Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Sparse, adaptive Smolyak algorithms for Bayesian inverse problems

Cl. Schillings and Ch. Schwab

Research Report No. 2012-37
November 2012

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

Sparse, Adaptive Smolyak Algorithms for Bayesian Inverse Problems

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. Based on the parametric deterministic formulation of Bayesian inverse problems with unknown input parameter from infinite dimensional, separable Banach spaces proposed in [28], we develop a practical computational algorithm whose convergence rates are provably higher than those of Monte-Carlo (MC) and Markov-Chain Monte-Carlo methods, in terms of the number of solutions of the forward problem. In the formulation of [28], the forward problems are parametric, deterministic elliptic partial differential equations, and the inverse problem is to determine the unknown, parametric deterministic coefficients from noisy observations comprising linear functionals of the solution.

Sparsity of the generalized polynomial chaos (gpc) representation of the posterior density being implied by sparsity assumptions on the class of the prior [28], we design, analyze and implement a class of adaptive, deterministic sparse tensor Smolyak quadrature schemes for the efficient approximate numerical evaluation of expectations under the posterior, given data. The proposed algorithm is based on a greedy, iterative identification of finite sets of most significant, "active" chaos polynomials in the the posterior density analogous to recently proposed algorithms for adaptive interpolation [7, 8]. Convergence rates for the quadrature approximation are shown, both theoretically and computationally, to depend only on the sparsity class of the unknown, but are bounded independently of the number of random variables activated by the adaptive algorithm.

Numerical results for a model problem of coefficient identification with point measurements in a diffusion problem confirm the theoretical results.

1. Introduction

The problems of calibration of partial differential equations, given large sets of noisy input data, and of prediction of responses, is a key problem in applied mathematics, statistics and in the sciences. Of particular interest in this context are predictive simulations and uncertainty quantification, i.e. the prediction of systems' responses and their uncertainty under the calibrated parameters, given noisy observational data. Currently, the most widely used numerical methods for the numerical treatment of these problems are sampling, ie. Monte-Carlo (MC) type algorithms, in particular the so-called Markov-Chain Monte-Carlo (MCMC) methods (eg. [19, 20, 25, 26]). MCMC methods can be interpreted as quadrature methods for the approximate numerical evaluation of the conditional expectations of system responses, given observational data. Being sampling algorithms in nature, MCMC based methods are limited to low convergence rates, but are amenable to parallelization.

It has been argued for some time (eg. [21, 23, 22]) that *deterministic* approaches which are based on gpc representations of the posterior measure, given data, could be more efficient than MCMC methods, at least in situations where the posterior admits such representations. Inspired by N. Wiener's "spectral" view of random fields [33], algorithms based on sparse, deterministic, gpc-based representation of random fields inputs and outputs of PDEs have undergone rapid computational and theoretical development in recent years, see [32, 27] and the references there. The principal motivation for these methods stems from the perspective to undercutting the asymptotic complexity of MC and MCMC methods. Their application to the numerical solution of inverse problems has been proposed recently (we refer to [21, 22, 23] for the first proposal of this approach), with promising numerical results in several problem classes of engineering interest.

In [28], it was shown for a model diffusion problem with unknown coefficient, that for rather general classes of priors taking values in function spaces, the posterior measure admits a density whose representation in terms of a countable number of so-called gpc coordinates is numerically sparse, i.e., a large number of coefficients in this representation are negligible at small thresholds. The result in [28] is not limited to linear, elliptic PDEs: analogous results hold for forward maps of a rather wide range of mathematical models (e.g. [15] for large, parametric systems arising in biological systems identification, and [14] and the references there for time-dependent PDEs).

These sparsity results opened the perspective of novel, *deterministic* computational approaches to inverse problems for identification of parameters in differential equations from noisy measurements. Specifically, we considered the case when *unknown resp. uncertain parameters in function spaces* are of interest. In this setting, statistical parameter estimation can be performed using Bayesian methods, suitably generalized to (infinite-dimensional) function space settings (see [31] and the references there for further details).

An alternative approach to Bayesian inverse problems in PDEs is via techniques of optimal control (e.g. [1]); however this does not lead naturally to quantification of uncertainty. A Bayesian approach to the inverse problem [19, 31] allows the observations to map a possibly simple prior probability distribution on the input parameters into a posterior distribution. This posterior distribution is typically much more complicated than the prior, involving many correlations and without a useable closed form. The posterior distribution completely quantifies the uncertainty in the system's response, under given prior and structural assumptions on the system and given observational data. It allows, in particular, the Bayesian statistical estimation of unknown system parameters and responses by integration with respect to the posterior measure, which is of interest in many applications.

MCMC methods probe this posterior probability distribution for the computation of estimates of uncertain system responses conditioned on given, noisy observation data. However, MCMC methods suffer from the same limitations of computational complexity as straightforward Monte Carlo methods. It is hence of interest to investigate whether sparse approximation techniques can be used to approximate the posterior density and conditional expectations given the data. In this paper we study this question in the context of a model elliptic inverse problem. Elliptic problems with random coefficients have provided an important class of model problems for the uncertainty quantification community, see, for example, [3, 27] and the references therein. In the context of inverse problems and noisy observational data, the corresponding elliptic problem arises naturally in the study of groundwater flow (see [24]) where hydrologists wish to determine the transmissivity (diffusion coefficient) from the head (solution of the elliptic PDE). The elliptic inverse problem hence provides a natural model problem within which to study sparse representations of the posterior distribution.

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. Section 3 then presents the concrete setting of a linear, elliptic PDE with unknown diffusion coefficient which we use here to develop our ideas. Section 4 recapitulates results from [28] on sparsity of the posterior density. Section 5 then contains the main results of the present paper: formulation and convergence analysis of a sparse, adaptive Smolyak quadrature approach. Several concrete families of univariate quadratures which can be used in the Smolyak construction are presented. Section 6 develops detailed numerical experiments which confirm the theoretical results and which, in particular, also allow to compare the performance of the proposed sparse quadrature methods in terms of the total number of PDE solutions necessary. Finally, in Section 7 we collect the major conclusions from the present work and indicate generalizations to other problem classes.

2. Bayesian Inverse Problems

Let $G : X \to R$ denote a "forward" response map from some separable Banach space X of unknown parameters into another separable Banach space R of responses. In the present paper, "responses" will comprise solutions of a (partial) differential equation model of the system to be described, for a given realization of an uncertain input function u. We equip X and R with norms $\|\cdot\|_X$ and with $\|\cdot\|_R$, respectively. In addition, we assume given $\mathcal{O}(\cdot) : R \to \mathbb{R}^K$ denoting a *bounded linear observation operator* on the space R of system responses, i.e. $\mathcal{O} \in R^*$ of the space R 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 $|\cdot|$.

We wish to determine the unknown data $u \in X$ from the noisy observations

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

where $\eta \in \mathbb{R}^K$ represents the observation noise. We assume that realization of the noise process is not known to us, but that it is a draw from the Gaussian measure $\mathcal{N}(0,\Gamma)$, for some positive (known) covariance operator Γ on $\mathbb{R}^K$. If we define $\mathcal{G} : X \to \mathbb{R}^K$ by $\mathcal{G} = \mathcal{O} \circ G$ then we may write the equation for the observations as

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

According to 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

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

where $|\cdot|_\Gamma = |\Gamma^{-\frac{1}{2}} \cdot |$ so that

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

In [31] it is shown that, under appropriate conditions on the forward and observation model $\mathcal{G}$ and the prior measure on u, the posterior distribution on u is absolutely continuous with respect to the prior with Radon-Nikodym derivative given by an infinite dimensional version of Bayes rule. Posterior uncertainty is then determined by integration of suitably chosen functions against this posterior. At the heart of the deterministic approach proposed and analyzed here lies the *reformulation of the forward problem with unknown stochastic input data as an infinite dimensional, parametric deterministic problem*. We are thus interested in expressing the posterior distribution in terms of a parametric representation of the unknown coefficient function u. To this end we assume that, under the prior distribution, this function admits a *parametric representation* of the form

$$u = \bar{a} + \sum_{j\in\mathbb{J}} y_j \psi_j \tag{2}$$

where $y = (y_j)_{j\in\mathbb{J}}$ is an i.i.d sequence of real-valued random variables $y_j \sim \mathcal{U}(-1/2, 1/2)$ and $\bar{a}$ and the ψ_j are elements of X. Here and throughout, $\mathbb{J}$ denotes a finite or countably infinite index set, i.e. either $\mathbb{J} = \{1, 2, ..., J\}$ or $\mathbb{J} = \mathbb{N}$. All assertions proved in the present paper hold in either case, and all bounds are in particular independent of the number J of parameters.

The basis for our Smolyak quadrature approach to Bayes' formula is a *parametric, deterministic representation of the derivative of the posterior measure with respect to the prior* μ_0. To fix notation, we denote by

$$U = (-1/2, 1/2)^{\mathbb{J}}$$

the space of all sequences $(y_j)_{j\in\mathbb{J}}$ of real numbers $y_j \in (-1/2, 1/2)$. Denoting the sub σ-algebra of Borel subsets on $\mathbb{R}$ which are also subsets of $(-1/2, 1/2)$ by $\mathcal{B}^1(-1/2, 1/2)$, the pair

$$(U, \mathcal{B}) = \left((-1/2, 1/2)^{\mathbb{J}}, \bigotimes_{j\in\mathbb{J}} \mathcal{B}^1(-1/2, 1/2) \right) \tag{3}$$

is a measurable space. With λ_1 denoting the Lebesgue measure on the real line $\mathbb{R}^1$, we equip $(U, \mathcal{B})$ with the uniform probability measure

$$\mu_0(dy) := \bigotimes_{j\in\mathbb{J}} \lambda_1(dy_j) \; .$$

Since the countable product of probability measures is again a probability measure, $(U, \mathcal{B}, \mu_0)$ is a probability space. *We assume in what follows that the prior measure on the uncertain input data, parametrized in the form* (2)*, is* $\mu_0(dy)$*.* We add in passing that unbounded parameter ranges as arise, e.g., in lognormal random diffusion coefficients in models for subsurface flow [24], can be treated by the techniques developed here, at the expense of additional technicalities. With U as in (3), we define $\Xi : U \to \mathbb{R}^K$ by

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

In the following theorem, from [28] (see also [31]), we view U as a bounded subset in $\ell^\infty(\mathbb{J})$, the Banach space of bounded sequences, and thereby introduce a notion of continuity in U.

Theorem 2.1. *Assume that* $\Xi : \bar{U} \to \mathbb{R}^K$ *is bounded and continuous. Then* $\mu^\delta(dy)$*, the distribution of* y *given* δ*, is absolutely continuous with respect to* $\mu_0(dy)$*. Furthermore, if*

$$\Theta(y) = \exp\big(-\Phi(u;\delta)\big)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j} , \tag{4}$$

then

$$\frac{d\mu^\delta}{d\mu_0}(y) = \frac{1}{Z}\Theta(y) \tag{5}$$

where the normalization constant Z *is given by*

$$Z = \int_U \Theta(y)\mu_0(dy) \; . \tag{6}$$

The goal of computation is, for given (noisy) data δ, to compute the expectation of a *prediction function* $\phi : X \to S$ (sometimes also referred to as *Quantity of Interest (QoI)*), taking values in some Banach space S. In particular the choices $\phi(u) = G(u)$ and $\phi(u) = G(u) \otimes G(u)$ facilitate computation of the mean and covariance of the response.

We associate with the function ϕ, the (infinite-dimensional) parametric mapping

$$\Psi(y) = \exp\big(-\Phi(u;\delta)\big)\phi(u)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j} : U \to S \; . \tag{7}$$

and define

$$Z' = \int_U \Psi(y)\mu_0(dy) = \int_U \exp\big(-\Phi(u;\delta)\big)\phi(u)\Big|_{u=\bar{a}+\sum_{j\in\mathbb{J}} y_j\psi_j}\mu_0(dy) \qquad (8)$$

so that the expectation of the quantity of interest is given by $Z'/Z \in S$. Thus our aim is to approximate Z' and Z. A widely used algorithm to approximate the infinite dimensional integrals Z, Z' is MCMC which has, in general, low convergence rates. Here, we propose a deterministic quadrature based on sparse, adaptive quadratures of Smolyak type. In the next sections we will study the elliptic problem and deduce, from known results concerning the parametric forward problem, the joint analyticity of the posterior density $\Theta(y)$, and also $\Psi(y)$, as a function of the parameter vector $y \in U$. From these results, we deduce *sharp estimates on size of domain of analyticity of* $\Theta(y)$ *(and* $\Psi(y)$*) as a function of each coordinate* y_j*,* $j \in \mathbb{N}$*.* We focus on the estimation of mean fields. The ensuing analysis can be extended to other choices of Ψ.

3. Model Parametric Elliptic Problem

3.1. Function Spaces

Let $D \subset \mathbb{R}^d$ be a bounded interval if $d = 1$ or a bounded Lipschitz domain in $\mathbb{R}^d$, $d \geq 2$, with Lipschitz boundary ∂D. Let further $\Big(H, (\cdot,\cdot), \|\cdot\|\Big)$ denote the Hilbert space $L^2(D)$ which we will identify throughout with its dual space, i.e. $H \simeq H^*$.

As in [28], we denote by $\Big(V, (\nabla\cdot, \nabla\cdot), \|\cdot\|_V\Big)$ the Hilbert space $H^1_0(D)$ (everything that follows will hold for rather general, elliptic problems with affine parameter dependence and "energy" space V). The dual space V^* of all continuous, linear functionals on V is isomorphic to the Banach space $H^{-1}(D)$ which we equip with the dual norm to V, denoted $\|\cdot\|_{-1}$. We shall assume for the (deterministic) data $f \in V^*$.

3.2. Forward Problem

In the bounded Lipschitz domain D for given $f \in L^2(D)$, we consider the following elliptic PDE:

$$-\nabla\cdot\big(u\nabla p\big) = f \quad \text{in} \quad D, \qquad p = 0 \quad \text{in} \quad \partial D. \qquad (9)$$

Given data $u \in L^\infty(D)$, a weak solution of (9) for any $f \in V^*$ is a function $p \in V$ which satisfies

$$\int_D u(x)\nabla p(x)\cdot\nabla q(x)dx = {}_V\langle q, f\rangle_{V^*} \;\; \text{for all} \;\; q \in V\,. \qquad (10)$$

Here ${}_V\langle\cdot,\cdot\rangle_{V^*}$ denotes the dual pairing between elements of V and V^*.

