathbb{J}}$ and, as we shall show ahead, also for $(p,\varepsilon)$-analyticity of the parametric solutions $q(\boldsymbol{y})$

**Theorem 2.4.** *Under Assumption 2.3, for every realization $\boldsymbol{y} \in U$ of the parameters, the affine parametric operator family $A(\boldsymbol{y})$ is boundedly invertible, uniformly with respect to the parameter sequence $\boldsymbol{y} \in U$. In particular, for the parametric bilinear form $\mathfrak{a}(\boldsymbol{y};\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ associated with $A(\boldsymbol{y}) \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ via*

$$\mathfrak{a}(\boldsymbol{y}; w, v) :=_{\mathcal{Y}} \langle v, A(\boldsymbol{y})w\rangle_{\mathcal{Y}'} \tag{16}$$

*there hold the uniform inf-sup conditions with $\mu = (1-\kappa)\mu_0$,*

$$\forall \boldsymbol{y} \in U : \quad \inf_{0\neq v\in\mathcal{X}} \sup_{0\neq w\in\mathcal{Y}} \frac{\mathfrak{a}(\boldsymbol{y};v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu \ , \ \inf_{0\neq w\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}} \frac{\mathfrak{a}(\boldsymbol{y};v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu \ . \tag{17}$$

*In particular, for every $f \in \mathcal{Y}'$ and for every $\boldsymbol{y} \in U$, the parametric operator equation*

$$\text{find } q(\boldsymbol{y}) \in \mathcal{X} : \quad \mathfrak{a}(\boldsymbol{y}; q(\boldsymbol{y}), v) = \langle f, v\rangle_{\mathcal{Y}\times\mathcal{Y}'} \quad \forall v \in \mathcal{Y} \tag{18}$$

*admits a unique solution $q(\boldsymbol{y}) = (A(\boldsymbol{y}))^{-1} f$ which is uniformly bounded over $U$, ie.*

$$\sup_{\boldsymbol{y}\in U} \|q(\boldsymbol{y})\|_{\mathcal{X}} \leq \frac{\|f\|_{\mathcal{Y}'}}{\mu} \ . \tag{19}$$

*If the observation functional $\mathcal{O} : \mathcal{X} \to \mathbb{R}^K$ comprises $K$ continuous, linear functionals $o_k \in \mathcal{X}'$, $k = 1,\ldots,K$, then*

$$\forall \boldsymbol{y} \in U : \quad |\mathcal{G}(\boldsymbol{y})| = |\mathcal{O}(q(\boldsymbol{y}))| \leq \frac{\|f\|_{\mathcal{Y}'}}{\mu} \Big(\sum_{k=1}^{K} \|o_k\|^2_{\mathcal{X}'}\Big)^{\frac{1}{2}} \ . \tag{20}$$

*The forward maps $q : U \to \mathcal{X}$ and $\mathcal{G} : U \to \mathbb{R}^K$ are globally Lipschitz.*

*Specifically (see [32, Lemma 3.3, Theorem 3.4]) if $q$ and $\tilde{q}$ are solutions of (4) with the same right hand side $f$ with operators $A(\boldsymbol{y})$ and $A(\boldsymbol{y}')$, respectively, then the forward solution map $\boldsymbol{y} \to q(\boldsymbol{y}) = (A(\boldsymbol{y}))^{-1} f$ is Lipschitz as a mapping from $U$ into $\mathcal{X}$, ie. there exists a constant $C > 0$ (depending only on $\kappa$ and $\mu_0$ in Assumption 2.3) such that for every $\boldsymbol{y}, \tilde{\boldsymbol{y}} \in U$ holds*

$$\|q(\boldsymbol{y}) - q(\tilde{\boldsymbol{y}})\|_{\mathcal{X}} \leq C\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|_{\ell^\infty}\|f\|_{\mathcal{Y}'} \ . \tag{21}$$

*Moreover, the parametric response map $U \ni \boldsymbol{y} \to \mathcal{G}(\boldsymbol{y}) := (\mathcal{O}\circ q)(\boldsymbol{y})$ is globally Lipschitz as a mapping from $\ell^\infty$ into $\mathbb{R}^K$, in the sense that*

$$|\mathcal{G}(\boldsymbol{y}) - \mathcal{G}(\tilde{\boldsymbol{y}})| \leq C\Big(\sum_{k=1}^{K} \|o_k\|^2_{\mathcal{X}'}\Big)^{\frac{1}{2}} \|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|_{\ell^\infty(\mathbb{J})}\|f\|_{\mathcal{Y}'} \ . \tag{22}$$

*Proof.* The assertions (19) and (20) are straightforward consequences of the inf-sup conditions (17). We therefore address the Lipschitz dependence (21) and (22). For any $\boldsymbol{y} \in U$, the equation $f = A(\boldsymbol{y})q(\boldsymbol{y})$ implies $\|q(\boldsymbol{y})\|^{-1}_{\mathcal{X}} \leq \|A(\boldsymbol{y})\|\|f\|_{\mathcal{Y}'}$.

For $\boldsymbol{y},\tilde{\boldsymbol{y}} \in U$, we write $A$ in place of $A(\boldsymbol{y})$ and $\tilde{A}$ in place of $A(\tilde{\boldsymbol{y}})$. We observe

$$A = A(\boldsymbol{y}) = A_0(I + \sum_{j\geq 1} y_j A_0^{-1} A_j) =: A_0(I-F) \;, \quad F := -\sum_{j\geq 1} y_j B_j \;.$$

We write analogously $\tilde{A} = A_0(I-\tilde{F})$.

Due to (13), $A_0$ is boundedly invertible, and $B_j := A_0^{-1}A_j \in \mathcal{L}(\mathcal{X},\mathcal{X})$ is well-defined for every $j \in \mathbb{J}$. By Assumption (14), we also have for every $\boldsymbol{y},\tilde{\boldsymbol{y}} \in U$ the following bound of the "fluctuation" operators $F,\tilde{F}$ (which express the relative deviation of the forward system from its nominal state $A_0$) $\|F\| \leq \|\boldsymbol{y}\|_\infty \sum_{j\geq 1} b_j \leq \kappa < 1$, and likewise $\|\tilde{F}\| \leq \kappa < 1$. Therefore, by a Neumann series argument, for all $\boldsymbol{y} \in U$, the inverses $(I-F)^{-1}$ and $(I-\tilde{F})^{-1}$ exist and are bounded by $1/(1-\|F\|) \leq 1/(1-\kappa) = 1/(1-\kappa)$. The inverses are given explicitly by the (norm-convergent in $\mathcal{L}(\mathcal{X},\mathcal{X})$) geometric (or Neumann) series

$$(I-F)^{-1} = \sum_{j\geq 0} F^j \;, \qquad (I-\tilde{F})^{-1} = \sum_{j\geq 0} \tilde{F}^j \;.$$

Therefore we may write, for every $\boldsymbol{y},\tilde{\boldsymbol{y}} \in U$, with $q$ and $\tilde{q}$ denoting the corresponding solutions,

$$\begin{aligned} q - \tilde{q} &= (A^{-1} - \tilde{A}^{-1})f = ((I-F)^{-1} - (I-\tilde{F})^{-1})A_0 f \\ &= \left[\textstyle\sum_{j\geq 1}(F^j - \tilde{F}^j)\right] A_0^{-1} f \;. \end{aligned} \tag{23}$$

We estimate

$$\begin{aligned} \Delta := \left\| \sum_{j\geq 1}(F^j - \tilde{F}^j) \right\| &= \left\| F\sum_{j\geq 1} F^{j-1} - F\sum_{j\geq 1}\tilde{F}^{j-1} + F\sum_{j\geq 1}\tilde{F}^{j-1} - \tilde{F}\sum_{j\geq 1}\tilde{F}^{j-1} \right\| \\ &\leq \|F\| \left\| \sum_{j\geq 1} F^{j-1} - \tilde{F}^{j-1} \right\| + \|F-\tilde{F}\| \left\| \sum_{j\geq 1} \tilde{F}^{j-1} \right\| \\ &\leq \|F\|\Delta + \|F-\tilde{F}\| \sum_{j\geq 0} \|\tilde{F}\|^j \;, \end{aligned}$$

from where we find

$$(1-\|F\|)\Delta \leq \|F-\tilde{F}\| \sum_{j\geq 0} \|\tilde{F}\|^j = \frac{\|F-\tilde{F}\|}{1-\|\tilde{F}\|}$$

or, finally, recalling that (14) implies $\|F\| \leq \kappa < 1$ and $\|\tilde{F}\| \leq \kappa < 1$,

$$\left\| \sum_{j\geq 1}(F^j - \tilde{F}^j) \right\| \leq \frac{\|F-\tilde{F}\|}{(1-\|\tilde{F}\|)(1-\|F\|)} \leq \frac{\|F-\tilde{F}\|}{(1-\kappa)^2} \;.$$

With (23) it follows that

$$\|q(\boldsymbol{y}) - q(\tilde{\boldsymbol{y}})\|_{\mathcal{X}} \leq \frac{\|\boldsymbol{y}-\tilde{\boldsymbol{y}}\|_\infty \|b\|_1}{(1-\kappa)^2} \|A_0^{-1} f\|_{\mathcal{X}} \leq \|\boldsymbol{y}-\tilde{\boldsymbol{y}}\|_\infty \frac{\kappa}{\mu_0(1-\kappa)^2} \|f\|_{\mathcal{Y}'}$$

which is (21). The assertion (22) then follows from the linearity and the boundedness of the observation functional $\mathcal{O}(\cdot) : \mathcal{X} \mapsto \mathbb{R}^K$. $\square$

To verify $(p,\varepsilon)$-analyticity, we consider a complex extension of the forward problem. To this end, we extend the spaces $\mathcal{X}$ and $\mathcal{Y}$ to spaces over $\mathbb{C}$.

We write, for every $j \in \mathbb{J}$, $z_j = y_j + \boldsymbol{i} w_j$ where $\boldsymbol{i} := \sqrt{-1}$ denotes the imaginary unit. We extend the forms $\mathfrak{a}_0(\cdot,\cdot)$ and $\mathfrak{a}_j(\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ in the usual fashion‡

to sesquilinear forms $\mathfrak{a}_j(\cdot,\cdot) : \mathcal{X} \times \mathcal{Y} \to \mathbb{C}$ so that

$$\mathfrak{a}_0(v,w) :={}_{\mathcal{Y}} \langle w, \overline{A_0 v}\rangle_{\mathcal{Y}'} \;, \quad \mathfrak{a}_j(v,w) :={}_{\mathcal{Y}} \langle w, \overline{A_j v}\rangle_{\mathcal{Y}'} \;, \quad v \in \mathcal{X},\; w \in \mathcal{Y}$$

with the overline denoting complex conjugation. Denoting by $|z_j| = (|y_j|^2 + |w_j|^2)^{1/2}$ the modulus of $z_j \in \mathbb{C}$, we impose the following complex analog of A1 in Assumption 2.3, (13): there exists a constant $\mu_0 > 0$ such that

$$\textbf{A1}\mathbb{C} \quad \inf_{0\neq v\in\mathcal{X}} \sup_{0\neq w\in\mathcal{Y}} \frac{\Re\mathfrak{a}_0(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0 \;, \inf_{0\neq w\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}} \frac{\Re\mathfrak{a}_0(v,w)}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq \mu_0 \;. \tag{24}$$

We recall from (A2) in Assumption 2.3 the definition of the sequence $b_j$ which remain unchanged in the complex variable setting.

By Proposition 2.1, Assumption **A1**$\mathbb{C}$ implies that the operator $A_0 \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ is boundedly invertible, and, for $z_j \in \mathbb{C}$, we may therefore consider the parametric operator family

$$A(\boldsymbol{z}) := A_0 + \sum_{j\in\mathbb{J}} z_j A_j \in \mathcal{L}(\mathcal{X},\mathcal{Y}'), \quad z \in \mathbb{C}^{\mathbb{J}} \tag{25}$$

or, equivalently, recalling $B_j := A_0^{-1}A_j$ for $j \in \mathbb{J}$,

$$B(\boldsymbol{z}) := I + \sum_{j\in\mathbb{J}} z_j A_0^{-1} A_j = I + \sum_{j\in\mathbb{J}} z_j B_j \in \mathcal{L}(\mathcal{X},\mathcal{X}), \quad z \in \mathbb{C}^{\mathbb{J}} \;. \tag{26}$$

**Lemma 2.5.** *Under Assumption A1$\mathbb{C}$, **A2** and **A3**, for every poly-radius $\rho$ such that*

$$\forall j : \rho_j > 1 \;, \quad \kappa < 1 - \varepsilon := \sum_{j\in\mathbb{J}} \rho_j b_j < 1 \tag{27}$$

*with $b_j := \|A_0^{-1}A_j\|$, the parametric operator equation*

$$\text{find } \; q(\boldsymbol{z}) : \quad A(\boldsymbol{z})q(\boldsymbol{z}) = f \quad \text{in} \quad \mathcal{Y}' \;, \quad \forall \boldsymbol{z} \in \mathcal{U}_\rho \tag{28}$$

*admits a unique solution for every $\boldsymbol{z} \in \mathcal{U}_\rho$ which satisfies the apriori estimate*

$$\sup_{z\in\mathcal{U}_\rho} \|q(\boldsymbol{z})\|_{\mathcal{X}} \leq \frac{\|f\|_{\mathcal{Y}'}}{\mu_0 \varepsilon} \;. \tag{29}$$

*Moreover, for every $\boldsymbol{z} \in \mathcal{U}_\rho$, the solution $q(\boldsymbol{z})$ is a holomorphic function taking values in $\mathcal{X}$ of each coordinate $z_j$ of $z \in \mathcal{U}_\rho$.*

*If, moreover, **A3** holds with some $0 < p < 1$, then the forward solution $q(\boldsymbol{z})$ is uniformly $(p,\varepsilon)$-analytic.*

‡ If $\mathcal{X}$ is a Hilbertspace, for $u_1, u_2, v_1, v_2 \in \mathcal{X}$ we set $u = u_1 + \boldsymbol{i}u_2$ and $v = v_1 + \boldsymbol{i}v_2$ with $i = \sqrt{-1}$. Then $u, v \in \mathcal{X}_{\mathbb{C}}$, the "complexified" version of the Hilbert space $\mathcal{X}$, which is a Hilbert space with the innerproduct $(u,v)_{\mathbb{C}} := (u_1,v_1) + (u_2,v_2) + \boldsymbol{i}[(u_2,v_1) - (v_1,v_2)]$. Linear operators $A \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ are extended via $A_{\mathbb{C}}u := Au_1 + \boldsymbol{i}Au_2$ and a bilinear form $a(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ to a sesquilinear form $a_{\mathbb{C}}(\cdot,\cdot)$ via $a_{\mathbb{C}}(u,v) := a(u_1,v_1) + a(u_2,v_2) + \boldsymbol{i}[a(u_1,v_2) + a(u_2,v_1)]$. We omit the subscript $\mathbb{C}$.

*Proof.* Using the bounded invertibility of $A_0$, we convert (28) to the equivalent form $B(\boldsymbol{z})q(\boldsymbol{z}) = g$ where $g := A_0^{-1}f \in \mathcal{X}$ and where $B(\boldsymbol{z})$ is as in (26).

We apply Proposition 2.1 and verify to this end the inf-sup conditions (2) for the parametric operator $B(\boldsymbol{z})$. To this end, for every $v \in \mathcal{X}$ and for every $\boldsymbol{z} \in \mathcal{U}_\rho$ we have

$$\Re\left( {}_{\mathcal{X}}\langle v, \overline{B(\boldsymbol{z})v}\rangle_{\mathcal{X}} \right) \quad \geq \quad \|v\|_{\mathcal{X}}^2 \left(1 - \sum_{j\in\mathbb{J}} |z_j| b_j\right) = \mu \|v\|_{\mathcal{X}}^2$$

where $\mu = 1 - \hat{\kappa} = 1 - (1-\varepsilon) = \varepsilon > 0$ by (27). Hence, for every $\boldsymbol{z} \in \mathcal{U}_\rho$, the operator $B(\boldsymbol{z})$ satisfies the uniform inf-sup conditions (2) with constant $\mu = \varepsilon > 0$. Therefore, for every $\boldsymbol{z} \in \mathcal{U}_\rho$ the parametric operator equation $B(\boldsymbol{z})q(\boldsymbol{z}) = g = A_0^{-1}f$ admits a unique solution $q(\boldsymbol{z}) \in \mathcal{X}$ which satisfies the a-priori estimate

$$\sup_{\boldsymbol{z}\in\mathcal{U}_\rho} \|q(\boldsymbol{z})\|_{\mathcal{X}} \leq \frac{\|g\|_{\mathcal{X}}}{\mu} = \frac{\|A_0^{-1}f\|_{\mathcal{X}}}{\mu} \leq \frac{\|f\|_{\mathcal{Y}'}}{\mu_0 \mu}$$

which is (29).

