ESAIM: M2AN 54 (2020) 1259–1307
https://doi.org/10.1051/m2an/2020003

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

CONVERGENCE RATES OF HIGH DIMENSIONAL SMOLYAK QUADRATURE

Jakob Zech[1,*] and Christoph Schwab[2]

Abstract. We analyse convergence rates of Smolyak integration for parametric maps $u : U \to X$ taking values in a Banach space X, defined on the parameter domain $U = [-1,1]^{\mathbb{N}}$. For parametric maps which are sparse, as quantified by summability of their Taylor polynomial chaos coefficients, dimension-independent convergence rates superior to N-term approximation rates under the same sparsity are achievable. We propose a concrete Smolyak algorithm to a priori identify integrand-adapted sets of active multiindices (and thereby unisolvent sparse grids of quadrature points) *via* upper bounds for the integrands' Taylor gpc coefficients. For so-called "$(\boldsymbol{b},\varepsilon)$-holomorphic" integrands u with $\boldsymbol{b} \in \ell^p(\mathbb{N})$ for some $p \in (0,1)$, we prove the dimension-independent convergence rate $2/p-1$ in terms of the number of quadrature points. The proposed Smolyak algorithm is proved to yield (essentially) the same rate in terms of the total computational cost for both nested and non-nested univariate quadrature points. Numerical experiments and a mathematical sparsity analysis accounting for cancellations in quadratures and in the combination formula demonstrate that the asymptotic rate $2/p-1$ is realized computationally for a moderate number of quadrature points under certain circumstances. By a refined analysis of model integrand classes we show that a generally large preasymptotic range otherwise precludes reaching the asymptotic rate $2/p-1$ for practically relevant numbers of quadrature points.

Mathematics Subject Classification. 65D30.

Received June 30, 2017. Accepted January 13, 2020.

1. Introduction

Let X be a Banach space, set $U = [-1,1]^{\mathbb{N}}$ and let μ be the infinite product (probability) measure $\bigotimes_{j\in\mathbb{N}} \lambda/2$ on U, where λ denotes the Lebesgue measure on $[-1,1]$. The efficient numerical approximation of formally infinite-dimensional integrals

$$\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}), \tag{1.1}$$

of strongly μ-measurable, parametric maps $u : U \to X$ is a key problem in computational uncertainty quantification ("UQ" for short). In computational UQ, the integrand function u in (1.1) is implicitly given as solution of a so-called *forward model*, typically an operator equation parametrized by a sequence $\boldsymbol{y} \in U$. The parameter sequences $\boldsymbol{y}$ can, for example, describe distributed uncertain constitutive relations or uncertain geometric

Keywords and phrases. Generalized polynomial chaos, Smolyak quadrature, sparsity, holomorphy.

[1] Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.

[2] Seminar for Applied Mathematics (SAM), Eidgenössische Technische Hochschule Zürich (ETHZ), Rämistrasse 101, 8092 Zürich, Switzerland.

*Corresponding author: jzech@mit.edu

shapes. Equation (1.1) then describes an "ensemble average" (with respect to μ) of the parametric solution, over all admissible realizations of the uncertainty.

The high (in this case infinite) dimension of the integration domain U demands the integrand to possess appropriate sparsity properties in order to make a numerical computation feasible, and overcome the so-called curse of dimensionality. For this reason, the integrand is typically assumed to be very smooth, *e.g.* to allow a bounded holomorphic extension into certain cylindrical subsets of $\mathbb{C}^{\mathbb{N}}$: here, as in [20], we consider parametric integrands which are holomorphic in cartesian products of discs with increasing radii. The rate at which those radii increase is a measure of the sparsity of the function, and as was observed in [20, 22, 30] governs the (dimension-independent) rate of convergence of the quadrature. These assumptions on the integrand are condensed in the notion of $(\boldsymbol{b}, \varepsilon)$-holomorphy for a positive sequence $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ and some $p \in (0, 1)$, see Definition 3.1 and also *cp.* [10–12]. This function class comprises in particular functions of the following type: Let Z and X be two complex Banach spaces and $(\psi_j)_{j\in\mathbb{N}} \subseteq Z$ such that $(\|\psi_j\|_Z)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$. Assume that $\mathfrak{u} : Z \to X$ is Fréchet differentiable (this can be weakened to Fréchet differentiability on a certain subset of Z). Then, as we show in Lemma 3.3, the function

$$u(\boldsymbol{y}) = \mathfrak{u}\left(\sum_{j\in\mathbb{N}} y_j\psi_j\right) \in X \qquad \boldsymbol{y} \in U \tag{1.2}$$

is $(\boldsymbol{b}, \varepsilon)$-holomorphic with $b_j = \|\psi_j\|_Z$. Functions of this type arise in the context of parametrized partial differential equations (PDEs) for a large variety of linear and nonlinear equations see for example [12, 13, 23, 25, 27]. Our new results, which imply the convergence rate $2/p - 1$ for the numerical approximation of (1.1), may consequently be applied to all such models.

