Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Sparsity in Bayesian Inversion of Parametric Operator Equations

Cl. Schillings and Ch. Schwab

Research Report No. 2013-17
October 2013

Revised: November 2013

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

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 [32] 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 [35, 32], 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.

In an infinite-dimensional parametric, deterministic description of distributed parameter uncertainty we prove regularity and sparsity of the posterior density in Bayesian inversion. These 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 the dimension of the parameter space. The 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 inference of responses of uncertain systems, in the presence of possibly large sets of observational data that are 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 given, noisy observational 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. [21, 23, 29, 30]). 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 [24], 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* can provide substantial improvements [20, 22]. 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, 20] 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 an infinite-dimensional space X of uncertain coefficient functions u which are distributed w.r. to a prior measure μ_0 on X 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 [32] 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 [27, 26, 25] and further analyzed in [20]. Here, as in our previous work [35, 32], we reformulate the Bayesian inversion problem as a deterministic, infinite-dimensional quadrature problem with respect to the posterior measure, given noisy observational data δ of a QoI ϕ, 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 [32] or by a Quasi Monte-Carlo quadrature rule such as a polynomial lattice rule, see [13, 14].

Under certain regularity assumptions on the covariance spectrum of the unknown system parameter, 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 [35, 32] 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 [35] and to [32], 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, 18]), and also to Bayesian inversion subject to uncertainty in coefficients, in loadings and in domains. Also, the Smolyak quadrature convergence result given in [32] is generalized herein: whereas in [32], the integrand functions were required to allow for analytic extensions into polydiscs, here this condition is weakened to analyticity in poly-ellipses 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 is shown in certain cases to allow global analytic continuation of parametric solutions of nonlinear problems, 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 [35]. 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. We denote by $\mathcal{X}$ and $\mathcal{Y}$ two reflexive Banach spaces over $\mathbb{R}$ 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 $\mathfrak{a}(\cdot,\cdot) : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}$ via (with ${}_{\mathcal{Y}}\langle\cdot,\cdot\rangle_{\mathcal{Y}'}$ denoting the $\mathcal{Y}\times\mathcal{Y}'$-duality pairing)

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

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. To this end, we extend the spaces $\mathcal{X}$ and $\mathcal{Y}$ to spaces over $\mathbb{C}$ and the form $\mathfrak{a}(\cdot,\cdot)$ in the usual fashion. ‡

2.1. Operator Equations with Uncertain Distributed Input

For a distributed, uncertain parameter $u \in X$, we consider a "forward" operator $\mathcal{A}(u;q)$ depending on $u \in X$ and acting on $q \in \mathcal{X}$ and taking values in $\mathcal{Y}'$. For the well-posedness of operator equations involving $\mathcal{A}(u;q)$, we assume the map $\mathcal{A}(u;q)$ to be boundedly invertible *locally, at a "nominal value"* $\langle u\rangle \in X$. In particular, then, for all u in a sufficiently small, open neighborhood (eg. $\tilde{X} = B_R(\langle u\rangle;X)$, a ball of radius $R > 0$ in X centered at $\langle u\rangle$) $\tilde{X} \subset X$ of $\langle u\rangle \in X$ the *forward problem*: for every $u \in \tilde{X} \subseteq X$, given $F(u) \in \mathcal{Y}'$, find $q(u) \in \mathcal{X}$ such that

$$\mathcal{A}(u;q) = A(u;q) - F(u) = 0 \quad \text{in} \quad \mathcal{Y}' \tag{2}$$

‡ 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 $\mathcal{X}$, is a Hilbert space with inner product $(u,v)_{\mathbb{C}} := (u_1,v_1) + (u_2,v_2) + \boldsymbol{i}[(u_2,v_1) - (v_1,v_2)]$. 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 $\mathbb{C}$.

should admit a unique solution. 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). Specifically, for every fixed $u \in \tilde{X}$, for every $F(u) \in \mathcal{Y}'$, there should exist a unique solution $q(u)$ of (2). The following proposition collects well-known sufficient conditions for well-posedness of (2). Its statement considers so-called *regular solutions* q_0 of (2): given $u \in \tilde{X}$, a solution $q_0 \in \mathcal{X}$ of (2) is a *regular solution of* (2) *at* $u \in \tilde{X}$ if, for this $u \in \tilde{X}$, the map $\mathcal{X} \ni q \mapsto \mathcal{A}(u;q)$ is Frechet differentiable with respect to q at $q_0 \in \mathcal{X}$ and if the differential $D_q\mathcal{A}(u;q_0) \in \mathcal{L}(\mathcal{X},\mathcal{Y}')$ is an isomorphism from $\mathcal{X}$ onto $\mathcal{Y}'$. For regular solutions, the differential $D_q\mathcal{A}(u;q_0)$ satisfies the so-called inf-sup conditions.

Proposition 2.1. *Assume that $\mathcal{Y}$ is reflexive and that, for every u in a subset $\tilde{X} \subseteq X$, the operator equation* (2) *admits a* regular solution $q(u) \in \mathcal{X}$. *Then the bilinear map $(\varphi,\psi) \mapsto_{\mathcal{Y}'} \langle D_q\mathcal{A}(u;q(u))\varphi,\psi\rangle_{\mathcal{Y}}$ is boundedly invertible,* uniformly with respect to $u \in \tilde{X}$. *Ie. that there exists a constant $\kappa > 0$ such that for every $u \in \tilde{X}$*

$$\|(D_q\mathcal{A})(u;q(u))\|_{\mathcal{L}(\mathcal{X},\mathcal{Y}')} = \sup_{0\neq\varphi\in\mathcal{X}} \sup_{0\neq\psi\in\mathcal{Y}} \frac{{}_{\mathcal{Y}'}\langle(D_q\mathcal{A})(u;q(u))\varphi,\psi\rangle_{\mathcal{Y}}}{\|\varphi\|_{\mathcal{X}}\|\psi\|_{\mathcal{Y}}} \leq \kappa^{-1}\,, \tag{3}$$

and

$$\inf_{0\neq\varphi\in\mathcal{X}} \sup_{0\neq\psi\in\mathcal{Y}} \frac{{}_{\mathcal{Y}'}\langle(D_q\mathcal{A})(u;q(u))\varphi,\psi\rangle_{\mathcal{Y}}}{\|\varphi\|_{\mathcal{X}}\|\psi\|_{\mathcal{Y}}} \geq \kappa > 0\,. \tag{4}$$

Under conditions (3), (4), for every $u \in \tilde{X} \subseteq X$, there exists a unique, regular solution $q(u)$ which is uniformly bounded with respect to $u \in \tilde{X}$ in the sense that there exists a constant $C(F,\tilde{X}) > 0$ such that

$$\sup_{u\in\tilde{X}} \|q(u)\|_{\mathcal{X}} \leq C(F,\tilde{X})\,. \tag{5}$$

2.2. Bayesian Inversion

By $\mathcal{G} : X \to Y$ we denote a "forward" response map from the Banach space X of unknown parameters u into a second Banach space Y 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 and analogously Y and $\mathcal{Y}$.

Then the equation (2) takes the form

$$\text{Given } u \in \tilde{X} \subseteq X, \text{ find } q \in \mathcal{X}: \quad A(u;q) = F(u) \quad \text{in} \quad \mathcal{Y}' \tag{6}$$

where the forcing function $F : \tilde{X} \mapsto \mathcal{Y}'$ is assumed to be known, and where the uncertain operator $A(u;\cdot) : \mathcal{X} \mapsto \mathcal{Y}'$ is assumed to be boundedly invertible, at least locally for uncertain input u in a sufficiently small neighborhood $\tilde{X}$ of a nominal input $\langle u\rangle \in X$ (eg. for $\|u - \langle u\rangle\|_X$ small enough) so that, for $u \in \tilde{X}$, the response of forward problem (6)

$$\tilde{X} \ni u \mapsto q(u) = G(u;F) \in \mathcal{X}$$

is well-defined and the uncertainty-to-observation map is given by

$$X \supseteq \tilde{X} \ni u \mapsto \mathcal{G}(u) := \mathcal{O}(G(u;F)) \in Y\ .$$

In order to not overburden notation, we simply write $q = G(u)$ for the uncertainty-to-solution map.

We also assume given a bounded, linear observation functional $\mathcal{O}(\cdot) \in \mathcal{L}(\mathcal{X}, Y)$, i.e. a *bounded linear observation operator* on the space $\mathcal{X}$ of system responses. We shall assume throughout that we are given finitely many observations, so that $Y = \mathbb{R}^K$ and $\mathcal{O} \in (\mathcal{X}^*)^K$, where $\mathcal{X}^*$ denotes 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 δ. Specifically, we assume the data δ to consist of observations of system responses corrupted by additive noise η, ie.

$$\delta = \mathcal{G}(u) + \eta = \mathcal{O}(G(u)) + \eta \in Y = \mathbb{R}^K \tag{7}$$

where $\eta \in \mathbb{R}^K$ represents Gaussian noise in the K-vector of observation functionals $\mathcal{O}(\cdot) = (o_k(\cdot))_{k=1}^K$. As in [37, 35, 32] and the references there, in the present paper we assume that the noise process η is Gaussian on $Y = \mathbb{R}^K$, i.e. η is a random vector $\eta \sim \mathcal{N}(0, \Gamma)$, for a positive definite covariance operator Γ on $\mathbb{R}^K$ (ie., a symmetric, positive definite $K \times K$ covariance matrix Γ) which we assume to be known. We then define the *uncertainty-to-observation map* of the system by $\mathcal{G} : X \to Y = \mathbb{R}^K$ 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

$$\Phi_\Gamma(u;\delta) = \frac{1}{2}\left((\delta - \mathcal{G}(u))^\top \Gamma^{-1} (\delta - \mathcal{G}(u))\right)\ . \tag{8}$$

In [37] 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 μ^δ of u given data δ is absolutely continuous with respect to the prior measure μ_0. In particular, then, the Radon-Nikodym derivative of the Bayesian posterior w.r. to the prior measure admits a bounded density Θ w.r. to the prior μ_0.

2.3. Uncertainty Parametrization

We parametrize the uncertain datum u in the forward equation (6). 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. To parametrize the uncertain input datum u, we assume in the present work that there exists a countable, unconditional Schauder 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

$$u = u(\boldsymbol{y}) := \langle u\rangle + \sum_{j\in\mathbb{J}} y_j \psi_j \in X \tag{9}$$

with unconditional convergence.

If, for example, X is a separable Hilbert space, then a Riesz basis (such as a biorthogonal wavelet basis) of X could serve as $\{\psi_j\}_{j\in\mathbb{N}}$ in (9), or $\{\psi_j\}_{j\in\mathbb{N}}$ in (9) could be chosen as Schauder basis (in separable Banach spaces X that admit such bases).

We refer to $u - \langle u\rangle$ in (9) as "fluctuation" of u about the nominal value $\langle u\rangle \in X$.