To show the strong holomorphy of the $\mathcal{X}$-valued function $q(\boldsymbol{z})$ with respect to all components $z_j$ of the parameter sequence $\boldsymbol{z}$, we observe that the power series argument in the proof of Theorem 2.4 extends verbatim to the complex setting.

Alternatively, a difference quotient argument as in [12] shows that the parametric mapping $q(\boldsymbol{z})$ is complex differentiable w.r. to $z_j$ at $\boldsymbol{z} \in \mathcal{U}_\rho$ for every $j \in \mathbb{J}$. Since the parameter dependence is affine, the complex derivative $\partial_{z_j} q(\boldsymbol{z})$ satisfies the parametric equation

$$A(\boldsymbol{z})(\partial_{z_j} q(\boldsymbol{z})) = -A_j q(\boldsymbol{z}) \;, \quad \boldsymbol{z} \in \mathcal{U}_\rho \;.$$

Since complex differentiability and the reflexivity of $\mathcal{X}$ imply strong holomorphy of the function $q(\boldsymbol{z})$, we conclude. The proof of $(p, \varepsilon)$-analyticity follows from the complex differentiability as in [12, 14, 10]. □

Given $\kappa = \sum_{j\geq 1} b_j < 1$ as in $A2$ of Assumption 2.3, call a poly-radius $\rho$ which satisfies (27) for some sufficiently small $\varepsilon > 0$ a *$\varepsilon$-admissible* poly-radius. Evidently, every $\varepsilon$-admissible poly-radius $\rho$ is also $\varepsilon'$-admissible for every $0 < \varepsilon' < \varepsilon < 1 - \kappa$.

## *2.6. Examples*

We illustrate the foregoing general concepts with several examples: diffusion problems with an isotropic, random coefficient function, in the stationary as well as in the time-dependent case.

*2.6.1. Elliptic Divergence-form Equations with Uncertain Coefficient* Let $D \subset \mathbb{R}^d$ be a bounded interval if $d = 1$ or a bounded domain in $\mathbb{R}^d$ when $d \geq 2$, with Lipschitz boundary $\partial D$. Let further $\left(H, (\cdot,\cdot), \|\cdot\|\right)$ denote the Hilbert space $L^2(D)$ which we will identify throughout with its dual space, i.e. $H \simeq H^*$. The stationary diffusion problem will fit in the abstract setting with the choices $\mathcal{X} = \mathcal{Y} = V = H^1_0(D)$. The dual space $\mathcal{X}' = V^*$ is isomorphic to the Hilbert space $H^{-1}(D)$. Unless explicitly stated otherwise, we shall assume for the (deterministic) data $f$ that $f \in V^*$. For given $f \in L^2(D)$, and for an uncertain diffusion tensor $u \in X = (L^\infty(D))^{d\times d}_{sym}$, we consider the stationary elliptic diffusion problem

$$Aq := -\nabla \cdot \big(u\nabla q\big) = f \quad \text{in} \quad D, \qquad q = 0 \quad \text{in} \quad \partial D. \tag{30}$$

A weak solution of (30) is $q \in \mathcal{X}$ which satisfies

$$\int_D u(x)\nabla q(x) \cdot \nabla \tilde{q}(x) dx = _{\mathcal{X}}\langle \tilde{q}, f\rangle_{\mathcal{X}^*} \ \text{ for all } \ \tilde{q} \in \mathcal{X}\ . \tag{31}$$

Here $_{\mathcal{X}}\langle \cdot, \cdot \rangle_{\mathcal{X}^*}$ denotes the dual pairing between elements of $V$ and $V^*$, and in (30) the uncertain diffusion coefficient $u(x) \in (L^\infty(D))^{d\times d}_{sym}$ admits the representation

$$u(x, \boldsymbol{y}) = \bar{a}(x) + \sum_{j\in\mathbb{J}} y_j \psi_j(x), \quad x \in D, \tag{32}$$

where $\bar{a} \in (L^\infty(D))^{d\times d}_{sym}$ and $(\psi_j)_{j\in\mathbb{J}} \subset (L^\infty(D))^{d\times d}_{sym}$ . The problem (30) formally fits into Assumption 2.3 if we choose in (12) the operators $A_0 = -\nabla \cdot (\bar{a}\nabla)$ and, for $j \geq 1$, $A_j = -\nabla \cdot (\psi_j \nabla)$. The validity of (17) is ensured by

**Assumption 2.6.** *For every $x \in D$ and for every $0 \neq \xi \in \mathbb{R}^d$ holds*

$$0 < a_{\text{MIN}} \leq \frac{\xi^\top u(x)\xi}{\xi^\top \xi} \leq a_{\text{MAX}} < \infty\ , \quad 0 < \bar{a}_{\min} \leq \frac{\xi^\top \bar{a}(x)\xi}{\xi^\top \xi} \leq \bar{a}_{max} < \infty \tag{33}$$

*and there exists $0 \leq \kappa < 1$ such that the tensors $\bar{a}$ and $\psi_j$ in (32) satisfy*

$$\sum_{j\in\mathbb{J}} \sup_{0\neq\xi\in\mathbb{R}^d} \left\| \frac{\xi^\top \psi_j(x)\xi}{\xi^\top \xi} \right\|_{L^\infty(D)} \leq \kappa \bar{a}_{\text{MIN}},$$

*with* $\overline{a}_{\text{MIN}} = \operatorname{ess\,inf}_{x\in D} \inf_{0\neq\xi\in\mathbb{R}^d} \frac{\xi^\top \bar{a}(x)\xi}{\xi^\top\xi} > 0$ .

Assumption 2.6 implies that the operator $A$ in (30) with coefficient (32) satisfies (2). Therefore, in particular, we have uniqueness of the response $q(\boldsymbol{y})$ of (31).

$$\|G(u)\|_{\mathcal{X}} = \|q\|_V \leq \frac{\|f\|_{V^*}}{a_{\text{MIN}}} \tag{34}$$

where we denoted by $\|q\|_{\mathcal{X}} = \|\nabla q\|_{L^2(D)}$ the norm in $V$. The inverse problem consists of determining the unknown diffusion coefficient $u$ from given noisy observation data $\delta$ in order to compute the expectation of a quantity of interest (47), given this data (see [32] for the proof). Analogous results hold for problems of linearized elasticity in $D$, where the response $q(\boldsymbol{y}) : D \mapsto \mathbb{R}^d$ is a displacement field, and the uncertain "coefficient" is the spatially inhomogeneous, parametric fourth order complicance tensor $\boldsymbol{C}_{klmn}(\boldsymbol{x}, \boldsymbol{y})$ in the Lame-Navier equations

$$\text{div}\sigma[q] = f \text{ in } D, \quad \sigma[q] = \boldsymbol{C}(\boldsymbol{x}, \boldsymbol{y})\epsilon[q], \epsilon[q] = \frac{1}{2}(\text{grad}q + \text{grad}q^\top) \in \mathbb{R}^{d\times d}_{sym}\ .$$

Here, $A(\boldsymbol{y})q = -\mathrm{div}(\boldsymbol{C}(\cdot,\boldsymbol{y})\epsilon[q])$, $X = (L^\infty(D))^{d^2\times d^2}_{sym}$, $V$ is a subspace of $H^1(D)^d$ with essential boundary conditions to be satisfied by $q(\boldsymbol{y})$. The uncertain compliance tensor $\boldsymbol{C}$ admits the expansion $\boldsymbol{C}(\boldsymbol{x},\boldsymbol{y}) = \langle \boldsymbol{C}\rangle(\boldsymbol{x}) + \sum_{j\in\mathbb{J}} y_j\psi_j(\boldsymbol{x})$ with suitable (fourth order) tensor functions $\psi_j \in X$ which could be, for example, principal components of a given covariance (eight order) tensor kernel function.

*2.6.2. Parabolic Problems with Uncertain Operator* The previous examples addressed the case when the abstract forward equation (18) is elliptic in nature. However, the general, parametric operator (12) in the forward equation (18) also accomodates parabolic problems as we shall show next. To this end, we denote by $B(\boldsymbol{y}) \in \mathcal{L}(V,V^*)$ a parametric operator pencil with affine parameter dependence (12) of an elliptic operator family such as those considered in Section 2.6.1. We further assume that we are given a second Hilbert space $H$ which we identify with its own dual $H^*$ which constitutes a Gel'fand evolution triple

$$V \subseteq H \simeq H^* \subseteq V^* \,. \tag{35}$$

For the parametric family $B(\boldsymbol{y})$ we assume the validity of a Garding in equality, i.e. that there exist a constant $\alpha > 0$ and a compact bilinear form $k(\cdot,\cdot) : V\times V \to \mathbb{R}$ such that

$$\forall \boldsymbol{y} \in U\,, \forall v \in V : \quad b(\boldsymbol{y};v,v) + k(v,v) \geq \alpha\|v\|_V^2\,. \tag{36}$$

For the (space-time) variational formulation of the evolution problems, for a given time horizen $0 < T < \infty$, we define the Bochner spaces

$$\mathcal{X} = L^2(0,T;V)\cap H^1(0,T;V^*),\ \ \mathcal{Y} = L^2(0,T;V)\times H\,. \tag{37}$$

Then the parametric evolution operator is, formally, given by $A(\boldsymbol{y}) := (\partial_t + B(\boldsymbol{y}), \iota_0)$ where $\iota_0$ denotes the time $t=0$ trace of the argument, i.e. $\iota_0 u = u(0)$. It follows from the continuous embedding $\mathcal{X} \subset C^0([0,T];H)$ that for every $v\in\mathcal{X}$, $\iota_0 v$ is well-defined as an element of $H$ and there holds the continuity estimate

$$\|\iota_0 v\|_H \leq C_T\|v\|_{\mathcal{X}},\ \text{where}\ \ \|v\|_{\mathcal{X}} := \left(\|v\|^2_{L^2(0,T;V)} + \|v\|^2_{H^1(0,T;V^*)}\right)^{1/2}\,.$$

In this case, the space-time variational formulation of the parametric forward model $A(\boldsymbol{y})p(\boldsymbol{y}) = f$ is given, for $v\in\mathcal{X}$ and for $w=(w_1,w_2)\in\mathcal{Y}$, by the bilinear form

$$\begin{aligned}\mathfrak{a}(\boldsymbol{y};v,w) &:= \int_0^T \left({}_V\langle w_1,\partial_t v\rangle_{V^*} + {}_V\langle w_1, B(\boldsymbol{y})v\rangle_{V^*}\right)dt + {}_H\langle v(0),w_2\rangle_H \\ &= \int_0^T \left({}_V\langle w_1,\partial_t v\rangle_{V^*} + b(\boldsymbol{y};v,w_1)\right)dt + {}_H\langle v(0),w_2\rangle_H\end{aligned} \tag{38}$$

and by the linear form

$$\ell(w) = \int_0^T \left({}_V\langle w_1(\cdot,t), f(\cdot,t)\rangle_{V^*}\right)dt + {}_H\langle w_2,u_0\rangle_H\,. \tag{39}$$

Note that then in the weak formulation

$$\forall \boldsymbol{y}\in U: \quad p(\boldsymbol{y})\in\mathcal{X} : \mathfrak{a}(\boldsymbol{y};p(\boldsymbol{y}),w) = \ell(w) \quad \forall w\in\mathcal{Y} \tag{40}$$

the initial condition $\iota_0 u = u_0$ has been imposed weakly. For the variational space-time formulation (40) it is once more known (see, eg., [31, Appendix]) that the parametric bilinear form satisfies the inf-sup conditions (17), provided that the parametric spatial operator $B(\boldsymbol{y})$ satisfies the Garding inequality (36).

*2.6.3. Elliptic Multiscale Problems with Uncertainty* In [17], a general framework for uncertainty modelling in elliptic divergence form equations with scale-separated, uncertain coefficients $a^\varepsilon(\boldsymbol{y};\boldsymbol{x}) = a(\boldsymbol{y};\boldsymbol{x},\varepsilon^{-1}\boldsymbol{x})$ where $0 < \varepsilon << 1$ is a *known* nondimensional length scale parameter and where $a(\boldsymbol{y};\boldsymbol{x},\xi)$ is independent of $\varepsilon$, 1-periodic w.r. to $\xi$ and depends on $\boldsymbol{y}$ once more in an affine fashion (see [17, Eqns. (1.7), (1.10)] for details). Such problems fit once more into the general framework of the present paper, *with all bounds in error estimates valid uniformly w.r. to $\varepsilon$ due to the use of two-scale convergence and the avoidance of homogenization formulas.*

*2.7. Parametric Bayesian Posterior*

Motivated by [32, 29], the basis for the presently proposed, adaptive deterministic quadrature approaches for Bayesian estimation via the computational realization of Bayes' formula is a *parametric, deterministic representation* of the derivative of the posterior measure with respect to the uniform prior measure $\mu_0$. The prior measure $\mu_0$ being uniform, we admit in (12) sequences $y$ which take values in the parameter domain $U = [-1,1]^{\mathbb{J}}$ . As explained in Section 2.3, we consider the parametric, deterministic forward problem in the probability space

$$(U, \mathcal{B}, \mu_0) . \tag{41}$$

*We assume throughout what follows that the prior measure on the uncertain input data, parametrized in the form* (12)*, is the uniform measure $\mu_0(d\boldsymbol{y})$.* We add in passing that unbounded parameter ranges as arise, e.g., in lognormal random diffusion coefficients in models for subsurface flow [25], can be treated by the techniques developed here, at the expense of additional technicalities. With the parameter domain $U$ as in (41) the parametric forward map $\Xi : U \to \mathbb{R}^K$ is given by

$$\Xi(\boldsymbol{y}) = \mathcal{G}(u)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j} . \tag{42}$$

The mathematical foundation of Bayesian inversion is Bayes' theorem. We present a version of it, from [32] (see also [34]). To do so, we view $U$ as unit ball in $\ell^\infty(\mathbb{J})$, the Banach space of bounded sequences taking values in $U$.

**Theorem 2.7.** *Assume that $\Xi : \bar{U} \to \mathbb{R}^K$ is bounded and continuous. Then $\mu^\delta(d\boldsymbol{y})$, the distribution of $y \in U$ given $\delta$, is absolutely continuous with respect to $\mu_0(d\boldsymbol{y})$, ie.*

$$\frac{d\mu^\delta}{d\mu_0}(bsy) = \frac{1}{Z}\Theta(\boldsymbol{y}) \tag{43}$$

*with the parametric Bayesian posterior* $\Theta$ *given by*

$$\Theta(\boldsymbol{y}) = \exp\bigl(-\Phi_\Gamma(u;\delta)\bigr)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j}, \tag{44}$$

*where the Bayesian potential* $\Phi_\Gamma$ *is as in* (6) *and the normalization constant* $Z$ *is given by*

$$Z = \mathbb{E}^{\mu_0}[1] = \int_U \Theta(\boldsymbol{y}) d\mu_0(\boldsymbol{y}) \; . \tag{45}$$

Computational Bayesian inversion is concerned with approximation of a "most likely" *system response* $\phi : X \to \mathcal{S}$ (sometimes also referred to as *Quantity of Interest (QoI)* which may take values in a Banach space $\mathcal{S}$) for given (noisy) observation data $\delta$ of the QoI $\phi$. In particular the choices $\phi(u) = G(u)$ (with $\mathcal{S} = \mathcal{X}$) and $\phi(u) = G(u) \otimes G(u)$ (with $\mathcal{S} = \mathcal{X} \otimes \mathcal{X}$) facilitate computation of the "most likely" (given the data $\delta$) mean and covariance of the system's response.

With the QoI $\phi$ we associate the (infinite-dimensional) parametric map

$$\begin{aligned} \Psi(\boldsymbol{y}) &= \Theta(\boldsymbol{y})\phi(u)\;|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j} \\ &= \exp\bigl(-\Phi_\Gamma(u;\delta)\bigr)\phi(u)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j} : U \to \mathcal{S} \; . \end{aligned} \tag{46}$$

Then the Bayesian estimate of the QoI $\phi$, given noisy data $\delta$, takes the form

$$\begin{aligned} \mathbb{E}^{\mu^\delta}[\phi] &= \tfrac{1}{Z}\int_{y\in U} \Psi(\boldsymbol{y})\mu_0(d\boldsymbol{y}) \\ &= \frac{1}{Z}\int_{y\in U} \exp\bigl(-\Phi_\Gamma(u;\delta)\bigr)\phi(u)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j}\mu_0(d\boldsymbol{y}) \; . \end{aligned} \tag{47}$$

Our aim is to approximate the expectations $Z'$ and $Z$ which, in the parametrization with respect to $\boldsymbol{y} \in U$, take the form of infinite-dimensional integrals with respect to the uniform prior $\mu_0(d\boldsymbol{y})$.