One possibility to numerically approximate the integral (1.1) is with a Monte Carlo method. Its advantage is that the convergence rate does not depend on the dimension of the integration domain. Its main disadvantage is the notoriously slow convergence rate of 1/2. For this reason, quasi Monte Carlo (QMC) methods exploiting the integrands' sparsity to attain higher order dimension-independent rates have been developed; we refer to [14,15], to the surveys [15,28] and to the references there. QMC quadrature is free from the curse of dimensionality, and additionally retains the Monte-Carlo feature of "embarrassingly parallel" integrand evaluation at the quadrature points. For high numbers of computationally intensive function evaluations (as is the case for numerical PDE solutions in the context of computational UQ) this becomes an important feature.

The present error analysis is based on so-called generalized polynomial chaos ("gpc" for short) expansions of the parametric integrand function. Expansions of gpc type have proved a valuable tool in regularity and sparsity analysis of countably-parametric functions taking values in a Banach space X; we refer to [10–12, 36] and to the survey [33] and the references there. The idea is to expand the integrand in a polynomial basis, and approximate the integral (1.1) with an interpolatory quadrature rule that is exact for the terms contributing most in the expansion. Such reasoning gives best N-term results, but in practice the optimal set of quadrature points is not known. The effectiveness of the method is due to the high smoothness of the integrand, which is why polynomial approximations converge very fast. We refer to [4, 17, 35] for a general description of sparse grid quadrature. For our proofs, as a basis we shall use the monomials, *i.e.* as in [11, 12, 36], we consider Taylor gpc expansions around $\mathbf{0} = (0, 0, \dots) \in U$. Unconditional convergence of such Taylor gpc expansion stipulates holomorphy of the integrand in polydiscs around $\mathbf{0}$. We choose the monomials for ease of presentation, but point out that holomorphy assumptions can be weakened by considering expansions in orthogonal bases such as the Legendre polynomials which merely require holomorphy on so-called Bernstein ellipses (*cp.* [12]). This results in more technical arguments, but also in weaker holomorphy assumptions, as shown in [37], see also Remark 2.17. The question remains on how to choose the quadrature points such that possibly few function evaluations result in a minimal error. In [18] an adaptive strategy has been proposed. The algorithm does not allow for parallel function evaluations in general however. Nonetheless, it delivers good results and has also been applied for parametrized PDEs, *e.g.* in [32]. In the case of *a priori* chosen quadrature points, the convergence for isotropic and anisotropic sparse grids was investigated in [2, 30], and more recently in [20, 22]. The last two papers can be considered as the closest to ours. Numerical experiments in these works often revealed much better convergence rates, than what the theoretical findings suggested, see in particular [22, 32].

The first aim of the present paper is to establish new, dimension-independent convergence rate bounds. These are stated in Theorem 4.3. This result will shed some new light on the previously observed discrepancy between the observed convergence rates, and the proven ones. As a general idea, we use *a priori* knowledge on size scaling of domains of holomorphic extension of the parametric integrand to estimate the norm of the Taylor coefficients. Based on these estimates, a sparse grid is constructed *a priori*. The crucial observation, allowing us to improve earlier estimates, is then the following: The linear term $y \mapsto y$ has integral 0 over $[-1, 1]$, and is integrated exactly by the midpoint rule (*i.e.* by an evaluation at $y = 0$ multiplied with the weight 1 corresponding to the probability measure $\lambda/2$). As a consequence, any polynomial in the multivariate Taylor expansion containing a linear term will always be integrated exactly by the Smolyak quadrature operator. This implies higher, dimension-independent convergence rates since the sequence of the remaining Taylor gpc coefficients has summability which is superior to the sequence of all Taylor gpc coefficients. Indeed, our new results improve previously established, dimension-independent convergence rates, by more than a factor two; see Remark 4.5 and Examples 5.2 and 5.3.