For the well-posedness of the forward problem, we shall work under

Assumption 3.1. *There exist constants* $0 < a_{\text{MIN}} \leq a_{\text{MAX}} < \infty$ *so that*

$$0 < a_{\text{MIN}} \leq u(x) \leq a_{\text{MAX}} < \infty, \quad x \in D, \qquad (11)$$

Under Assumption 3.1, the Lax-Milgram Lemma ensures the existence and uniqueness of the response p of (10). Thus, in the notation of the previous section, $R = V$ and $G(u) = p$. Moreover, this variational solution satisfies the a-priori estimate

$$\|G(u)\|_V = \|p\|_V \leq \frac{\|f\|_{V^*}}{a_{\text{MIN}}} \,. \tag{12}$$

We assume that the observation function $\mathcal{O} : V \to \mathbb{R}^K$ comprises K linear functionals $o_k \in V^*$, $k = 1, \ldots, K$. In the notation of the previous section, we denote by $X = L^\infty(D)$ the Banach space in which the unknown input parameter u takes values. It follows that

$$|\mathcal{G}(u)| \leq \frac{\|f\|_{V^*}}{a_{\text{MIN}}} \Big(\sum_{k=1}^{K} \|o_k\|_{V^*}^2 \Big)^{\frac{1}{2}} \,.$$

3.3. Structural Assumptions on Diffusion Coefficient

As discussed in section 2 we introduce a parametric representation of the unknown u via an affine representation with respect to y, which means that the parameters y_j are the coefficients of the function u in the formal series expansion

$$u(x,y) = \bar{a}(x) + \sum_{j \in \mathbb{J}} y_j \psi_j(x), \quad x \in D, \tag{13}$$

where $\bar{a} \in L^\infty(D)$ and $(\psi_j)_{j\in\mathbb{J}} \subset L^\infty(D)$. We are interested in the effect of approximating the solutions input parameter $u(x,y)$, by truncation of the series expansion (13) in the case $\mathbb{J} = \mathbb{N}$, and on the corresponding effect on the forward (resp. observational) map $G(u(\cdot))$ (resp. $\mathcal{G}(u(\cdot))$) to the family of elliptic equations with the above input parameters. In the decomposition (13), we have the choice to either normalize the basis (e.g., assume they all have norm one in some space) or to normalize the parameters. It is more convenient for us to do the latter. This leads us to the following assumptions which shall be made throughout:

i) For all $j \in \mathbb{J} : \psi_j \in L^\infty(D)$ and $\psi_j(x)$ is defined for all $x \in D$,

ii) $$y = (y_1, y_2, \ldots) \in U = [-1/2, 1/2]^{\mathbb{J}},$$
i.e. the parameter vector y in (13) belongs to the ball of radius $1/2$ of the sequence space $\ell^\infty(\mathbb{J})$,

iii) for each $u(x,y)$ to be considered, (13) holds for every $x \in D$ and every $y \in U$.

We will, on occasion, use (13) with $\mathbb{J} \subset \mathbb{N}$, as well as with $\mathbb{J} = \mathbb{N}$ (in the latter case the additional Assumption 3.2 below has to be imposed). In either case, we will work throughout under the assumption that the ellipticity condition (11) holds uniformly for $y \in U$.

Uniform Ellipticity Assumption: *there exist* $0 < a_{\text{MIN}} \leq a_{\text{MAX}} < \infty$ *such that for all* $x \in D$ *and for all* $y \in U$

$$0 < a_{\text{MIN}} \leq u(x,y) \leq a_{\text{MAX}} < \infty. \tag{14}$$

We refer to assumption (14) as $\mathbf{UEA}(a_{\text{MIN}}, a_{\text{MAX}})$ in the following. In particular, $\mathbf{UEA}(a_{\text{MIN}}, a_{\text{MAX}})$ implies $a_{\text{MIN}} \leq \bar{a}(x) \leq a_{\text{MAX}}$ for all $x \in D$, since we can choose $y_j = 0$ for all $j \in \mathbb{N}$. Also observe that the validity of the lower and upper inequality in (14) for all $y \in U$ are respectively equivalent to the conditions that

$$\frac{1}{2}\sum_{j\in\mathbb{J}} |\psi_j(x)| \leq \bar{a}(x) - a_{\text{MIN}}, \quad x \in D,$$

and

$$\frac{1}{2}\sum_{j\in\mathbb{J}} |\psi_j(x)| \leq a_{\text{MAX}} - \bar{a}(x), \quad x \in D\,.$$

We shall require in what follows a quantitative control of the relative size of the fluctuations in the representation (13). To this end, we shall impose

Assumption 3.2. *The functions $\bar{a}$ and ψ_j in* (13) *satisfy*

$$\frac{1}{2}\sum_{j\in\mathbb{J}} \|\psi_j\|_{L^\infty(D)} \leq \frac{\kappa}{1+\kappa}\overline{a}_{\text{MIN}},$$

with $\overline{a}_{\text{MIN}} = \min_{x\in D} \bar{a}(x) > 0$ and $\kappa > 0$.

Assumption 3.1 is then satisfied by choosing (eg. [9, 10])

$$a_{\text{MIN}} := \overline{a}_{\text{MIN}} - \frac{\kappa}{1+\kappa}\overline{a}_{\text{MIN}} = \frac{1}{1+\kappa}\overline{a}_{\text{MIN}}.$$

3.4. Inverse Problem

The inverse problem consists of determining the unknown diffusion coefficient u from given noisy observation data δ in order to compute the expectation of a quantity of interest (8), given this data. As was shown in [28], the forward maps $G : X \to V$ and $\mathcal{G} : X \to \mathbb{R}^K$ are Lipschitz. Specifically (see [28, Lemma 3.3, Theorem 3.4]) if p and $\tilde{p}$ are solutions of (10) with the same right hand side f and with coefficients u and $\tilde{u}$, respectively, and if these coefficients both satisfy Assumption 3.1 then the forward solution map $u \to p = G(u)$ is Lipschitz as a mapping from X into V with Lipschitz constant defined by

$$\|p - \tilde{p}\|_V \leq \frac{\|f\|_{V^*}}{a_{\text{MIN}}^2}\|u - \tilde{u}\|_{L^\infty(D)}.$$

Moreover the forward solution map can be composed with the observation operator to prove that the observed response map $u \to \mathcal{G}(u) = (\mathcal{O} \circ G)(u)$ is Lipschitz as a mapping from X into $\mathbb{R}^K$ with Lipschitz constant defined by

$$|\mathcal{G}(u) - \mathcal{G}(\tilde{u})| \leq \frac{\|f\|_{V^*}}{a_{\text{MIN}}^2}\Big(\sum_{k=1}^{K} \|o_k\|_{V^*}^2\Big)^{\frac{1}{2}}\|u - \tilde{u}\|_{L^\infty(D)}.$$

and there holds (see [28] for the proof)

Theorem 3.3. *Under the* $\mathbf{UEA}(a_{\text{MIN}}, a_{\text{MAX}})$ *and Assumption 3.2 it follows that the Bayesian posterior measure* $\mu^\delta(dy)$ *on* y *conditioned on noisy observation data* δ *is absolutely continuous with respect to the prior measure* $\mu_0(dy)$ *with Radon-Nikodym derivative given explicitly by* (4) *and* (5).

4. Sparse Polynomial Chaos Approximations of the Posterior

Building on Theorem 3.3, we now present from [28] sparsity results for the posterior density $\Theta(z)$, viewed as a holomorphic functional over $z \in \mathbb{C}^{\mathbb{J}}$, in so-called *polynomial chaos representations*. Exactly the same results on analyticity and on N-term approximation of $\Psi(z)$ hold, cp. [28]. 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 $\Theta(y)$ in U in Section 5 ahead, we analyze also N-term truncated *monomial gpc-approximations* of $\Theta(y)$.

4.1. gpc Representations of Θ

With the index set $\mathbb{J}$ from the parametrization (13) of the input, we associate the countable index set

$$\mathcal{F} = \{\nu \in \mathbb{N}_0^{\mathbb{J}} : |\nu|_1 < \infty\}$$

of multiindices where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For $\nu \in \mathcal{F}$, we denote by $\mathbb{I}_\nu = \{j \in \mathbb{N} : \nu_j \neq 0\} \subset \mathbb{N}$ the "support" of $\nu \in \mathcal{F}$, i.e. the *finite* set of indices of nonvanishing entries of $\nu \in \mathcal{F}$ and by $\aleph(\nu) := \#\mathbb{I}_\nu < \infty$, $\nu \in \mathcal{F}$ the "support size" of ν, i.e. the cardinality of $\mathbb{I}_\nu$.

For the deterministic approximation of the posterior density $\Theta(y)$ in (4) we shall use tensorized polynomial bases similar to what is done in so-called "polynomial chaos" expansions of random fields. We shall consider two particular polynomial bases, Legendre and monomial bases.

4.1.1. Legendre Expansions of Θ Since we assumed that the prior measure $\mu_0(dy)$ is built by tensorization of the uniform probability measures on $(-1/2, 1/2)$, we build the bases by tensorization as follows: let $L_k(z_j)$ denote the k^{th} Legendre polynomial of the variable $z_j \in \mathbb{C}$, normalized such that

$$\int_{-1/2}^{1/2} (L_k(t))^2 dt = 1, \quad k = 0, 1, 2, ... \tag{15}$$

Since $L_0 \equiv 1$, the Legendre polynomials L_k in (15) can be tensorized on the (possibly infinite-dimensional) parameter domains U via

$$L_\nu(z) = \prod_{j \in \mathbb{J}} L_{\nu_j}(z_j), \quad z \in \mathbb{C}^{\mathbb{J}},\ \nu \in \mathcal{F}\ . \tag{16}$$

The normalization (15) implies that the polynomials $L_\nu(z)$ in (16) 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 (16) is the product of only finitely many nontrivial polynomials. It moreover implies that the set of tensorized Legendre polynomials

$$\mathbb{P}(U, \mu_0(dy)) := \{L_\nu : \nu \in \mathcal{F}\} \tag{17}$$

forms a countable orthonormal basis in $L^2(U, \mu_0(dy))$. This observation suggests, by virtue of the square integrability discussed below, the use of mean square convergent gpc-expansions to represent 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 $\mu_0(dy)$ over U, i.e. $\Theta \in L^2\big(U, \mu_0(dy)\big)$. Moreover, if the function $\phi(\cdot) : U \to S$ in (7) is bounded, then

$$\int_U \|\Psi(y)\|_S^2 \mu_0(dy) < \infty, \tag{18}$$

i.e. $\Psi \in L^2\big(U, \mu_0(dy); S\big)$.

Remark 4.1. It is a consequence of (12) that in the case where $\phi(u) = G(u) = p \in V$ we have $\|\Psi(y)\|_V \leq \|f\|_{V^*}/a_{\text{MIN}}$ for all $y \in U$. Thus $\Psi \in L^2\big(U, \mu_0(dy); S\big)$ 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. $\square$

Since $\mathbb{P}(U, \mu_0(dy))$ in (17) is a countable orthonormal basis of $L^2(U, \mu_0(dy))$, the density $\Theta(y)$ of the posterior measure given data $\delta \in Y$, and the posterior reweighted pressure $\Psi(y)$ can be represented in $L^2(U, \mu_0(dy))$ by (parametric and deterministic) generalized Legendre polynomial chaos expansions. We first address the representation of the scalar valued function $\Theta(y)$.

$$\Theta(y) = \sum_{\nu \in \mathcal{F}} \theta_\nu L_\nu(y) \quad \text{in} \quad L^2(U, \rho(dy)) \tag{19}$$

where the gpc expansion coefficients θ_ν are defined by

$$\theta_\nu = \int_U \Theta(y) L_\nu(y) \mu_0(dy) \ , \quad \nu \in \mathcal{F} \ .$$

By Parseval's equation and the normalization (15), it follows immediately from (19) and (18) with Parseval's equality that the second moment of the posterior density with respect to the prior

$$\|\Theta\|^2_{L^2(U, \mu_0(dy))} = \sum_{\nu \in \mathcal{F}} |\theta_\nu|^2 \tag{20}$$

is finite.

4.1.2. Monomial Expansions of Θ We next consider expansions of the posterior density Θ with respect to monomials

$$y^\nu = \prod_{j\geq 1} y_j^{\nu_j}, \quad y \in U, \quad \nu \in \mathcal{F} .$$

Once more, the infinite product is well-defined since, for every $\nu \in \mathcal{F}$, it contains only $\aleph(\nu)$ many nontrivial factors. In the subsequent section, the analyticity of $\Theta(y)$ in U will be shown, based on the results in [7, 28] which implies that in U, the density $\Theta(y)$ can be represented by an unconditionally convergent (in U) monomial expansion about the section $y = 0 \in U$ with uniquely determined Taylor coefficients $\tau_\nu \in V$ which coincide, by uniqueness of the analytic continuation, with the Taylor coefficients of Θ at $0 \in U$:

$$\forall y \in U : \quad \Theta(y) = \sum_{\nu\in\mathcal{F}} \tau_\nu y^\nu \ , \quad \tau_\nu := \frac{1}{\nu!}\partial_y^\nu \Theta(y) \mid_{y=0} \ . \tag{21}$$

4.2. N-term gpc Approximations of Θ

The efficient numerical evaluation of expectations Z' as in (8) under the posterior requires evaluation of the integrals (6) and (8). Our strategy is to approximate these integrals by truncating the spectral respresentation (19), as well as a similar expression for $\Psi(y)$, to a finite number N of significant terms, and to estimate the error incurred by doing so. We first introduce a class of subsets $\Lambda \subset \mathcal{F}$ which contain, on the one hand, N-term gpc approximations which converge at the best possible rates afforded by Θ and, on the other hand, allow for convenient recursive construction in the quadrature algorithm. As in [7, 15], we shall use the notion of *monotone sets of multiindices*.

Defintion 4.2. *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 $i \in \mathbb{J}\backslash\{j\}$.*

Note that for monotone index sets $\Lambda_N \subset \mathcal{F}$ properties (M1) and (M2) in Definition 4.2 imply

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

By (20), the coefficient sequence $(\theta_\nu)_{\nu\in\mathcal{F}}$ must necessarily decay. If this decay is sufficiently strong, possibly high convergence rates of N-term approximations of the integrals (6), (8) occur.

4.2.1. $L^2(U;\mu_0)$ *Approximation* Denote by $\Lambda_N \subset \mathcal{F}$ a set of indices $\nu \in \mathcal{F}$ corresponding to N largest gpc coefficients $|\theta_\nu|$ in (19), and denote by

$$\Theta_{\Lambda_N}(y) := \sum_{\nu\in\Lambda_N} \theta_\nu L_\nu(y) \tag{22}$$

the Legendre expansion (19) truncated to this set of indices. Using a standard result of Stechkin with $q = 2$ (see (3.13) in [10]), Paseval's equation (20) and $0 < \sigma \leq 1$ we obtain for all N

$$\|\Theta(y) - \Theta_{\Lambda_N}(y)\|_{L^2(U,\mu_0(dy))} \leq N^{-s}\|(\theta_\nu)\|_{\ell^\sigma(\mathcal{F})}, \quad s := \frac{1}{\sigma} - \frac{1}{2}\,. \tag{23}$$

We infer from (23) that *a mean-square convergence rate $s > 1/2$ of the approximate posterior density Θ_{Λ_N} can be achieved* provided that $(\theta_\nu) \in \ell^\sigma(\mathcal{F})$ *for some* $0 < \sigma < 1$.