Many choices for the functions ψ_j in (9) are conceivable; among them are standard spline bases, but also Karhunen-Loève eigenfunctions. If the uncertain input datum u is an X-valued random field u in arbitary domains D, and for general covariance kernels, the ψ_j must be obtained numerically, eg. by Fast Multipole Methods together with a Krylov subspace iteration, cp. [36]. With y_j denoting the coordinate variables, the parametrization (9) is deterministic. In order to place (7), (9) into the (probabilistic) Bayesian setting of [37], 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 μ_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 λ^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 (9) becomes a random field, with μ_0 charging the possible realizations of u. As indicated in [12, 35, 32], *holomorphy of uncertainty parametrization* (9) *with respect to the parameter sequence* $\boldsymbol{y}$ can be used to derive sparsity results for this posterior.

2.4. $(\boldsymbol{b}, p, \varepsilon)$-Holomorphy

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 $\mathrm{dist}(\partial\mathcal{E}_r, [-1, 1]) = r - 1$ and that in the limit $r \downarrow 1$, $\mathcal{E}_r$ degenerates to $[-1, 1]$.

Definition 2.2. *Given a sequence $\boldsymbol{b} \in \ell^p(\mathbb{N})$ for some 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 $(\boldsymbol{b}, p, \varepsilon)$-holomorphic if*

$(\boldsymbol{b}, p, \varepsilon) : 1$ *(well-posedness of the forward problem)*
for each $\boldsymbol{y} \in U$, there exists a unique realization $u(\boldsymbol{y}) \in \tilde{X} \subseteq X$ of the uncertainty and a unique solution $q(\boldsymbol{y}) \in \mathcal{X}$ of the forward problem (6). The parametric solution satisfies the a-priori estimate

$$\forall \boldsymbol{y} \in U : \quad \|q(\boldsymbol{y})\|_{\mathcal{X}} \leq C_0(\boldsymbol{y}) \tag{10}$$

where $U \ni \boldsymbol{y} \mapsto C_0(\boldsymbol{y}) \in L^1(U; \mu_0)$; we say that (6) is uniformly well-posed *if in (10) the bound C_0 does not depend on $\boldsymbol{y}$.*

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

$$\sum_{j \in \mathbb{J}} (\rho_j - 1) b_j \leq \varepsilon\ , \tag{11}$$

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

$$\forall \boldsymbol{z} \in \mathcal{E}_\rho : \quad \|q(\boldsymbol{z})\|_{\mathcal{X}} \leq C_\varepsilon(\boldsymbol{\rho}) \tag{12}$$

where $\boldsymbol{y} := \Re(\boldsymbol{z}) \in \bigotimes_{j \in \mathbb{J}} [-\rho_j, \rho_j] \subset \mathbb{R}^{\mathbb{J}}$.

In the case that the sequence $\boldsymbol{b} \in \ell^p(\mathbb{N})$ is clear from the context, we shall call the parametric family simply (p, ε)-holomorphic.

The following result, proved in [10], shows that $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the solution map $\boldsymbol{y} \mapsto q(\boldsymbol{y})$ follows from $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the maps A and F in (2).

Theorem 2.3. *Assume that for $\varepsilon > 0$ and for some $0 < p < 1$, there exists a positive sequence $\boldsymbol{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$, and two constants $0 < r \leq R < \infty$ and a constant $M < \infty$ independent of $u \in U$ such that the following holds:*

(i) For any sequence $\rho := (\rho_j)_{j \geq 1}$ of numbers strictly greater than 1 that satisfies (11) with the given value of ε, the maps $\mathfrak{a}$ and F admit extensions that are holomorphic with respect to every variable on a set of the form $\mathcal{O}_{\boldsymbol{\rho}} := \bigotimes_{j \geq 1} \mathcal{O}_{\rho_j}$, where, for each index $j \in \mathbb{J}$, $\mathcal{O}_{\rho_j} \subset \mathbb{C}$ denotes an open set containing $\tilde{\mathcal{O}}_{\rho_j}$.

(ii) These extensions satisfy for all $\boldsymbol{z} \in \mathcal{O}_\rho$ the uniform continuity conditions

$$\sup_{w \in \mathcal{Y} \setminus \{0\}} \frac{|f(\boldsymbol{z}; w)|}{\|w\|_{\mathcal{Y}}} \leq M, \qquad \sup_{v \in \mathcal{X} \setminus \{0\}, w \in \mathcal{Y} \setminus \{0\}} \frac{|\mathfrak{a}(\boldsymbol{z}; v, w)|}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \leq R, \tag{13}$$

where f denotes the corresponding linear form of F, and the uniform inf-sup conditions: there exists $r > 0$ such that for every $z \in \mathcal{O}_\rho$ there hold the uniform inf-sup conditions

$$\inf_{v\in\mathcal{X}\setminus\{0\}} \sup_{w\in\mathcal{Y}\setminus\{0\}} \frac{|\mathfrak{a}(z;v,w)|}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} \geq r \quad \textit{and} \qquad \inf_{w\in\mathcal{Y}\setminus\{0\}} \sup_{v\in\mathcal{X}\setminus\{0\}} \frac{|\mathfrak{a}(z;v,w)|}{\|v\|_{\mathcal{X}}\|w\|_{\mathcal{Y}}} > r \,. \tag{14}$$

Then, the nonlinear parametric operator $\mathcal{A}(u;q) = A(u;q) - F(u)$ in (2) satisfies the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy assumptions with the same p and ε and with the same sequence $\boldsymbol{b}$.

2.5. 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, which will be considered in the numerical experiments. We emphasize, however, that the setting (2) and the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of parametric solutions shown in Theorem 2.3 extend to a much larger range of problems; further concrete examples are given in [10].

2.5.1. Parabolic Problems with Uncertain Operator The general, parametric operator in the forward equation (6) accomodates parabolic problems with uncertain coefficients, 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 (9) of an elliptic operator. 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

$$V \subseteq H \simeq H^* \subseteq V^* \,. \tag{15}$$

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

$$\forall \boldsymbol{y} \in U \,, \forall v \in V : \quad \mathfrak{b}(\boldsymbol{y}; v, v) + k(v,v) \geq \alpha \|v\|_V^2 \,. \tag{16}$$

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

$$\mathcal{X} = L^2(0,T;V) \cap H^1(0,T;V^*), \;\; \mathcal{Y} = L^2(0,T;V) \times H \,. \tag{17}$$

Then the parametric evolution operator is, formally, given by $A(\boldsymbol{y}) := (\partial_t + B(\boldsymbol{y}), \iota_0)$ where ι_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 $A(\boldsymbol{y})q(\boldsymbol{y}) = f$ is given, for $v \in \mathcal{X}$ and for $w = (w_1, w_2) \in \mathcal{Y}$, by the bilinear form

$$\begin{aligned} \mathfrak{a}(\boldsymbol{y}; v, w) &:= \int_0^T \left({}_V\langle w_1, \partial_t v\rangle_{V^*} + {}_V\langle w_1, B(\boldsymbol{y})v\rangle_{V^*} \right) dt + {}_H\langle v(0), w_2\rangle_H \\ &= \int_0^T \left({}_V\langle w_1, \partial_t v\rangle_{V^*} + \mathfrak{b}(\boldsymbol{y}; v, w_1) \right) dt + {}_H\langle v(0), w_2\rangle_H \end{aligned} \tag{18}$$

and by the linear form

$$\mathfrak{l}(w) = \int_0^T \left({}_V\langle w_1(\cdot, t), f(\cdot, t)\rangle_{V^*} \right) dt + {}_H\langle w_2, u_0\rangle_H \ . \tag{19}$$

Note that then in the weak formulation

$$\forall \boldsymbol{y} \in U : \quad q(\boldsymbol{y}) \in \mathcal{X} : \mathfrak{a}(\boldsymbol{y}; q(\boldsymbol{y}), w) = \mathfrak{l}(w) \quad \forall w \in \mathcal{Y} \tag{20}$$

the initial condition $\iota_0 u = u_0$ has been imposed weakly. For the variational space-time formulation (20) it is once more known (see, eg., [34, Appendix]) that the parametric bilinear form $\mathfrak{a}(\boldsymbol{y}; \cdot, \cdot)$ satisfies the (uniform w.r. to $\boldsymbol{y} \in U$) inf-sup conditions (14), provided that the parametric spatial operator $B(\boldsymbol{y})$ in (18) satisfies the Garding inequality (16).

2.5.2. Elliptic Multiscale Problems with Uncertainty In [19], 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 ε, 1-periodic w.r. to ξ and depends on $\boldsymbol{y}$ once more in an affine fashion (see [19, 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 ε due to the use of two-scale convergence and the avoidance of homogenization formulas.*

2.6. Parametric Bayesian Posterior

Motivated by [35, 32], 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 μ_0. The prior measure μ_0 being uniform, we admit in (9) 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

$$(U, \mathcal{B}, \mu_0) \ . \tag{21}$$

We assume throughout what follows that the prior measure on the uncertain input data, parametrized in the form (9), 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 [28], can be treated by the techniques developed here, at the expense of additional technicalities. With the parameter domain U as in (21) the parametric forward map $\Xi : U \to \mathbb{R}^K$ is given by

$$\Xi(\boldsymbol{y}) = \mathcal{G}(u)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j} . \tag{22}$$

The mathematical foundation of Bayesian inversion is Bayes' theorem. We present a version of it, from [35] (see also [37]). 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 2.4. *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 δ, is absolutely continuous with respect to $\mu_0(d\boldsymbol{y})$, ie.*

$$\frac{d\mu^\delta}{d\mu_0}(\boldsymbol{y}) = \frac{1}{Z}\Theta(\boldsymbol{y}) \tag{23}$$

with the parametric Bayesian posterior $\Theta(\boldsymbol{y})$ given by

$$\Theta(\boldsymbol{y}) = \exp\big(-\Phi_\Gamma(u;\delta)\big)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j}, \tag{24}$$

where the Bayesian potential Φ_Γ is as in (8) and the normalization constant Z is given by

$$Z_\Gamma = \int_U \Theta(\boldsymbol{y})d\mu_0(\boldsymbol{y}) . \tag{25}$$

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) observational data δ of the QoI ϕ. 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 δ) mean and covariance of the system's response.