In the next section we establish the joint analyticity of the posterior densities $\Theta(\boldsymbol{y})$ and $\Psi(\boldsymbol{y})$, as a function of the parameter sequence $\boldsymbol{y} \in U$. Following [32], we then deduce sharp estimates on size of domain of analyticity of the forward solution $q(\boldsymbol{y})$ and of the densities $\Theta(\boldsymbol{y})$ and $\Psi(\boldsymbol{y})$ as a function of each coordinate $y_j$, $j \in \mathbb{N}$. These will then be used to infer sparsity of gpc expansions which, in turn, are the basis for $N$-term approximation rates as well as of convergence rates for various quadrature methods.

## 3. Sparsity of the Forward Solution

As shown in [32, 29] for scalar, isotropic diffusion problems, *dimension-independent convergence rates* of numerical approximations of integrals like (45), (47) are based on *sparsity results* for the posterior density $\Theta$ which arises in Bayes' theorem. In the present section, we establish such sparsity results in the general setting of our affine parametric uncertainty model, ie. the operator equation (12). As in [12, 14], the sparsity results will be based on analytic dependence of the forward solution $q(\boldsymbol{y})$ of the parametric operator equation (18), with precise bounds on the size of domain of analyticity.

*3.1. Sparsity*

Sparsity of the dependence of the forward solution $q(\boldsymbol{y})$ on the parameter sequence $\boldsymbol{y}$ is a consequence of the $(p, \varepsilon)$-analyticity established in Section 2.5: we approximate the parametric solution $q(\boldsymbol{y})$ by partial sums of tensorized Taylor and Legendre series. As was shown in [11, 12, 8, 14], $(p, \varepsilon)$-analyticity of the forward solution $q(\boldsymbol{y})$ implies that such expansions are *sparse*. Sparsity of tensorized Taylor expansions requires $(p, \varepsilon)$-analyticity of the forward map on the *polydiscs* $\mathcal{U}_\rho$ (as in [14]), whereas $(p, \varepsilon)$-analyticity of $q(\boldsymbol{y})$ on the (smaller) poly-ellipses $\mathcal{E}_\rho$ (as in [12]) suffices for sparsity of Legendre expansions.

Unconditional convergence and $p$-sparsity of forward maps are available for various Legendre and Tschebyscheff expansions, also for nonaffine parameter dependence, and for certain nonlinear operator equations (see, eg., [12, 14, 10]). To define the Legendre polynomial chaos expansions, we introduce the univariate Legendre polynomials $L_k(z_j)$ of degree $k^{th}$ of the variable $z_j \in \mathbb{C}$, normalized such that

$$\int_{-1}^{1} (L_k(t))^2 \frac{dt}{2} = 1, \quad k = 0, 1, 2, ... \tag{48}$$

Since $L_0 \equiv 1$, the Legendre polynomials $L_k$ in (48) can be tensorized on the parameter domains $U$ via

$$L_\nu(\boldsymbol{z}) = \prod_{j \in \mathbb{J}} L_{\nu_j}(z_j), \quad \boldsymbol{z} \in \mathbb{C}^{\mathbb{J}},\ \nu \in \mathcal{F}\ . \tag{49}$$

The set of tensorized Legendre polynomials

$$\mathbb{L}(U) = \{L_\nu : \nu \in \mathcal{F}\} \tag{50}$$

forms a countable orthonormal basis in $L^2(U, \mu_0)$.

This observation suggests, by virtue of the square integrability discussed below, approximations by truncated mean square convergent gpc-expansions such as

$$q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} q_\nu L_\nu(\boldsymbol{y})\ , \quad \boldsymbol{y} \in U\ . \tag{51}$$

For the statement of sparsity in the response map, we shall approximate the parametric solution $q(\boldsymbol{y})$ of (18) and also the Bayesian posterior density in terms of $N$-term truncations of the Taylor or Legendre series (51).

Truncations of tensorized Taylor and Legendre expansions take the form of partial sums over finite sets $\Lambda_N \subset \mathcal{F}$ of indices of cardinality at most $N$. We shall say that a sequence $(\Lambda_N)_{N\geq 1} \subset \mathcal{F}$ of index sets *exhausts* $\mathcal{F}$, if for every finite subset $\Lambda \subset \mathcal{F}$ there exists $N_0(\Lambda)$ such that for all $N \geq N_0$, $\Lambda \subset \Lambda_N$. We recall that, by Lemma 2.5, under assumptions $A1\mathbb{C}$, **A2** and **A3**, the parametric forward solution $q(\boldsymbol{y})$ of the forward equation (4) with the affine-parametric operator family (12) is $(p, \varepsilon)$-analytic on a family $\mathcal{U}_\rho$ of polydiscs (and therefore also on a family $\mathcal{E}_\rho$ of poly-ellipses).

The sparsity results which follow are based on establishing $p$-summability of (the $\mathcal{X}$-norms of) Legendre coefficients of the parametric forward solutions $q(\boldsymbol{y})$ and, in the next section, also of the Bayesian posterior density. The $p$-summability (with exponent $p$ as in the sparsity assumption (9) of Definition 2.2) will imply convergence rates of best $N$-term truncations of generalized polynomial chaos (gpc for short) expansions. In general, however, sets $\Lambda_N \subset \mathcal{F}$ of $N$ largest gpc coefficients could be quite arbitrary. In view of numerical approximations it is important to have further information about their structure. For general, $(p,\varepsilon)$-analytic, parametric mappings, it was shown in [8] that partial sums of (51) with summation over nested sequences of so-called *monotone* index sets $\Lambda_N \subset \mathcal{F}$ of cardinality at most $N$ already achieve the convergence rates of best $N$-term approximations, albeit with a possibly worse constant (cp. [8, Remarks 2.2 and 2.3]).

**Definition 3.1.** *(Monotone Index Sets) A subset $\Lambda_N \subset \mathcal{F}$ of finite cardinality $N$ is called* monotone *if (M1) $\{0\} \subset \Lambda_N$ and if (M2) $\forall 0 \neq \nu \in \Lambda_N$ it holds that $\nu - e_j \in \Lambda_N$ for all $j \in \mathbb{I}_\nu$, where $e_j \in \{0,1\}^{\mathbb{J}}$ denotes the index vector with $1$ in position $j \in \mathbb{J}$ and $0$ in all other positions $i \in \mathbb{J}\backslash\{j\}$.*

Properties (M1) and (M2) in Definition 3.1 imply for monotone $\Lambda_N \subset \mathcal{F}$

$$\mathbb{L}_{\Lambda_N}(U) = \mathrm{span}\{y^\nu : \nu \in \Lambda_N\} = \mathrm{span}\{L_\nu : \nu \in \Lambda_N\}\ .$$

Closely related to the notion of monotone index sets is the notion of *monotone majorant* which was introduced in [8] (see also [9, 10]).

**Definition 3.2.** *A* monotone majorant *of a sequence $(a_\nu)_{\nu\in\mathcal{F}} \subset \mathcal{X}$ is a sequence $\boldsymbol{a}^* = (\boldsymbol{a}^*_\nu)_{\nu\in\mathcal{F}} \subset \mathbb{R}$ which is defined by $\boldsymbol{a}^*_\nu := \sup_{\mu\geq\nu} \|a_\nu\|_{\mathcal{X}}\ ,\quad \nu \in \mathcal{F}$ . Here, $\mu \geq \nu$ for $\mu,\nu \in \mathcal{F}$ means that $\mu_j \geq \nu_j$ for all $j \in \mathbb{J}$.*

The monotone majorant depends on the norm $\|\circ\|_{\mathcal{X}}$ on $\mathcal{X}$ and

$$\|a\|_{\ell^p_m(\mathcal{F};\mathcal{X})} = \|\boldsymbol{a}^*\|_{\ell^p(\mathcal{F})}\ .$$

Sets $\Lambda_N$ of $N$ largest coefficients of monotone majorants can be chosen to be monotone sets (cp. [8, Remark 2.2]). Further, if $\Lambda \subset \mathcal{F}$ is any monotone set, $\nu \in \Lambda$ and $\mu \leq \nu$ imply that $\mu \in \Lambda$.

### 3.2. Sparse Legendre Expansions

We recall for $\nu \in \mathcal{F}$ the definition (49) of the tensorized Legendre polynomials $L_\nu(\boldsymbol{y})$; the normalization (48) differs slightly from the classical one, where

$$P_k(1) = 1,\ \|P_k\|_{L^\infty(-1,1)} = 1\ ,\quad k = 0, 1, \ldots \tag{52}$$

Also for the system $(P_k)_{k\geq 0}$, $P_0 \equiv 1$, and hence the (formally countable) tensor product polynomials contain for each $\nu \in \mathcal{F}$ only finitely many nontrivial factors. Hence,

$$P_\nu(\boldsymbol{y}) := \prod_{j\in\mathbb{J}} P_{\nu_j}(y_j),\ \ \nu \in \mathcal{F}$$

is meaningful. We also note that due to the $L^2(U;\mu_0)$-orthonormality of the $L_\nu$, we may expand every $q(\boldsymbol{y}) \in L^2(U,\mu_0)$

$$q(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} q_\nu^L L_\nu(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} q_\nu^P P_\nu(\boldsymbol{y}) \tag{53}$$

where

$$\|q\|^2_{L^2(U,\mu_0;\mathcal{X})} = \sum_{\nu\in\mathcal{F}} \|q_\nu^L\|^2_{\mathcal{X}} < \infty, \;\; q_\nu^L := \int_U q(\boldsymbol{y}) L_\nu(\boldsymbol{y}) d\mu_0(\boldsymbol{y})$$

and where the coefficient sequences $q_\nu^L$ and $q_\nu^P$ in (53) are related by

$$q_\nu^L = \left(\prod_{j\in\mathbb{J}}(1+2\nu_j)\right)^{1/2} q_\nu^P\,, \quad \nu\in\mathcal{F}\,. \tag{54}$$

**Lemma 3.3.** *If the parametric forward map $q(\boldsymbol{y})$ is $(p,\varepsilon)$-analytic in a poly-ellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$, then for every $\nu\in\mathcal{F}$ there holds, for every $\rho$ as in (9) in Definition 2.2, the estimate*

$$\|q_\nu^L\|_{\mathcal{X}} \le \frac{\|f\|_{\mathcal{Y}'}}{\mu_0\varepsilon} \prod_{j\in\mathbb{J},\nu_j\neq 0} \rho_j^{-\nu_j}\,. \tag{55}$$

*Proof.* As in [14, Theorem 9, Proposition 5], the result follows by using Rodrigues' formula and partial integration. □

Due to $\|q_\nu^P\|_{\mathcal{X}} \le \|q_\nu^L\|_{\mathcal{X}}$, the summability of the sequence $(\|q_\nu^L\|_{\mathcal{X}})_{\nu\in\mathcal{F}}$ directly implies the summability of $(\|q_\nu^P\|_{\mathcal{X}})_{\nu\in\mathcal{F}}$. The next result, from [12, 14] specifies the type of convergence in the Legendre expansions (53), and also quantifies sparsity in the sequences $\{q_\nu^L : \nu\in\mathcal{F}\}$ and $\{q_\nu^P : \nu\in\mathcal{F}\}$ of Legendre coefficients. Its proof is analogous to the arguments in [12, 10, 14]. The estimate of the Legendre coefficients of Lemma 3.3 allows to construct a monotone majorant $\boldsymbol{q}^* = (\boldsymbol{q}_\nu^*)_{\nu\in\mathcal{F}}$ of the sequence $(\|q_\nu^L\|_{\mathcal{X}})_{\nu\in\mathcal{F}}$ and thus obtain that $(\|q_\nu^P\|_{\mathcal{X}})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ reasoning as in the proof of [8, Theorem 2.4].

**Theorem 3.4.** *Assume that the parametric forward solution map $q(\boldsymbol{y})$ admits a $(p,\varepsilon)$-analytic extension to the poly-ellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$ with $\rho$ satisfying condition (9) in Definition 2.2. Then the following holds.*

*(i) the Legendre series (53) converge unconditionally, in $L^2(U,\mu_0;\mathcal{X})$ resp. in $L^\infty(U,\mu_0;\mathcal{X})$, to $q$,*

*(ii) with $0<p<1$ as in (9) of Definition 2.2, the sequence $(q_\nu^L)_{\nu\in\mathcal{F}}$ of Legendre coefficients admits a monotone majorant $\boldsymbol{q}^* = (\boldsymbol{q}_\nu^*)_{\nu\in\mathcal{F}}$ which is p-summable in norm, in the sense that*

$$C(p,\boldsymbol{q}) := \|\boldsymbol{q}\|_{\ell^p_m(\mathcal{F})} = \left(\sum_{\nu\in\mathcal{F}} |\boldsymbol{q}_\nu^*|^p\right)^{1/p} < \infty\,.$$

*Denoting, for every $N \in \mathbb{N}$ by $\Lambda_N^L \subset \mathcal{F}$ a set of $N$ largest coefficients of the monotone majorant $\boldsymbol{q}^*$ of the Legendre expansion* (53)*, there holds the error bound*

$$\left\| q(\cdot) - \sum_{\nu \in \Lambda_N^L} q_\nu^L L_\nu(\cdot) \right\|_{L^2(U,\mu_0;\mathcal{X})} \leq C(p,\boldsymbol{q}) N^{-(1/p-1/2)} . \tag{56}$$

*(iii) Likewise, denoting by $\Lambda_N^P \subset \mathcal{F}$ a set of $N$ largest (in $\mathcal{X}$-norm) terms of the monotone majorant $\boldsymbol{q}$ of the sequence of Legendre coefficients $q_\nu^P \in \mathcal{X}$ in the Legendre expansions* (53)*, there holds the error bound*

$$\sup_{\boldsymbol{y} \in U} \left\| q(\boldsymbol{y}) - \sum_{\nu \in \Lambda_N^P} q_\nu^P P_\nu(\boldsymbol{y}) \right\|_{\mathcal{X}} \leq C(p,\boldsymbol{v}) N^{-(1/p-1)} . \tag{57}$$

## 4. Sparsity of the Posterior Density $\Theta$

For operator equations (4) with operators $A(u;q)$ with parametric uncertainty which produce parametric solutions which are $(p,\varepsilon)$-analytic in the sense of Definition 2.2, in Theorem 3.4 the representation of the forward solution in unconditionally convergent Legendre polynomial chaos expansions was presented, with coefficient sequences which admit $p$-sparse, monotone majorants. In the present section, we show corresponding results also for the Bayesian posterior density $\Theta(\boldsymbol{y})$ which was defined in (43), (44).

### *4.1. $(p,\varepsilon)$-Analyticity of $\Theta$*

Our verification of $(p,\varepsilon)$-analyticity of $\Theta$ will be based on verifying $(p,\varepsilon)$-analyticity for the parametric posterior density $\Theta(\boldsymbol{y})$ defined in (43), (44). Once this is established, sparsity and $N$-term approximation results for $\Theta$ will follow similarly as for the parametric solution $q(\boldsymbol{y})$ of (4). As in [29], we then infer convergence rates for Smolyak quadratures from $N$-term approximation rates for truncated tensorized Legendre approximation rates for the posterior density $\Theta$.

**Theorem 4.1.** *Consider the Bayesian inversion of the parametric operator equation* (4) *with uncertain input $u \in X$, parametrized by the sequence $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}} \in U$. Assume further that the corresponding forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is $(p,\varepsilon)$-analytic for some $0 < p < 1$ and $\delta > 0$.*

*Then the Bayesian posterior density $\Theta(\boldsymbol{y})$ is, as a function of the parameter $\boldsymbol{y}$, likewise $(p,\varepsilon)$-analytic,* with the same $p$ and the same $\delta$.

*The modulus of the holomorphic extension of the Bayesian posterior $\Theta(\boldsymbol{y})$ over the polyellipse $\mathcal{E}_\rho$ is bounded by $C\exp(b^2\|\Gamma^{-1}\|)$ with $\Gamma > 0$ denoting the positive definite covariance matrix in the additive, Gaussian observation noise model* (5)*. Here, the constants $b, C > 0$ in the bound of the modulus $\sup_{\boldsymbol{z}\in\partial\mathcal{E}_\rho} |\Theta(\boldsymbol{z})|$ depend on the condition number of the uncertainty-to-observation map $\mathcal{G}(\cdot) = (\mathcal{O} \circ G)(\cdot)$ but are independent of $\Gamma$ in* (5).