4.2.2. $L^1(U;\mu_0)$ and Pointwise Approximation of Θ The analyticity of $\Theta(y)$ in U implies that $\Theta(y)$ can be represented by the Taylor exansion (21). This expansion is unconditionally summable for all $y \in U$: for any sequence $(\Lambda_N)_{N\in\mathbb{N}} \subset \mathcal{F}$ which exhausts $\mathcal{F}$ ‡, the corresponding sequence of N-term truncated partial Taylor sums

$$T_{\Lambda_N}(y) := \sum_{\nu\in\Lambda_N} \tau_\nu y^\nu\,, \quad \tau_\nu := \frac{1}{\nu!}\partial_y^\nu \Theta(y)\mid_{y=0} \tag{24}$$

converges pointwise in U to Θ. Since for $y \in U$ and $\nu \in \mathcal{F}$ we have $|y^\nu| = \prod_{j\geq 1}|y^{\nu_j}| \leq \prod_{j\geq 1}|2^{-\nu_j}| = 2^{-|\nu|_1}$, for any $\Lambda_N \subset \mathcal{F}$ of cardinality not exceeding N holds

$$\sup_{y\in U}\left|\Theta(y) - T_{\Lambda_N}(y)\right| = \sup_{y\in U}\left|\sum_{\nu\in\mathcal{F}\backslash\Lambda_N} \tau_\nu y^\nu\right| \leq \sum_{\nu\in\mathcal{F}\backslash\Lambda_N} 2^{-|\nu|_1}|\tau_\nu|\,.$$

Similarly, we have

$$\|\Theta - T_{\Lambda_N}\|_{L^1(U,\mu_0)} = \left\|\sum_{\nu\in\mathcal{F}\backslash\Lambda_N} \tau_\nu y^\nu\right\|_{L^1(U,\mu_0)} \leq \sum_{\nu\in\mathcal{F}\backslash\Lambda_N} |\tau_\nu|\,\|y^\nu\|_{L^1(U,\mu_0)}\,.$$

For $\nu \in \mathcal{F}$, we calculate

$$\|y^\nu\|_{L^1(U,\mu_0)} = \int_{y\in U}|y^\nu|\mu_0(dy) = \prod_{j\in\mathbb{I}_\nu}\int_{y_j=-\frac{1}{2}}^{\frac{1}{2}}|y_j|^{\nu_j}\lambda_1(dy_j) = \prod_{j\in\mathbb{I}_\nu}\frac{1}{(\nu_j+1)2^{\nu_j}}$$

so that we find

$$\|\Theta - T_{\Lambda_N}\|_{L^1(U,\mu_0)} \leq \sum_{\nu\in\mathcal{F}\backslash\Lambda_N}\frac{|\tau_\nu|}{\prod_{j\in\mathbb{I}_\nu}(\nu_j+1)2^{\nu_j}} = \sum_{\nu\in\mathcal{F}\backslash\Lambda_N}\frac{1}{(\nu+1)!2^\nu}|\partial_y^\nu\Theta(y)\mid_{y=0}|\,.$$

In the following, we will establish the summability of the coefficient sequence $(\tau_\nu)_{\nu\in\mathcal{F}}$ in (21) in order to derive convergence rates of N-term approximations of the integrals (6), (8). The theoretical results derived in the remainder of this section ensure the existence of a sequence of monotone index sets $(\Lambda_N)_{N\geq 1}$ for which the quadrature operator will recover optimal convergence rates (as compared with the

‡ A sequence $(\Lambda_N)_{N\in\mathbb{N}} \subset \mathcal{F}$ of index sets Λ_N whose cardinality does not exceed N exhausts $\mathcal{F}$ if any finite $\Lambda \subset \mathcal{F}$ is contained in all Λ_N for $N \geq N_0$ with N_0 sufficiently large.

best N-term benchmark). Thus, an important issue to be addressed concerns the identification of the sequence of monotone index sets $(\Lambda_N)_{N\geq 1}$ leading to optimal convergence rates. In Subsection 5.3, we will present a greedy strategy which adaptively identifies the most significant polynomials in the gpc representation of the posterior density.

4.3. Sparsity of the Posterior Density Θ

The analysis in the previous section shows that the convergence rate of the truncated gpc-type approximations (22), (24) on the parameter space U is determined by the σ-summability of the corresponding coefficient sequences $(|\theta_\nu|)_{\nu\in\mathcal{F}}$, $(|\tau_\nu|)_{\nu\in\mathcal{F}}$. We now show that summability (and, hence, sparsity) of Legendre and Taylor coefficient sequences in the expansions (19), (21) is determined by that of the sequence $(\|\psi_j\|_{L^\infty(D)})_{j\in\mathbb{N}}$ in the input's fluctuation expansion (13). Throughout, Assumptions 3.1 and 3.2 will be required to hold.

We now impose a *sparsity requirement for the unknown coefficient function u in the forward problem* (9). It is formalized in terms of decay of the ψ_j in (2) by imposing σ-summability in the following form.

Assumption 4.3. *There exists* $0 < \sigma < 1$ *such that for the parametric representations* (13), (2) *it holds that*

$$\sum_{j=1}^{\infty} \|\psi_j\|^{\sigma}_{L^\infty(D)} < \infty \,.$$

Under assumption 4.3, it was shown in [28] that the posterior density $\Theta(y)$ in (4) admits unconditionally convergent polynomial chaos expansions *with the same sparsity in the coefficient sequences*.

The strategy of establishing sparsity of the sequences $(|\theta_\nu|)_{\nu\in\mathcal{F}}$, $(|\tau_\nu|)_{\nu\in\mathcal{F}}$ in [28] is based on estimating the sequences by Cauchy's integral formula applied to the analytic continuation of Θ. Sparsity of polynomial chaos expansions of the posterior density Θ in turn implies best N-term approximation rates.

4.4. Best N-term Convergence Rates

Theorem 4.4. *([28]) If Assumptions 3.1, 3.2 and 4.3 hold then there exists a* monotone *sequence* $(\Lambda_N)_{N\in\mathbb{N}} \subset \mathcal{F}$ *of index sets with cardinality not exceeding N (depending σ and on the data δ) such that the corresponding N-term truncated gpc Legendre expansions* Θ_{Λ_N} *in* (22) *satisfy*

$$\|\Theta - \Theta_{\Lambda_N}\|_{L^2(U,\mu_0(dy))} \leq N^{-(\frac{1}{\sigma}-\frac{1}{2})} \|(\theta_\nu)\|_{\ell^\sigma(\mathcal{F};\mathbb{R})} \,.$$

Likewise, for $q = 1, \infty$ *and for every* $N \in \mathbb{N}$, *there exist* monotone *sequences* $(\Lambda_N)_{N\in\mathbb{N}} \subset \mathcal{F}$ *of index sets (depending, in general, on σ, q and the data) whose cardinality does not exceed N such that the N-term truncated Taylor sums* (24) *converge with rate* $1/\sigma - 1$, *i.e.*

$$\|\Theta - T_{\Lambda_N}\|_{L^q(U,\mu_0(dy))} \leq N^{-(\frac{1}{\sigma}-1)} \|(\tau_\nu)\|_{\ell^\sigma(\mathcal{F};\mathbb{R})} \,.$$

Here, for $q = \infty$ the norm $\| \circ \|_{L^\infty(U;\mu_0)}$ is the supremum over all $y \in U$.

5. Sparse Polynomial Quadrature

5.1. Univariate Quadrature and Tensorization

By $(Q^k)_{k\geq 0}$ we denote a sequence of univariate quadrature formulas associated with the quadrature points $(z_j^k)_{j=0}^{n_k} \subset (-1/2, 1/2)$ with $z_j^k \in [-\frac{1}{2}, \frac{1}{2}]\,, \forall j, k$ and $z_0^k = 0\,, \forall k$ and weights $w_j^k\,, 0 \leq j \leq n_k,\ \forall k \in \mathbb{N}_0$.

The univariate quadrature operators associated with the sequences $(z_j^k)_{j=0}^{n_k}$ are of the form

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

where g is a function $\mathrm{g} : [-\frac{1}{2}, \frac{1}{2}] \mapsto \mathcal{S}$, taking values in some Banach space $\mathcal{S}$. In the following, we will work under several assumptions on the quadrature formulas:

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* $I(v_k) = \int_{[-\frac{1}{2},\frac{1}{2}]} v_k(y)\lambda_1(dy)$.

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

Note that Assumption 5.1 (ii) implies $Q^k(1) = 1$ for all k. With the convention that Q^{-1} is the null operator,
i.e.

$$Q^{-1} = 0 \tag{25}$$

and $z_0^0 = 0, w_0^0 = 1$, so that

$$Q^0(\mathrm{g}) = \mathrm{g}(0) \tag{26}$$

we define the univariate quadrature difference operator by

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

Then, the univariate quadrature formula Q^k can be rewritten as telescoping sum

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

where $\mathcal{Z}^k = \{z_j^k : 0 \leq j \leq n_k\} \subset [-\frac{1}{2}, \frac{1}{2}]$ denotes the set of points corresponding to Q^k. We emphasize that so far, arbitrary sequences of quadrature points $(z_j^k)_{j=0}^{n_k}$ can be used (satisfying, however, Assumptions A.1 and A.2). In particular, the ensuing Smolyak construction is not limited to univariate families $(Q^k)_{k\geq 0}$ of quadratures which are based on nested sequences of quadrature points.

For $\nu \in \mathcal{F}$, tensorized multivariate operators are defined by

$$\mathcal{Q}_\nu = \bigotimes_{j\geq 1} Q^{\nu_j}\,, \qquad \Delta_\nu = \bigotimes_{j\geq 1} \Delta_{\nu_j}\,. \tag{27}$$

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 $g_t(\hat{y}) = g(y), y = (\ldots, y_{i-1}, t, y_{i+1}, \ldots)\,, i > 1$ and $y = (t, y_2, \ldots)\,, i = 1$, see [8].

By (25) and (26), since $\nu \in \mathcal{F}$, the countable product quadrature $\mathcal{Q}_\nu$ in (27) is well-defined and corresponds to "anchoring" the integrals at $y_j = 0$ for all $j \notin \mathbb{I}_\nu$, which means that u (and hence p) is independent of these y_j (cf. (13)).

5.2. Sparse Quadrature Operator

For any finite monotone set $\Lambda \subset \mathcal{F}$, the quadrature operator is defined by

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

with associated collocation grid

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

Theorem 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 $\mathbb{P}_{\Lambda_N} = \mathrm{span}\{y^\nu : \nu \in \Lambda_N\}$, 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 $I(g) = \int_U g(y)\mu_0(dy)$.

Proof. We proceed by induction over N.

For $N = 1$, the index set is given by $\Lambda_1 = \{0_\mathcal{F}\}$ (due to the monotonicity of Λ). Given $g \in \mathbb{S}_{\Lambda_1} = \mathrm{span}\{y^\nu : \nu \in \Lambda_1\} \otimes \mathcal{S} = \mathrm{span}\{y^{0_\mathcal{F}}\} \otimes \mathcal{S}$, it follows $g \equiv const. \in \mathcal{S}$ which implies

$$\mathcal{Q}_{\Lambda_1}(g) = \bigotimes_{j\geq 1} \Delta_0(g) = \bigotimes_{j\geq 1} Q^0(g) = g(z_{0_\mathcal{F}}) = I(g)\,.$$

Induction step $N-1 \to N$: For $N > 1$, we assume for any monotone index set Λ_{N-1} it has been shown that

$$\mathcal{Q}_{\Lambda_{N-1}}(g) = I(g), \qquad \forall g \in \mathbb{S}_{\Lambda_{N-1}} = \mathrm{span}\{y^\nu : \nu \in \Lambda_{N-1}\} \otimes \mathcal{S} \ .$$

We denote by $\vartheta \in \mathcal{F} \setminus \Lambda_{N-1}$ the new index so that $\Lambda_N = \Lambda_{N-1} \cup \{\vartheta\}$. Then,

$$\begin{aligned}\mathbb{S}_{\Lambda_N} &= \mathrm{span}\{y^\nu : \nu \in \Lambda_N\} \otimes \mathcal{S} \\ &= \mathrm{span}\{y^\nu : \nu \in \Lambda_{N-1}\} \otimes \mathcal{S} \oplus \mathrm{span}\{y^\vartheta\} \otimes \mathcal{S} \ .\end{aligned}$$

Hence, $g \in \mathbb{S}_{\Lambda_N}$ can be uniquely expressed as $g = g_1 + g_2$ with $g_1 \in \mathbb{S}_{\Lambda_{N-1}}$ and $g_2 \in \mathbb{S}_\vartheta$. Then, we have

$$\begin{aligned}\mathcal{Q}_{\Lambda_N}(g) &= \sum_{\nu \in \Lambda_N} \Delta_\nu(g) = \sum_{\nu \in \Lambda_N} \Delta_\nu(g_1 + g_2) \\ &= \sum_{\nu \in \Lambda_N} \Delta_\nu(g_1) + \sum_{\nu \in \Lambda_N} \Delta_\nu(g_2) = \quad \mathrm{I.} \quad + \quad \mathrm{II.}\end{aligned}$$

The induction hypothesis gives

$$\mathrm{I.} = \mathrm{I}(\mathrm{g}_1).$$

In the following, we show $\mathrm{II.} = \mathrm{I}(\mathrm{g}_2)$ by case distinction.

We first consider the case $|\vartheta| = 1$, i.e. the case when a new dimension is added to the sparse grid Λ_{N-1}. For simplicity of exposition, we assume $\vartheta_1 = 1,\ \vartheta_j = 0,\ j \geq 2$. Therefore, g_2 is of the form

$$g_2(y) = c \cdot y_1,$$

where $c \in \mathcal{S}$ is a constant factor, i.e.

$$g_2 = \bigotimes_{j \geq 1} g_2^j, \qquad g_2^1 \in \mathrm{span}\{y_1\} \otimes \mathcal{S}\ ,\ g_2^j \equiv 1\ , \quad \text{for } j \geq 2\ .$$

We have

$$\begin{aligned}\mathrm{II.} &= \sum_{\nu \in \Lambda_{N-1}} \Delta_\nu(g_2) + \Delta_\vartheta(g_2) \\ &= \sum_{\nu \in \Lambda_{N-1}} (\Delta_{\nu_1} \otimes \bigotimes_{j \geq 2} \Delta_{\nu_j})(g_2) + \Delta_\vartheta(g_2) \\ &= \sum_{\nu \in \Lambda_{N-1}} \Delta_{\nu_1}(g_2^1) \cdot \bigotimes_{j \geq 2} \Delta_{\nu_j}(\otimes_{j \geq 2} g_2^j) + \Delta_\vartheta(g_2) \\ &= Q^0(g_2^1) \sum_{\nu \in \Lambda_{N-1}} \bigotimes_{j \geq 2} \Delta_{\nu_j}(\otimes_{j \geq 2} g_2^j) + \Delta_{\vartheta_1}(g_2^1) \cdot \bigotimes_{j \geq 2} \Delta_{\vartheta_j}(\otimes_{j \geq 2} g_2^j) \\ &= I(\otimes_{j \geq 2} g_2^j) \cdot (Q^0(g_2^1) + Q^1(g_2^1) - Q^0(g_2^1)) \\ &= I(g_2)\ .\end{aligned}$$

We next consider the case $|\vartheta| > 1$:

The function g_2 can expressed by $g_2 = \otimes_{j\geq 1} g_2^j, \quad g_2^j \in \text{span}\{y_j^{\vartheta_j}\} \otimes \mathcal{S}$.

$$\begin{aligned}
\text{II.} &= \sum_{\nu\in\Lambda_{N-1}} \Delta_\nu(g_2) + \Delta_\vartheta(g_2) \\
&= \sum_{\nu\in\Lambda_{N-1}} \Delta_\nu(g_2) + \bigotimes_{j\geq 1}(\Delta_{\vartheta_j}(g_2^j)) \\
&= \sum_{\nu\in\Lambda_{N-1}} \bigotimes_{j\geq 1} \Delta_{\nu_j}(g_2) + \bigotimes_{j\geq 1}((Q^{\vartheta_j} - Q^{\vartheta_j - 1})(g_2^j)) \\
&= \sum_{\nu\in\Lambda_{N-1}} \bigotimes_{j\geq 1}(Q^{\nu_j}(g_2^j) - Q^{\nu_j - 1}(g_2^j)) + \bigotimes_{j\geq 1}(I(g_2^j) - Q^{\vartheta_j - 1}(g_2^j))
\end{aligned}$$

Due to the monotonicity of Λ_N, $\vartheta_j - 1 \in \Lambda_{N-1}$ and with $Q^\nu = \sum_{\mu\leq\nu}\Delta_\mu$, it holds

$$\text{II.} = \mathrm{I}(\mathrm{g}_2)$$

Thus, we have

$$\mathcal{Q}_{\Lambda_N}(g) = \sum_{\nu\in\Lambda_N}\Delta_\nu(g) = \sum_{\nu\in\Lambda_N}\Delta_\nu(g_1) + \sum_{\nu\in\Lambda_N}\Delta_\nu(g_2) = I(g_1) + I(g_2) = I(g)\ .$$

□

We are now in position to state the main theorem in this section. It takes the form of a convergence result for quadrature of functions of countably many variables, fulfilling the following assumptions.