With the QoI ϕ we associate the (infinite-dimensional) parametric map

$$\begin{aligned}\Psi(\boldsymbol{y}) &= \Theta(\boldsymbol{y})\phi(u)\ |_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j} \\ &= \exp\big(-\Phi_\Gamma(u;\delta)\big)\phi(u)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j} : U \to \mathcal{S} .\end{aligned} \tag{26}$$

Then the Bayesian estimate of the QoI ϕ, given noisy data δ, takes the form

$$\begin{aligned}\mathbb{E}^{\mu^\delta}[\phi] &= \tfrac{Z'_\Gamma}{Z_\Gamma} = \tfrac{1}{Z_\Gamma}\textstyle\int_{y\in U}\Psi(\boldsymbol{y})\mu_0(d\boldsymbol{y}), \\ Z'_\Gamma &= \int_{y\in U}\exp\big(-\Phi_\Gamma(u;\delta)\big)\phi(u)\Big|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} y_j\psi_j}\mu_0(d\boldsymbol{y}) .\end{aligned} \tag{27}$$

Our aim is to approximate the expectations Z'_Γ and Z_Γ 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 [35], 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 [35, 32] for scalar, isotropic diffusion problems, *dimension-independent convergence rates* of numerical approximations of integrals like (25), (27) are based on *sparsity results* for the posterior density Θ which arises in Bayes' theorem. In the present section, we establish such sparsity results in the general setting of the operator equation (9). As in [12, 16], the sparsity results will be based on analytic dependence of the forward solution $q(\boldsymbol{y})$ of the parametric operator equation (2), 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 $(\boldsymbol{b}, p, \varepsilon)$-holomorphy established in Theorem 2.3, we approximate the parametric solution $q(\boldsymbol{y})$ by partial sums of tensorized Legendre series. As was shown in [11, 12, 8, 16], $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the forward solution $q(\boldsymbol{y})$ implies that such expansions are *sparse*. Sparsity of tensorized Taylor expansions requires $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the forward map on the *polydiscs* $\mathcal{U}_\rho$ (as in [16]), whereas $(\boldsymbol{b}, p, \varepsilon)$-holomorphy 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, 16, 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

$$\int_{-1}^{1} (L_k(t))^2 \frac{dt}{2} = 1, \quad k = 0, 1, 2, ... \tag{28}$$

Since $L_0 \equiv 1$, the Legendre polynomials L_k in (28) can be tensorized on the parameter domains U via

$$L_\nu(\boldsymbol{z}) = \prod_{j \in \mathbb{J}} L_{\nu_j}(z_j), \quad \boldsymbol{z} \in \mathbb{C}^{\mathbb{J}},\ \nu \in \mathcal{F}\ . \tag{29}$$

Here, $\mathcal{F}$ denotes the countable set $\mathbb{N}_0^{\mathbb{J}}$ of sequences of multi-indices $\boldsymbol{\nu} = (\nu_j)_{j \in \mathbb{J}}$ which are summable: $\sum_{j \in \mathbb{J}} \nu_j < \infty$. The set of tensorized Legendre polynomials

$$\mathbb{L}(U) = \{L_\nu : \nu \in \mathcal{F}\} \tag{30}$$

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

$$q(\boldsymbol{y}) = \sum_{\nu \in \mathcal{F}} q_\nu L_\nu(\boldsymbol{y})\ , \quad \boldsymbol{y} \in U\ . \tag{31}$$

For the statement of sparsity in the response map, we shall approximate the parametric solution $q(\boldsymbol{y})$ of (2) and also the Bayesian posterior density in terms of N-term truncations of the Legendre series (31).

Truncations of tensorized 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 Theorem 2.3, the parametric forward solution $q(\boldsymbol{y})$ of the forward equation (6) is $(\boldsymbol{b}, p, \varepsilon)$-holomorphic 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 the same exponent p as in the sparsity assumption (11) 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, $(\boldsymbol{b}, p, \varepsilon)$-holomorphic, parametric mappings, it was shown in [8] that partial sums of (31) 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 (or lower) 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 $\Lambda_N \subset \mathcal{F}$, that

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

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 on $\mathcal{X}$ since $\|a\|_{\ell^p_m(\mathcal{F};\mathcal{X})} = \|\boldsymbol{a}^*\|_{\ell^p(\mathcal{F})}$. Sets Λ_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 $\mu \in \Lambda$.

§ Also referred to in the literature as *lower index sets*

3.2. Sparse Legendre Expansions

We recall for $\nu \in \mathcal{F}$ the definition (29) of the tensorized Legendre polynomials $L_\nu(\boldsymbol{y})$; the normalization (28) differs slightly from the classical one, where

$$P_k(1) = 1,\ \|P_k\|_{L^\infty(-1,1)} = 1\ , \quad k = 0, 1, \ldots \tag{32}$$

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_ν, we may expand every $q(\boldsymbol{y}) \in L^2(U,\mu_0)$

$$q(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} q_\nu^L L_\nu(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} q_\nu^P P_\nu(\boldsymbol{y}) \tag{33}$$

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_ν^L and q_ν^P in (33) are related by

$$q_\nu^P = \left(\prod_{j\in\mathbb{J}} (1+2\nu_j) \right)^{1/2} q_\nu^L\ , \quad \nu \in \mathcal{F}\ . \tag{34}$$

Lemma 3.3. *If the parametric forward map $q(\boldsymbol{y})$ is $(\boldsymbol{b}, p, \varepsilon)$-holomorphic in a poly-ellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$, then for every $\nu \in \mathcal{F}$ there holds, for every ρ as in* (11) *in Definition 2.2, the estimate*

$$\|q_\nu^P\|_{\mathcal{X}} \leq C_\epsilon \prod_{j\in\mathbb{J},\nu_j\neq 0} (2\nu_j+1) \frac{\pi\rho_j}{2(\rho_j-1)} \rho_j^{-\nu_j}\ . \tag{35}$$

The proof can be found in [10, Lemma 3.1].

Due to $\|q_\nu^L\|_{\mathcal{X}} \leq \|q_\nu^P\|_{\mathcal{X}}$, the summability of the sequence $(\|q_\nu^P\|_{\mathcal{X}})_{\nu\in\mathcal{F}}$ directly implies the summability of $(\|q_\nu^L\|_{\mathcal{X}})_{\nu\in\mathcal{F}}$. The next result, from [12, 16] specifies the type of convergence in the Legendre expansions (33), 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, 16]. 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 $(\boldsymbol{b}, p, \varepsilon)$-holomorphic extension to the poly-ellipse $\mathcal{E}_\rho \subset \mathbb{C}^{\mathbb{J}}$ with ρ satisfying condition* (11) *in Definition 2.2. Then the following holds.*

(i) the Legendre series (33) *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 (11) 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

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

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 (33), there holds the error bound

$$\left\| q(\cdot) - \sum_{\nu\in\Lambda_N^L} q_\nu^L L_\nu(\cdot) \right\|_{L^2(U,\mu_0;\mathcal{X})} \leq C(p,\boldsymbol{q}) N^{-(1/p-1/2)}\,. \tag{36}$$

(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 (33), there holds the error bound

$$\sup_{\boldsymbol{y}\in U} \left\| q(\boldsymbol{y}) - \sum_{\nu\in\Lambda_N^P} q_\nu^P P_\nu(\boldsymbol{y}) \right\|_{\mathcal{X}} \leq C(p,\boldsymbol{v}) N^{-(1/p-1)}\,. \tag{37}$$

4. Sparsity of the Posterior Density Θ

For operator equations (6) with operators $A(u;q)$ with parametric uncertainty which produce parametric solutions which are $(\boldsymbol{b},p,\varepsilon)$-holomorphic 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 (23), (24).

4.1. $(\boldsymbol{b},p,\varepsilon)$-Holomorphy of Θ

Our verification of $(\boldsymbol{b},p,\varepsilon)$-holomorphy of the posterior density Θ will be based on verifying $(\boldsymbol{b},p,\varepsilon)$-holomorphy for the parametric posterior density $\Theta(\boldsymbol{y})$ defined in (23), (24). Once this is established, sparsity and N-term approximation results for Θ will follow similarly as for the parametric solution $q(\boldsymbol{y})$ of (6). As in [32], we then infer convergence rates for Smolyak quadratures from N-term approximation rates for truncated tensorized Legendre approximation rates for the posterior density Θ.

Theorem 4.1. *Consider the Bayesian inversion of the parametric operator equation (6) 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 $(\boldsymbol{b},p,\varepsilon)$-holomorphic for some positive sequence $\boldsymbol{b} = (b_j)_{j\geq 1} \in \ell^p(\mathbb{N})$, $0 < p < 1$, and some $\delta > 0$, with respect to poly-ellipses $\mathcal{E}_{\boldsymbol{\rho}}$. Then the Bayesian posterior density $\Theta(\boldsymbol{y})$ is, as a function of the parameter $\boldsymbol{y}$, likewise $(\boldsymbol{b},p,\varepsilon)$-holomorphic,* with the same $\boldsymbol{b}$, p and the same *δ.*

The modulus of the holomorphic extension of the Bayesian posterior $\Theta(\boldsymbol{y})$ over the poly-ellipse $\mathcal{E}_\rho$ 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 (7). Here, the constants $b, C > 0$ in the bound of the modulus $\sup_{z\in\partial\mathcal{E}_\rho}|\Theta(z)|$ depend on the condition number of the uncertainty-to-observation map $\mathcal{G}(\cdot) = (\mathcal{O}\circ G)(\cdot)$ but are independent of Γ in (7).

Proof. By Theorem 2.3, the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the operator implies that the parametric forward solution map $q(\boldsymbol{y})$ admits a holomorphic extension, denoted $q(\boldsymbol{z})$, to any poly-ellipses $\boldsymbol{z} \in \mathcal{E}_\rho$ whose poly-radius ρ satisfies (11) with the sequence $\boldsymbol{b} \in \ell^p(\mathbb{J})$.

We consider first the case when there is only a single parameter, ie. the case 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 (8) is, in this case, given by (assuming that the data δ, $\langle u\rangle$ and ψ 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 $q(\boldsymbol{z}) \in \mathcal{X}$, the response function $z \mapsto \mathcal{G}(\langle u\rangle + z\psi)$ is holomorphic in $\mathcal{E}_{\rho_1}$. Therefore, the complex extension of Φ_Γ, ie.

$$\mathcal{E}_{\rho_1} \ni z \mapsto \Phi_\Gamma(u(\boldsymbol{z});\delta) := \frac{1}{2}\left(\delta - \mathcal{G}(\langle u\rangle + z\psi)\right)^\top \Gamma^{-1}\left(\delta - \mathcal{G}(\langle u\rangle + z\psi)\right) \tag{38}$$

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

$$\Theta(\boldsymbol{z}) = \exp\left(-\Phi_\Gamma(u;\delta)|_{u=\langle u\rangle+\sum_{j\in\mathbb{J}} z_j\psi_j}\right) . \tag{39}$$

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 (39) 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 (39) 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 (39) as a function of the parameters $\boldsymbol{z}$ over the poly-ellipses $\mathcal{E}_\rho$ of holomorphy, with the semiaxes ρ as in (11). 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

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

where $\|\circ\|_\Gamma$ denotes the covariance-weighted Euclidean norm in $\mathbb{C}^K$. Based on the definition (38), and on the fact that $\delta\in\mathbb{R}^K$, and the definition (7) of $\mathcal{G}(\cdot)$, we find

$$\forall z\in\mathcal{E}_\rho: \quad |\Theta(\boldsymbol{z})| \leq \exp\left(\frac{1}{2}\mathrm{Im}\left(\mathcal{G}(u(\boldsymbol{z}))\right)^\top \Gamma^{-1}\mathrm{Im}\left(\mathcal{G}(u(\boldsymbol{z}))\right)\right)\,. \tag{40}$$

Since the map $\mathcal{E}_\rho \ni z \mapsto \mathcal{G}(u(\boldsymbol{z}))$ does not depend on the observation noise covariance Γ, a bound for the modulus $\sup_{z\in\mathcal{E}_\rho}|\Theta(\boldsymbol{z})|$ which is explicit in terms of Γ can be inferred from (40). 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. [35]. 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 Θ. In the following, we analyze the convergence rate of *N-term truncated Legendre gpc-approximations of* Θ 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.