*Proof.* By Lemma 2.5, the $(p,\varepsilon)$-analyticity of the operator implies that the forward solution map $q(\boldsymbol{y})$ admits a holomorphic extension, denoted $q(\boldsymbol{z})$, to $\boldsymbol{z} \in \mathcal{E}_\rho$ with $\rho$ as in (9).

We consider first the case there is only a single parameter, ie. that $\mathbb{J} = \{1\}$. Then $\rho = \{\rho_1\}$ and we may write $u = \langle u \rangle + z\psi \in X$, with $z \in \mathcal{E}_{\rho_1} \subset \mathbb{C}$ and by assumption the foward map is holomorphic with respect to $z \in \mathcal{E}_{\rho_1}$.

The unique holomorphic extension of the Bayesian potential $\Phi_\Gamma(u;\delta)$ defined in (6) is, in this case, given by (assuming that the data $\delta$, $\langle u \rangle$ and $\psi$ are real-valued)

$$\Phi_\Gamma(\langle u \rangle + z\psi; \delta) = \frac{1}{2}\left(\delta - \mathcal{G}(\langle u \rangle + z\psi)\right)^\top \Gamma^{-1}\left(\delta - \mathcal{G}(\langle u \rangle + z\psi)\right) \,.$$

By the holomorphy of $q(z) \in \mathcal{X}$, the response function $z \mapsto \mathcal{G}(\langle u \rangle + z\psi)$ is holomorphic in $\mathcal{E}_{\rho_1}$. Therefore, the complex extension of $\Phi_\Gamma$, ie.

$$\mathcal{E}_{\rho_1} \ni z \mapsto \Phi_\Gamma(u(z);\delta) := \frac{1}{2}\left(\delta - \mathcal{G}(\langle u \rangle + z\psi)\right)^\top \Gamma^{-1}\left(\delta - \mathcal{G}(\langle u \rangle + z\psi)\right) \tag{58}$$

is holomorphic in $\mathcal{E}_{\rho_1}$, being a quadratric polynomial of $\mathcal{G}(\langle u \rangle + z\psi)$.

The preceding argument immediately generalizes to any coordinate $y_j$ for $j \in \mathbb{J} \subseteq \mathbb{N}$ so that we infer that the Bayesian potential

$$\Phi_\Gamma(u;\delta)\mid_{\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j} = \frac{1}{2}\left(\delta - \mathcal{G}\left(\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j\right)\right)^\top \Gamma^{-1}\left(\delta - \mathcal{G}\left(\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j\right)\right)$$

is holomorphic on the polyellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$. Hence, also the Bayesian posterior admits a holomorphic extension to $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$ which is given by

$$\Theta(\boldsymbol{z}) = \exp\left(-\Phi_\Gamma(u;\delta)|_{u=\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j}\right) \,. \tag{59}$$

By the holomorphy of the Bayesian potential $\Phi_\Gamma(u;\delta)\mid_{\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j}$ with respect to the parameters $\boldsymbol{z}$, the extension $\Theta(\boldsymbol{z})$ in (59) is, as composition of a holomorphic function with the entire, analytic function $\exp(\cdot)$, holomorphic on $\mathcal{E}_\rho$ and, therefore, $\Theta(\boldsymbol{z})$ in (59) is the unique analytic continuation of the Bayesian posterior $\Theta(\boldsymbol{y})$ from $U$ to $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$.

It remains to deduce bounds on the modulus of this holomorphic continuation of the posterior density $\Theta(\boldsymbol{z})$ in (59) as a function of the parameters $\boldsymbol{z}$ over the polyellipses $\mathcal{E}_\rho$ of holomorphy, with the semiaxes $\rho$ as in (9). Recalling the definition $\mathcal{G}(\cdot) = (\mathcal{O} \circ G)(\cdot)$, we find

$$\mathcal{G}(u)\mid_{u=\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j} = (o_k(G(u)))_{k=1}^K = (o_k(q(\boldsymbol{z})))_{k=1}^K \in \mathbb{C}^K \,.$$

This implies that the modulus of the posterior density $\Theta(\boldsymbol{z})$ can be bounded as

$$\sup_{\boldsymbol{z}\in\mathcal{E}_\rho} |\Theta(\boldsymbol{z})| \le \exp\left(\sup_{\boldsymbol{z}\in\mathcal{E}_\rho} \frac{1}{2}\left\|\delta - \mathcal{G}\left(\langle u \rangle + \sum_{j\in\mathbb{J}} z_j\psi_j\right)\right\|_\Gamma^2\right)$$

where $\| \circ \|_\Gamma$ denotes the covariance-weighted Euclidean norm in $\mathbb{C}^K$. Based on the definition (58), and on the fact that $\delta \in \mathbb{R}^K$, and the definition (5) of $\mathcal{G}(\cdot)$, we find

$$\forall z \in \mathcal{E}_\rho : \quad |\Theta(\boldsymbol{z})| \leq \exp\left(\frac{1}{2}\mathrm{Im}\left(\mathcal{G}(u(\boldsymbol{z}))\right)^\top \Gamma^{-1}\mathrm{Im}\left(\mathcal{G}(u(\boldsymbol{z}))\right)\right) . \tag{60}$$

Since the map $\mathcal{E}_\rho \ni z \mapsto \mathcal{G}(u(\boldsymbol{z}))$ does not depend on the observation noise covariance $\Gamma$, a bound for the modulus $\sup_{z\in\mathcal{E}_\rho} |\Theta(\boldsymbol{z})|$ which is explicit in terms of $\Gamma$ can be inferred from (60). This establishes the asserted dependence of the modulus of $\Theta(\boldsymbol{z})$ over $\mathcal{E}_\rho$ and completes the proof. □

Exactly the same results on analyticity and on $N$-term approximation of $\Psi(\boldsymbol{z})$ hold, cp. [32]. We omit details for reasons of brevity of exposition and confine ourselves to establishing rates of convergence of $N$-term truncated representations of the posterior density $\Theta$. In the following, we analyze the convergence rate of *N-term truncated Legendre gpc-approximations of* $\Theta$ and, with the aim of an *adaptive sparse quadrature approximation* to efficiently evaluate the expectation of interest with respect to the posterior $\Theta(\boldsymbol{y})$ in $U$ in Section 5 ahead, we analyze also $N$-term truncated *monomial gpc-approximations* of $\Theta(\boldsymbol{y})$. For the deterministic approximation of the posterior density $\Theta(\boldsymbol{y})$ in (44) we shall use tensorized polynomial bases similar to what is done in so-called "polynomial chaos" expansions of random fields.

### *4.2. Sparse Legendre Expansions of* $\Theta$

Since we assumed that the prior measure $\mu_0$ is built by tensorization of the uniform probability measures $\frac{1}{2}\lambda^1$ on $[-1,1]$, the normalization (48) implies that the polynomials $L_\nu(\boldsymbol{z})$ in (49) are well-defined for any $z \in \mathbb{C}^{\mathbb{J}}$ since the finite support of each element of $\nu \in \mathcal{F}$ implies that $L_\nu$ in (49) is the product of only finitely many nontrivial polynomials. This observation suggests, by virtue of the square integrability discussed below, the use of mean square convergent gpc-expansions and their truncations to represent and approximate the densities $\Theta$ and $\Psi$. Such expansions can also serve as a basis for sampling of these quantities with draws that are equidistributed with respect to the prior $\mu_0$. In particular, the density $\Theta : U \to \mathbb{R}$ is square integrable with respect to the prior $\mu_0$ over $U$, i.e. $\Theta \in L^2(U, \mu_0)$. Moreover, if the QoI $\phi(\cdot) : U \to \mathcal{S}$ in (46) is bounded, then

$$\int_U \|\Psi(\boldsymbol{y})\|^2_{\mathcal{S}} d\mu_0(\boldsymbol{y}) < \infty, \tag{61}$$

i.e. $\Psi \in L^2(U, \mu_0; \mathcal{S})$.

*Remark* 4.2. If the QoI is the parametric solution, $\mathcal{S} = \mathcal{X}$ ie. when $\phi(u) = G(u) = q(\boldsymbol{y}) \in \mathcal{X}$, we have $\|\Psi(\boldsymbol{y})\|_V \leq C\|f\|_{V^*}$ for all $\boldsymbol{y} \in U$, where the constant $C$ is independent of the data $\delta$. Thus $\Psi \in L^2(U, \mu_0; \mathcal{S})$ holds for calculation of the expectation of the pressure under the posterior distribution on $u$. Indeed the assertion holds for all moments of the pressure, the concrete examples which we concentrate on here. □

Since $\mathbb{L}(U,\mu_0)$ in (50) is a countable orthonormal basis of $L^2(U,\mu_0)$, the density $\Theta(\boldsymbol{y})$ of the posterior measure given data $\delta \in Y$, and the posterior reweighted pressure $\Psi(\boldsymbol{y})$ can be represented in $L^2(U,\mu_0)$ by (parametric and deterministic) generalized Legendre polynomial chaos expansions. We first address the representation of the scalar valued function $\Theta(\boldsymbol{y})$.

$$\Theta(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} \theta^L_\nu L_\nu(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} \theta^P_\nu P_\nu(\boldsymbol{y}) \quad \text{in} \quad L^2(U,\mu_0) \tag{62}$$

where the gpc expansion coefficients $\theta^L_\nu$ and $\theta^P_\nu$ are defined by (cf. also (54))

$$\theta^L_\nu = \int_U \Theta(\boldsymbol{y}) L_\nu(\boldsymbol{y})\mu_0(d\boldsymbol{y})\;, \quad \theta^L_\nu = \left(\prod_{j\in\mathbb{J}}(1+2\nu_j)\right)^{1/2} \theta^P_\nu\;, \quad \nu\in\mathcal{F}\;.$$

By Parseval's equation and the normalization (48), it follows immediately from (62) and (61) with Parseval's equality that the second moment of the posterior density with respect to the prior is finite and can be expressed as

$$\|\Theta\|^2_{L^2(U,\mu_0} = \sum_{\nu\in\mathcal{F}} |\theta^L_\nu|^2 = \|\boldsymbol{\theta}^L\|^2_{\ell^2(\mathcal{F})}\;. \tag{63}$$

*4.3. Monotone N-term Approximation of $\Theta$ in $L^2(U,\mu_0)$ and $L^\infty(U,\mu_0)$*

For every $N\in\mathbb{N}$, denote by $\Lambda^L_N\subset\mathcal{F}$ a set of indices $\nu\in\mathcal{F}$ corresponding to $N$ largest $\theta^*_\nu$ of the monotone majorant $\boldsymbol{\theta}^L$ of the Legendre coefficient sequence $(\theta^L_\nu)_{\nu\in\mathcal{F}}$ in (62), and denote by

$$\Theta^L_{\Lambda_N}(\boldsymbol{y}) := \sum_{\nu\in\Lambda^L_N} \theta^L_\nu L_\nu(\boldsymbol{y}) \tag{64}$$

the corresponding $N$-term truncated Legendre expansion (62) of the posterior. Then, with $0<p<1$ in the $(p,\varepsilon)$-analyticity of the parametric forward solution, there is a sequence $\{\Lambda_N\}_{N\geq 0}$ of nested, monotone index sets $\Lambda_N\subset\mathcal{F}$ which exhausts $\mathcal{F}$, with $\#(\Lambda_N)\leq N$ and which are such that there hold the $N$-term approximation results

$$\|\Theta(\boldsymbol{y}) - \Theta^L_{\Lambda_N}(\boldsymbol{y})\|_{L^2(U,\mu_0)} \leq CN^{-s}\|\theta^L\|_{\ell^p_m(\mathcal{F})},\;\; s:=\frac{1}{p}-\frac{1}{2}\;. \tag{65}$$

Likewise, denoting by $\Lambda^P_N\subset\mathcal{F}$ a set of indices $\nu\in\mathcal{F}$ corresponding to $N$ largest (in $\|\circ\|_{\mathcal{X}}$-norms) of the coefficients of the monotone majorant $\boldsymbol{\theta}^P$ of the Legendre coefficient sequence $(\theta^P_\nu)_{\nu\in\mathcal{F}}$ in (62), and denote by

$$\Theta^P_{\Lambda_N}(\boldsymbol{y}) := \sum_{\nu\in\Lambda^P_N} \theta^P_\nu L_\nu(\boldsymbol{y}) \tag{66}$$

the corresponding $N$-term truncated Legendre expansion (62) of the posterior. Then, with $0<p<1$ in the $(p,\varepsilon)$-analyticity of the parametric forward solution, there hold the $N$-term approximation results

$$\|\Theta(\boldsymbol{y}) - \Theta^P_{\Lambda_N}(\boldsymbol{y})\|_{L^\infty(U,\mu_0)} \leq CN^{-s}\|\theta^P\|_{\ell^p_m(\mathcal{F})},\;\; s:=\frac{1}{p}-1\;. \tag{67}$$

In (67) and (65), the constant $C \geq 1$ depends on $s$ and on the covariance $\Gamma > 0$ in the additive observation noise $\eta$ in (5), but is independent of $N$. We refer to [30] for an analysis of the limit $\Gamma \to 0$.

## 5. Sparse Adaptive Smolyak Quadrature

### 5.1. Univariate Quadrature and Tensorization

We consider a sequence $(Q^k)_{k\geq 0}$ of univariate quadrature formulas of the form

$$Q^k(\mathrm{g}) = \sum_{i=0}^{n_k} w_i^k \cdot \mathrm{g}(z_i^k)\,,$$

associated with the quadrature points $(z_j^k)_{j=0}^{n_k} \subset [-1,1]$ with $z_j^k \in [-1,1]\,, \forall j,k$ and $z_0^k = 0\,, \forall k$ and weights $w_j^k\,, 0 \leq j \leq n_k$, $\forall k \in \mathbb{N}_0$, where g is a function $\mathrm{g} : [-1,1] \mapsto \mathcal{S}$, taking values in some Banach space $\mathcal{S}$. We impose the following assumptions on the sequence $(Q^k)_{k\geq 0}$.

**Assumption 5.1.**

*(i)* $(I - Q^k)(v_k) = 0\,, \quad \forall v_k \in \mathbb{S}_k := \mathbb{P}_k \otimes \mathcal{S}\,, \mathbb{P}_k = \mathrm{span}\{y^j : j \in \mathbb{N}_0, j \leq k\}$, *with* $I(v_k) = \int_{[-1,1]} v_k(\boldsymbol{y}) \frac{\lambda_1(d\boldsymbol{y})}{2}$.

*(ii)* $w_j^k > 0\,, \qquad 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0$.

Defining the univariate quadrature difference operator by

$$\Delta_j = Q^j - Q^{j-1}, \qquad j \geq 0\,.$$

with $Q^{-1} = 0$, $Q^k$ can be rewritten as telescoping sum

$$Q^k = \sum_{j=0}^{k} \Delta_j\,,$$

where $\mathcal{Z}^k = \{z_j^k : 0 \leq j \leq n_k\} \subset [-1,1]$ denotes the set of points corresponding to $Q^k$. Following [29], we introduce the tensorized multivariate operators

$$\mathcal{Q}_\nu = \bigotimes_{j\geq 1} Q^{\nu_j}\,, \qquad \Delta_\nu = \bigotimes_{j\geq 1} \Delta_{\nu_j}\,. \tag{68}$$

for $\nu \in \mathcal{F}$ with associated set of multivariate points $\mathcal{Z}^\nu = \times_{j\geq 1} \mathcal{Z}^{\nu_j} \in U$. The tensorization can be defined inductively: for a $\mathcal{S}$-valued function $g$ defined on $U$,

- If $\nu = 0_\mathcal{F}$, then $\Delta_\nu g = Q^\nu g$ denotes the integral over the constant polynomial with value $g(z_{0_\mathcal{F}}) = g(0_\mathcal{F})$.
- If $0_\mathcal{F} \neq \nu \in \mathcal{F}$, then denoting by $\hat{\nu} = (\nu_j)_{j\neq i}$

$$Q^\nu g = Q^{\nu_i}(t \mapsto \bigotimes_{j\geq 1} Q^{\hat{\nu}_j} g_t)\,, \qquad i \in \mathbb{I}_\nu$$

and

$$\Delta_\nu g = \Delta_{\nu_i}(t \mapsto \bigotimes_{j\geq 1} \Delta_{\hat{\nu}_j} g_t)\,, \qquad i \in \mathbb{I}_\nu\,,$$

for $g \in \mathcal{Z}$, $g_t$ is the function defined on $\mathcal{Z}^{\mathbb{N}}$ by $g_t(\hat{\boldsymbol{y}}) = g(\boldsymbol{y}), \boldsymbol{y} = (\ldots, y_{i-1}, t, y_{i+1}, \ldots)\,, i > 1$ and $\boldsymbol{y} = (t, y_2, \ldots)\,, i = 1$, see [9, 29].