Assumption 5.3.
Let $g : U \mapsto S$ denote a bounded, continuous function of countably many variables $y_1, y_2, ...$ which is defined on $U = [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}$ and which admits an extension to the complex domain, i.e. $g : \mathcal{U} \mapsto S$, with $\mathcal{U} := \bigotimes_{j\geq 1}\{z_j \in \mathbb{C} : |z_j| \leq \frac{1}{2}\}$, and $\mathcal{S}$ is replaced by its extension to the coefficient field $\mathbb{C}$. This extension (again denoted by g) satisfies:

(i) *g admits an analytic extension to the polydisc $\mathcal{U}_\rho := \{z = (z_j)_{j\geq 1} \in \mathbb{C}^{\mathbb{N}} : |z_j| \leq \frac{1}{2}\rho_j, j \in \mathbb{N}\}$ for a vector $\rho = (\rho_j)_{j\geq 1}$ of radii $\rho_j > 1$.*

(ii) *The function $g : \mathcal{U}_\rho \mapsto \mathcal{S}$ satisfies an a priori estimate*

$$\sup_{z\in\mathcal{U}_\rho} \|g(z)\|_{\mathcal{S}} \leq B(\rho)\,,$$

where the ρ-dependent bound satisfies for every $J \in \mathbb{N}$

$$B(\rho_1, \dots, \rho_J, 1, \dots) \leq B_0 \prod_{j=1}^{J} e^{\alpha_j \rho_j}$$

for some constant B_0 independent of ρ and J for certain positive real numbers α_j.

(iii) *The poly-radii ρ satisfy*

$$\sum_{j\geq 1}\rho_j L_j < \infty$$

for some fixed sequence $(L_j)_{j\geq 1}$ of positive real numbers and it holds

$$(\alpha_j)_{j\in\mathbb{N}}, (L_j)_{j\in\mathbb{N}} \in l^\sigma(\mathbb{N}), \qquad 0<\sigma<1\,.$$

The sequence ρ is called admissible poly-radii. We say that an admissible sequence of poly-radii ρ is (b,r)-admissible, if $\sum_{j\geq 1}\rho_j\alpha_j \leq b-r$ with a constant b and some $r>0$.

The Assumption 5.3 implies that g can be represented by the Taylor expansion,

$$g(z) = \sum_{\nu\in\mathcal{F}} \tau_\nu z^\nu\,, \quad \tau_\nu := \frac{1}{\nu!}\partial_z^\nu g(z)\big|_{z=0}\,, \qquad \forall z\in\mathcal{U}_\rho\,,$$

which is unconditionally summable in $\mathcal{U}_\rho$ and for any (not necessarily monotone sequence) $(\Lambda_N)_{N\in\mathbb{N}}\subset\mathcal{F}$ which exhausts $\mathcal{F}$, the corresponding sequence of N-term truncated partial Taylor sums

$$T_{\Lambda_N}(g(z)) := \sum_{\nu\in\Lambda_N}\tau_\nu z^\nu \tag{28}$$

converges pointwise in $\mathcal{U}_\rho$ to g. In [15], it is shown in this general setting that the sequence $(\tau_\nu)_{\nu\in\mathcal{F}}$ of Taylor coefficients $\tau_\nu = \frac{1}{\nu!}\partial^\nu g(0)\in\mathcal{S}$ of the function g is σ-summable, i.e. $(\|\tau_\nu\|_\mathcal{S})_{\nu\in\mathcal{F}}\in l^\sigma(\mathcal{F})$. Due to Stechkin's Lemma (see e.g. [10], Section 3.3), we have

$$\sup_{z\in\mathcal{U}_\rho}\|g(z)-T_{\Lambda_N}(g(z))\|_\mathcal{S} \leq C\,N^{-s}\,, \quad s:=\frac{1}{\sigma}-1\,.$$

Further, this rate is achieved even if the sequence $(\Lambda_N)_{N\in\mathbb{N}}$ of index sets are constrained to be monotone index sets, cf. [7].

We will now show that there exists a sequence $(\Lambda_N)_{N\geq 1}$ of monotone sets $\Lambda_N\subset\mathcal{F}$ such that $\#\Lambda_N\leq N$ and such that, for some $C>0$ independent of N,

$$\|I(g)-\mathcal{Q}_{\Lambda_N}(g)\|_\mathcal{S}\leq CN^{-s}\,, \qquad s=\frac{1}{\sigma}-1\,,$$

with $I(g)=\int_U g(y)\mu_0(dy)$. Since $\mathcal{Q}_{\Lambda_N}$ is exact for every $\Upsilon_N\in\mathbb{S}_{\Lambda_N}$, we have

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

where we used $|\!|\!|I|\!|\!|=\mu_0(U)=1$ and $|\!|\!|\mathcal{Q}_{\Lambda_N}|\!|\!| =: C_{\mathcal{Q}_{\Lambda_N}}$. Using the triangle inequality gives

$$C_{\mathcal{Q}_{\Lambda_N}} = |\!|\!|\sum_{\nu\in\Lambda_N}\bigotimes_{j\geq 1}\Delta_{\nu_j}|\!|\!| \leq \sum_{\nu\in\Lambda_N}\prod_{j\geq 1}|\!|\!|\Delta_{\nu_j}|\!|\!|\,,$$

with

$$\begin{aligned}|\!|\!|\Delta_{\nu_j}|\!|\!| &= \sup_{0\neq \mathrm{g}\in C([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}\frac{\|\Delta_{\nu_j}\mathrm{g}\|_\mathcal{S}}{\|\mathrm{g}\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}} \leq |\!|\!|Q^{\nu_j}|\!|\!| + |\!|\!|Q^{\nu_j-1}|\!|\!|\\ &=: c_{\nu_j}+c_{\nu_j-1}\,,\end{aligned}$$

with $c_k = 1\,, k \geq 0$ (note that $\|\!|Q^k|\!\| = 1$ by Assumption 5.1 (ii)) and $c_{-1} := 0$. Thus, we get the bound

$$C_{\mathcal{Q}_{\Lambda_N}} \leq \sum_{\nu \in \Lambda_N} \prod_{j \geq 1} (c_{\nu_j} + c_{\nu_j - 1})\,. \tag{29}$$

Bound (29) is pessimistic in the tensor product case, since in this case the exact value of the constant $C_{\mathcal{R}_\nu}$ is given by

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

Lemma 5.4. *The constant $C_{\mathcal{Q}_\Lambda}$ satisfies*

$$C_{\mathcal{Q}_\Lambda} \leq \#\Lambda^{\log_2 3}\,.$$

Proof. For $\nu \in \Lambda$, we have

$$\prod_{j \geq 1} (c_{\nu_j} + c_{\nu_j - 1}) = \prod_{j \in \mathbb{I}\nu} (1 + 1) = 2^{\#\mathbb{I}_\nu}\,.$$

In order to derive a bound of the maximum cardinality of $\mathbb{I}_\nu\,,\ \nu \in \Lambda$ for arbitrary monotone index sets Λ, we consider a binary monotone index set $\tilde{\Lambda}$ of the following form

$$\tilde{\Lambda} \subset \{\nu \in \{0,1\}^{\mathbb{J}} : |\nu|_1 < \infty\} \quad \text{and} \quad \#\tilde{\Lambda} = \#\Lambda.$$

The support of each index $\nu \in \Lambda$ can be bounded by $\max_{\nu \in \tilde{\Lambda}} \#\mathbb{I}_\nu$ due to the monotonicity of Λ and $\tilde{\Lambda}$. Therefore, it holds

$$\max_{\nu \in \Lambda} \#\mathbb{I}_\nu \leq \max_{\nu \in \tilde{\Lambda}} \#\mathbb{I}_\nu.$$

The monotonicity of the index set $\tilde{\Lambda}$ implies that all predecessors of each index $\nu \in \tilde{\Lambda}$ are contained in $\tilde{\Lambda}$, i.e. it holds that $\nu - e_j \in \tilde{\Lambda}$ for all $j \in \mathbb{I}_\nu$, cf. Definition 4.2. Therefore, we can associate the binary monotone index set $\tilde{\Lambda}$ with a complete binary tree, where the number of leaf nodes corresponds to the cardinality of the index set $\tilde{\Lambda}$ and a path from the root to the leaf nodes represents an index $\nu \in \tilde{\Lambda}$. Thus, the maximum cardinality of $\mathbb{I}_\nu$ with $\nu \in \tilde{\Lambda}$ can be bounded by the height of a complete binary tree. We have $\max_{\nu \in \Lambda} \#\mathbb{I}_\nu \leq \lfloor \log_2 \#\Lambda \rfloor$ and hence, we may estimate

$$\begin{aligned} C_{\mathcal{Q}_\Lambda} = \sum_{\nu \in \Lambda} \prod_{j \geq 1} (c_{\nu_j} + c_{\nu_j - 1}) &\leq \sum_{k=0}^{\lfloor \log_2 \#\Lambda \rfloor} \binom{\lfloor \log_2 \#\Lambda \rfloor}{k} \cdot 2^k \\ &= (1+2)^{\lfloor \log_2 \#\Lambda \rfloor} = 3^{\lfloor \log_2 \#\Lambda \rfloor} \leq \#\Lambda^{\log_2 3}\,. \end{aligned}$$

□

The preceding result allows to derive a convergence estimate of the form

$$\|I(g) - \mathcal{Q}_{\Lambda_N}(g)\|_{\mathcal{S}} \leq C N^{-s + \log_2 3}\,, \qquad s = \frac{1}{\sigma} - 1\,.$$

Relating the quadrature error with the Taylor coefficients τ_ν of $u = \sum_{\nu \in \mathcal{F}} \tau_\nu y^\nu$, the approximation error can be bounded as follows:

Lemma 5.5. *Assume that $T_{\Lambda_N} g(y) = \sum_{\nu\in\mathcal{F}} \tau_\nu y^\nu$ in the sense of unconditional convergence in $L^\infty(U, \mathcal{S})$. Then, we have*

$$\|I(g) - \mathcal{Q}_\Lambda(g)\|_{\mathcal{S}} \le 2 \cdot \sum_{\nu\notin\Lambda} p_\nu \|\tau_\nu\|_{\mathcal{S}}$$

for any monotone set Λ, where $p_\nu := \prod_{j\ge1}(1+\nu_j)^2$.

Proof. Due to the unconditional convergence of the Taylor series, we have

$$\mathcal{Q}_\Lambda(g) = \mathcal{Q}_\Lambda(\sum_{\nu\in\mathcal{F}} \tau_\nu y^\nu) = \sum_{\nu\in\mathcal{F}} \tau_\nu \mathcal{Q}_\Lambda(y^\nu) = \sum_{\nu\in\Lambda} \tau_\nu I(y^\nu) + \sum_{\nu\notin\Lambda} \tau_\nu \mathcal{Q}_{\Lambda\cap\mathcal{R}_\nu}(y^\nu)\,.$$

Therefore,

$$I(g) - \mathcal{Q}_\Lambda(g) = \sum_{\nu\notin\Lambda} \tau_\nu (I - \mathcal{Q}_{\Lambda\cap\mathcal{R}_\nu})(y^\nu)\,,$$

which results in

$$\|I(g) - \mathcal{Q}_\Lambda(g)\|_{\mathcal{S}} \le \sum_{\nu\in\mathcal{F}\setminus\Lambda} \|\tau_\nu\|_{\mathcal{S}}(1 + C_{\mathcal{Q}_{\Lambda\cap\mathcal{R}_\nu}}) \le 2 \cdot \sum_{\nu\notin\Lambda} \|\tau_\nu\|_{\mathcal{S}} \underbrace{(\#\mathcal{R}_\nu)^2}_{=p_\nu}\,.$$

□

Theorem 5.6. *If Assumption 5.3 is satisfied, $(p_\nu\|\tau_\nu\|_{\mathcal{S}})_{\nu\in\mathcal{F}} \in l^\sigma(\mathcal{F})$.*

Proof. Due to the summability of $(\alpha_j)_{j\in\mathbb{N}}$ and $(L_j)_{j\in\mathbb{N}}$, it holds

$$\sum_{j\ge1} \alpha_j \le b - r$$

for some sufficiently large constant b and $r > 0$, so that we may choose an integer $J_0 \in \mathbb{N}$ such that

$$\sum_{j>J_0} \gamma_j \le \frac{1}{e^2}\frac{r}{12}$$

with $\gamma_j = \max(\alpha_j, L_j)$, where it is assumed that the indexing is chosen such that the sequence $(\gamma_j)_{j\ge1}$ is non-decreasing. We follow [15] and present the details. We set $E := \{j : 1 \le J_0\}$ and $F := \mathbb{N} \setminus E$ and choose $\kappa > 1$ such that

$$(\kappa - 1)\sum_{j\le J_0} \gamma_j \le \frac{r}{6}\,.$$

For each $\nu \in \mathcal{F}$, we denote by ν_E and ν_F the restrictions of ν on E and F and define the sequence $\rho = \rho(\nu)$ by

$$\rho_j = \kappa\,, \quad j \in E\,; \qquad \rho_j = \max(1, \frac{r\nu_j}{4|\nu_F|\gamma_j}) + e^2\,, \quad j \in F\,,$$

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

Then, the sequence ρ is $(\frac{r}{2})$-admissible, since

$$
\begin{aligned}
\sum_{j\geq 1}\rho_j\alpha_j &= \kappa\sum_{j\leq J_0}\alpha_j + \sum_{j>J_0}\left(\max(1,\frac{r\nu_j}{4|\nu_F|\gamma_j}) + e^2\right)\alpha_j \\
&\leq (\kappa-1)\sum_{j\leq J_0}\gamma_j + \sum_{j\leq J_0}\alpha_j + \sum_{j>J_0}(1+\frac{r\nu_j}{4|\nu_F|\gamma_j})\alpha_j + \sum_{j>J_0}e^2\alpha_j \\
&\leq \frac{r}{6} + \sum_{j\geq 1}\alpha_j + \frac{r}{4} + \frac{r}{12} \leq b - \frac{r}{2}\,.
\end{aligned}
$$