4.2. Sparse Legendre Expansions of Θ

Since we assumed that the prior measure μ_0 is built by tensorization of the uniform probability measures $\frac{1}{2}\lambda^1$ on $[-1,1]$, the normalization (28) implies that the polynomials $L_\nu(\boldsymbol{z})$ in (29) 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_ν in (29) 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 Θ and Ψ. Such expansions can also serve as a basis for sampling of these quantities with draws that are equidistributed with respect to the prior μ_0. In particular, the density $\Theta: U\to\mathbb{R}$ is square integrable with respect to the prior μ_0 over U, i.e. $\Theta\in L^2(U,\mu_0)$. Moreover, if the QoI $\phi(\cdot): U\to\mathcal{S}$ in (26) is bounded, then

$$\int_U \|\Psi(\boldsymbol{y})\|_\mathcal{S}^2 d\mu_0(\boldsymbol{y}) < \infty, \tag{41}$$

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 $\|\Psi(\boldsymbol{y})\|_V \leq CM/r$ for all $\boldsymbol{y}\in U$, where the constant C is independent of the data δ. 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 $\mathbb{L}(U)$ in (30) 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 $\Theta(\boldsymbol{y})$.

$$\Theta(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} \theta_\nu^L L_\nu(\boldsymbol{y}) = \sum_{\nu\in\mathcal{F}} \theta_\nu^P P_\nu(\boldsymbol{y}) \quad \text{in} \quad L^2(U,\mu_0) \tag{42}$$

where the gpc expansion coefficients θ_ν^L and θ_ν^P are defined by (cf. also (34))

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

By Parseval's equation and the normalization (28), it follows immediately from (42) and (41) that the second moment of the posterior density with respect to the prior is finite and can be expressed as

$$\|\Theta\|^2_{L^2(U,\mu_0)} = \sum_{\nu\in\mathcal{F}} |\theta_\nu^L|^2 = \|\boldsymbol{\theta}^L\|^2_{\ell^2(\mathcal{F})} \;. \tag{43}$$

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

For every $N\in\mathbb{N}$, denote by $\Lambda_N^L \subset \mathcal{F}$ a set of indices $\nu\in\mathcal{F}$ corresponding to N largest θ_ν^* of the monotone majorant $\boldsymbol{\theta}^L$ of the Legendre coefficient sequence $(\theta_\nu^L)_{\nu\in\mathcal{F}}$ in (42), and denote by

$$\Theta^L_{\Lambda_N}(\boldsymbol{y}) := \sum_{\nu\in\Lambda_N^L} \theta_\nu^L L_\nu(\boldsymbol{y}) \tag{44}$$

the corresponding N-term truncated Legendre expansion (42) of the posterior. Then, with $0 < p < 1$ in the $(\boldsymbol{b},p,\varepsilon)$-holomorphy 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 is such that there exists a constant $C>0$ such that for all $N\in\mathbb{N}$ holds

$$\|\Theta(\boldsymbol{y}) - \Theta^L_{\Lambda_N}(\boldsymbol{y})\|_{L^2(U,\mu_0)} \leq CN^{-s}\|\theta^L\|_{\ell^p_m(\mathcal{F})}, \;\; s := \frac{1}{p} - \frac{1}{2} \;. \tag{45}$$

Likewise, denoting by $\Lambda_N^P \subset \mathcal{F}$ 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 $(\theta_\nu^P)_{\nu\in\mathcal{F}}$ in (42), and denote by

$$\Theta^P_{\Lambda_N}(\boldsymbol{y}) := \sum_{\nu\in\Lambda_N^P} \theta_\nu^P L_\nu(\boldsymbol{y}) \tag{46}$$

the corresponding N-term truncated Legendre expansion (42) of the posterior. Then, with $0 < p < 1$ in the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the parametric forward solution, there hold the N-term approximation results

$$\|\Theta(\boldsymbol{y}) - \Theta^{P}_{\Lambda_N}(\boldsymbol{y})\|_{L^\infty(U,\mu_0)} \leq CN^{-s}\|\theta^P\|_{\ell^p_m(\mathcal{F})}, \quad s := \frac{1}{p} - 1\,. \tag{47}$$

In (47) and (45), the constant $C \geq 1$ depends on s and on the covariance $\Gamma > 0$ in the additive observation noise η in (7), but is independent of N. We refer to [33] for detailed investigation of the scaling 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 [32], we introduce the tensorized multivariate operators

$$\mathcal{Q}_\nu = \bigotimes_{j\geq 1} Q^{\nu_j}\,, \qquad \Delta_\nu = \bigotimes_{j\geq 1} \Delta_{\nu_j}\,. \tag{48}$$

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 $y = (t, y_2, \ldots)\,, i = 1$, see [9, 32].

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

For the proof, we refer to [32, Theorem 4.2].

We will now establish convergence rates for the approximation of the expectation of QoI with respect to the posterior, given data δ, based on the (p,ε)-holomorphy 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

$$C_{\mathcal{Q}_{\Lambda_N}} \le \sum_{\nu\in\Lambda_N}\prod_{j\ge 1}(c_{\nu_j} + c_{\nu_j - 1}) \le \#\Lambda^{\log_2 3} \tag{49}$$

with $c_k = 1\,, k \ge 0$ (note that $\|Q^k\| = 1$ by Assumption 5.1 (ii)) and $c_{-1} := 0$, see [32, Lemma 4.4]. In the tensor product case where, for $\boldsymbol{\nu} \in \mathcal{F}$, the constant $C_{\mathcal{R}_\nu}$ is given by

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

so that bound (49) is pessimistic in this case.

The quadrature error for the normalization constant (25) and the quantity Z' (27) can be bounded by relating the error with the Legendre coefficients θ_ν^P of $\Theta = \sum_{\nu\in\mathcal{F}} \theta_\nu^P P_\nu(\boldsymbol{y})$ and ψ_ν^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*

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

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 [32, Lemma 4.5].

Theorem 5.4. *If the forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is $(\boldsymbol{b}, p, \varepsilon)$-holomorphic 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 Λ_N^θ, Λ_N^ψ 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 $s = 1/p - 1$,*

$$|I[\Theta] - \mathcal{Q}_{\Lambda_N^\theta}[\Theta]| \le C^1 N^{-s}\,, \tag{50}$$

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

$$\|I[\Psi] - \mathcal{Q}_{\Lambda_N^\psi}[\Psi]\|_{\mathcal{S}} \le C^2 N^{-s} \quad . \tag{51}$$

Proof. The proof proceeds in two steps: first, we will construct a sequence $\boldsymbol{\rho}$ of poly-radii in the sense of (11) based on the estimate of the Legendre coefficients in Lemma 3.3. Then, we use the resulting estimate to prove $(\gamma_\nu|\theta_\nu^P|)_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$ and $(\gamma_\nu\|\psi_\nu^P\|_{\mathcal{S}})_{\nu\in\mathcal{F}} \in \ell^p(\mathcal{F})$, 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 [10, 9, 12, 32], and present the details. Due to the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy of the forward solution map, the parametric posterior density $\Theta(\boldsymbol{y})$ defined in (24), being a composition of the holomorphic parametric forward solution map with an exponential function and a (quadratic) polynomial, admits likewise an extension to the complex domain. It remains to verify the quantitative bounds on the size of domains of analytic continuation. Let $B > 0$ a fixed constant and $J_0 \ge 1$ be an integer such that

$$\sum_{j>J_0} b_j \le \frac{\varepsilon}{4B}\,,$$

and define the constant

$$\kappa = 1 + \frac{\epsilon}{4\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})}} > 1 \,.$$

We set $E := \{j : 1 \le J_0\}$ and $F := \mathbb{N} \setminus E$ and denote for each $\nu \in \mathcal{F}$ by ν_E and ν_F the restrictions of ν to E and F. Then, we define the sequence $\boldsymbol{\rho} = \boldsymbol{\rho}(\nu)$ by

$$\rho_j = \kappa \,, \quad j \in E \,; \qquad \rho_j = \kappa + B + \frac{\varepsilon \nu_j}{2b_j(1+|\nu_F|)} \,, \quad j \in F \,,$$

with $|\nu_F| = \sum_{j>J_0} \nu_j$ (with the convention $\frac{\nu_j}{|\nu_F|} = 0$, if $|\nu_F| = 0$). The sequence $\boldsymbol{\rho}$ depends on the multi-index $\boldsymbol{\nu} \in \mathcal{F}$ and satisfies

$$\sum_{j \ge 1} (\rho_j - 1) b_j = (\kappa - 1) \sum_{j \le J_0} b_j + \sum_{j > J_0} \left(\kappa + B + \frac{\varepsilon \nu_j}{2b_j(1+|\nu_F|)} - 1 \right) b_j \le \varepsilon \,.$$

Using the estimate of the Legendre coefficients in Lemma 3.3 and the fact that $\frac{\pi\rho_j}{2(\rho_j-1)} \ge \frac{\pi\rho_i}{2(\rho_i-1)}, \forall j < i$, we have

$$\gamma_\nu |\theta_\nu^P| \le C_\varepsilon \left(\prod_{j \le J_0, \nu_j \ne 0} C_\kappa (1+\nu_j)^2 (2\nu_j+1) \kappa^{-\nu_j} \right) \cdot \left(\prod_{j > J_0, \nu_j \ne 0} C_\kappa (1+\nu_j)^2 (2\nu_j+1) \rho^{-\nu_j} \right) ,$$

with $C_\kappa := \frac{\pi\kappa}{2(\kappa-1)}$. Similar to the proof of Theorem 2.2 in [10], it follows

$$\gamma_\nu |\theta_\nu^P| \le C_\epsilon \cdot \alpha_E(\nu) \cdot \beta_F(\nu) \,,$$

where $\alpha_E(\nu) := \prod_{j \le J_0, \nu_j \ne 0} C_\kappa C_{2,\kappa} \kappa^{-\nu_j/2}$, $\beta_F(\nu) := \prod_{j > J_0, \nu_j \ne 0} (mC_\kappa)^{\nu_j} \rho_j^{-\nu_j}$ using $(1+\nu_j)^2(2\nu_j+1)\kappa^{-\nu_j} \le C_{2,\kappa}\kappa^{-\nu_j/2}$ for some constant $C_{2,\kappa} > 0$ and $C_\kappa(1+\nu_j)^2(2\nu_j + 1) \le (mC_\kappa)^{\nu_j}$ for some $m > 1$.