*5.2. Sparse Quadrature Operator*

Based on the definitions in the previous subsection, we will now introduce the sparse quadrature operator

$$\mathcal{Q}_\Lambda = \sum_{\nu\in\Lambda} \Delta_\nu = \sum_{\nu\in\Lambda} \bigotimes_{j\geq 1} \Delta_{\nu_j}\,,$$

for any finite monotone set $\Lambda \subset \mathcal{F}$ with associated collocation grid

$$\mathcal{Z}_\Lambda = \cup_{\nu\in\Lambda} \mathcal{Z}^\nu\,.$$

**Lemma 5.2.** *For any monotone index set $\Lambda_N \subset \mathcal{F}$, the sparse quadrature $\mathcal{Q}_{\Lambda_N}$ is exact for any polynomial $g \in \mathbb{S}_{\Lambda_N}$, i.e. there holds*

$$\mathcal{Q}_{\Lambda_N}(g) = I(g), \qquad \forall g \in \mathbb{S}_{\Lambda_N} := \mathbb{P}_{\Lambda_N} \otimes \mathcal{S}\,,$$

*with* $\mathbb{P}_{\Lambda_N} = \operatorname{span}\{y^\nu : \nu \in \Lambda_N\} = \operatorname{span}\{P_\nu : \nu \in \Lambda_N\}$ *i.e.* $\mathbb{S}_{\Lambda_N} = \operatorname{span}\left\{\sum_{\nu\in\Lambda_N} s_\nu y^\nu : s_\nu \in \mathcal{S}\right\}$*, and* $I(g) = \int_U g(\boldsymbol{y}) d\mu_0(\boldsymbol{y})$.

For the proof, we refer to [29, Theorem 4.2].

We will now establish convergence rates for the approximation of the expectation of QoI with respect to the posterior, given data $\delta$, based on the $(p,\varepsilon)$-analyticity results presented in sections 3 and 4. In particular, we will prove the existence of two sequences $(\Lambda_N^1)_{N\geq 1}$, $(\Lambda_N^2)_{N\geq 1}$ of monotone index sets $\Lambda_N^{1,2} \subset \mathcal{F}$ such that $\#\Lambda_N^{1,2} \leq N$ which exhaust $\mathcal{F}$ and such that, for some $C^1, C^2 > 0$ independent of $N$,

$$|I(\Theta) - \mathcal{Q}_{\Lambda_N^1}(\Theta)| \leq C^1 N^{-s}\,, \qquad s = \frac{1}{p} - 1\,,$$

with $I(\Theta) = \int_U \Theta(\boldsymbol{y}) d\mu_0(\boldsymbol{y})$ and

$$\|I[\Psi] - \mathcal{Q}_{\Lambda_N^2}[\Psi]\|_{\mathcal{S}} \leq C^2 N^{-s}\,, \qquad s = \frac{1}{p} - 1\,,$$

with $I[\Psi] = \int_U \Psi(\boldsymbol{y}) d\mu_0(\boldsymbol{y})$, respectively. By Lemma 5.2, we have

$$\begin{aligned}\|(I - \mathcal{Q}_{\Lambda_N})(g)\|_{\mathcal{S}} &= \|(I - \mathcal{Q}_{\Lambda_N})(g - \Upsilon_N)\|_{\mathcal{S}} \\ &\leq (|\!|\!| I |\!|\!| + |\!|\!| \mathcal{Q}_{\Lambda_N} |\!|\!|) \cdot \inf_{\Upsilon_n \in \mathbb{S}_{\Lambda_N}} \|g - \Upsilon_n\|_{L^\infty(U;\mathcal{S})} \\ &\leq (1 + C_{\mathcal{Q}_{\Lambda_N}}) \cdot C N^{-s}\,,\end{aligned}$$

since $\|\!|I|\!\|=\mu_0(U)=1$ and $\|\!|\mathcal{Q}_{\Lambda_N}|\!\|=: C_{\mathcal{Q}_{\Lambda_N}}$, for a $\mathcal{S}$-valued function $g$ on $U$. Then

$$C_{\mathcal{Q}_{\Lambda_N}} \leq \sum_{\nu\in\Lambda_N}\prod_{j\geq 1}(c_{\nu_j}+c_{\nu_j-1}) \leq \#\Lambda^{\log_2 3} \tag{69}$$

with $c_k = 1\,, k\geq 0$ (note that $\|\!|Q^k|\!\|=1$ by Assumption 5.1 (ii)) and $c_{-1}:=0$, see [29, Lemma 4.4]. In the tensor product case, the exact value of the constant $C_{\mathcal{R}_\nu}$ is given by

$$C_{\mathcal{R}_\nu}=\prod_{j\geq 1}c_{\nu_j}=1\,,\qquad \text{with } \mathcal{R}_\nu=\{\mu\in\mathcal{F}:\mu\leq\nu\},$$

so that bound (69) is pessimistic in this case.

The quadrature error for the normalization constant (45) and the quantity $Z'$ (47) can be bounded by relating the error with the Legendre coefficients $\theta_\nu^P$ of $\Theta=\sum_{\nu\in\mathcal{F}}\theta_\nu^P P_\nu(\boldsymbol{y})$ and $\psi_\nu^P$ of $\Psi=\sum_{\nu\in\mathcal{F}}\psi_\nu^P P_\nu(\boldsymbol{y})$ as follows:

**Lemma 5.3.** *Assume for a $\mathcal{S}$-valued function $g$ on $U$ that $g(\boldsymbol{y})=\sum_{\nu\in\mathcal{F}}g_\nu^P P_\nu(\boldsymbol{y})$ in the sense of unconditional convergence in $L^\infty(U,\mathcal{S})$. Then, we have*

$$\|I(g)-\mathcal{Q}_\Lambda(g)\|_{\mathcal{S}} \leq 2\cdot\sum_{\nu\notin\Lambda}\gamma_\nu\|g_\nu^P\|_{\mathcal{S}}$$

*for any monotone set $\Lambda\subset\mathcal{F}$, where $\gamma_\nu:=\prod_{j\in\mathbb{J}}(1+\nu_j)^2$.*

For the proof, we refer to [29, Lemma 4.5].

**Theorem 5.4.** *If the forward solution map $U\ni\boldsymbol{y}\mapsto q(\boldsymbol{y})$ is $(p,\varepsilon)$-analytic for some $0<p<1$ and $\varepsilon>0$, then $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}}\in l_m^p(\mathcal{F})$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}}\in l_m^p(\mathcal{F})$. Denoting by $\Lambda_N^\theta$, $\Lambda_N^\psi$ the sets of $N$-largest terms of the monotone majorants of $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}}$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$, respectively, then there holds the error bound for $s=1/p-1$,*

$$|I[\Theta]-\mathcal{Q}_{\Lambda_N^\theta}[\Theta]|\leq C^1N^{-s}\,, \tag{70}$$

*with $I[\Theta]=\int_U\Theta(\boldsymbol{y})d\mu_0(\boldsymbol{y})$ and, with $I[\Psi]=\int_U\Psi(\boldsymbol{y})d\mu_0(\boldsymbol{y})$,*

$$\|I[\Psi]-\mathcal{Q}_{\Lambda_N^\psi}[\Psi]\|_{\mathcal{S}}\leq C^2N^{-s}\quad. \tag{71}$$

*Proof.* The proof proceeds in two steps: first, we will construct a $\frac{\varepsilon}{2}$-admissible sequence $\rho$ in the sense of (27) based on the estimate of the Legendre coefficients in Lemma 3.3. Then, we use the resulting estimate to prove $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}}\in l^p(\mathcal{F})$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}}\in l^p(\mathcal{F})$, respectively and construct a $l^p$-summable monotone majorant of $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}}$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$. We follow [9, 12, 29], and present the details. Due to the $(p,\varepsilon)$-analyticity of the forward solution map, it holds

$$\sum_{j\geq 1}b_j\leq 1-\varepsilon$$

for some $0<\varepsilon\leq 1$. In order to construct a $\frac{\varepsilon}{2}$-admissible sequence $\rho$, we choose a constant $1<\kappa\leq 2$ such that

$$(\kappa-1)\sum_{j\geq 1}b_j\leq\frac{\varepsilon}{6}\,.$$

and an integer $J_0$ as the smallest integer such that

$$\sum_{j>J_0} b_j \leq \frac{\varepsilon}{12e^2}\,.$$

We set $E := \{j : 1 \leq J_0\}$ and $F := \mathbb{N} \setminus E$ and denote denote by $\nu_E$ and $\nu_F$ the restrictions of $\nu$ on $E$ and $F$ for each $\nu \in \mathcal{F}$. Then, we define the sequence $\rho = \rho(\nu)$ by

$$\rho_j = \kappa\,, \quad j \in E\,; \qquad \rho_j = \frac{\varepsilon \nu_j}{4|\nu_F| b_j} + e^2\,, \quad j \in F\,,$$

with $|\nu_F| = \sum_{j>J_0} \nu_j$ (with the convention $\frac{\nu_j}{|\nu_F|} = 0$, if $|\nu_F| = 0$). The sequence $\rho$ satisfies

$$\begin{aligned}\sum_{j\geq 1} \rho_j b_j &= \kappa \sum_{j\leq J_0} b_j + \sum_{j>J_0} \frac{\varepsilon \nu_j}{4|\nu_F| b_j} b_j + e^2 \sum_{j>J_0} b_j \\ &= (\kappa - 1) \sum_{j\leq J_0} b_j + \sum_{j\leq J_0} b_j + \frac{\varepsilon}{4} + e^2 \sum_{j>J_0} b_j \\ &\leq \frac{\varepsilon}{6} + \sum_{j\geq 1} b_j + \frac{\varepsilon}{4} + e^2 \sum_{j>J_0} b_j \leq \frac{\varepsilon}{2} + \sum_{j\geq 1} b_j \leq 1 - \frac{\varepsilon}{2}\,,\end{aligned}$$

and thus, is $(\frac{\varepsilon}{2})$-admissible. Similar to the proof of Theorem 4.2 in [9], we have

$$p_\nu |\theta_\nu^P| \leq C_{\frac{\varepsilon}{2}} \left( \prod_{j\leq J_0} \frac{(1+\nu_j)^2}{\kappa^{\nu_j}} \right) \left( \prod_{j>J_0} \frac{(1+\nu_j)^2}{\rho^{\nu_j}} \right)\,,$$

and it follows

$$p_\nu |\theta_\nu^P| \leq C \cdot \alpha(\nu_E) \cdot \beta(\nu_F)\,,$$

where $\alpha(\nu_E) := \prod_{j\leq J_0} \eta^{\nu_j}$, $\beta(\nu_F) := \prod_{j>J_0} \left(\frac{|\nu_F| d_j}{\nu_j}\right)^{\nu_j}$ with $\eta := \frac{1+\kappa}{2\kappa}$ and $d_j := 4e^2 b_j/\varepsilon$. It holds $\sum_{j>J_0} d_j < \frac{1}{3}$. The summability of $(\gamma_\nu |\theta_\nu^P|)_{\nu\in\mathcal{F}}$ follows then as in Subsection 3.2 in [12]. The same argument implies also the summability of $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$.

With the same argument as in [9] Section 4.2, it holds that $(\gamma_\nu |\theta_\nu^P|)_{\nu\in\mathcal{F}} \in l_m^p(\mathcal{F})$ and $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}} \in l_m^p(\mathcal{F})$, since the sequence $(\alpha(\nu_E) \cdot \beta(\nu_F))_{\nu\in\mathcal{F}}$ is monotonically decreasing. Exactly the same analysis allows to bound the quadrature error of the density $\Psi$. □

### *5.3. Adaptive Smolyak Construction of Monotone Index Sets*

We now discuss the adaptive construction of a sequence of monotone index sets $(\Lambda_N)_{N\geq 1}$ which are, in general, not equal to sets generated by $N$-term approximations of monotone envelopes, but which yield in practice approximations of the Bayesian estimates which converge with rate $s = 1/p$ (rather than $1/p - 1$ as predicted in the theoretical error bounds). The idea is to successively identify the index set $\Lambda_N$ corresponding to the $N$ largest contributions of the sparse quadrature operator to the approximation of the integral $Z$ and $Z'$, i.e. to the $N$ largest

$$\|\Delta_\nu(\Xi)\|_{\mathcal{S}} = \|\bigotimes_{j\geq 1} \Delta_{\nu_j}(\Xi)\|_{\mathcal{S}}\,, \quad \nu \in \mathcal{F}$$

with $\Xi = \Theta$, $\mathcal{S} = \mathbb{R}$ or $\Xi = \Psi$, $\mathcal{S} = \mathcal{X}$, minimizing the approximation error (70) and (71), respectively (cf. [29, 9, 15, 13]).

Following [13, 9, 8], We use a greedy-type strategy based on finite sets of reduced neighbors defined by

$$\mathcal{N}(\Lambda) := \{\nu \notin \Lambda : \nu - e_j \in \Lambda, \forall j \in \mathbb{I}_\nu \text{ and } \nu_j = 0\,, \forall j > j(\Lambda) + 1\}$$

for any monotone set $\Lambda \subset \mathcal{F}$, where $j(\Lambda) = \max\{j : \nu_j > 0 \text{ for some } \nu \in \Lambda\}$. This approach attempts to control the global approximation error by locally collecting indices of the current set of reduced neighbors with the largest error contributions. In the following, the resulting algorithm to adaptively construct the monotone index set $\Lambda$ in the Smolyak quadrature is summarized. We refer to [29, 9, 8, 15, 13] for more details.

---

**function** ASG
    Set $\Lambda_1 = \{0\}\,, k = 1$ and compute $\Delta_0(\Xi)$.
    Determine the set of reduced neighbors $\mathcal{N}(\Lambda_1)$.
    Compute $\Delta_\nu(\Xi)\,, \forall \nu \in \mathcal{N}(\Lambda_1)$.
    **while** $\sum_{\nu \in \mathcal{N}(\Lambda_k)} \|\Delta_\nu(\Xi)\|_{\mathcal{S}} > tol$ **do**
        Select $\nu$ from $\mathcal{N}(\Lambda_k)$ with largest $\|\Delta_\nu\|_{\mathcal{S}}$ and set $\Lambda_{k+1} = \Lambda_k \cup \{\nu\}$.
        Determine the set of reduced neighbors $\mathcal{N}(\Lambda_{k+1})$.
        Compute $\Delta_\nu(\Xi)\,, \forall \nu \in \mathcal{N}(\Lambda_{k+1})$.
        Set $k = k + 1$.
    **end while**
**end function**

---

The sparse quadrature operator is constructed based on the following univariate sequences $(z_j^k)_{j=0}^{n_k}$ of quadrature points

- Clenshaw-Curtis (CC),
  $z_j^k = -\cos\left(\frac{\pi j}{n_k - 1}\right)\,, j = 0, \ldots, n_k - 1, \text{if } n_k > 1$
  and $z_0^k = 0\,, \text{if } n_k = 1$ with $n_0 = 1$ and $n_k = 2^k + 1$, for $k \geq 1$,

- $\Re$-Leja sequence (RL), projection on $[-1, 1]$ of a Leja sequence for the complex unit disk initiated at $i$, i.e.
  $z_0^k = 0\,, z_1^k = 1\,, z_2^k = -1\,, \text{if } j = 0, 1, 2$ and
  $z_j^k = \Re(\hat{z})$, with $\hat{z} = \operatorname{argmax}_{|z|\leq 1} \prod_{l=1}^{j-1} |z - z_l^k|\,, j = 3, \ldots, n_k, \text{if } j \text{ odd}\,,$
  $z_j^k = -z_{j-1}^k\,, j = 3, \ldots, n_k, \text{if } j \text{ even}\,,$ with $n_k = 2 \cdot k + 1$, for $k \geq 0$, see [6].

The positivity assumption on the quadrature weights 5.1 (ii) is not satisfied in the case of the Leja sequence. However, Theorem 5.4 can be generalized to these quadrature formulas due to the moderate, algebraic growth of the Lebesgue constants (cp. [29, 5, 6, 7]). The following result is shown as in [29, Lemma 4.10].