Hence, we have

$$
p_\nu\|\tau_\nu\|_{\mathcal{S}} \leq C_{\frac{r}{2}}\prod_{j\leq J_0}\frac{(1+\nu_j)^2}{\kappa^{\nu_j}}\prod_{j>J_0}\frac{(1+\nu_j)^2}{\rho^{\nu_j}}\,.
$$

Similar to the proof of Theorem 4.2 in [8], we get

$$
p_\nu\|\tau_\nu\|_{\mathcal{S}} \leq C\cdot\alpha(\nu_E)\cdot\beta(\nu_F)\,,
$$

where

$$
\alpha(\nu_E) := \prod_{j\leq J_0}\left(\frac{1+\kappa}{2\kappa}\right)^{\nu_j}\,,\quad \beta(\nu_F) := \prod_{j>J_0}\left(\frac{|\nu_F|d_j}{\nu_j}\right)^{\nu_j}
$$

with $d_j := \frac{4e^2\gamma_j}{r}$. It holds $\sum_{j>J_0} d_j < \frac{1}{3}$. The desired result follows then with Subsection 3.2 in [10]. □

Theorem 5.7. *Let Assumption 5.3 be satisfied, then there exists a sequence* $(\Lambda_N)_{N\geq 1}$ *of monotone sets* $\Lambda_N \subset \mathcal{F}$ *such that* $\#\Lambda_N \leq N$ *and*

$$
\|I(g) - \mathcal{Q}_{\Lambda_N}(g)\|_{\mathcal{S}} \leq CN^{-s}\,,\qquad s = \frac{1}{\sigma} - 1\,.
$$

Proof. The l^σ-summability of $(p_\nu\|\tau_\nu\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$ implies the existence of a sequence $(\Lambda_N)_{N\geq 1}$ of sets $\Lambda_N \subset \mathcal{F}$ with $\#\Lambda_N \leq n$ such that

$$
\sum_{\nu\notin\Lambda_N} p_v\|\tau_v\|_{\mathcal{S}} \leq CN^{-s}\,,\qquad s = \frac{1}{\sigma} - 1\,.
$$

With the same argument as in [8] Section 4.2, the sequence $(\Lambda_N)_{N\geq 1}$ can be chosen to be monotone and nested. □

Based on the results discussed in Sections 3 and 4, we next show that the functions of the forward problem $G(\cdot)\,,\mathcal{G}(\cdot)$, the potential $\Phi(u(\cdot);\delta)$ as well as the posterior $\Theta(\cdot)$ and $\Psi(\cdot)$ of the model parametric elliptic problem defined in Section 3 fulfill Assumption 5.3, and, hence, the quantities of interest Z and Z' can be efficiently approximated with convergence rate s.

Since the sparsity and quadrature error analysis is based on arguments from complex analysis, we extend (for purposes of error analysis only) all functions to the complex domain. Accordingly, we replace $\mathbf{UEA}(a_{\text{MIN}}, a_{\text{MAX}})$ by a complex-valued counterpart: **Uniform Ellipticity Assumption in $\mathbb{C}$:** *there exist* $0 < a_{\text{MIN}} \leq a_{\text{MAX}} < \infty$ *such that for all* $x \in D$ *and for all* $z \in \mathcal{U}$

$$0 < a_{\text{MIN}} \leq \Re(u(x,z)) \leq |u(x,z)| \leq a_{\text{MAX}} < \infty.$$

We refer to [9, 10] for a detailed derivation of the ensuing results.

Lemma 5.8. *Under* $\mathbf{UEAC}(\mathbf{a}_{\text{MIN}}, \mathbf{a}_{\text{MAX}})$ *and if* $(\|\psi_j\|_{L^\infty(D)})_{j\leq 1} \in l^\sigma(\mathbb{N})$*, the solution* $G : U \to V$ *of the parametric forward problem (9) and* $\mathcal{G} : U \to \mathbb{R}^k$ *satisfy the Assumption 5.3.*

Proof. In [10], it is shown that $\mathbf{UEAC}(\mathbf{a}_{\text{MIN}}, \mathbf{a}_{\text{MAX}})$ implies analyticity of the V-valued forward map $z \mapsto G(u(z))$ in $\mathcal{U}_\rho := \{z = (z_j)_{j\geq 1} \in \mathbb{C}^{\mathbb{N}} : |z_j| \leq \frac{1}{2}\rho_j, j \in \mathbb{N}\}$ for a vector $\rho = (\rho_j)_{j\geq 1}$ of $(\Re(\bar{a}(x)), r)$-admissible radii $\rho_j > 1$ satisfying the estimate

$$\sup_{z\in\mathcal{U}_\rho} \|G(z)\|_V \leq \frac{\|f\|_{V^*}}{r},$$

where $0 < r < a_{\text{MIN}}$. Choosing $\alpha_j = \frac{1}{2}|\psi_j(x)|, x \in D$, $L_j = \frac{1}{2}\|\psi_j\|_{L^\infty(D)}$ and $b = \Re(\bar{a}(x))$, 5.3 (i)-(v) are fulfilled.

For every bounded linear observation operator $\mathcal{O} \in V^*$, $\mathcal{G}$ admits an analytic continuation to the domain $\mathcal{U}_\rho$ with

$$\sup_{z\in\mathcal{U}_\rho} |\mathcal{G}(z)| \leq \frac{\|f\|_{V^*}}{r} \sum_{k=1}^{K} \|o_k\|_{V^*},$$

cf. [28]. With the same choice of $\alpha_j = \frac{1}{2}|\psi_j(x)|, x \in D$, $L_j = \frac{1}{2}\|\psi_j\|_{L^\infty(D)}$ and with $b = \Re(\bar{a}(x))$, $\mathcal{G}$ satisfies 5.3 (i)-(v). □

Lemma 5.9. *Under* $\mathbf{UEAC}(\mathbf{a}_{\text{MIN}}, \mathbf{a}_{\text{MAX}})$ *and if* $(\|\psi_j\|_{L^\infty(D)})_{j\geq 1} \in l^\sigma(\mathbb{N})$*, the functions* $\Phi(u(\cdot);\delta), \Theta$ *and* Ψ *(with* $\phi(u(z)) = G(u(z))^m$*) satisfy the Assumption 5.3.*

Proof. As was shown in [28], due to $\mathbf{UEAC}(\mathbf{a}_{\text{MIN}}, \mathbf{a}_{\text{MAX}})$ and $(\|\psi_j\|_{L^\infty(D)})_{j\leq 1} \in l^\sigma(\mathbb{N})$, it follows with Lemma 5.8 that G and $\mathcal{G}$ are analytic in a domain $\mathcal{U}_\rho := \{z = (z_j)_{j\geq 1} \in \mathbb{C}^{\mathbb{N}} : |z_j| \leq \frac{1}{2}\rho_j, j \in \mathbb{N}\}$ for a vector $\rho = (\rho_j)_{j\geq 1}$ of $(\Re(\bar{a}(x)), r)$-admissible radii $\rho_j > 1, 0 < r < a_{\text{MIN}}$, so that $\Phi(u(\cdot);\delta)$ is analytic as a composition of a quadratic (hence analytic) and an analytic function. Further, the domain of the analytic continuation of the potential $\Phi(u(\cdot);\delta)$ coincides with the domain of holomorphy of G and $\mathcal{G}$, respectively. Based on Lemma 5.8, as in [28], we can derive the following bound

$$\sup_{z\in\mathcal{U}_\rho} |\Phi(u(z);\delta)| \leq C(\Gamma,\delta) \frac{\|f\|^2_{V^*}}{r^2} \sum_{k=1}^{K} \|o_k\|^2_{V^*},$$

with a constant $C(\Gamma,\delta)$ depending on the observations δ and with the covariance operator Γ, $0 < r < a_{\text{MIN}}$. Then, due to the analyticity of the exponential function, it follows that the posterior density Θ admits analytic continuations to the same domain of holomorphy as G and it holds

$$\sup_{z\in\mathcal{U}_\rho} |\Theta(z)| \le \exp\left(C(\Gamma,\delta)\frac{\|f\|_{V^*}^2}{r^2}\sum_{k=1}^{K}\|o_k\|_{V^*}^2 \right) .$$

Thus, the analyticity of Ψ, again with the same domain of holomorphy, follows from the analyticity of the product of two analytic functions. The analytic continuation of Ψ admits the bounds

$$\sup_{z\in\mathcal{U}_\rho} \|\Theta(z)(p(z))^m\|_{V^{(m)}} \le \frac{\|f\|_{V^*}^m}{r^m}\exp\left(C(\Gamma,\delta)\frac{\|f\|_{V^*}^2}{r^2}\sum_{k=1}^{K}\|o_k\|_{V^*}^2 \right) .$$

Choosing again $\alpha_j = \frac{1}{2}|\psi_j(x)|\,, x\in D$, $L_j = \frac{1}{2}\|\psi_j\|_{L^\infty(D)}$ and $b = \Re(\bar{a}(x))$, 5.3 (i)-(v) are fulfilled for $\Phi(u(\cdot);\delta), \Theta$ and Ψ. □

5.3. *Adaptive Construction of the Monotone Index Set*

We now discuss the adaptive construction of the monotone index set $(\Lambda_N)_{N\ge1}$. The results in the previous subsection ensure the existence of a nested sequence of monotone index sets $(\Lambda_N)_{N\ge1}$, which exhausts $\mathcal{F}$, such that the sparse quadrature operator $\mathcal{Q}_{\Lambda_N} = \sum_{\nu\in\Lambda_N}\Delta_\nu$ applied to the function Θ and Ψ given by (6) and (7) to compute the infinite dimensional parametric integrals of the form

$$Z = \int_{y\in U}\Theta(y)\mu_0(dy)$$

and

$$Z' = \int_{y\in U}\Psi(y)\mu_0(dy)$$

gives

$$|Z - \mathcal{Q}_{\Lambda_N}(\Psi)| \le C_Z N^{-s}\,, \qquad s = \frac{1}{\sigma} - 1\,, \tag{30}$$

$$\|Z' - \mathcal{Q}_{\Lambda_N}(\Psi)\|_{V^{(m)}} \le C_{Z'} N^{-s}\,, \qquad s = \frac{1}{\sigma} - 1\,, \tag{31}$$

respectively, with $\#\Lambda_N \le N$.

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 and Z', i.e. the N largest

$$\|\Delta_\nu(\mathcal{X})\|_{\mathcal{S}} = \|\bigotimes_{j\ge1}\Delta_{\nu_j}(\mathcal{X})\|_{\mathcal{S}}\,, \quad \nu\in\mathcal{F}$$

with $\mathcal{X} = \Theta\,,\ \mathcal{S} = \mathbb{R}$ or $\mathcal{X} = \Psi\,,\ \mathcal{S} = V^{(m)}$, minimizing the approximation error (30) and (31), respectively (cf. [8, 16, 13]).

Note that the approach leads to a nested but not necessarily monotone index set Λ and, further, that a priori knowledge on the size of $\|\Delta_\nu(\Psi)\|_{V^{(m)}}$ is required for this approach which is generally not available.

Therefore, following [8], we consider a (finite) set of reduced neighbors

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

for any monotone set Λ, where $j(\Lambda) = \max\{j : \nu_j > 0 \text{ for some } \nu \in \Lambda\}$. The index set Λ is adaptively chosen as follows:

---

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

This greedy strategy attempts to control the global approximation error by locally collecting indices of the current set of reduced neighbors with the largest error contributions. The convergence of the algorithm with a convergence rate comparable to that of an optimal choice has not yet been proven and there exist test cases, where this approach may fail to converge (cf. [8]). However, excellent results based on this approach could be observed in many numerical experiments, see e.g.[8, 7, 16, 13].

In order to construct the sparse quadrature operator, we consider three choices for the univariate sequence $(z_j^k)_{j=0}^{n_k}$ of quadrature points

- Clenshaw-Curtis (CC),

  $$z_j^k = -cos\left(\frac{\pi j}{n_k - 1}\right)\,, j = 0, \dots, n_k - 1, \text{if } n_k > 1 \text{ and}$$
  $$z_0^k = 0\,, \text{if } n_k = 1$$

  with $n_0 = 1$ and $n_k = 2^k + 1$, for $k \geq 1$

- symmetrized Leja abscissas (L),

  $$z_0^k = 0\,, z_1^k = 1\,, z_2^k = -1\,, \text{if } j = 0, 1, 2 \text{ and}$$

$$z_j^k = \text{argmax}_{z\in[-1,1]} \prod_{l=1}^{j-1} |z - z_l^k|\,, j = 3, \ldots, n_k, \text{if } j \text{ odd}\,,$$
$$z_j^k = -z_{j-1}^k\,, j = 3, \ldots, n_k, \text{if } j \text{ even}$$

with $n_k = 2 \cdot k + 1$, for $k \geq 0$.

- $\mathfrak{R}$-Leja sequence (RL),
projection on $[-1, 1]$ of a Leja sequence for the complex unit disk initiated at i

$$z_0^k = 0\,, z_1^k = 1\,, z_2^k = -1\,, \text{if } j = 0, 1, 2 \text{ and} \tag{32}$$
$$z_j^k = \mathfrak{R}(\hat{z}), \text{ with } \hat{z} = \text{argmax}_{|z|\leq 1} \prod_{l=1}^{j-1} |z - z_l^k|\,, j = 3, \ldots, n_k, \text{if } j \text{ odd}\,,$$
$$z_j^k = -z_{j-1}^k\,, j = 3, \ldots, n_k, \text{if } j \text{ even}\,,$$

with $n_k = 2 \cdot k + 1$, for $k \geq 0$, see [5].

The Clenshaw-Curtis points fulfill Assumption 5.1, whereas the quadrature weights based on the Leja sequences are not all positive, that means Assumption 5.1 (ii) is not satisfied. However, the positivity assumption on the quadrature weights can be weakened to the case of $\mathfrak{R}$-Leja points due to the moderate, algebraic growth of the Lebesgue constants of the univariate quadrature operators (cp. [4, 5, 6]).

Lemma 5.10. *Let $\mathcal{Q}_\Lambda^{RL}$ denote the sparse quadrature operator for any monotone set Λ based on the univariate quadrature formulas associated with the Leja sequence defined by (32). Under Assumption 5.3, it follows that there exists a sequence $(\Lambda_N)_{N\geq 1}$ of monotone sets Λ_N such that $\#\Lambda_N \leq N$ and*

$$\|I(g) - \mathcal{Q}_{\Lambda_N}^{RL}(g)\|_{\mathcal{S}} \leq CN^{-s}\,, \qquad s = \frac{1}{\sigma} - 1\,. \tag{33}$$

Proof. We denote by $\mathcal{I}_{RL}^k$ the univariate polynomial interpolation operator of the form

$$\mathcal{I}_{RL}^k(g) = \sum_{i=0}^{n_k} \mathrm{g}\left(\frac{z_i^k}{2}\right) \cdot l_i^k\,,$$

associated with the Leja sequence $(z_j^k)_{j=0}^{n_k}$ defined by (32), where g is a function $\mathrm{g} : [-\frac{1}{2}, \frac{1}{2}] \mapsto \mathcal{S}$, taking values in some Banach space $\mathcal{S}$ and $l_i^k(y) := \prod_{i=0 i\neq j}^{n_k} \frac{2y - z_i}{z_j - z_i}$ are the Lagrange polynomials. Since we have

$$\mathcal{I}_{RL}^k(v_k) = v_k\,, \quad \forall v_k \in \mathbb{S}_k := \mathbb{P}_k \otimes \mathcal{S}\,, \mathbb{P}_k = \text{span}\{y^j : j \in \mathbb{N}_0, j \leq k\}\,,$$

the univariate quadrature operators Q_{RL}^k associated with the sequences defined by (32) satisfy

$$(I - Q_{RL}^k)(v_k) = (I - I[\mathcal{I}_{RL}^k])(v_k) = I(v_k - \mathcal{I}_{RL}^k(v_k)) = 0$$

$\forall v_k \in \mathbb{S}_k := \mathbb{P}_k \otimes \mathcal{S}\,, \mathbb{P}_k = \text{span}\{y^j : j \in \mathbb{N}_0, j \leq k\}$. Hence, Assumption 5.1 (i) is fulfilled.