The p-summability of $(\gamma_\nu |\theta_\nu^P|)_{\nu \in \mathcal{F}}$ follows then with the same arguments as in the proof of [10, Theorem 2.2]. If the constant B is chosen large enough, monotone envelope can be constructed following the lines of [10, Theorem 2.2], so that it holds $(\gamma_\nu |\theta_\nu^P|)_{\nu \in \mathcal{F}} \in \ell_m^p(\mathcal{F})$.

Exactly the same analysis allows to bound the quadrature error of the density Ψ, i.e. of $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{S}})_{\nu \in \mathcal{F}} \in \ell_m^p(\mathcal{F})$. □

Remark 5.5. Theorem 5.4 ensures the existence of two monotone sequences of index sets, such that the quantities Z'_Γ and Z_Γ can be approximated with convergence rate $1/p - 1$. In particular, the result implies that there exists *one common set* Λ_N for the approximation of both quantities Z'_Γ and Z_Γ , which can be constructed by the union of Λ_N^1 and Λ_N^2 defined by (50) and (51), respectively.

Based on the approximation result presented in Theorem 5.4, the error of the expectation $\mathbb{E}^{\mu^\delta}[\phi]$ conditioned on given the observational data δ can be bounded in the following way:

Lemma 5.6. *Assume that the forward solution map $U \ni \boldsymbol{y} \mapsto q(\boldsymbol{y})$ is $(\boldsymbol{b}, p, \varepsilon)$-holomorphic for some $0 < p < 1$ and $\varepsilon > 0$. Then, there exists a constant $C_{\delta,\Gamma} > 0$ which is independent of N (but which depends, in general, exponentially on $\Gamma > 0$, cp. Theorem 4.1 ahead) such that there holds the error bound for $s = 1/p - 1$*

$$\Big\|\mathbb{E}^{\mu^\delta}[\phi] - \frac{\mathcal{Q}_{\Lambda_N^\psi}[\Psi]}{\mathcal{Q}_{\Lambda_N^\theta}[\Theta]}\Big\|_{\mathcal{S}} \leq C_{\delta,\Gamma} N^{-s} \tag{52}$$

for sufficiently large N.

Proof. The error of the expected value $\mathbb{E}^{\mu^\delta}[\phi]$ of the QoI ϕ, conditioned on given observational data δ can be written as

$$\begin{aligned}\Big\|\mathbb{E}^{\mu^\delta}[\phi] - \frac{\mathcal{Q}_{\Lambda_N^\psi}[\Psi]}{\mathcal{Q}_{\Lambda_N^\theta}[\Theta]}\Big\|_{\mathcal{S}} &= \Big\|\frac{Z'}{Z} - \frac{\mathcal{Q}_{\Lambda_N^\psi}[\Psi]}{\mathcal{Q}_{\Lambda_N^\theta}[\Theta]}\Big\|_{\mathcal{S}} \\ &= \Big\|\frac{1}{Z}(Z' - \mathcal{Q}_{\Lambda_N^\psi}[\Psi]) + \frac{1}{Z}\frac{\mathcal{Q}_{\Lambda_N^\psi}[\Psi]}{\mathcal{Q}_{\Lambda_N^\theta}[\Theta]}(Z - \mathcal{Q}_{\Lambda_N^\theta}[\Theta])\Big\|_{\mathcal{S}}.\end{aligned}$$

Using the crude estimates $Z \geq \exp(-\frac{1}{2}\|\Gamma^{-1}\|(|\delta| + \sup_{u\in\tilde{X}}|\mathcal{G}(u)|)^2)$ and $Z' \leq \exp(\frac{1}{2}\|\Gamma^{-1}\|(|\delta| + \sup_{u\in\tilde{X}}|\mathcal{G}(u)|)^2)\sup_{u\in\tilde{X}}\|\phi(u)\|_{\mathcal{S}}$, in particular, if $\phi(u) = G(u)$, $\sup_{u\in\tilde{X}}\|\phi(u)\|_{\mathcal{S}} \leq C(F, \tilde{X})$, where $C(F, \tilde{X})$ defined by (5), we have

$$\left|\frac{1}{Z}\right| \leq C_{\delta,\Gamma}$$

with $C_{\delta,\Gamma} = 1/\exp(-\frac{1}{2}\|\Gamma^{-1}\|(|\delta| + C(F, \tilde{X})\big(\sum_{k=1}^K \|o_k\|_{\mathcal{X}'}\big))^2)$ and for N large enough

$$\left\|\frac{\mathcal{Q}_{\Lambda_N^\psi}[\Psi]}{\mathcal{Q}_{\Lambda_N^\theta}[\Theta]}\right\|_{\mathcal{S}} \leq C_{\delta,\Gamma,\phi}$$

due to the approximation results (50) and (51). Combining these estimates with (50) and (51) completes the proof. □

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 Λ_N corresponding to the N largest contributions of the sparse quadrature operator to the approximation of the integral Z_Γ and Z'_Γ, i.e. to 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 (50) and (51), respectively (cf. [32, 9, 17, 15]).

Following [15, 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 Λ in the Smolyak quadrature is summarized. We refer to [32, 9, 8, 17, 15] 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 ν 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 $k \geq 1$,

- $\mathfrak{R}$-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
  $z_j^k = \mathfrak{R}(\hat{z}),$ with $\hat{z} = \text{argmax}_{|z| \leq 1} \prod_{l=1}^{j-1} |z - z_l^k|\,, j = 3, \ldots, n_k, \text{if } j \text{ odd}\,,$
  $z_j^k = -z_{j-1}^k\,, j = 3, \ldots, n_k, \text{if } j \text{ even}\,,$ 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. [32, 5, 6, 7]). The following result is shown as in [32, Lemma 4.10].

Proposition 5.7. *Let $\mathcal{Q}_\Lambda^{RL}$ denote the sparse quadrature operator for any monotone set Λ 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* $(\boldsymbol{b}, p, \varepsilon)$*-holomorphic for some* $0 < p < 1$ *and* $\varepsilon > 0$*, then* $(\gamma_\nu |\theta_\nu^P|)_{\nu \in \mathcal{F}} \in \ell_m^p(\mathcal{F})$ *and* $(\gamma_\nu \|\psi_\nu^P\|_{\mathcal{S}})_{\nu \in \mathcal{F}} \in \ell_m^p(\mathcal{F})$*. 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

$$\begin{aligned} &\partial_t q(t,x) - \operatorname{div}(u(x) \nabla q(t,x)) = f(t,x) \quad (t,x) \in T \times D\,, \\ &q(0,x) = 0 \quad x \in D\,, \\ &q(t,0) = q(t,1) = 0 \quad t \in T\,, \end{aligned} \tag{53}$$

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

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

with $D_j = [(j-1)\frac{1}{64}, j\frac{1}{64})$, $\boldsymbol{y} = (y_j)_{j=1,\dots,64}$, $X = \cup_{j=1}^{64} C^0(\overline{D_j})$ and $\alpha_j = \frac{0.9}{j^\zeta}, \zeta = 2, 3, 4$.

For a given realization of $u(x)$, the forward problem (53) 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 δ, the goal of computation is the conditioned expectation $\mathbb{E}^{\mu^\delta}[\phi]$ of the QoI $\phi(u) = G(u)$ given by

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

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

$$Z_\Gamma = \int_U \exp\big(-\Phi(u;\delta)\big)\Big|_{u = \langle u \rangle + \sum_{j=1}^{64} y_j \psi_j} d\mu_0(\boldsymbol{y}).$$

The noisy observational data is computed as a single realization of

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

with $\eta \sim \mathcal{N}(0, \Gamma)$ and $\mathcal{G} : X \to Y = \mathbb{R}^K$, with $K = 1, 3, 9$. The noise $\eta = (\eta_j)_{j=1,\dots,K}$ 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 δ is then computed according to (7) by the sum of the observed solution and a realization of the additive noise η.

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 Λ 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_Γ with respect to the cardinality of the index set Λ based on the sequence CC.

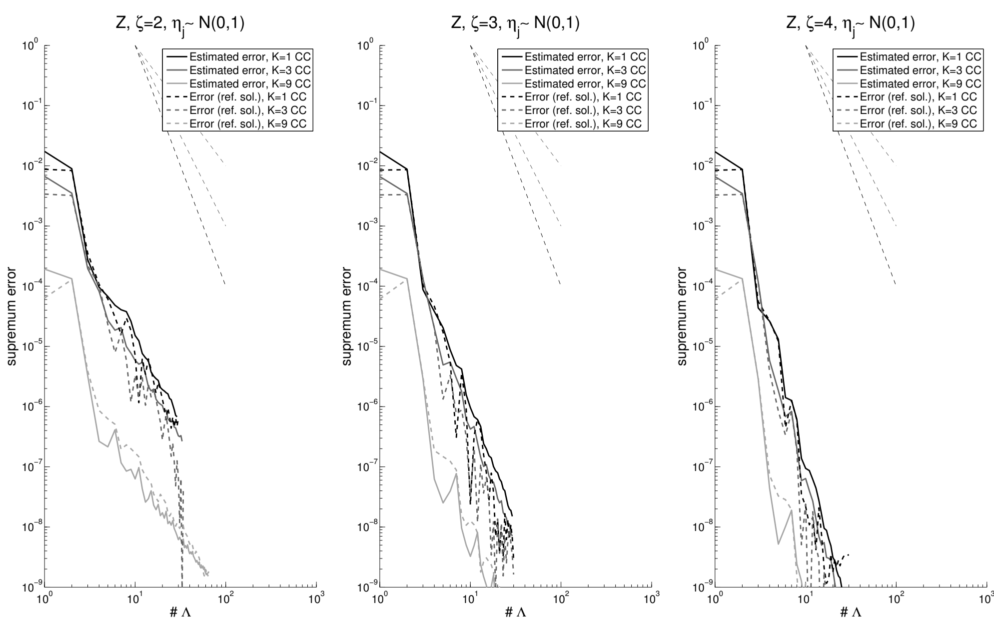

Figure 1. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta =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