The second contribution concerns a novel *a priori* construction of gpc index sets which we prove to provide near optimal, dimension-independent convergence rates. Whereas many authors consider the number of quadrature points as a measure for the work, in fact, due to its structure based on differences of tensor product quadratures, the actual cost of the Smolyak algorithm does not in general behave linearly in the number of quadrature points. The mentioned convergence rates are proven with respect to the total number of quadrature points in case of nested point sets such as Leja points. In addition, we show that essentially the same rate can be obtained also for non-nested point sets, such as the Gauss points. Finally, this rate is also proven in terms of the total number of floating point operations. The precise statements are given in Theorem 2.16 and in a bit more generality in Theorem 4.3. The proven rates are asymptotic, and might not always be observable in the range of "small" numbers of quadrature points that are realizable in practice, as our numerical experiments and further analysis of particular model parametric integrand families in Section 5 reveal.

Structure of the paper

In Section 2 we first set up notation and state a few assumptions used throughout. Subsequently the Smolyak algorithm is recalled, and we present a short complexity analysis. This then provides sufficient preliminaries to state our main result in Theorem 2.16.

In Section 3 we formalize the concept of $(\boldsymbol{b}, \varepsilon)$-holomorphic, parametric maps from the parameter domain U into a complex Banach space X. Maps of this type admit unconditionally convergent Taylor gpc expansions, with a specific decay of the Taylor gpc coefficients $(t_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq X$. In Section 3.3, we prove novel summability results for certain subsequences of $(t_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$. These results quantify the effect of cancellations of Taylor gpc coefficients due to symmetries in the Smolyak quadrature operators. As they are abstract sequence summation results, they play a role also in more general gpc approximation results. The main result of the section is Theorem 3.14.

In Section 4, we prove a convergence result for the Smolyak algorithm in Theorem 4.3. The algebraic convergence rate is stated in terms of the number of function evaluations for both nested and non-nested quadrature points, and additionally in terms of the number of required floating point operations. Additionally, we provide explicit constructions of suitable sets of multiindices, which allows to *a priori* devise a sparse-grid. This provides an algorithm for which the integrand can be evaluated at all quadrature points in parallel.

Section 5 is devoted to numerical experiments. We give more details on the implementation in Section 5.1. In Section 5.2 we provide a precise description of the proposed algorithm. As already mentioned above, a large preasymptotic range is observed in certain situations. This is numerically investigated in Section 5.3, and we give (heuristic) arguments why it occurs. Finally, in Section 5.4 the convergence of our algorithm is tested for two exemplary real valued functions.

2. Smolyak algorithm and main result

2.1. Notation

Throughout we let $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. The symbol C will stand for a generic, positive constant independent of any quantities determining the asymptotic behaviour of an estimate. It may change even within the same formula.

Multiindices are denoted by $\boldsymbol{\nu} = (\nu_j)_{j\in\mathbb{N}} \in \mathbb{N}_0^{\mathbb{N}}$. The notation $\operatorname{supp}\boldsymbol{\nu}$ stands for the *support* of the multiindex, *i.e.* the set $\{j \in \mathbb{N} : \nu_j \neq 0\}$. For the *total order* of a multiindex we write $|\boldsymbol{\nu}| := \sum_{j\in\mathbb{N}} \nu_j$ and introduce the countable sets

$$\mathcal{F} := \{\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} : |\boldsymbol{\nu}| < \infty\} \qquad \text{and} \qquad \mathcal{F}_k := \{\boldsymbol{\nu} \in \mathcal{F} : \nu_j \geq k \ \forall j \in \operatorname{supp}\boldsymbol{\nu}\} \tag{2.1}$$

for all $k \in \mathbb{N}$. In particular $\mathcal{F} = \mathcal{F}_1$. Note that $\mathcal{F}$ consists of all finitely supported multiindices in $\mathbb{N}_0^{\mathbb{N}}$. For two multiindices $\boldsymbol{\nu}$, $\boldsymbol{\mu} \in \mathcal{F}$, by $\boldsymbol{\mu} \leq \boldsymbol{\nu}$ we mean mean $\mu_j \leq \nu_j$, for all $j \geq 1$. A subset $\Lambda \subseteq \mathcal{F}$ will be called *downward closed* if for every $\boldsymbol{\nu} \in \Lambda$ it holds $\{\boldsymbol{\mu} \in \mathcal{F} : \boldsymbol{\mu} \leq \boldsymbol{\nu}\} \subseteq \Lambda$.