In [4, 5], an asymptotic bound

of the Lebesgue constants for Lagrange interpolation in Leja points is proved. It takes the form

$$|||\mathcal{I}_{RL}^k||| = \sup_{0\neq g\in C([-\frac{1}{2},\frac{1}{2}];\mathcal{S})} \frac{\|\mathcal{I}_{RL}^k(\mathrm{g})\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}}{\|\mathrm{g}\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}} \leq Ck^3 \log k\ .$$

This bound has been improved in [6] to

$$|||\mathcal{I}_{RL}^k||| \leq 3(k+1)^2 \log(k+1)\ .$$

Therefore, we have

$$\begin{aligned} |||Q_{RL}^k||| &= \sup_{0\neq g\in C([-\frac{1}{2},\frac{1}{2}];\mathcal{S})} \frac{\|Q_{RL}^k(\mathrm{g})\|_{\mathcal{S}}}{\|\mathrm{g}\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}} \\ &\leq \sup_{0\neq g\in C([-\frac{1}{2},\frac{1}{2}];\mathcal{S})} \frac{\|\mathcal{I}_{RL}^k(\mathrm{g})\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}}{\|\mathrm{g}\|_{L^\infty([-\frac{1}{2},\frac{1}{2}];\mathcal{S})}} \leq 3(k+1)^2 \log(k+1)\ . \end{aligned}$$

Since $|||Q_{RL}^k||| \leq (k+1)^\theta$ for some $\theta \geq 1$, we obtain from Lemma 3.1 and Lemma 4.1 in [8]

$$\|I(g) - \mathcal{Q}_\Lambda^{RL}(g)\|_{\mathcal{S}} \leq 2 \cdot \sum_{\nu\notin\Lambda} p_\nu \|\tau_\nu\|_{\mathcal{S}}$$

for any monotone set Λ, where $p_\nu := \prod_{j\geq 1}(1+\nu_j)^{\theta+1}$ and $T_{\Lambda_N} g(y) = \sum_{\nu\in\mathcal{F}} \tau_\nu y^\nu$ denotes the Taylor expansion of g, cp. (28). The summability of the sequence $(p_\nu\|\tau_\nu\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$, i.e. $(p_\nu\|\tau_\nu\|_{\mathcal{S}})_{\nu\in\mathcal{F}} \in l^\sigma(\mathcal{F})$, can be proven with the same arguments as in Theorem 5.6, which implies the existence of a sequence $(\Lambda_N)_{N\geq 1}$ of sets $\Lambda_N \subset \mathcal{F}$ with $\#\Lambda_N \leq N$ such that (33) holds. □

Remark 5.11. As the proof of Lemma 5.10 indicates, Assumption 5.1 (ii) can be relaxed to the condition $|||Q^k||| \leq Ck^\theta$ for a given $\theta \geq 1$. The polynomial bound of the univariate quadrature operators allows to derive a similar algebraic bound for the operator norm of the sparse quadrature quadrature $\mathcal{Q}_\Lambda$, for any monotone set Λ, of the form $|||\mathcal{Q}_\Lambda||| \leq (\#\Lambda)^{\theta+1}$, cp. Lemma 3.1 in [8]. By slight modifications of the arguments presented in Section 5, the quadrature error can be related to the Taylor coefficients of the underlying function g denoted by $T_{\Lambda_N} g(y) = \sum_{\nu\in\mathcal{F}} \tau_\nu y^\nu$ in the following way: $\|I(g) - \mathcal{Q}_\Lambda(g)\|_{\mathcal{S}} \leq 2 \cdot \sum_{\nu\notin\Lambda} p_\nu \|\tau_\nu\|_{\mathcal{S}}$ for *any* monotone set Λ, with $p_\nu := \prod_{j\geq 1}(1+\nu_j)^{\theta+1}$ and with u satisfying Assumption 5.3. The l^σ summability of the sequence $(p_\nu\|\tau_\nu\|_{\mathcal{S}})_{\nu\in\mathcal{F}}$, which follows with the same arguments as in the proof of Theorem 5.6, implies then the desired convergence result $\|I(g) - \mathcal{Q}_{\Lambda_N}(g)\|_{\mathcal{S}} \leq CN^{-s}$ with $s = \frac{1}{\sigma} - 1$.

For $0 < \sigma < 2/3$, this analysis yields $s > 1/2$, i.e. the rate of convergence is, asymptotically, superior to the rate afforded by MCMC, in terms of the number N of solutions of the forward problem. As we shall see in the numerical experiments below, in fact even $s > 1/\sigma$ appears to hold. This, in turn, would imply the rate $s > 1/2$ even for $\sigma < 2$.

6. Numerical Experiments

We consider the model parametric elliptic boundary value problem

$$-\mathrm{div}(u\nabla p) = f \quad \text{in } D := [0,1]\,,\ p = 0 \quad \text{in } \partial D\,,$$

with $f(x) = 100 \cdot x$. The diffusion coefficient is defined as

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

with $D_j = [(j-1)\frac{1}{64}, j\frac{1}{64}]$, $y = (y_j)_{j=1,\ldots,64}$ and $\alpha_j = \frac{1.8}{j^\zeta}, \zeta = 2,3,4$. The forward problem is numerically solved by a finite element method using continuous, piecewise linear ansatz functions on a uniform mesh with meshwidth $h = 2^{-18}$. LAPACK's DPTSV routine is used to compute the solution of the resulting symmetric positive definite tridiagonal system.

The goal of computation is, for given (noisy) data δ,

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

with $\eta \sim \mathcal{N}(0,\Gamma)$ and $\mathcal{G} : L^\infty(D) \to \mathbb{R}^K$, with $K = 2^{N_K} - 1\,,\ N_K = 2,3,4$, to compute the expectation of the observed solution of the forward model, i.e. our aim is to approximate

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

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

$$Z = \int_U \exp\bigl(-\Phi(u;\delta)\bigr)\Big|_{u=\bar{a}+\sum_{j=1}^{64} y_j\psi_j}\mu_0(dy)\,,$$

so that the expectation of interest is given by Z'/Z. The noise $\eta = (\eta_j)_{j=1,\ldots,K}$ is assumed to be independent and identically distributed with $\eta_j \sim \mathcal{N}(0,1)$, $\eta_j \sim \mathcal{N}(0,0.5^2)$ and $\eta_j \sim \mathcal{N}(0,0.1^2)$, respectively. The observation operator $\mathcal{O}$ consists of K system responses at K equispaced observation points in D with spacing $\tau_N^{\mathcal{O}} = 2^{-N}$ at $x_k = \frac{k}{2^{N_K}}, k = 1,\ldots,2^{N_K}-1$, $o_k(\cdot) = \delta(\cdot - x_k)$ with $K = 2^{N_K} - 1\,,\ N_K = 2,3,4$.

Figures 1, 2 and 3 show the quadrature error of the normalization constant Z with respect to the cardinality of the monotone index set Λ_N, which is adaptively determined by Agorithm 5.3. The quadrature error is estimated by the contributions of the reduced neighbor set of the current index set Λ_N, cp. Section 5.3. The results are based on the three choices CC, L and RL of the univariate sequences $(z_j^k)_{j=0}^{n_k}$. Furthermore, a variation of the noise, i.e. $\eta_j \sim \mathcal{N}(0,1)$, $\eta_j \sim \mathcal{N}(0,0.5^2)$ and $\eta_j \sim \mathcal{N}(0,0.1^2)$, as well as of the parameter ζ is considered in order to investigate the convergence behavior of the proposed approach.

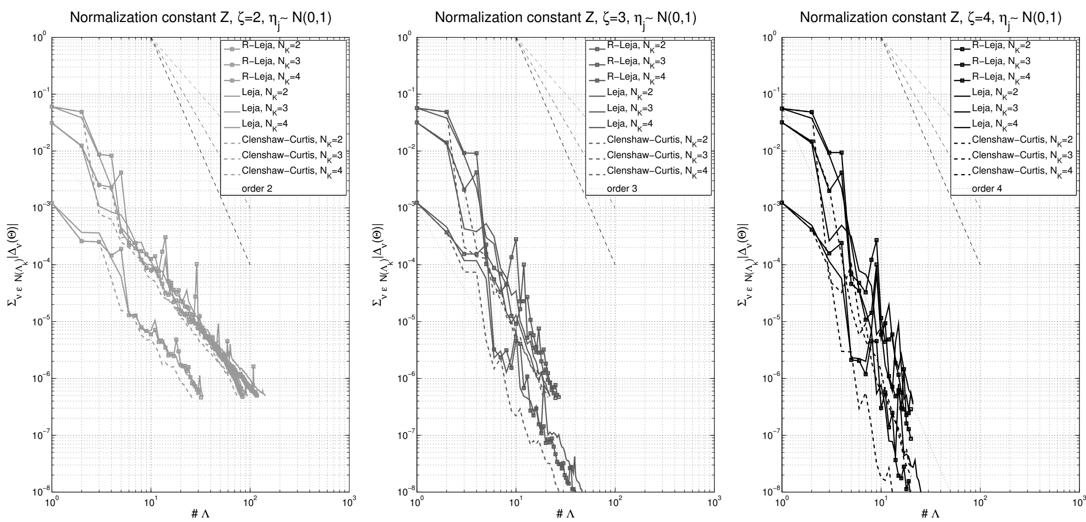

Figure 1. Comparison of the error curves of the normalization constant Z with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

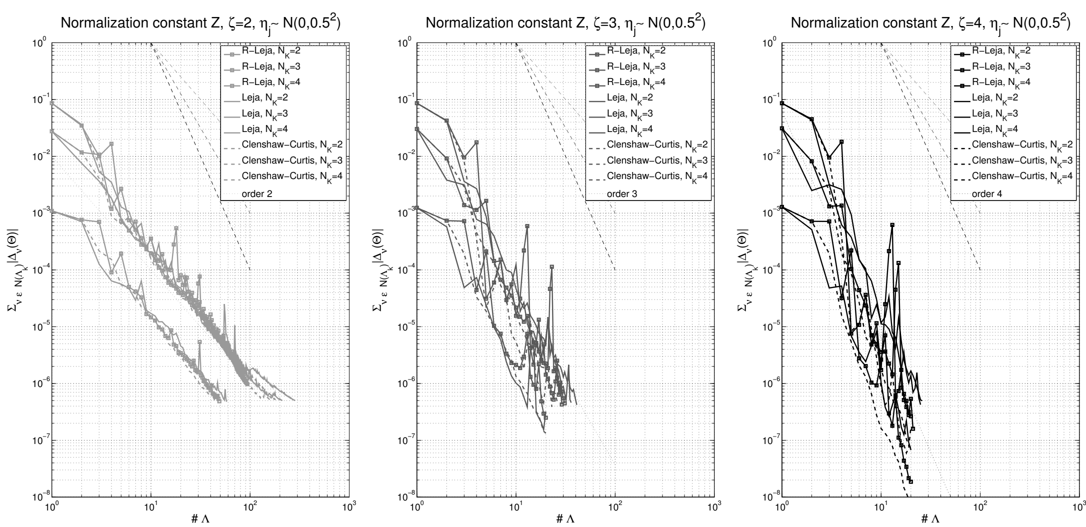

Figure 2. Comparison of the error curves of the normalization constant Z with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

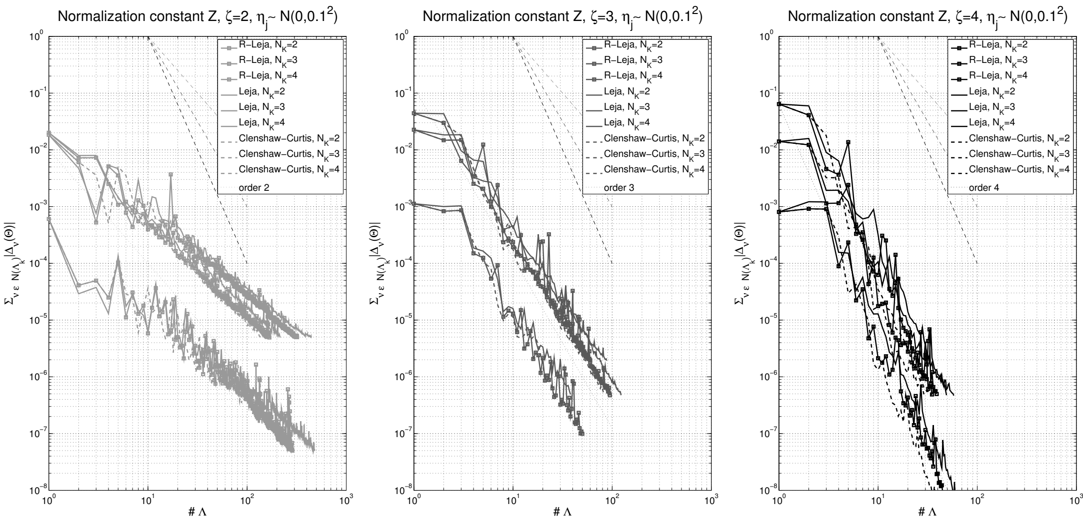

Figure 3. Comparison of the error curves of the normalization constant Z with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1\,,\ N_K = 2,3,4,\ \eta \sim \mathcal{N}(0, 0.1^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

We can observe a convergence order of $(\#\Lambda)^{-\zeta}$ for the three choices of the univariate sequences $(z_j^k)_{j=0}^{n_k}$, which is independent of the number of observation points K and of the variance of the noise. In the case of Clenshaw-Curtis points, the results indicate an even higher convergence order of $(\#\Lambda)^{-\zeta-1}$. However, comparing the error curves with respect to the number of PDE solves needed, we notice that, in the case of Clenshaw-Curtis points, a higher number of quadrature points is needed to achieve the same accuracy as in the case of Leja points (due to the exponential growth of the number of Clenshaw-Curtis points), cp. Figures 4, 5 and 6, respectively.