**Proposition 5.5.** *Let $\mathcal{Q}_\Lambda^{RL}$ denote the sparse quadrature operator for any monotone set $\Lambda$ based on the univariate quadrature formulas associated with the $\mathfrak{R}$-Leja sequence. If the forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is $(p,\varepsilon)$-analytic for some $0 < p < 1$ and $\varepsilon > 0$, then $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}} \in l_m^p(\mathcal{F})$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}} \in l_m^p(\mathcal{F})$. Furthermore, there exist two sequences $(\Lambda_N^{RL,1})_{N\geq 1}$, $(\Lambda_N^{RL,2}))_{N\geq 1}$ of monotone index sets $\Lambda_N^{RL,i} \subset \mathcal{F}$ such that $\#\Lambda_N^{RL,i} \leq N$, $i = 1,2$, and such that, for some $C^1, C^2 > 0$ independent of $N$, with $s = \frac{1}{p} - 1$,*

$$|I[\Theta] - \mathcal{Q}_{\Lambda_N^{RL,1}}[\Theta]| \leq C^1 N^{-s}\,,$$

*where $I[\Theta] = \int_U \Theta(\boldsymbol{y})d\mu_0(\boldsymbol{y})$ and, with $I(\Psi) = \int_U \Psi(\boldsymbol{y})d\mu_0(\boldsymbol{y})$, there holds*

$$\|I[\Psi] - \mathcal{Q}_{\Lambda_N^{RL,2}}[\Psi[\|_{\mathcal{S}} \leq C^2 N^{-s}\,.$$

## 6. Numerical Experiments

We consider the following parametric, parabolic problem

$$\begin{aligned}
&\partial_t q(t,x) - \operatorname{div}(u(x)\nabla q(t,x)) = f(t,x) \quad (t,x) \in T \times D\,,\\
&q(0,x) = 0 \quad x \in D\,,\\
&q(t,0) = q(t,1) = 0 \quad t \in T\,,
\end{aligned} \tag{72}$$

with $f(t,x) = 100 \cdot tx$, $D = (0,1)$ and $T = (0,1)$. The uncertain coefficient $u$ is parametrized as

$$u(x,y) = \bar{a} + \sum_{j=1}^{64} y_j\psi_j\,, \text{where } \bar{a} = 1 \text{ and } \psi_j = \alpha_j\chi_{D_j}$$

with $D_j = [(j-1)\frac{1}{64}, j\frac{1}{64}]$, $\boldsymbol{y} = (y_j)_{j=1,\ldots,64}$ and $\alpha_j = \frac{0.9}{j^\zeta}$, $\zeta = 2,3,4$.

For a given realization of $u(x)$, the forward problem (72) is numerically solved by a backward Euler scheme in time with uniform time step $h_T = 2^{-11}$ and by a finite element method using continuous, piecewise linear ansatz functions in space on a uniform mesh with meshwidth $h_D = 2^{-11}$. The solution of the linear system in each time step is computed by LAPACK's DPTSV routine.

For given noisy observational data $\delta$, the goal of computation is the conditioned expectation $\mathbb{E}^{\mu^\delta}[\phi]$ of the QoI $\phi(u) = G(u)$ given by

$$Z' = \int_U \exp\big(-\Phi(u;\delta)\big)\phi(u)\Big|_{u=\bar{a}+\sum_{j=1}^{64} y_j\psi_j} d\mu_0(\boldsymbol{y})\,,$$

with $\phi(u) = \mathcal{G}(u)$, $\mathcal{S} = \mathcal{X}$ and with the normalization constant $Z$ given by

$$Z = \int_U \exp\big(-\Phi(u;\delta)\big)\Big|_{u=\bar{a}+\sum_{j=1}^{64} y_j\psi_j} d\mu_0(\boldsymbol{y})\,.$$

The noisy observational data is computed as a single realization of

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

with $\eta \sim \mathcal{N}(0,\Gamma)$ and $\mathcal{G} : L^\infty(D) \to \mathbb{R}^K$, with $K = 1,3,9$. The noise $\eta = (\eta_j)_{j=1,\ldots,K}$ in the measurements is assumed to be independent and normally distributed with $\eta_j \sim \mathcal{N}(0,1)$ and $\eta_j \sim \mathcal{N}(0,0.1^2)$. The observation operator $\mathcal{O}$ consists of $K$ system responses at $K$ observation points in $T \times D$ at $t_i = \frac{i}{2^{N_{K,T}}}, i = 1,\ldots,2^{N_{K,T}} - 1$ and $x_j = \frac{j}{2^{N_{K,D}}}, k = 1,\ldots,2^{N_{K,D}} - 1$, $o_k(\cdot,\cdot) = \delta(\cdot - t_k)\delta(\cdot - x_k)$ with $K = 1$, $N_{K,D} = 1, N_{K,T} = 1$, $K = 3$, $N_{K,D} = 2, N_{K,T} = 1$, $K = 9$, $N_{K,D} = 2, N_{K,T} = 2$. The numerical results presented below are based on synthetic noisy observational data, i.e. for a given realization of $u(x)$, the forward problem is solved with meshwidth $h_T = h_D = 2^{-12}$, the data $\delta$ is then computed according to (5) by the sum of the observed solution and a realization of the additive noise $\eta$.

In the following, we will compare the results of the proposed adaptive algorithm with a reference solution computed by the Smolyak algorithm with a fixed number of indices, $\#\Lambda = 1500$, i.e. altogether the number of PDE solves for the computation of the reference solution is in the range of $6149 - 18721$, depending on the adaptively determined set $\Lambda$ of active Smolyak details. The algorithm is used in the 64 dimensional parameter space, i.e. the dimension is not adaptively controlled in the case of the reference solution. Therefore, the set of reduced neighbours coincides with the set of neighbours.

Figure 1 and 2 show the quadrature error of the normalization constant Z with respect to the cardinality of the index set $\Lambda$ based on the sequence CC.

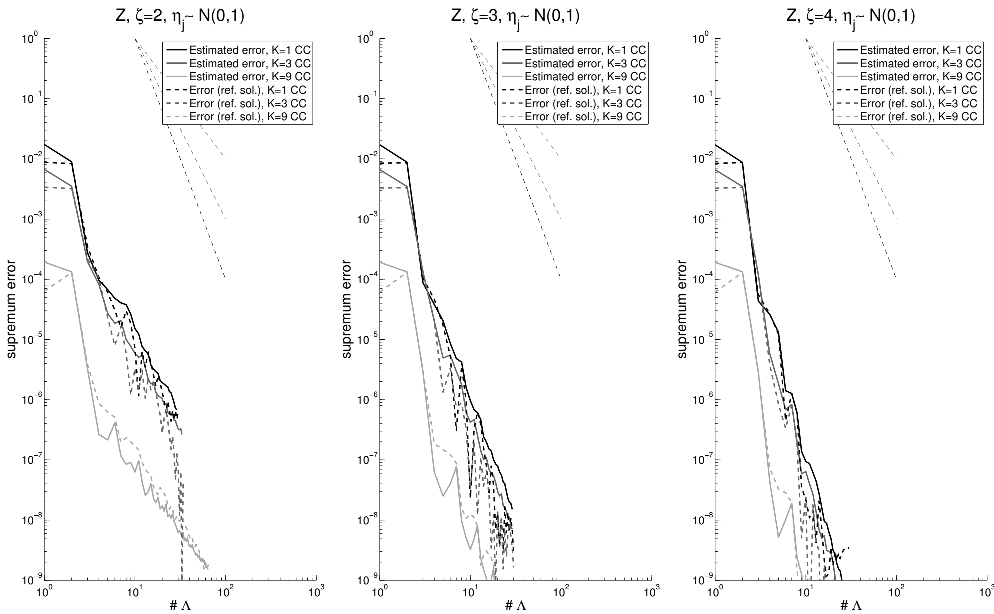


**Figure 1.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the cardinality of the index set $\Lambda_N$ based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

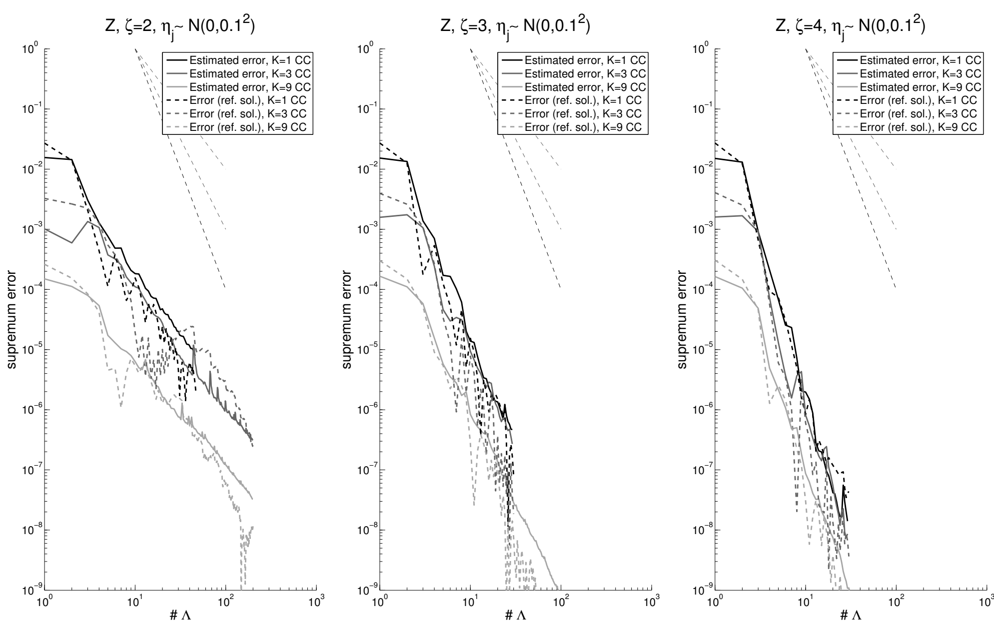


**Figure 2.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the cardinality of the index set $\Lambda_N$ based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

The corresponding, estimated error curves and error curves computed by the reference solution of the normalization constant $Z$ based on the sequence RL are displayed in Figure 3 and 4.

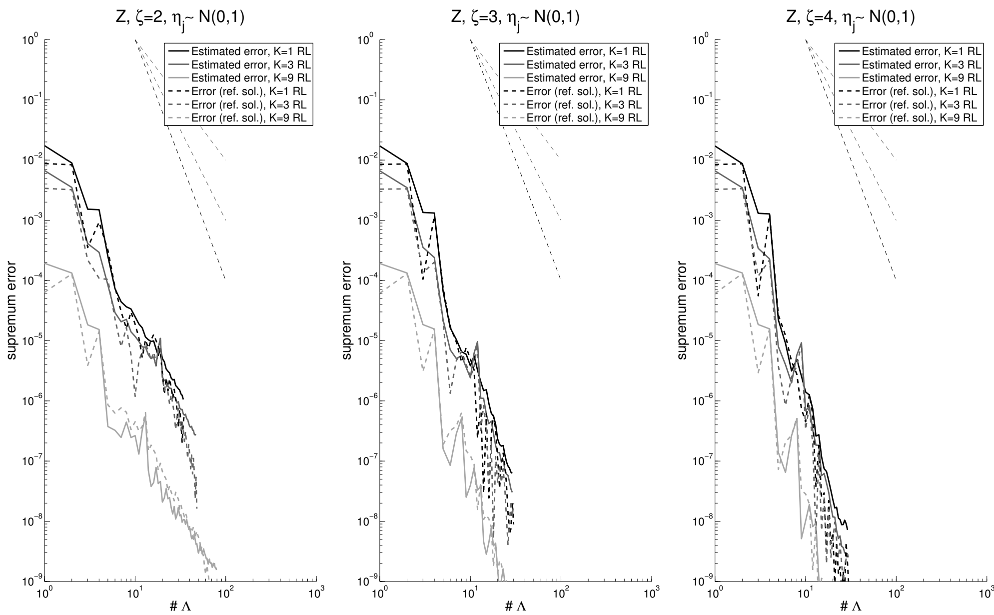


**Figure 3.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the cardinality of the index set $\Lambda_N$ based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

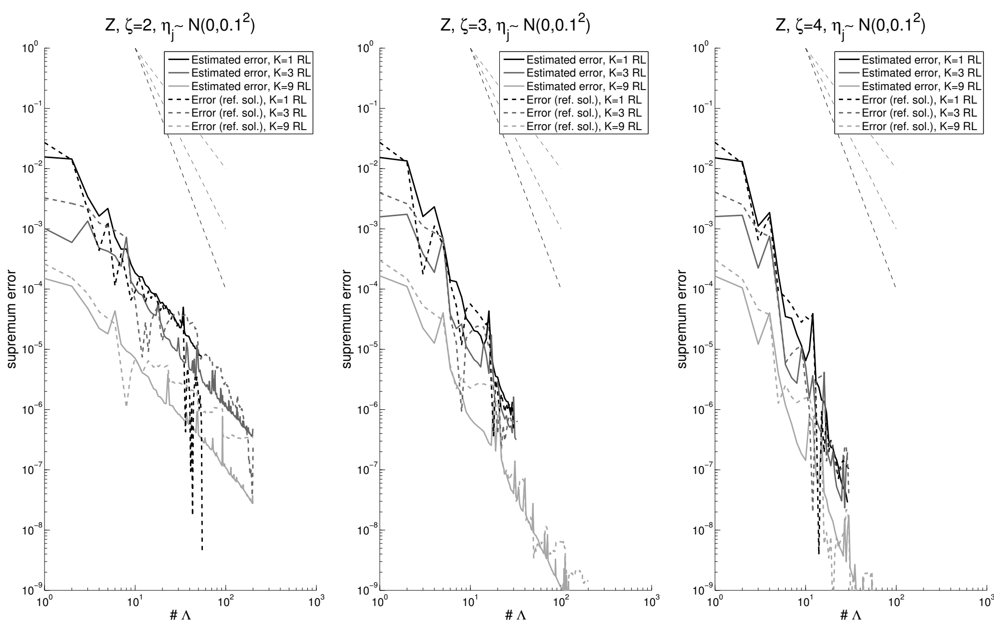


**Figure 4.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the cardinality of the index set $\Lambda_N$ based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

We observe that the estimated error by the adaptive algorithm provides a good indicator, so that the proposed algorithm shows an optimal performance with respect to the convergence rates. The theoretical convergence rate can be observed for all values of the parameter $\zeta$ controlling the sparsity class of the unknown coefficient $u$. Further, the Clenshaw-Curtis points show a better convergence behaviour with respect to the cardinality of the index set $\Lambda$ than the Leja points. This behaviour could be already observed in the elliptic test case, cp. [29] and can be attributed to the exponential growth of the number of quadrature points within the order of CC sequences. As Figure 5 and Figure 6 exemplarily show, this effect is not any more observable in the error curves of the normalization constant with respect to the number of PDE solves needed.

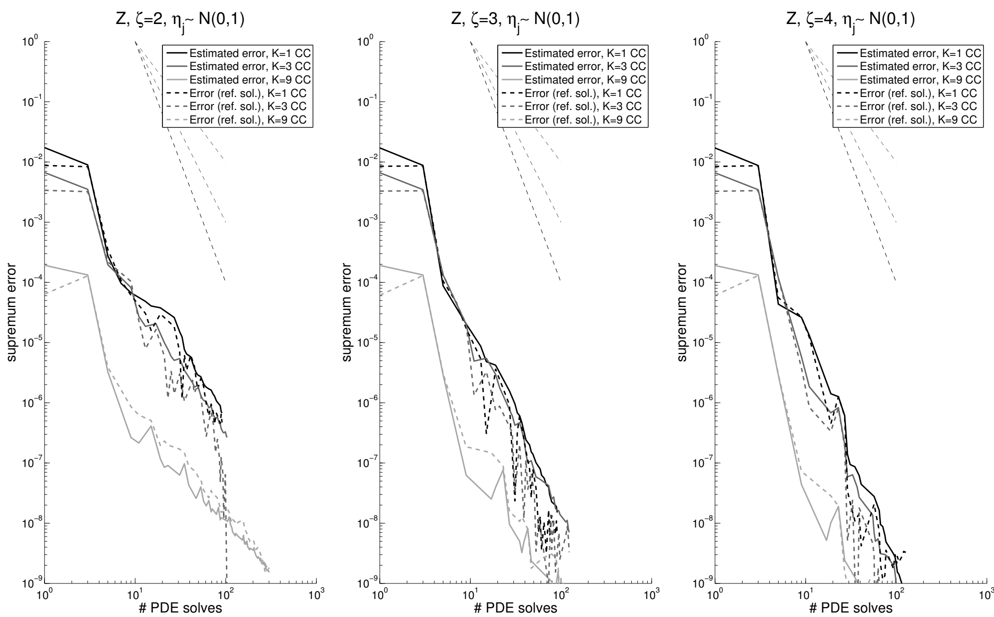


**Figure 5.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the number of PDE solves needed based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

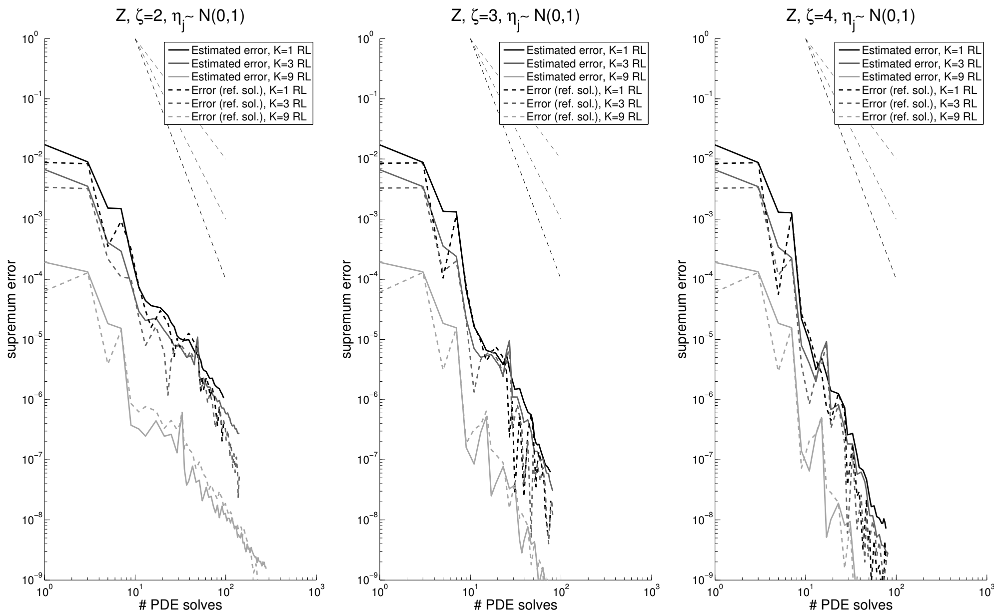


**Figure 6.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the number of PDE solves needed based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

The same convergence behavior for the approximation of the quantity $Z'$ can be observed, cp. Figure 7 - Figure 10 showing the error curves with respect to the cardinality of the index set $\Lambda$.