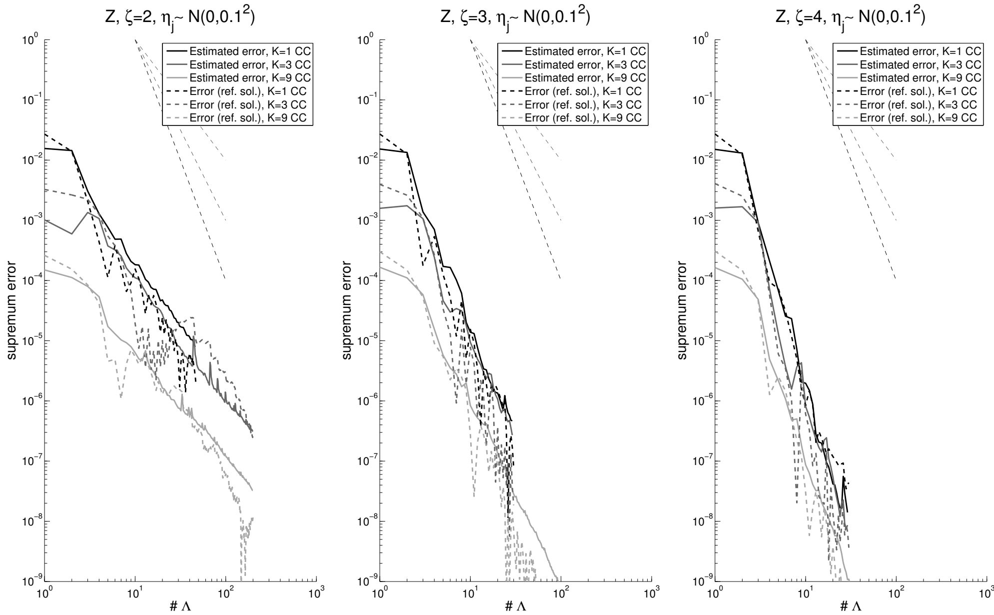

Figure 2. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (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_Γ based on the sequence RL are displayed in Figure 3 and 4.

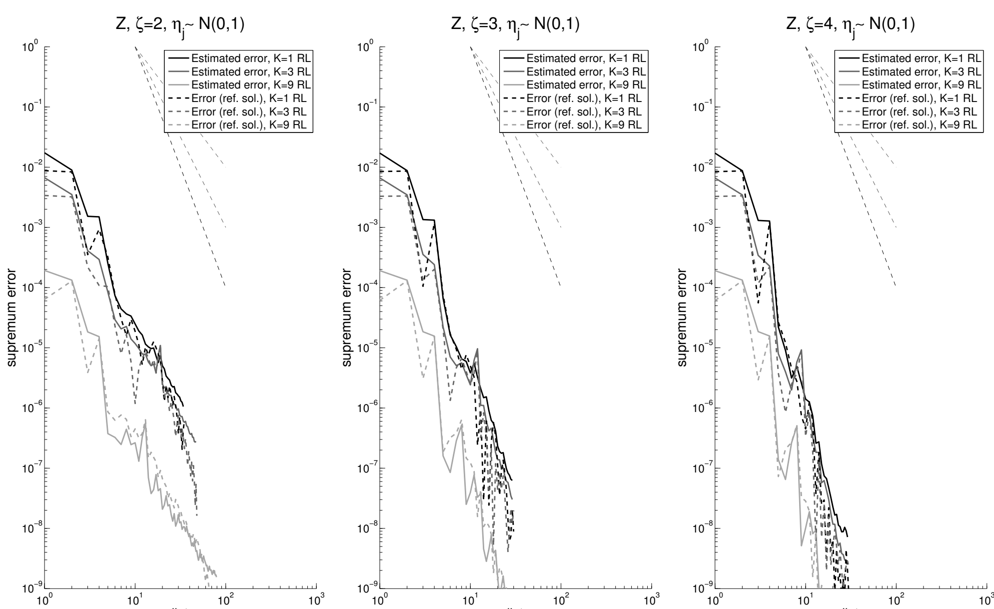

Figure 3. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta =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

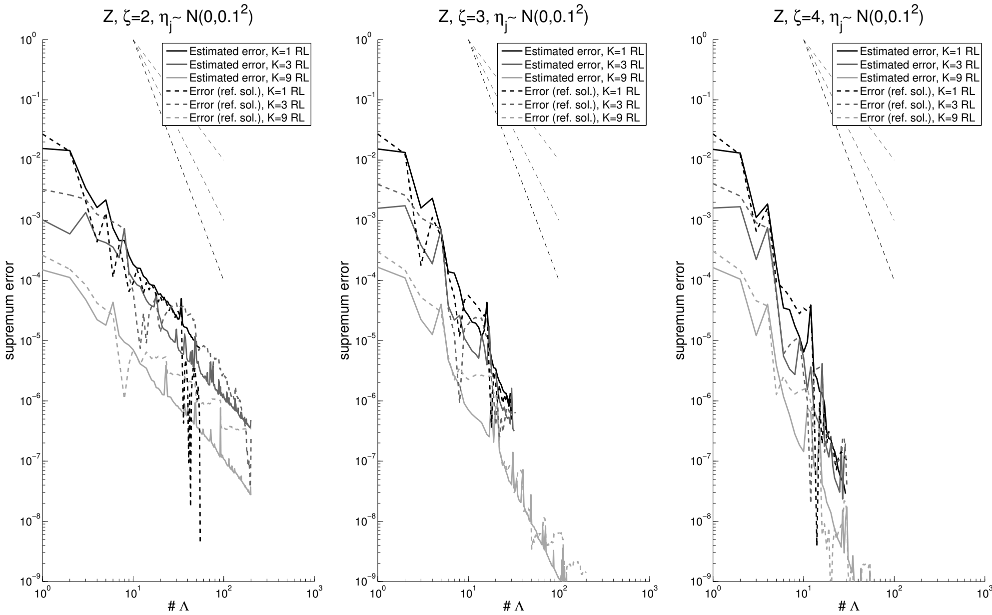

Figure 4. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (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 ζ 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 Λ than the Leja points. This behaviour could be already observed in the elliptic test case, cp. [32]. It can be attributed to the exponential growth of the number of quadrature points within the hierarchy of CC sequences. As Figure 5 and Figure 6 exemplarily show, this offset disappears in the error curves of the normalization constant with respect to the number of PDE solves.

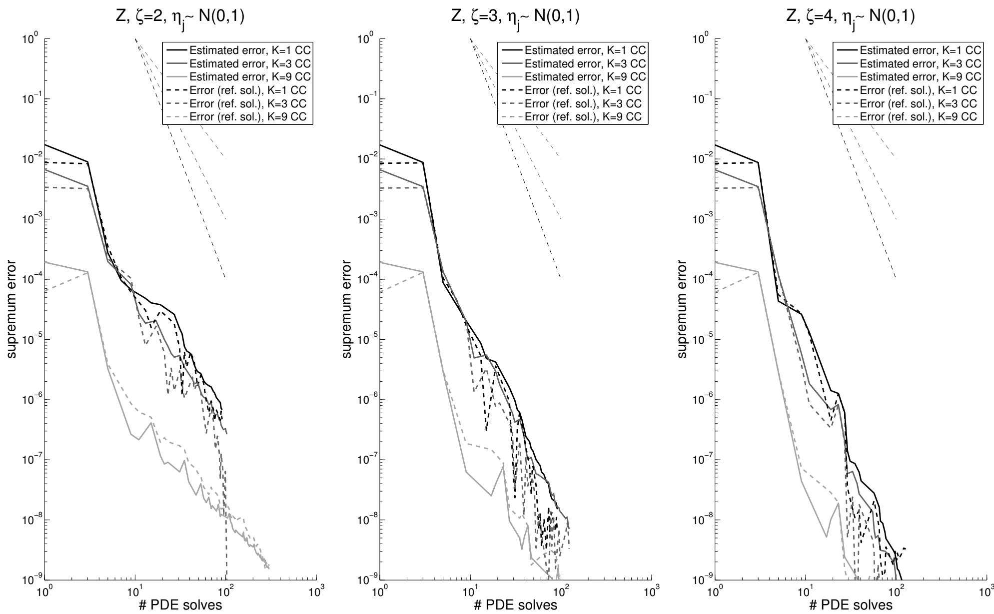

Figure 5. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the number of PDE solves needed based on the sequence CC with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.)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

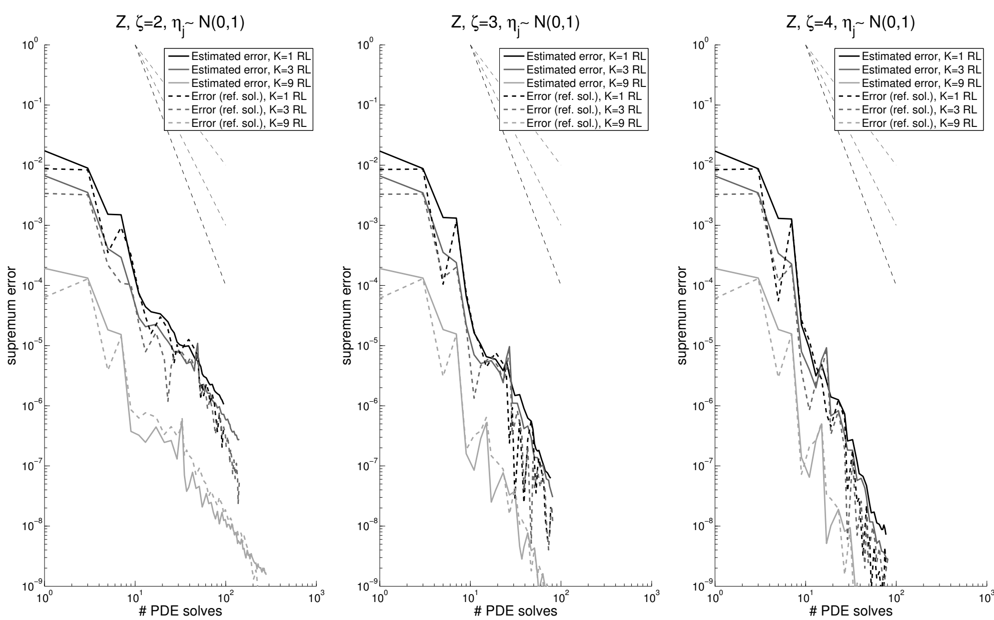

Figure 6. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the number of PDE solves needed based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) 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'_Γ can be observed, cp. Figure 7 - Figure 10 showing the error curves with respect to the cardinality of the index set Λ.