For $p > 0$ we let $\ell^p(\mathcal{F}_k)$ be the space of $\mathbb{R}$-valued sequences $\mathbf{a} = (a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}$, satisfying

$$\|\mathbf{a}\|_{\ell^p(\mathcal{F}_k)} := \left(\sum_{\boldsymbol{\nu}\in\mathcal{F}_k} |a_{\boldsymbol{\nu}}|^p\right)^{1/p} < \infty.$$

Similarly, $\ell^p(\mathbb{N})$ is defined for sequences indexed over $\mathbb{N}$. By a *decreasing rearrangement* $(a_j^*)_{j\in\mathbb{N}}$ of a sequence $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}$, we mean that there exists a bijection $\pi : \mathbb{N} \to \mathcal{F}_k$ such that $a_j^* = a_{\pi(j)}$ for all $j \in \mathbb{N}$, and additionally $a_j^* \geq a_{j+1}^*$ for all $j \in \mathbb{N}$.

As a topology on $\mathbb{C}^{\mathbb{N}}$ we choose the product topology, and any subset such as $[-1,1]^{\mathbb{N}}$ is equipped with the subspace topology. For a ball of radius $r > 0$ in $\mathbb{C}$ we write $B_r^{\mathbb{C}} := \{z \in \mathbb{C} : |z| < r\} \subseteq \mathbb{C}$. Furthermore, if $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$, then $B_{\boldsymbol{\rho}}^{\mathbb{C}} := \bigtimes_{j\in\mathbb{N}} B_{\rho_j}^{\mathbb{C}} \subseteq \mathbb{C}^{\mathbb{N}}$. Moreover, the parameter set $U = [-1,1]^{\mathbb{N}}$ endowed with the Borel product sigma algebra and the uniform product probability measure $\mu := \bigotimes_{j\in\mathbb{N}} \lambda/2$ is a probability space. Here, λ denotes the Lebesgue measure on $[-1,1]$. With this topology, for a Banach space X we write $C^0(U, X)$ for the space of (bounded) continuous functions mapping from U to X. Denoting the norm on X by $\|\cdot\|_X$, we let

$$\|u\|_{C^0(U,X)} := \sup_{\boldsymbol{y}\in U} \|u(\boldsymbol{y})\|_X.$$

Similarly we define $C^0([-1,1], X)$, and in case $X = \mathbb{R}$ we simply write $C^0([-1,1]) := C^0([-1,1], X)$ and $C^0(U) := C^0(U, X)$. Elements of $\mathbb{C}^{\mathbb{N}}$ are denoted by boldface characters such as $\boldsymbol{y} = (y_j)_{j\in\mathbb{N}} \in [-1,1]^{\mathbb{N}}$. For $\boldsymbol{\nu} \in \mathcal{F}$, standard multivariate notations $\boldsymbol{y}^{\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}} y_j^{\nu_j}$ and $\boldsymbol{\nu}! = \prod_{j\in\mathbb{N}} \nu_j!$ will be employed.

For a complex Banach space $(X, \|\cdot\|_X)$, $x \in X$ and $\epsilon > 0$, as above we write $B_\epsilon^X := \{z \in X : \|z\|_X < \epsilon\}$. A function u mapping from an open subset of $\mathbb{C}^{\mathbb{N}}$ to X will be called separately holomorphic, if it is holomorphic in each variable. For such a function we denote by

$$\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{y}) = \frac{\partial^{|\boldsymbol{\nu}|}}{\partial y_1^{\nu_1} \partial y_2^{\nu_2} \cdots} u(\boldsymbol{y})$$

the partial derivatives of u w.r.t. the multiindex $\boldsymbol{\nu} \in \mathcal{F}$ where $|\boldsymbol{\nu}| < \infty$. We write X' for the topological dual space of X (*i.e.* the continuous *linear* functionals). The space of bounded linear maps between two Banach spaces X and Y is denoted by $L(X, Y)$.

Finally, for a set A we denote by $|A|$ the cardinality of the set.

2.2. Smolyak quadrature

Let in the following X be a Banach space and $u : U \to X$ a pointwise defined function. For $n \in \mathbb{N}_0$, let $(\chi_{n;j})_{j=0}^n \subseteq [-1,1]$ be a sequence of pairwise distinct points in $[-1,1]$. The Smolyak algorithm is built on a family of univariate quadrature rules $Q_n : C^0([-1,1],X) \to X$ that we assume to be interpolatory quadrature rules with quadrature points $(\chi_{n;0},\dots,\chi_{n;n})$, and w.r.t. the probability measure $\lambda/2$ on $[-1,1]$. That is, for all $f \in C^0([-1,1],X)$