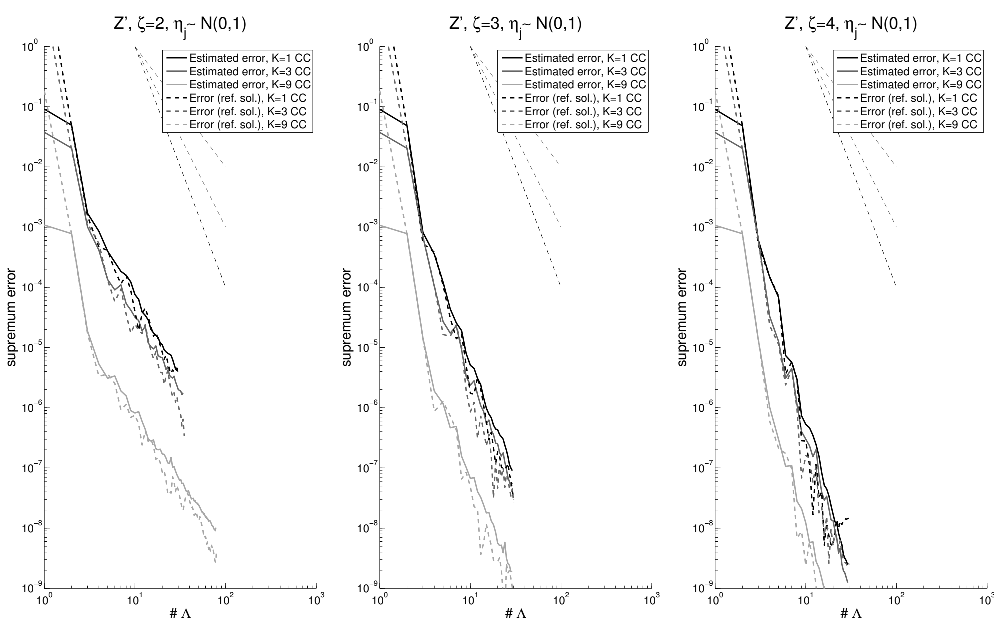


**Figure 7.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity $Z'$ with respect to the number of PDE solves needed based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

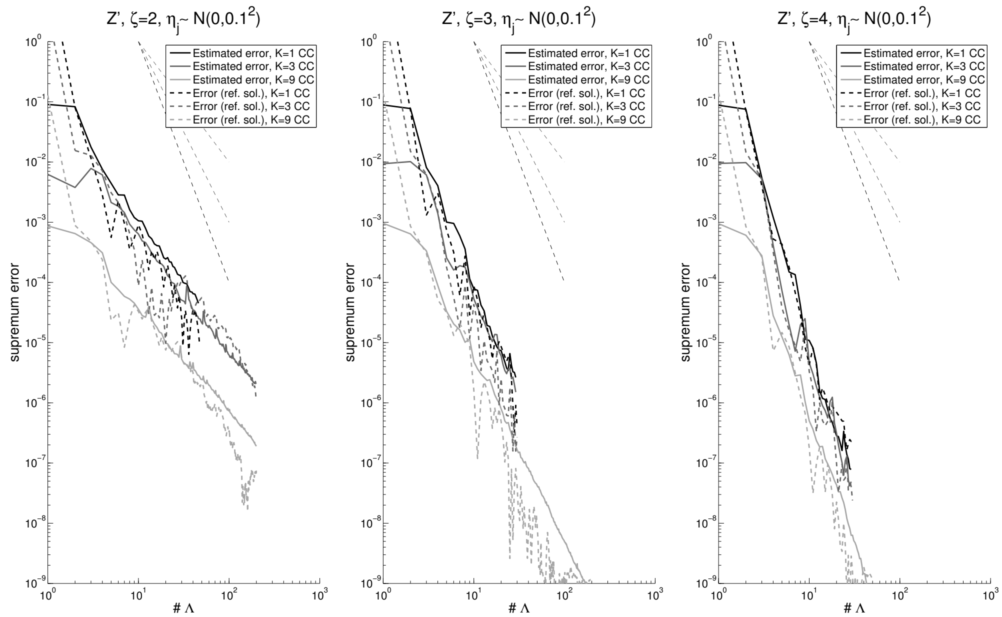


**Figure 8.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity $Z'$ with respect to the number of PDE solves needed based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\#\mathbb{J} = 64$ and $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

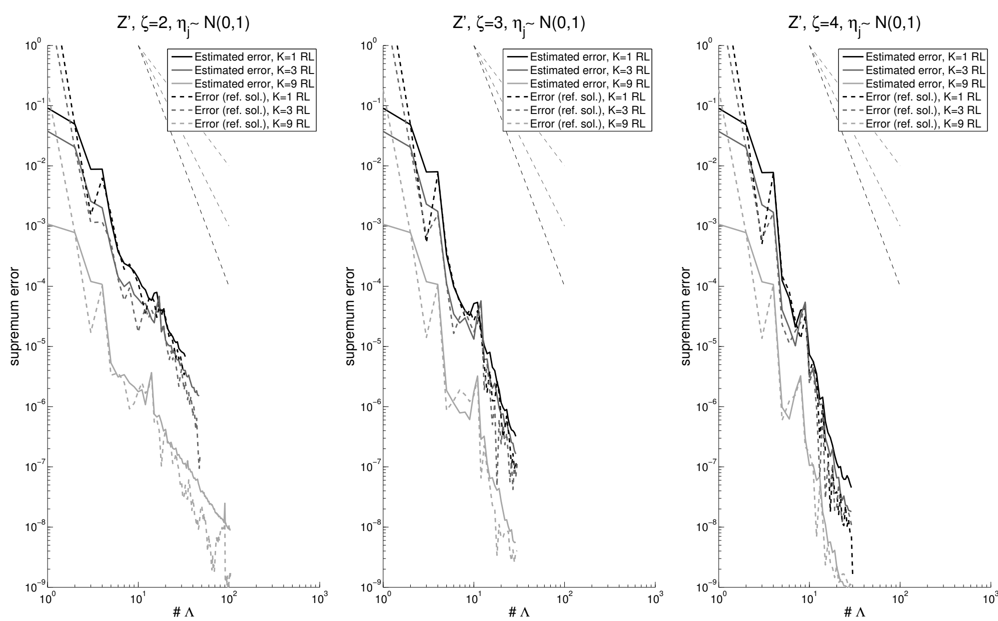


**Figure 9.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity $Z'$ with respect to the number of PDE solves needed based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

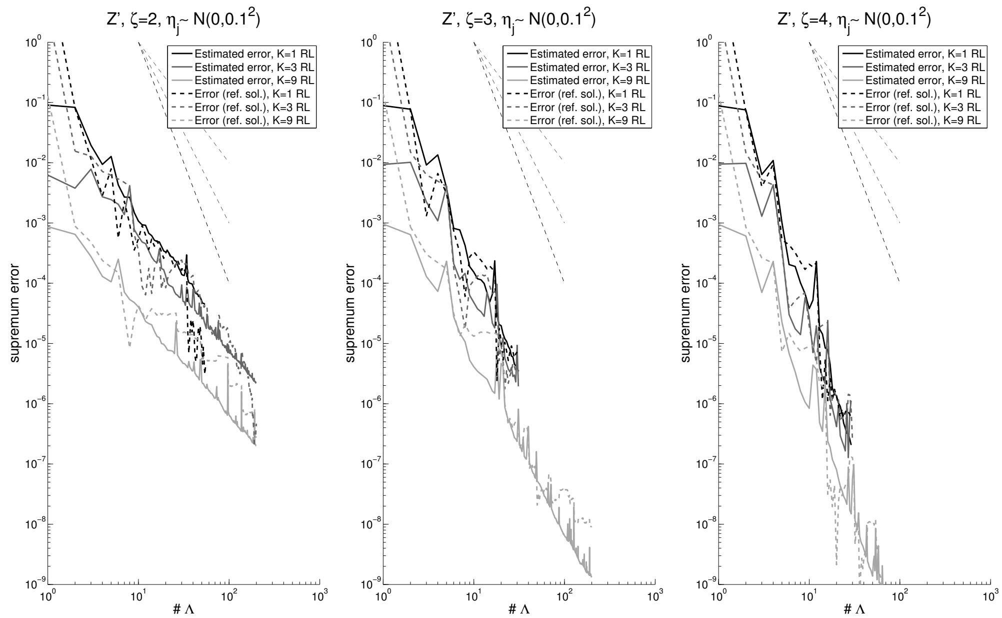


**Figure 10.** Comparison of the estimated error and actual error. Curves computed by Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity $Z'$ with respect to the number of PDE solves needed based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

In order to numerically verify the dimension robust behavior of the proposed algorithm, we will finally investigate the convergence rates of the model parametric parabolic problem (72) in the 128 dimensional parameter case, i.e. the uncertain coefficient $u$ is parametrized by

$$u(x, y) = \bar{a} + \sum_{j=1}^{128} y_j \psi_j \,, \text{where } \bar{a} = 1 \text{ and } \psi_j = \alpha_j \chi_{D_j}$$

with $D_j = [(j-1)\frac{1}{128}, j\frac{1}{128}]$, $\boldsymbol{y} = (y_j)_{j=1,\dots,128}$ and $\alpha_j = \frac{0.6}{j^\zeta}, \zeta = 2, 3, 4$.

The doubling of the number of parameters has no effect on the observed convergence rates, cp. Figure 11 and 12, this observation is consistent with the theoretical results derived in Theorem 5.4.

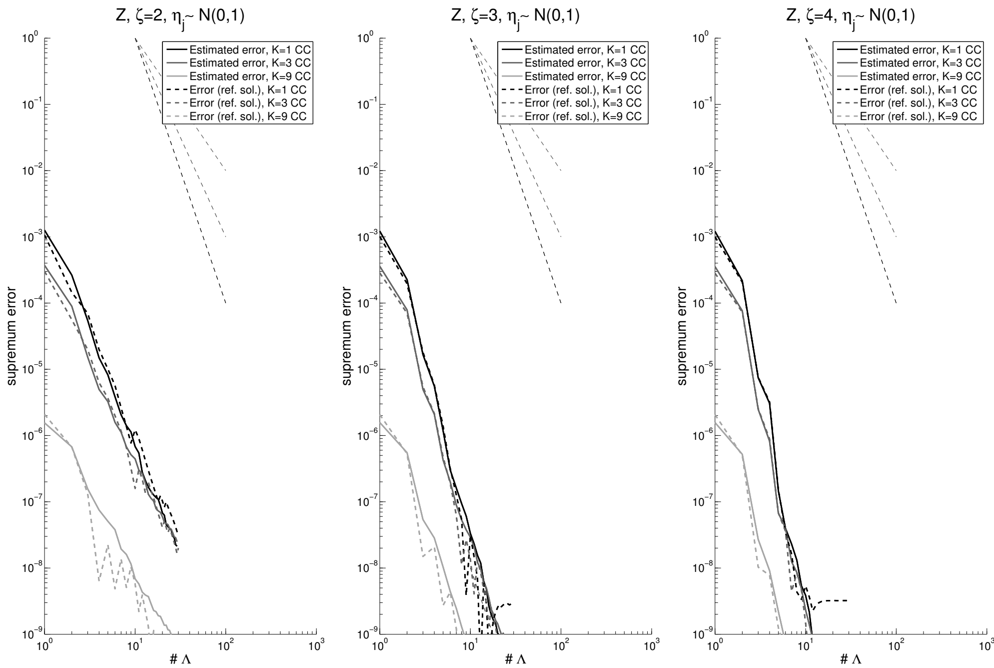


**Figure 11.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant $Z$ with respect to the cardinality of the index set $\Lambda_N$ based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 128$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

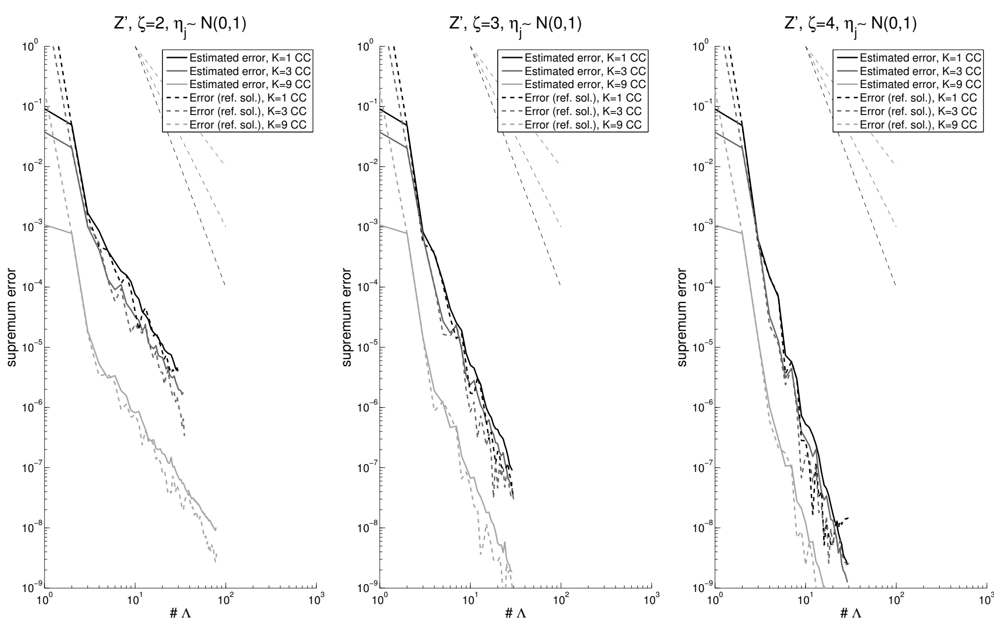


**Figure 12.** Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity $Z'$ with respect to the number of PDE solves needed based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 128$ and $h_T = h_D = 2^{-11}$ for the reference and the adaptively computed solution.

In summary, for the parametric, parabolic evolution problem with random coefficients, our theoretical results could be numerically verified, and the experimentally observed convergence rates are even slightly better. Further, the variation of the number of observation points as well as the variation of the observational noise do not influence the convergence behaviour of the proposed method. The convergence only depends on the sparsity class of the unknown coefficient $u$ and is independent on the dimension of the underlying parameter space.

## 7. Discussion and Conclusions

We consider the Bayesian inversion for classes of operator equations with distributed uncertainties $u$ taking values in a Banach space $X$. We showed sparsity of coefficient sequences in polynomial chaos representations of the Bayesian posterior density $\Theta$ for parametrizations of the uncertain forward solution map of the system in terms of possibly countably many variables $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$, provided that the parametric responses $q(\boldsymbol{y})$ satisfy the $(p,\varepsilon)$-analyticity condition in Definition 2.2 with some $0 < p < 1$.