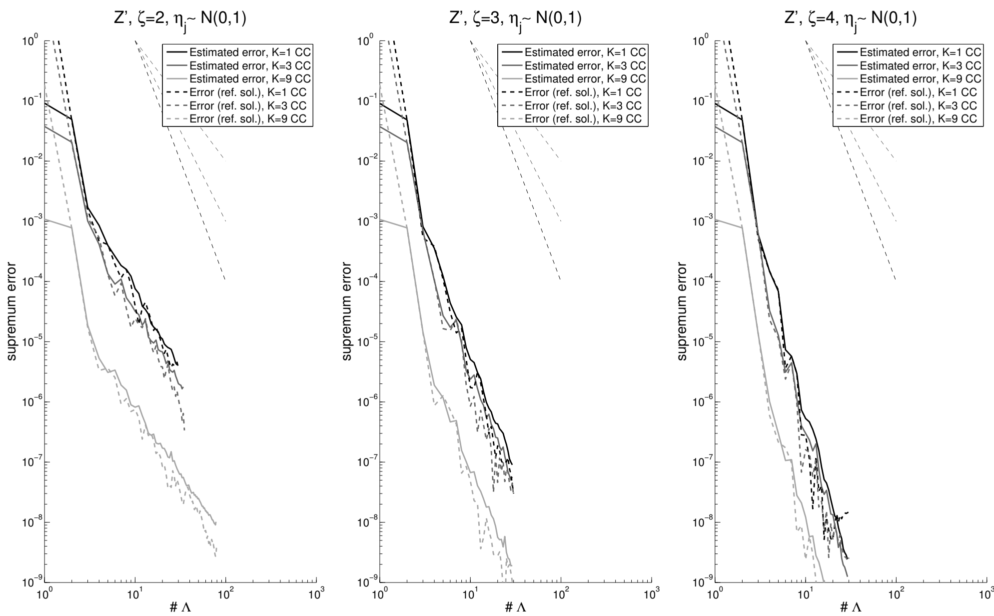

Figure 7. Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z'_Γ with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1,3,9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta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

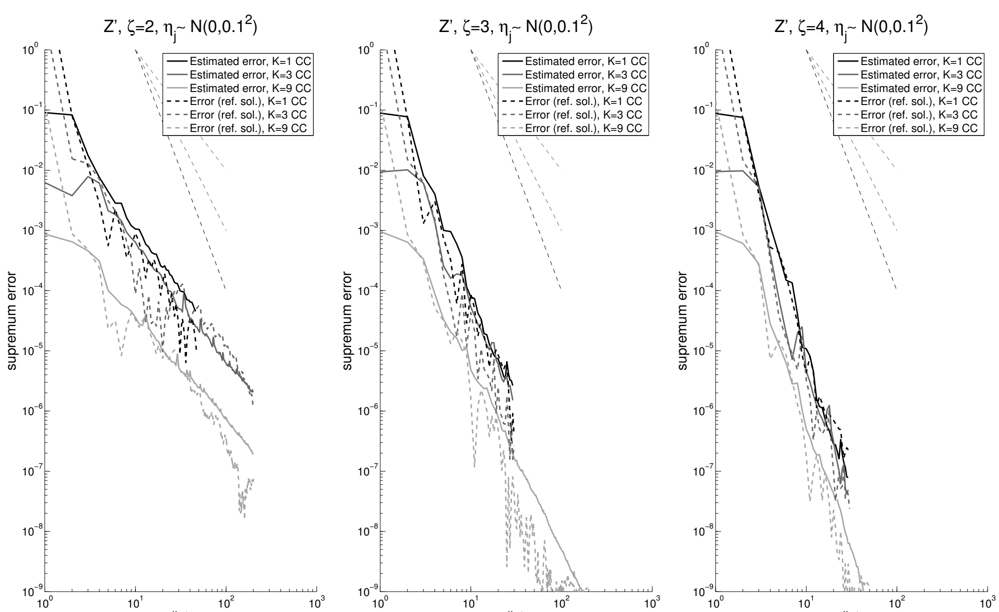

Figure 8. Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z'_Γ with respect to the cardinality of the index set Λ_N needed based on the sequence CC with $K = 1,3,9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\#\mathbb{J}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

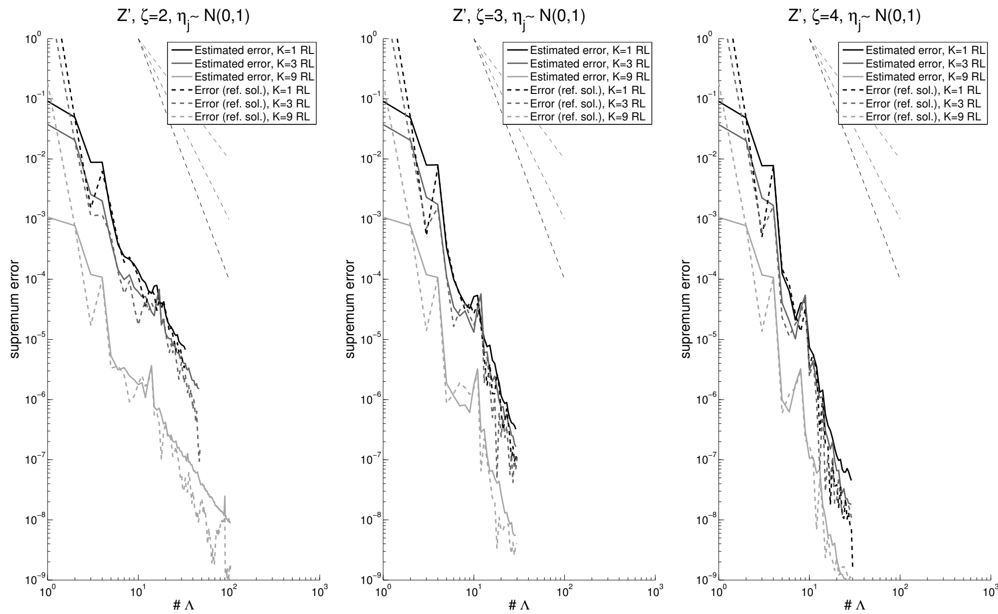

Figure 9. Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z'_Γ with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and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

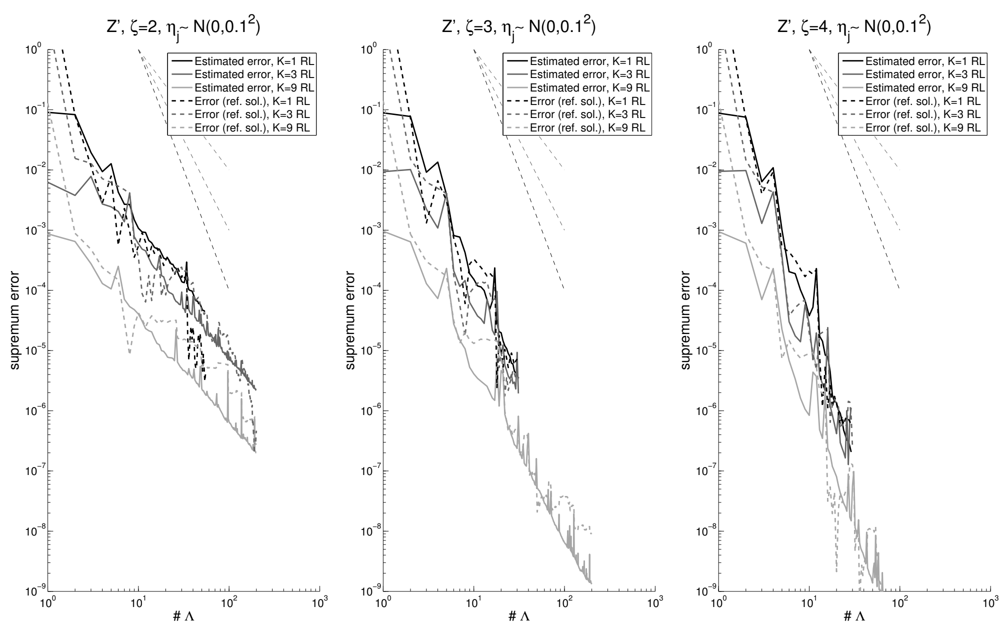

Figure 10. Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z'_Γ with respect to the cardinality of the index set Λ_N based on the sequence RL with $K = 1, 3, 9$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 64$ and $h_T = h_D = 2^{-11}$ for reference and 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 (53) in the 128 dimensional parameter case, i.e. the uncertain coefficient u is parametrized by

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

with $D_j = [(j-1)\frac{1}{128}, j\frac{1}{128}]$, $\boldsymbol{y} = (y_j)_{j=1,\ldots,128}$ 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.

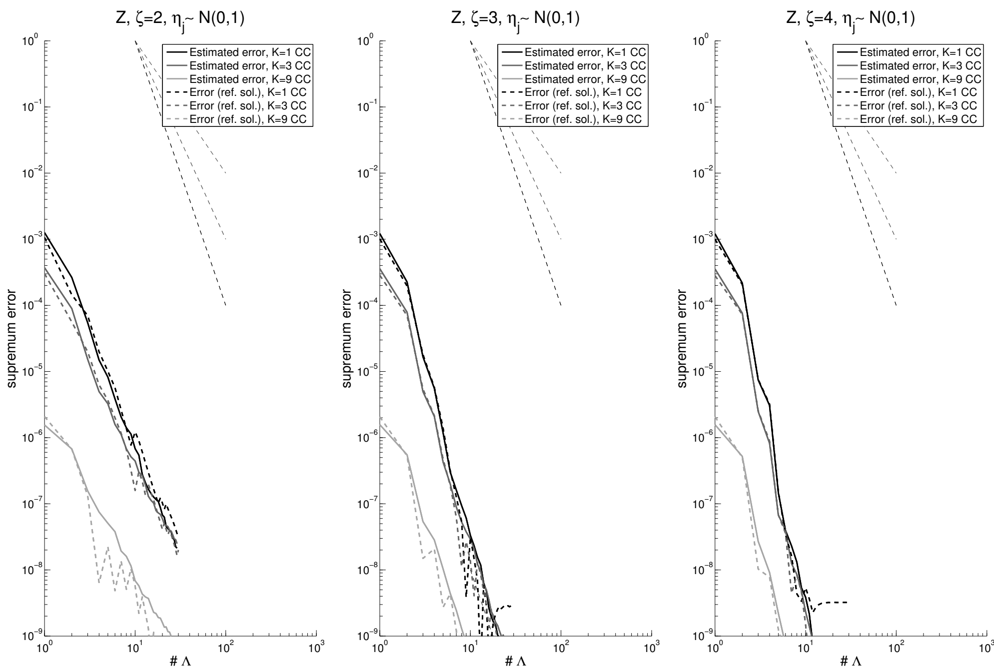

Figure 11. Comparison of the estimated error and actual error. Curves computed by the reference solution of the normalization constant Z_Γ with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1,3,9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$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