$$Q_n f = \sum_{j=0}^{n} f(\chi_{n;j})\alpha_{n;j} \qquad \text{where} \qquad \alpha_{n;j} = \frac{1}{2}\int_{-1}^{1}\prod_{\substack{i=0\\ i\neq j}}^{n}\frac{y-\chi_{n;i}}{\chi_{n;j}-\chi_{n;i}}\,\mathrm{d}y, \tag{2.2}$$

with an empty product denoting the constant unit function, *i.e.* $\alpha_{0;0} = 1$. We interpret Q_n in the following both as an operator mapping from $C^0([-1,1],X) \to X$ and $C^0([-1,1]) \to \mathbb{R}$ (recall that $C^0([-1,1]) = C^0([-1,1],\mathbb{R})$). The definition of Q_n implies $Q_n w = \int_{-1}^{1} w(y)\,\mathrm{d}\lambda(y)/2$ for all polynomials w of degree at most n. Note that in general the quadrature weights $\alpha_{n;j}$ of Q_n can be negative. Throughout we assume that there exists $\vartheta \in [0,\infty)$ such that the condition of the univariate quadratures Q_n is polynomially bounded according to

$$\forall n \in \mathbb{N}_0: \qquad \sup_{0\neq f\in C^0([-1,1])}\frac{|Q_n f|}{\|f\|_{C^0([-1,1])}} \le (n+1)^{\vartheta}. \tag{2.3}$$

To introduce the Smolyak quadrature, first define $Q_{-1} := 0$. For every $\boldsymbol{\nu} \in \mathcal{F}$ set $Q_{\boldsymbol{\nu}} := \bigotimes_{j\in\mathbb{N}} Q_{\nu_j}$, *i.e.* for $u : U \to X$

$$Q_{\boldsymbol{\nu}} u = \sum_{\{\boldsymbol{\mu}\in\mathcal{F}\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}} u((\chi_{\nu_j;\mu_j})_{j\in\mathbb{N}})\prod_{j\in\mathbb{N}}\alpha_{\nu_j;\mu_j} = \sum_{\{\boldsymbol{\mu}\in\mathcal{F}\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}} u((\chi_{\nu_j;\mu_j})_{j\in\mathbb{N}})\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\alpha_{\nu_j;\mu_j}, \tag{2.4}$$

where an empty product equals 1 by convention. For a downward closed index set $\Lambda \subseteq \mathcal{F}$ of finite cardinality, the Smolyak quadrature Q_Λ is defined by

$$Q_\Lambda := \sum_{\boldsymbol{\nu}\in\Lambda}\bigotimes_{j\in\mathbb{N}}(Q_{\nu_j} - Q_{\nu_j-1}).$$

By induction over $d = |\operatorname{supp}\boldsymbol{\nu}|$, it is easily verified that Q_Λ allows the representation

$$Q_\Lambda = \sum_{\boldsymbol{\nu}\in\Lambda}\varsigma_{\Lambda,\boldsymbol{\nu}}Q_{\boldsymbol{\nu}} \qquad \text{where} \qquad \varsigma_{\Lambda,\boldsymbol{\nu}} := \sum_{\{\mathbf{e}\in\{0,1\}^{\mathbb{N}}\,:\,\boldsymbol{\nu}+\mathbf{e}\in\Lambda\}}(-1)^{|\mathbf{e}|}. \tag{2.5}$$

We also refer to $(\varsigma_{\Lambda,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda}$ as the "combination coefficients". The latter representation of Q_Λ in (2.5) is preferred in implementations, since it skips evaluations of $Q_{\boldsymbol{\nu}}$ for all $\boldsymbol{\nu} \in \Lambda$ with $\varsigma_{\Lambda,\boldsymbol{\nu}} = 0$.

2.3. Number of function evaluations

2.3.1. Quadrature points

Denote in the following the array of univariate sampling points

$$\boldsymbol{\chi} = ((\chi_{n;j})_{j=0}^{n})_{n\in\mathbb{N}_0}. \tag{2.6}$$

By (2.4) and (2.5) the computation of $Q_\Lambda u$ requires to evaluate u at all points in

$$\operatorname{pts}(\Lambda,\boldsymbol{\chi}) := \left\{(\chi_{\nu_j;\mu_j})_{j\in\mathbb{N}} \;:\; \boldsymbol{\nu}\in\Lambda,\ \varsigma_{\Lambda,\boldsymbol{\nu}}\neq 0,\ \boldsymbol{\mu}\le\boldsymbol{\nu}\right\} \subseteq U. \tag{2.7}$$

Definition 2.1. The univariate points $\boldsymbol{\chi} = ((\chi_{n;j})_{j=0}^{n})_{n\in\mathbb{N}_0} \subset [-1,1]$ are called *nested* if there exists a sequence $(\chi_j)_{j\in\mathbb{N}_0}$ such that $\chi_{n;j} = \chi_j$ for every $j \in \{0,\dots,n\}$ and every $n \in \mathbb{N}_0$. Otherwise, $\boldsymbol{\chi}$ are called *non-nested*.