This analyticity condition is valid for a wide range of PDE problems.

We showed that a certain type of degree and dimension adaptive Smolyak quadrature can, in principle, achieve convergence rate $N^{-(1/p-1)}$ where $N$ denotes the number quadrature points; numerical experiments indicate that even the higher rate $N^{-1/p}$ is achieved by the proposed deterministic quadrature methods, *provided that covariance* $\Gamma > 0$ *of the observation noise is not small.*

In the case of observation noise with variance $\Gamma \to 0$, the bound (60) reveals that the constants in the bounds on the Legendre coefficients $\theta^P_\nu$ in the gpc expansions (62) and, via (63), also the constants $C > 0$ in the error bounds (65), (67) and, in turn, also the constants $C^i$ in the Smolyak quadrature error estimates (70), (71), depend on $\Gamma$ as $C \sim \exp(b/\Gamma)$ for some constant $b > 0$. We also not that the *convergence rates* in (70), (71) are not affected by the size of $\Gamma$. In our numerical experiments, we observe this dependence on $\Gamma$, which renders our approach infeasible for small values of $\Gamma$. This is due to concentration effects in the integrand functions of the integrals $Z_\Gamma$ and $Z'_\Gamma$ in (47), (45) for small values of $\Gamma$. Since the integrals (47), (45) are nonoscillatory, as $\Gamma \to 0^+$, all contributions to the integrals $Z_\Gamma$ and $Z'_\Gamma$ in (47), (45) come from the vicinity of points $\boldsymbol{y}^0 \in U$ where the potential $\Phi_\Gamma(\boldsymbol{y};\delta)$ is minimal, and the asymptotics of $Z_\Gamma$ and $Z'_\Gamma$ as $\Gamma \to 0^+$ can be analyzed by Laplace's method. Specifically, assuming that the number $K$ of observations equals one to simplify notation, we define

$$S(\boldsymbol{y}) := -\Phi_\Gamma(\boldsymbol{y};\delta) = -\frac{1}{2}\Gamma^{-1}(r(\boldsymbol{y}))^2\,, \quad \Gamma > 0$$

where the residuum $r(\boldsymbol{y}) := \mathcal{G}(\boldsymbol{y}) - \delta \geq 0$ of the uncertainty-to-observation map is independent of $\Gamma$ and a smooth function of the coordinates $y_j$ of $\boldsymbol{y} \in U$. Assume, moreover, that the dimension $U$ (resp. the set $\mathbb{J}$) is finite, $\#(\mathbb{J}) = J < \infty$ (which could be achieved by dimension truncation in the parametric representation (7) of the uncertain input $u$).

Since $U = [-1,1]^J$ is compact, the continuous function $S(\boldsymbol{y}) \leq 0$ attains its maximum on $U$ in a point $\boldsymbol{y}^0 \in U$, say. Two cases can occur: $\boldsymbol{y}^0 \in \mathrm{int}(U)$ and $\boldsymbol{y}^0 \in \partial U$. Assume the former, ie. $\mathrm{dist}(\boldsymbol{y}^0, \partial U) > 0$. Then $\boldsymbol{y}^0$ is a critical point of $S(\cdot)$, and there holds the first order necessary condition

$$0 = \nabla_{\boldsymbol{y}} S(\boldsymbol{y}^0) \Longleftrightarrow r(\boldsymbol{y}) \nabla_{\boldsymbol{y}} r(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{y}^0} = 0 \, . \tag{73}$$

Again two cases can occur: either ("consistent case") $r(\boldsymbol{y}^0) = 0$ in which case the observed noise-free data is the exact system response for the realization $u = u(\boldsymbol{y}^0)$ of the uncertainty, or ("inconsistent case") $r(\boldsymbol{y}^0) \neq 0$. In the latter case, $S(\boldsymbol{y}^0) < 0$ and (73) implies $\nabla_{\boldsymbol{y}} r(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{y}^0} = 0$ i.e. that $\boldsymbol{y}^0$ is a critical point of the residuum. Assume that $\boldsymbol{y}^0$ is nondegenerate, so that $S(\boldsymbol{y}_0) < 0$ is a local maximum of $S$ (and, hence, a local minimum of the potential $\Phi_\Gamma$) and the Hessian $S_{\boldsymbol{yy}}(\boldsymbol{y}^0)$ is negative definite.

Asymptotic analysis of $Z_\Gamma$ and of $Z'_\Gamma$ via Laplace's method shows, as $\Gamma \to 0^+$,

$$Z'_\Gamma = \exp(\Gamma^1 S(\boldsymbol{y}^0))(2\pi\Gamma)^{J/2} \frac{\phi(\boldsymbol{y}^0) + O(\Gamma)}{\sqrt{|\det(S_{\boldsymbol{yy}}(\boldsymbol{y}^0))|}}$$

and likewise for $Z_\Gamma$ with $\phi(\boldsymbol{y})$ replace by 1. Under the provision of nondegeneracy of the Hessian $S_{yy}(\boldsymbol{y}_0)$, the Bayesian estimate (47) admits an *asymptotic expansion with respect to small observation noise variance* $\Gamma$,

$$\mathbb{E}^{\mu^\delta}[\phi] = \frac{Z'_\Gamma}{Z_\Gamma} \sim \phi(\boldsymbol{y}^0) + \sum_{k \geq 1} a_k \Gamma^k \quad \text{as} \quad \Gamma \to 0^+ \, .$$

In the case that the maximum is degenerate or that $\boldsymbol{y}^0 \in \partial U$ analogous arguments can be applied; we refer to [30]. Apart from being of interest in its own right (it indicates that in the limit of noise-free observations the expected response $\phi(\boldsymbol{y}^0)$ occurs at a realization which is a (nonlinear) least square minimizer of the uncertainty for the determination of which deterministic methods from optimization are available) the precise asymptotic information afforded by the asymptotic analysis will also allow the regularization of the integrand functions $\Theta(\boldsymbol{y})$ and $\Psi(\boldsymbol{y})$ in (47) and (45). This will be presented in [30].

For $\Gamma > 0$ not necessarily small, we showed in particular for parametric operator equations whose solutions $q(\boldsymbol{y})$ are $(p, \varepsilon)$-analytic, that in inverse problems for such operator equations, under parametric uncertainty, the density of the Bayesian posterior measure with respect to a uniform prior $\mu_0$ on the parametrization space $U$ of the uncertainty is, likewise, $(p, \varepsilon)$-analytic on $U$. This implies that the adaptive Smolyak quadrature algorithms presented herein can, in principle, achieve convergence rate $s = 1/p - 1$ for the approximation of the Bayesian posterior, which rate is superior to the Monte-Carlo rate $s = 1/2$ for $p < 2/3$. In numerical experiments, we observed the rate $s = 1/p$ indicating convergence rates which are even higher than the MC rate 1/2 whenever $p < 2$.

We assumed in the present paper that the uncertainties $u \in X$ were charged with a *uniform* prior measure $\mu_0$ which assigns equal probability to all relizations of each coordinate $y_j$ in the uncertainty parametrization (7): we worked within the probability space $(U, \mathcal{B}, \mu_0)$. However, all results and algorithms generalize straightforwardly also the more general setting where $U = \prod_{j\in\mathbb{J}} \Gamma_j$ with $\Gamma_j \subset \mathbb{R}$ compact, with $\frac{1}{2}\lambda^1$ replaced by the probability measures $\rho_j(y_j)dy_j$ with $\int_{-1}^1 \rho_j(\xi)d\xi = 1$. In this case, the families $\{Q^k\}_{k\geq 0}$ of univariate quadratures on which the Smolyak construction in Section 5 was based will be replaced by coordinate-dependent families $\{Q^{k,j}\}_{k\geq 0}$, $j \in \mathbb{J}$, such as, for example, Gaussian quadratures with weight function $\rho_j$ which are tailored to the prior with respect to coordinate $y_j$ in the parametric representation (7) of the distributed uncertainty $u \in X$.

The extension of the present theory to $U = \mathbb{R}^{\mathbb{N}}$ which arises, for example, in the context of lognormal Gaussian models for the uncertain input $u$, will require technical modifications; however, the adaptive Smolyak algorithm for fast, deterministic Bayesian estimation presented in Section 5 ahead does generalize to this case. See, eg., [28].

So far, we assumed that the forward problems are solved numerically with high accuracy so that the discretization error is negligible with respect to the quadrature error; the present error analysis allows, however, to adapt the discretization error of the forward problem to the expected significance of its contribution to the Bayesian estimate, leading to substantial reduction in overall computational complexity. We refer to [15] for first numerical experiments on this in the context of adaptive solution of parametric initial value problems. The use of the presently proposed Smolyak quadrature scheme in connection with efficient processing of large sets of data $\delta$ will be dealt with in a forthcoming work.

**Acknowledgments**

This work is supported by Swiss National Science Foundation (SNF) and by the European Research Council (ERC) under FP7 Grant AdG247277.

**References**

[1] Abdulle A, Barth A and Schwab C, Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs, Report 2012-29, Seminar for Applied Mathematics, ETH Zürich (in review).

[2] Barth A, Schwab C and Zollinger N, Multi-Level Monte Carlo Finite Element Method for Elliptic PDEs with Stochastic Coefficients Numerische Mathematik **119**, Number 1, pages 123-161 (2011).

[3] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Verlag, Berlin, 1991.

[4] Bui-Thanh T and Ghattas O 2012 Analysis of the Hessian for inverse scattering problems. Part II: Inverse medium scattering of acoustic waves *Inverse Problems* **28** 055002

[5] Calvi J-P and Phung Van M 2011 On the Lebesgue constant of Leja sequences for the unit disk and its applications to multivariate interpolation *Journal of Approximation Theory* **163** 608-622

[6] Calvi J-P and Phung Van M 2012 Lagrange interpolation at real projections of Leja sequences for the unit disk *Proceedings of the American Mathematical Society* (to appear) (uncorrected preprint available on ArXiv)

[7] Chkifa A 2013 On the Lebesgue constant of Leja sequences for the unit disk Journal of Approximation Theory (2013).

[8] Chkifa A, Cohen A, DeVore R and Schwab C 2011 Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs *Report 2011-44, Seminar for Applied Mathematics, ETH Zürich, Switzerland http : //www.sam.math.ethz.ch/reports/*2011/44 (to appear in M2AN 2012)

[9] Chkifa A, Cohen A and Schwab C 2012 High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs *Report 2012-NN, Seminar for Applied Mathematics, ETH Zürich, Switzerland http : //www.sam.math.ethz.ch/reports/*2012 (to appear in Journ. Found. Comp. Math. 2013).

[10] Chkifa A, Cohen A and Schwab C 2013 Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs (in preparation) (2013).

[11] Cohen A, DeVore R and Schwab C 2010 Convergence rates of best $N$-term Galerkin approximations for a class of elliptic SPDEs *Journ. Found. Comp. Math.* **10** 615-646

[12] Cohen A, DeVore R and Schwab C 2011 Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs *Analysis and Applications* **9** 1-37

[13] Gerstner T and Griebel M 2003 Dimension-adaptive tensor-product quadrature *Computing* **71** 65-87

[14] Hansen M and Schwab C 2011 Analytic regularity and best $N$-term approximation of high dimensional parametric initial value problems *Report 2011-64, Seminar for Applied Mathematics, ETH Zürich, Switzerland* `http://www.sam.math.ethz.ch/reports/2011/64` *(to appear in Vietnam Journ. Mathematics (2013))*

[15] Hansen M, Schillings C and Schwab C 2013 Sparse approximation algorithms for high dimensional parametric initial value problems *Proceedings 5th Conf. High Performance Computing, Hanoi* Springer LNCSE (to appear).

[16] Ha Hoang V and Schwab C Regularity and Generalized Polynomial Chaos Approximation of Parametric and Random Second-Order Hyperbolic Partial Differential Equations Analysis and Applications (Singapore) (2012) **10**(3)

[17] Ha Hoang V and Schwab C Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale PDEs, Analysis and Applications (Singapore) **11**(1) (2013) 1350001.

[18] Ha Hoang V, Schwab C and Stuart A M 2012 Sparse MCMC gpc Finite Element methods for Bayesian inverse problems *Inverse Problems (to appear)*

[19] Kaipio J and Somersalo E 2005 *Statistical and computational inverse problems* Applied Mathematical Sciences vol 160 (Springer)

[20] Liu J 2001 *Monte Carlo Strategies in Scientific Computing* Springer Texts in Statistics (Springer)

[21] Martin J, Wilcox L C, Burstedde C and Ghattas O 2012 A stochastic Newton MCMC method for large-scale statistical inverse problems with application to seismic inversion *SIAM Journal on Scientific Computing* **34** 1460-1487

[22] Marzouk Y M, Najm H N and Rahn L A 2007 Stochastic spectral methods for efficient Bayesian solution of inverse problems *Journ. Comp. Phys.* **224** 560-586

[23] Marzouk Y M and Xiu D 2009 A Stochastic Collocation Approach to Bayesian Inference in Inverse Problems *Communications in Comp. Phys.* **6** 826-847

[24] Marzouk Y M and Najm H N2099 Dimensionality reduction and polynomial chaos acceleration

of Bayesian inference in inverse problems *Journal of Computational Physics* **228** 1862-1902

[25] McLaughlin D and Townley L R 1996 A reassessment of the groundwater inverse problem *Water Resour. Res.* **32** 1131-1161

[26] Robert C P and Casella G C 1999 *Monte Carlo Statistical Methods* Springer Texts in Statistics (Springer)

[27] Roberts G O and Sherlock C 2012 Optimal Scaling of Random Walk Metropolis algorithms with discontinuous target densities `http://www.imstat.org/aap/future-papers.html` (to appear in Ann. Appl. Proba.)

[28] Schillings Cl 2013 A Note on Sparse, Adaptive Smolyak Quadratures for Bayesian Inverse Problems *Report 2013-06, Seminar for Applied Mathematics, ETH Zürich, Switzerland* `http://www.sam.math.ethz.ch/reports/2013/06` (to appear in Oberwolfach Reports (2013)).

[29] Schillings Cl and Schwab Ch 2013 Sparse, Adaptive Smolyak Algorithms for Bayesian Inverse Problems, *Inverse Problems (to appear)*

[30] Schillings Cl and Schwab Ch 2013 Scaling Limits in Computational Bayesian Inversion of Parametric and Stochastic Operator Equations (in preparation)

[31] Schwab Ch and Stevenson R S *Space-Time adaptive wavelet methods for parabolic evolution equations*, Math. Comp. **78** (2009), no. 267, 1293 - 1318

[32] Schwab Ch and Stuart A M 2012 Sparse Deterministic Approximation of Bayesian Inverse Problems *Inverse Problems* **28** 045003

[33] Schwab Ch and Todor R A, Karhunen-Loève Approximation of Random Fields by Generalized Fast Multipole Methods, Journal of Computational Physics **217** (2006), 100-122.

[34] Stuart A M 2010 Inverse problems: a Bayesian approach *Acta Numerica* **19** 451-559

# Recent Research Reports

| Nr. | Authors/Title |
|---|---|
| 2013-07 | A. Paganini and M. López-Fernández<br>Efficient convolution based impedance boundary condition |
| 2013-08 | R. Hiptmair and C. Jerez-Hanckes and J. Lee and Z. Peng<br>Domain Decomposition for Boundary Integral Equations via Local Multi-Trace Formulations |
| 2013-09 | C. Gittelson and R. Andreev and Ch. Schwab<br>Optimality of Adaptive Galerkin methods for random parabolic partial differential equations |
| 2013-10 | M. Hansen and C. Schillings and Ch. Schwab<br>Sparse Approximation Algorithms for High Dimensional Parametric Initial Value Problems |
| 2013-11 | F. Mueller and Ch. Schwab<br>Finite Elements with mesh refinement for wave equations in polygons |
| 2013-12 | R. Kornhuber and Ch. Schwab and M. Wolf<br>Multi-Level Monte-Carlo Finite Element Methods for stochastic elliptic variational inequalities |
| 2013-13 | X. Claeys and R. Hiptmair and E. Spindler<br>A Second-Kind Galerkin Boundary Element Method for Scattering at Composite Objects |
| 2013-14 | I.G. Graham and F.Y. Kuo and J.A. Nichols and R. Scheichl and Ch. Schwab and I.H. Sloan<br>Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficient |
| 2013-15 | A. Lang and Ch. Schwab<br>Isotropic Gaussian random fields on the sphere: regularity, fast simulation, and stochastic partial differential equations |
| 2013-16 | P. Grohs and H. Hardering and O. Sander<br>Optimal A Priori Discretization Error Bounds for Geodesic Finite Elements |