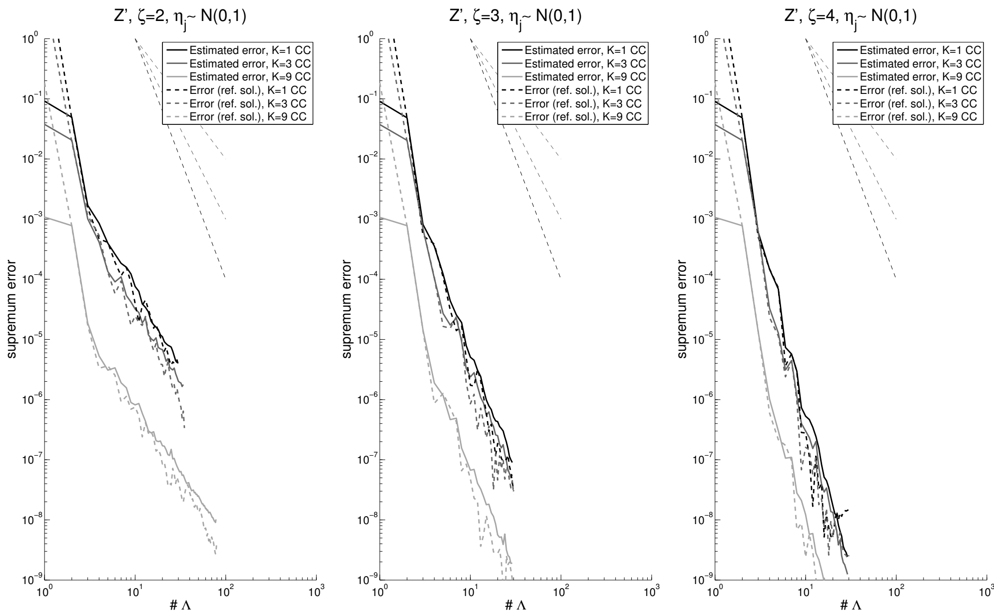

Figure 12. Comparison of the estimated error and actual error. Curves computed by the reference solution of the quantity Z'_Γ with respect to the cardinality of the index set Λ_N based on the sequence CC with $K = 1,3,9$, $\eta \sim \mathcal{N}(0,1)$ and with $\zeta = 2$ (l.), $\zeta = 3$ (m.) and $\zeta = 4$ (r.), $\#\mathbb{J} = 128$ and $h_T = h_D = 2^{-11}$ for reference and 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 Θ for parametrizations of the uncertain forward solution map of the system in terms of possibly countably many variables $\boldsymbol{y} = (y_j)_{j\in\mathbb{J}}$ parametrizing u, provided that the parametric responses $q(\boldsymbol{y})$ satisfy the $(\boldsymbol{b}, p, \varepsilon)$-holomorphy condition in Definition 2.2 with some $0 < p < 1$. This analyticity condition is valid for a wide range of PDE problems, see [10].

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 the covariance* $\Gamma > 0$ *of the observation noise is not small.*

In the case of observation noise with variance $\Gamma \to 0$, the bound (40) reveals that the constants in the bounds on the Legendre coefficients θ^P_ν in the gpc expansions (42) and, via (43), also the *constants* $C > 0$ in the error bounds (45), (47) and, in turn, also the constants C^i in the Smolyak quadrature error estimates (50), (51), depend on Γ as $C \sim \exp(b\|\Gamma^{-1}\|)$ for some constant $b > 0$. We also note that the *convergence rates* in (50), (51) are not affected by the size of Γ. In our numerical experiments, we indeed observe this dependence on Γ, which renders our approach infeasible for small values of Γ. This is due to concentration effects in the integrand functions of the integrals Z_Γ and Z'_Γ in (27), (25) for small values of Γ. Since the integrals (27), (25) are nonoscillatory, as $\Gamma \to 0^+$, all contributions to the integrals Z_Γ and Z'_Γ in (27), (25) 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_Γ and Z'_Γ 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 Γ 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, $\#(\mathbb{J}) = J < \infty$ (achieved by dimension truncation in the parametric representation (9) of the uncertain input u, see [33]).

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 (which point is also referred to as "map point" in the literature). We distinguish two cases: $\boldsymbol{y}^0 \in \mathrm{int}(U)$ 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

$$0 = (\nabla_{\boldsymbol{y}} S)(\boldsymbol{y}^0) \Longleftrightarrow r(\boldsymbol{y})(\nabla_{\boldsymbol{y}} r)(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{y}^0} = 0\,. \tag{54}$$

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 (54) 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 Φ_Γ) and the Hessian $S_{\boldsymbol{yy}}(\boldsymbol{y}^0)$ is negative definite. Then, an asymptotic analysis of Z_Γ and of Z'_Γ via Laplace's method shows (cp. [33])

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

and likewise for Z_Γ with $\phi(\boldsymbol{y})$ replaced by 1. Under the provision of nondegeneracy of the Hessian $S_{yy}(\boldsymbol{y}_0)$, the Bayesian estimate (27) thus admits an *asymptotic expansion with respect to small observation noise variance* Γ (cp. [33])

$$\mathbb{E}^{\mu^\delta}[\phi] = \frac{Z'_\Gamma}{Z_\Gamma} \sim a_0 + a_1\Gamma + a_2\Gamma^2 + ... \,. \tag{55}$$

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 $\boldsymbol{y}^0$ which is a (nonlinear) least square minimizer of the Bayesian potential. For the determination of this minimizer related to the "MAP" estimate, well-developed computational methods from nonlinear optimization are available) the information on the structure of the integrand function which is afforded by the asymptotic analysis will also allow the regularization of the integrand functions $\Theta(\boldsymbol{y})$ and $\Psi(\boldsymbol{y})$ in (27) and (25). Since the minimum $\boldsymbol{y}^0$ is, as a rule, *degenerate*, some form of *regularization* must be employed. Mathematical details and algorithmic aspects will be addressed in [33]; there, also the effect of degeneracies in the Bayesian potential on the asymptotic expansion (55) are considered.

For $\Gamma > 0$ not necessarily small, we showed in particular for parametric operator equations whose solutions $q(\boldsymbol{y})$ are $(\boldsymbol{b}, p, \varepsilon)$-holomorphic, that in inverse problems for such operator equations, under parametric uncertainty, the density of the Bayesian posterior measure with respect to a uniform prior μ_0 on the parametrization space U of the uncertainty is, likewise, $(\boldsymbol{b}, p, \varepsilon)$-holomorhic on U.

We assumed in the present paper that the uncertainties $u \in X$ were charged with a *uniform* prior measure μ_0 which assigns equal probability to all relizations of each coordinate y_j in the uncertainty parametrization (9): 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 ρ_j which are tailored to the prior with respect to coordinate y_j in the parametric representation (9) 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., [31].

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, in addition, 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 [17] for first numerical experiments on this in the context of adaptive solution of parametric initial value problems.

Acknowledgments

This work is supported by Swiss National Science Foundation (SNF) and by the European Research Council (ERC) under FP7 Grant AdG247277. All simulations in this work were performed on the BRUTUS computing cluster at ETH Zürich.

References

[1] Abdulle A, Barth A and Schwab C 2012 Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs, *Report 2012-29, Seminar for Applied Mathematics, ETH Zürich* (To appear in SIAM J. Multiscale Modeling and Analysis)

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

[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 *Jorn. Approx. 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 *Proc. Amer. Math. Soc.* **140** 4271-4284

[7] Chkifa A 2013 On the Lebesgue constant of Leja sequences for the unit disk *Journ. Approx. Theory*

[8] Chkifa A, Cohen A, DeVore R and Schwab C 2011 Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs *ESAIM: Mathematical Modelling and Numerical Analysis* **47** 253-280

[9] Chkifa A, Cohen A and Schwab C 2013 High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs *Journ. Found. Comp. Math.* 1-33

[10] Chkifa A, Cohen A and Schwab C 2013 Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs *Report 2013-25, Seminar for Applied Mathematics, ETH Zürich*

[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] Dick J 2004 On the convergence rate of the component-by-component construction of good lattice rules *Journ. Complexity* **20** 493-522

[14] Dick J, Pillichshammer F and Waterhouse B J 2008 The construction of good extensible rank-1 lattices *Math. Comp.* **77** 2345-2374

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

[16] Hansen M and Schwab C 2011 Analytic regularity and best N-term approximation of high dimensional parametric initial value problems *Vietnam Journ. Math.* **1** 1-35

[17] 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)

[18] Ha Hoang V and Schwab C 2012 Regularity and generalized polynomial chaos approximation of parametric and random second-order hyperbolic partial differential equations *Analysis and Applications (Singapore)* **10** 295-326

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

[20] Ha Hoang V, Schwab C and Stuart A M 2012 Sparse MCMC gpc Finite Element methods for Bayesian inverse problems *Report 2012-23, Seminar for Applied Mathematics, ETH Zürich*

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

[22] Ketelsen C and Scheichl R and Teckentrup A 2013 A hierarchical multilevel Markov-Chain Monte-Carlo algorithm with applications to uncertainty quantification in subsurface flow Preprint 2013.

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

[24] 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 Journ Sci. Comp.* **34** 1460-1487

[25] 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

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

[27] Marzouk Y M and Najm H N 2009 Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems *Journ. Comp. Phys.* **228** 1862-1902

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

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

[30] 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.)

[31] Schillings C 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))

[32] Schillings C and Schwab C 2013 Sparse, Adaptive Smolyak Algorithms for Bayesian Inverse Problems, *Inverse Problems* **29** 065011

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

[34] Schwab C and Stevenson R S 2009 *Space-Time adaptive wavelet methods for parabolic evolution equations*, *Math. Comp.* **78** 1293 - 1318

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

[36] Schwab C and Todor R A 2006 Karhunen-Loève Approximation of Random Fields by Generalized Fast Multipole Methods, *Journ. Comp. Phys.* **217** 100-122

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

Recent Research Reports

2013-16 P. Grohs and H. Hardering and O. Sander
Optimal a priori discretization error bounds for geodesic finite elements

2013-15 A. Lang and Ch. Schwab
Isotropic Gaussian random fields on the sphere: regularity, fast simulation, and stochastic partial differential equations

2013-14 I.G. Graham and F.Y. Kuo and J.A. Nichols and R. Scheichl and Ch. Schwab and I.H. Sloan
Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficient

2013-13 X. Claeys and R. Hiptmair and E. Spindler
A second-kind Galerkin boundary Element Method for scattering at composite objects

2013-12 R. Kornhuber and Ch. Schwab and M. Wolf
Multi-Level Monte-Carlo Finite Element Methods for stochastic elliptic variational inequalities

2013-11 F. Müller and Ch. Schwab
Finite Elements with mesh refinement for wave equations in polygons

2013-10 M. Hansen and C. Schillings and Ch. Schwab
Sparse approximation algorithms for high dimensional parametric initial value problems

2013-09 C. Gittelson and R. Andreev and Ch. Schwab
Optimality of Adaptive Galerkin methods for random parabolic partial differential equations

2013-08 R. Hiptmair and C. Jerez-Hanckes and J. Lee and Z. Peng
Domain decomposition for boundary integral equations via local multi-trace formulations

2013-07 A. Paganini and M. López-Fernández
Efficient convolution based impedance boundary condition