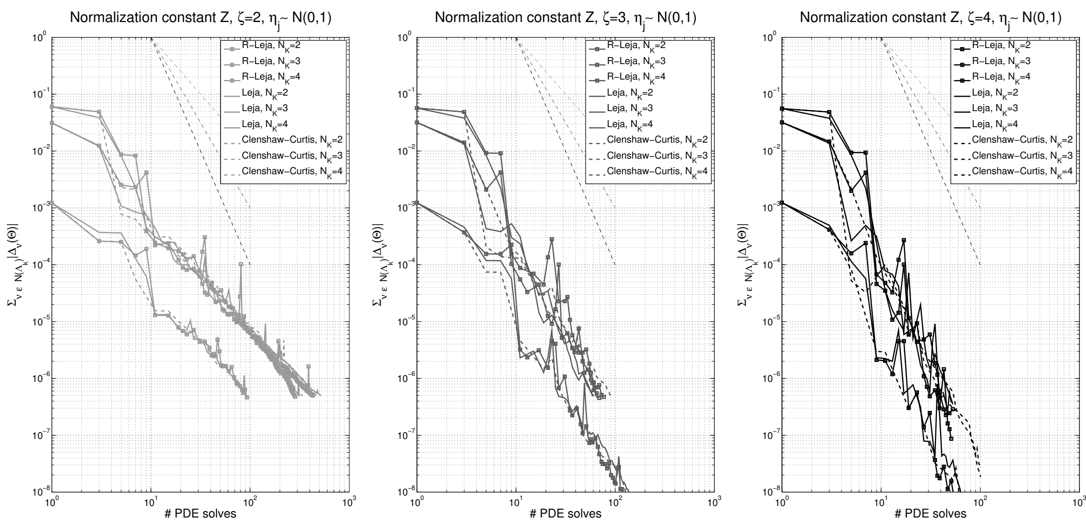

Figure 4. Comparison of the error curves of the normalization constant Z with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

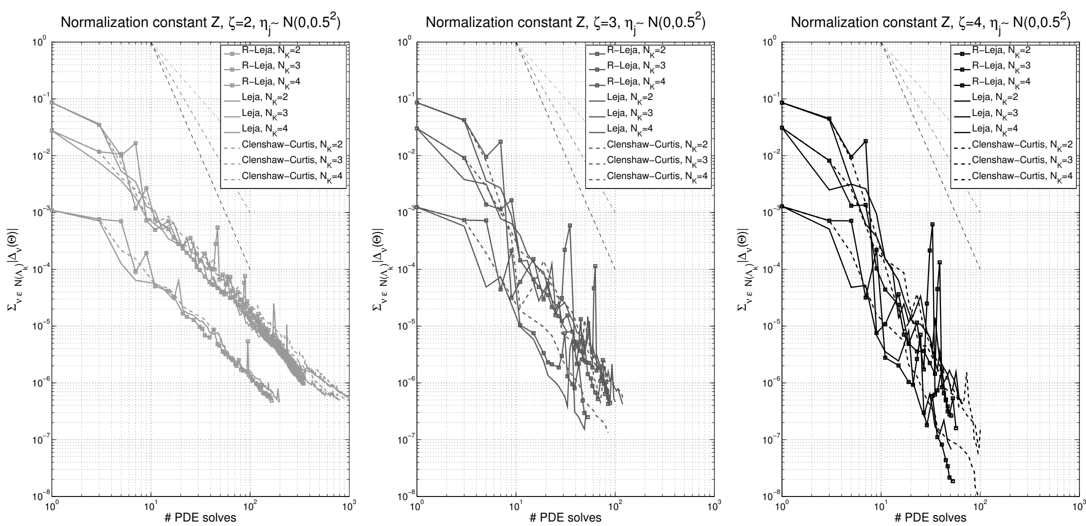

Figure 5. Comparison of the error curves of the normalization constant Z with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

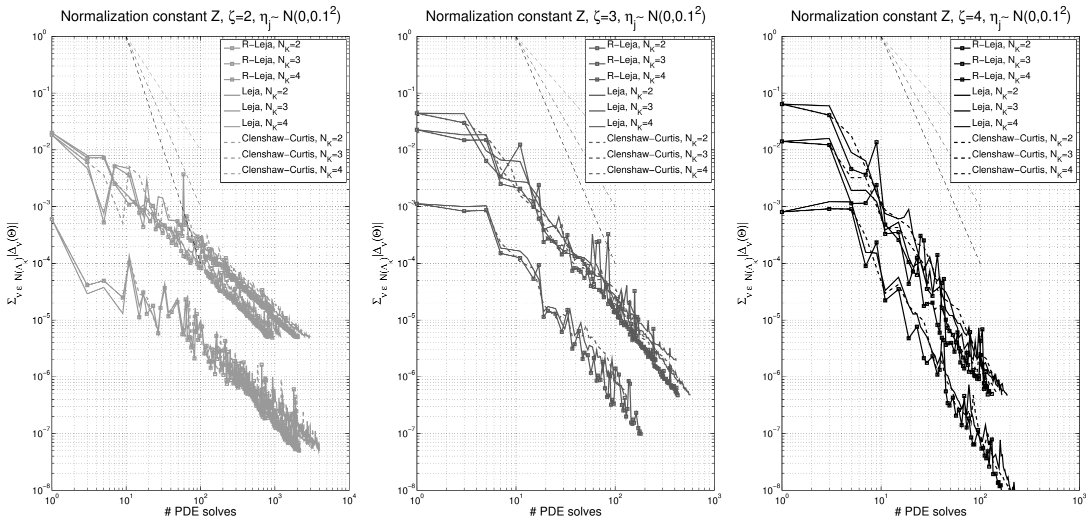

Figure 6. Comparison of the error curves of the normalization constant Z with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

The same observations can be made for the approximation of the quantity Z'. The following figures display the maximum error of the quantity Z' with respect to the cardinality of the current index set Λ_n based on the sequences CC, L and RL with variations of the variance of the noise as well as of the number of observations points K.

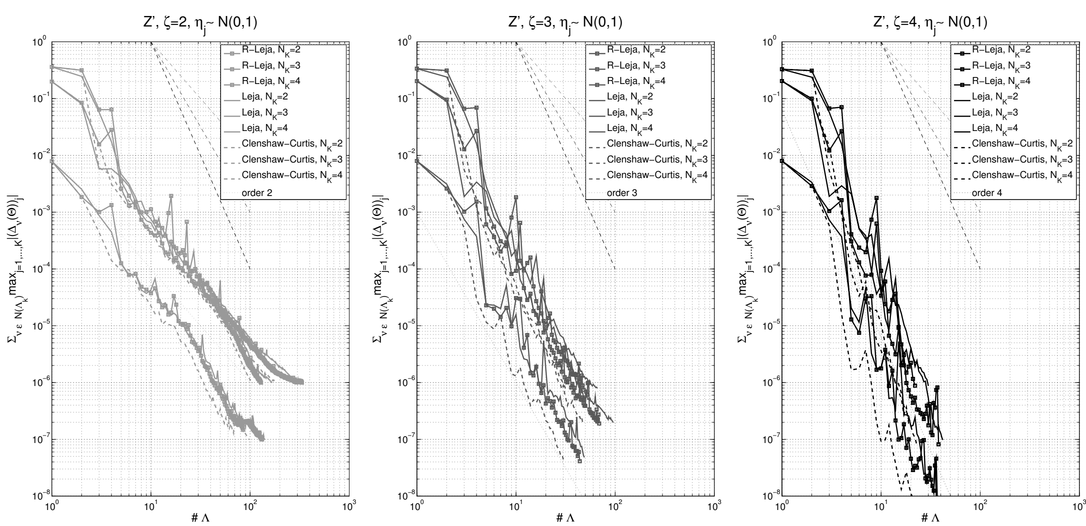

Figure 7. Comparison of the L^∞ error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

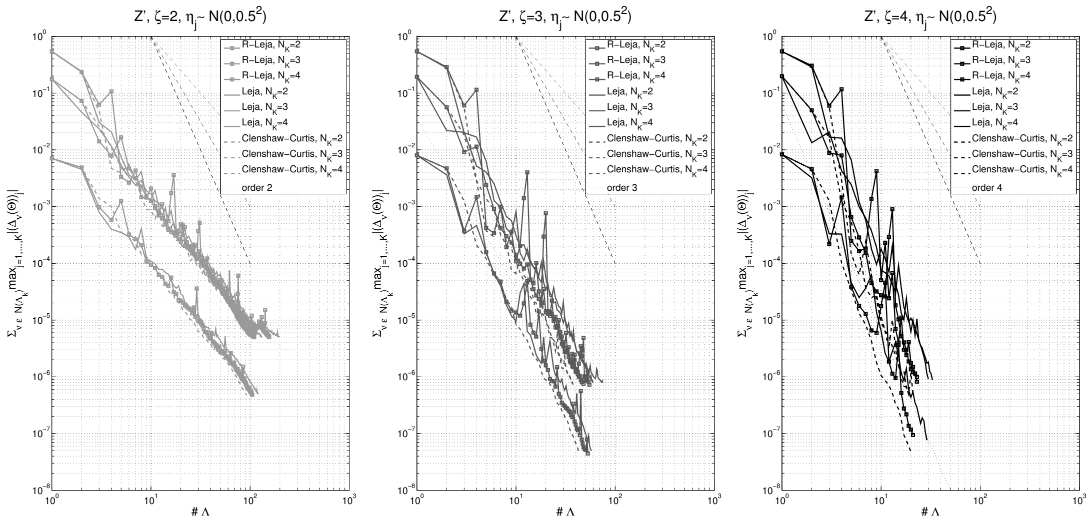

Figure 8. Comparison of the L^∞ error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

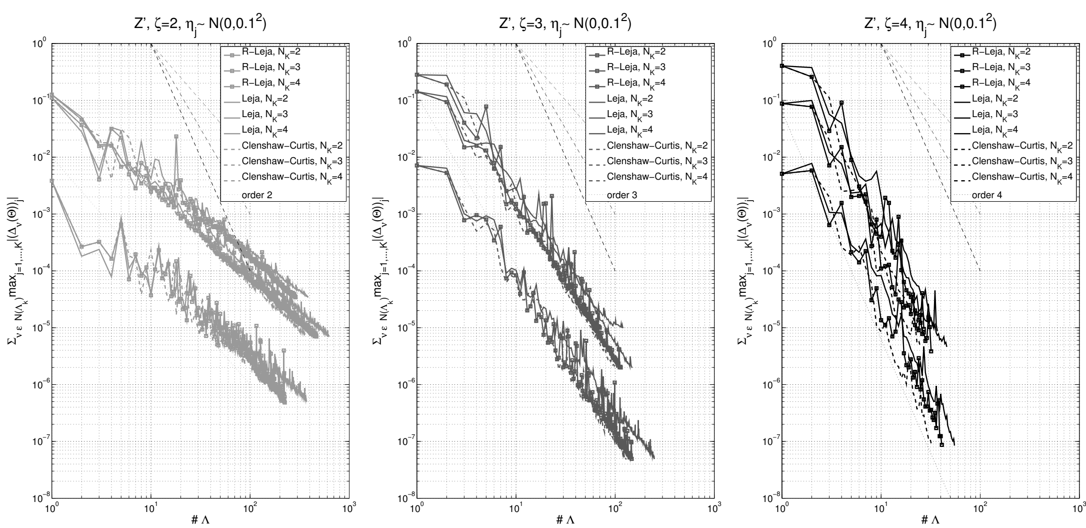

Figure 9. Comparison of the L^∞ error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

Especially for large values of the parameter ζ indicating the sparsity of the unknown parameter, the Clenshaw-Curtis quadrature seems to be superior to the Leja based quadrature. However, comparing the error with respect to total number of boundary value problems which have been solved, a similar behavior as in the approximation of the normalization constant Z can be observed, see Figures 10, 11 and 12.

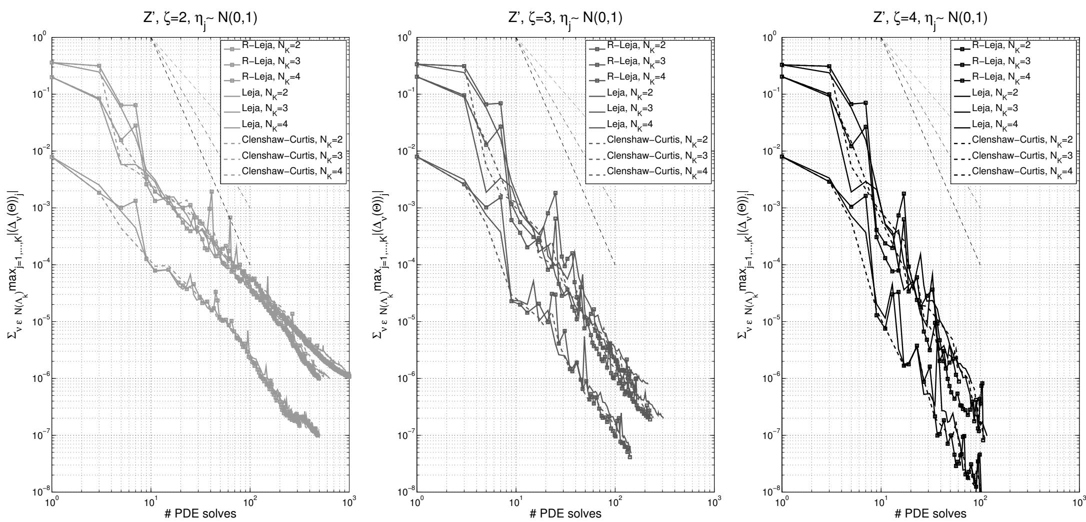

Figure 10. Comparison of the L^∞ error curves of the quantity Z' with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

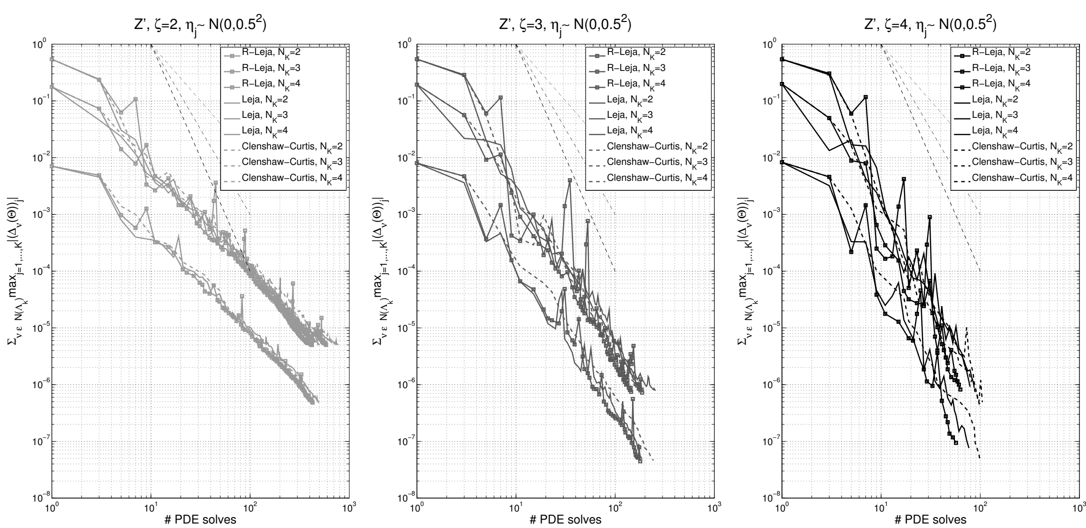

Figure 11. Comparison of the L^∞ error curves of the quantity $Z'$with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

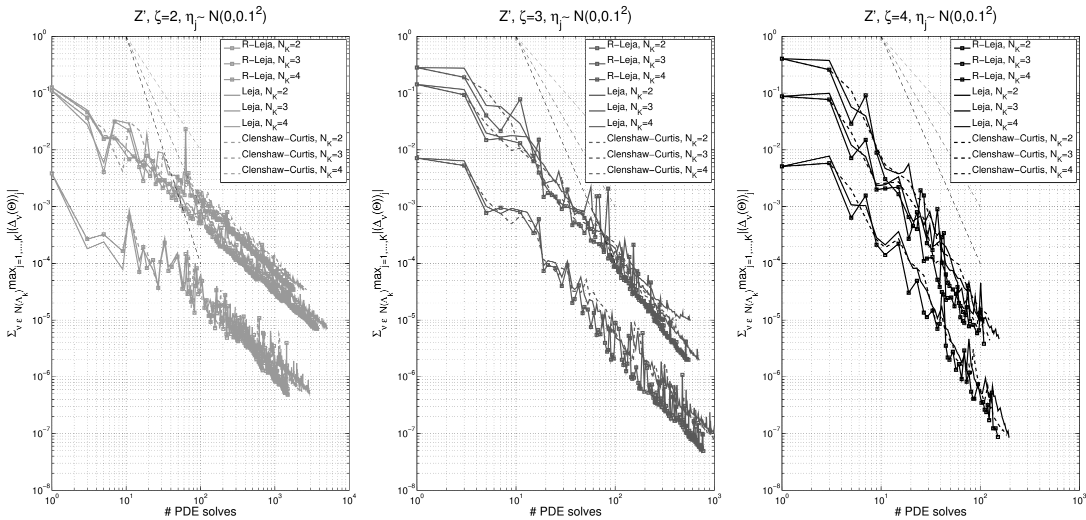

Figure 12. Comparison of the L^∞ error curves of the quantity Z' with respect to the number of PDE solves needed based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

Due to the exponential growth of the number of quadrature points with the order of CC sequences, the Leja based sequences show a better performance especially in cases with high sparsity of the unknown parameters.

Finally, we investigate the convergence order with respect to the L^2 error.

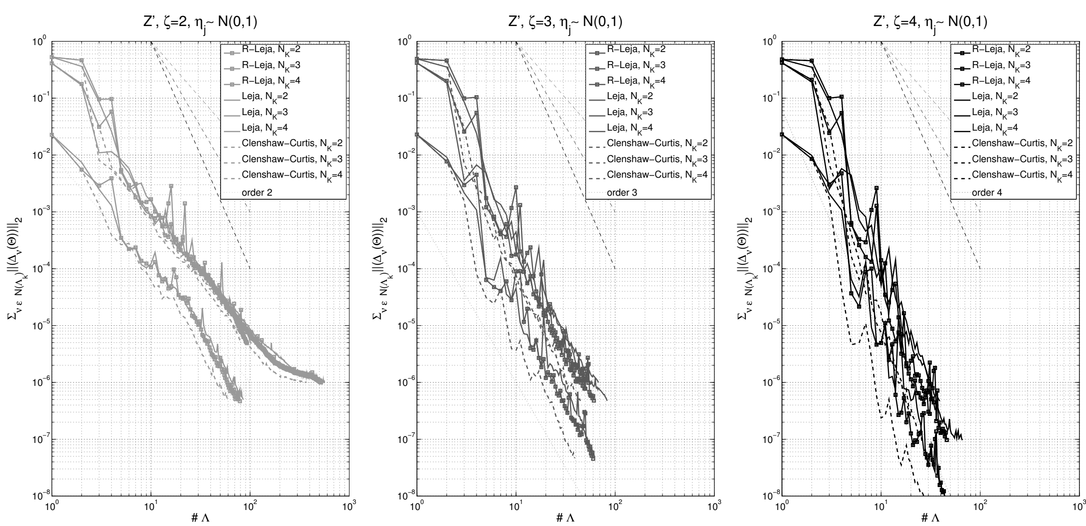

Figure 13. Comparison of the L^2 error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 1)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

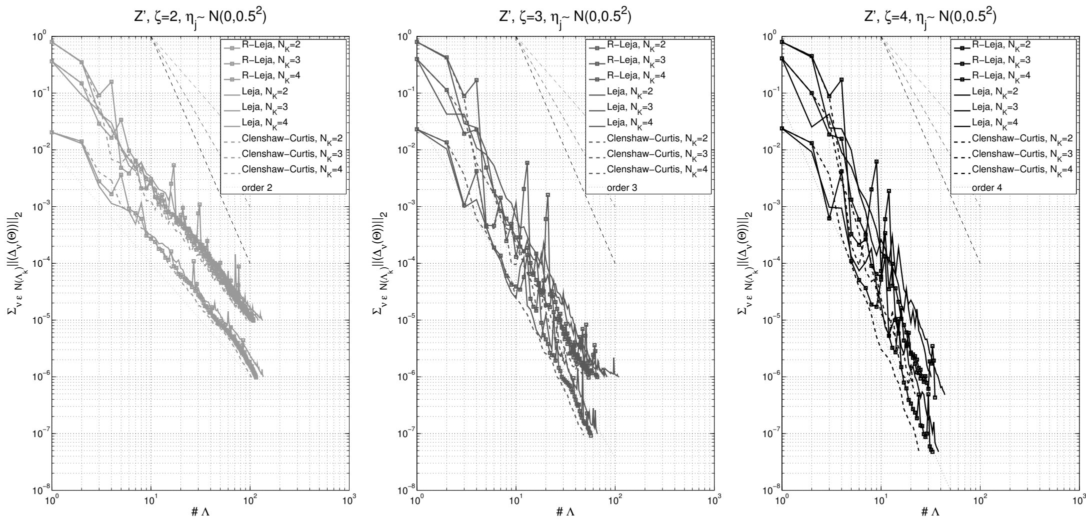

Figure 14. Comparison of the L^2 error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.5^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

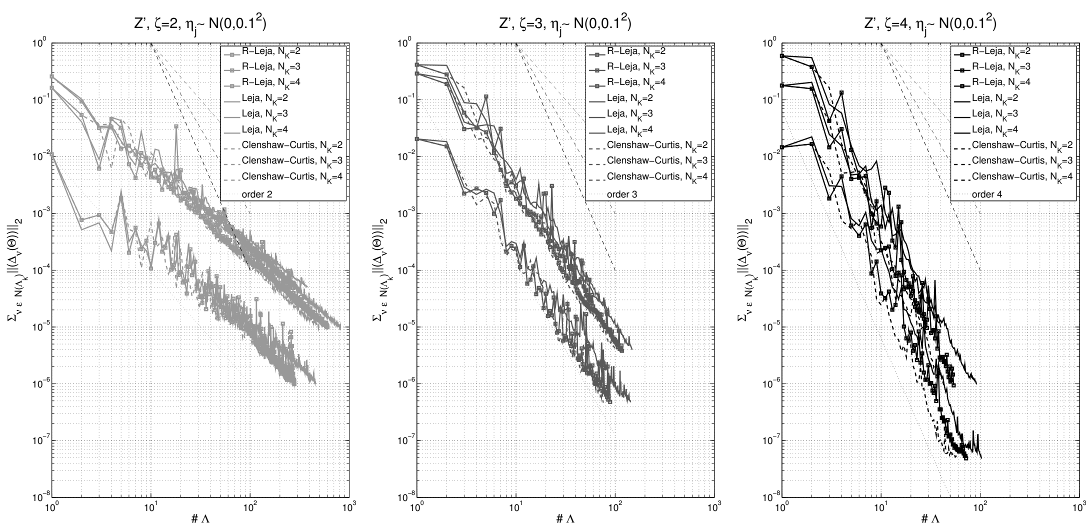

Figure 15. Comparison of the L^2 error curves of the quantity Z' with respect to the cardinality of the index set Λ_n based on the sequences CC, L and RL with $K = 2^{N_K} - 1$, $N_K = 2, 3, 4$, $\eta \sim \mathcal{N}(0, 0.1^2)$ and $\zeta = 2$ (left), $\zeta = 3$ (middle) and $\zeta = 4$ (right).

The following major observations may be drawn from the error curves presented in this section:

- The proposed algorithm to adaptively construct the monotone index set Λ shows a robust performance for the three choices of the univariate sequence $(z_j^k)_{j=0}^{n_k}$ of quadrature points. Furthermore, the variations of the number of observation points as well as of the noise do not affect the convergence order. This behavior is consistent with the theoretical results discussed in Section 5.

- The error curves with varying sparsity of the input parameter indicate a convergence order of $(\#\Lambda)^{-\zeta}$ for the three choices CC, L and RL of the univariate sequences $(z_j^k)_{j=0}^{n_k}$. The theoretical result derived in Theorem 5.7 suggests a convergence order of $(\#\Lambda)^{-s}$ with $s < \zeta - 1$, which is one order smaller than the observed convergence order in the numerical results. In addition, we notice that the sparse quadrature based on Clenshaw-Curtis points seems to outperform the Leja sequences. However, comparing the error with respect to the total number of boundary value problems which have been solved, the Leja based sequences show a better performance especially in cases with $\zeta = 4$, which can be attributed to the exponential growth of the number of quadrature points within the order of CC sequences.

7. Conclusions

For Bayesian inverse problems in partial differential equations, we have presented and analyzed a new class of sparse, adaptive tensor quadrature methods for the efficient numerical evaluation of the expectations of system responses under the posterior measure given noisy data δ. The approach is based on the reformulation of this expectation as an infinite dimensional iterated integral with respect to the posterior density with respect to the prior measure. The analysis was developed for prior measures which are countable product measures of the univariate uniform probability measure. The proposed sparse, adaptive quadrature algorithms were shown mathematically to converge with a dimension-independent rate depending only on the sparsity in the density of the posterior measure. This, in turn, is completely determined by the sparsity in model for the unknown density u, by the result in [28]. Numerical experiments confirmed the predicted performance of the algorithms, and indicated that the convergence rates are, in fact, one order better than can be shown by current error analysis. The adaptive quadrature schemes are based on families of numerical integration formulas in $\mathbb{R}^1$. Among several choices of univariate families of quadrature rules, nested families interpolatory quadrature Newton-Cotes formulas based on Leja abscissas in $(-1, 1)$ were found to be optimal, in terms of accuracy versus number of forward solves. The (algebraic in the order quadrature) ill-conditioning due to the appearance of negative quadrature weights in such quadrature formulas did not cause instability in the overall algorithm.

The Smolyak quadrature algorithm being a collocation algorithm in nature is nonintrusive and can be combined with any given PDE solver for the forward problem.

The convergence and error analysis for the quadrature algorithm was developed here in the specific case of a linear, elliptic PDE with unknown, inhomogeneous coefficient for which sparsity of the PDEs response map had recently been shown in [9, 10]. Similar sparsity results are, however, available also for several other large classes of forward models: we mention only large systems of parametric initial value ODEs [15, 16] semilinear elliptic PDEs, parabolic and hyperbolic evolution PDEs. The present analysis, as well as [28], was based on uniform prior measure. All components of the analysis and the algorithm can, however, be generalized to the case of nonuniform priors with separable density $\rho(y)$ whose factor densities $\rho_j(y_j)$ have compact supports in the coordinates y_j.

Acknowledgments

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

References

[1] Banks H T and Kunisch K 1989 *Estimation techniques for distributed parameter systems* (Birkhäuser)

[2] Bieri M, Andreev R and Schwab C 2010 Sparse tensor discretization of elliptic SPDEs *SIAM J. Sci. Comp.* **31** 4281-4304

[3] Babuška I, Tempone R and Zouraris G E 2004 Galerkin finite element approximations of stochastic elliptic partial differential equations *SIAM J. Numer. Anal.* **42** 800–825

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

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

[6] Chkifa A 2012 On the Lebesgue constant of Leja sequences for the unit disk (submitted to Journal of Approximation Theory)

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

[8] Chkifa A, Cohen A and Schwab C 2012 High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs *Report 2012-NN, Seminar for Applied Mathematics, ETH Zürich, Switzerland* $http://www.sam.math.ethz.ch/reports/2012$

[9] 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

[10] 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

[11] Cotter S L, Dashti M and Stuart A M 2010 Approximation of Bayesian inverse problems in differential equations *SIAM J. Numer. Anal.* **48** 322-345

[12] DeVore R 1998 Nonlinear approximation *Acta Numerica* **7** 51-150

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

[14] Gittelson C J 2011 Adaptive wavelet methods for elliptic partial differential equations with random operators *Report 2011-37, Seminar for Applied Mathematics, ETH Zürich, Switzerland* `http://www.sam.math.ethz.ch/reports/2011/37`

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

[16] Hansen M, Schillings C and Schwab C 2012 Sparse approximation algorithms for high

dimensional parametric initial value problems *Proceedings 5th Conf. High Performance Computing, Hanoi*

[17] Ha Hoang V, Schwab C and Stuart A M 2012 Sparse MCMC gpc Finite Element methods for Bayesian inverse problems (in review)

[18] Hairer M, Stuart A M and Voss J 2007 Analysis of SPDEs arising in path sampling, part II: The nonlinear case *Annals of Applied Probability* **17** 1657-1706

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

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

[21] 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

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

[23] Marzouk Y M and Najm H N2099 Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems *Journal of Computational Physics* **228** 1862-1902

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

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

[26] 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.)

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

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

[29] Spanos P D and Ghanem R 1989 Stochastic finite element expansion for random media *J. Eng. Mech.* **115** 1035-1053

[30] Spanos P D and Ghanem R 2003 *Stochastic Finite Elements: A Spectral Approach* (Dover)

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

[32] Todor R A and Schwab C 2007 Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients *IMA J. Num. Anal.* **27** 232-261

[33] Wiener N 1938 The homogeneous chaos *American Journal of Mathematics* **60** 897-936

Research Reports

No. Authors/Title

12-37 *Cl. Schillings and Ch. Schwab*
Sparse, adaptive Smolyak algorithms for Bayesian inverse problems

12-36 *R. Hiptmair and L. Kielhorn*
BETL A generic boundary element template library

12-35 *S. Mishra, N.H. Risebro, Ch. Schwab and S. Tokareva*
Numerical solution of scalar conservation laws with random flux functions

12-34 *R. Hiptmair, Ch. Schwab and C. Jerez-Hanckes*
Sparse tensor edge elements

12-33 *R. Hiptmair, C. Jerez-Hanckes and S. Mao*
Extension by zero in discrete trace spaces: Inverse estimates

12-32 *A. Lang, S. Larsson and Ch. Schwab*
Covariance structure of parabolic stochastic partial differential equations

12-31 *A. Madrane, U.S. Fjordholm, S. Mishra and E. Tadmor*
Entropy conservative and entropy stable finite volume schemes for multi-dimensional conservation laws on unstructured meshes

12-30 *G.M. Coclite, L. Di Ruvo, J. Ernest and S. Mishra*
Convergence of vanishing capillarity approximations for scalar conservation laws with discontinuous fluxes

12-29 *A. Abdulle, A. Barth and Ch. Schwab*
Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs

12-28 *E. Fonn, Ph. Grohs and R. Hiptmair*
Hyperbolic cross approximation for the spatially homogeneous Boltzmann equation

12-27 *P. Grohs*
Wolfowitz's theorem and consensus algorithms in Hadamard spaces

12-26 *H. Heumann and R. Hiptmair*
Stabilized Galerkin methods for magnetic advection

12-25 *F.Y. Kuo, Ch. Schwab and I.H. Sloan*
Multi-level quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients

12-24 *St. Pauli, P. Arbenz and Ch. Schwab*
Intrinsic fault tolerance of multi level Monte Carlo methods