Lemma 2.2. *Let $\Lambda \subseteq \mathcal{F}$ be finite and downward closed. For nested points $\boldsymbol{\chi}$ holds $|\mathrm{pts}(\Lambda,\boldsymbol{\chi})| = |\Lambda|$.*

Lemma 2.2 for nested points is easily verified. In the general case (of possibly non-nested points), due to $|\{\boldsymbol{\mu} \in \mathcal{F} : \boldsymbol{\mu} \le \boldsymbol{\nu}\}| = \prod_{j\in\mathbb{N}}(1+\nu_j)$, it follows immediately from (2.7) that

$$|\mathrm{pts}(\Lambda,\boldsymbol{\chi})| \le \sum_{\{\boldsymbol{\nu}\in\Lambda\,:\,\varsigma_{\Lambda,\boldsymbol{\nu}}\neq 0\}} \prod_{j\in\mathbb{N}} (1+\nu_j). \tag{2.8}$$

Note that we have an equality in (2.8) in case

$$\{\chi_{n,j} \,:\, 0 \le j \le n\} \cap \{\chi_{m,j} \,:\, 0 \le j \le m\} = \emptyset \qquad \forall n \neq m. \tag{2.9}$$

To obtain good bounds on $|\mathrm{pts}(\Lambda,\boldsymbol{\chi})|$ for non-nested points $\boldsymbol{\chi}$, we will devise multiindex sets Λ for which certain combination coefficients vanish, *i.e.* $\varsigma_{\Lambda,\boldsymbol{\nu}} = 0$ for certain $\boldsymbol{\nu} \in \Lambda$. This is exploited in the next example, which demonstrates that the right-hand side of (2.8) can be smaller for larger multiindex sets.

Example 2.3. For $n \in \mathbb{N}$ consider the two sets of (two dimensional) multiindices

$$\Lambda_1(n) := \{(\nu_1,\nu_2) \in \mathbb{N}_0^2 \,:\, \nu_1+\nu_2 < 2n\}, \qquad \Lambda_2(n) := \Lambda_1(n) \cup \{(2j-1, 2(n-j)+1) \,:\, j=1,\dots,n\},$$

and we also use the shortcuts $\Lambda_j = \Lambda_j(n)$ in the following. For $n = 3$ these sets are shown in Figure 1. One checks that (with $\varsigma_{\Lambda,\boldsymbol{\nu}}$ defined in (2.5))

$$\varsigma_{\Lambda_1,\boldsymbol{\nu}} = \begin{cases} 1 & \text{if} \quad \nu_1+\nu_2 = 2n-1 \\ -1 & \text{if} \quad \nu_1+\nu_2 = 2n-2 \\ 0 & \text{otherwise,} \end{cases} \qquad \varsigma_{\Lambda_2,\boldsymbol{\nu}} = \begin{cases} 1 & \text{if} \quad (\nu_1+\nu_2 = 2n) \wedge (\nu_1,\nu_2 \text{ are odd}) \\ -1 & \text{if} \quad (\nu_1+\nu_2 = 2n-2) \wedge (\nu_1,\nu_2 \text{ are odd}) \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

Therefore with

$$A(N) := \sum_{\{\boldsymbol{\nu}\in\mathbb{N}^2\,:\,\nu_1+\nu_2=N+1\}} \nu_1\nu_2 = \sum_{j=1}^{N} j(N+1-j) = \frac{N^3}{6} + O(N^2)$$

it holds

$$\sum_{\{\boldsymbol{\nu}\in\Lambda_1\,:\,\varsigma_{\Lambda_1,\boldsymbol{\nu}}\neq 0\}} \prod_{j=1}^{2} (1+\nu_j) = A(2n) + A(2n-1) = \frac{16}{6}n^3 + O(n^2)$$

whereas

$$\sum_{\{\boldsymbol{\nu}\in\Lambda_2\,:\,\varsigma_{\Lambda_2,\boldsymbol{\nu}}\neq 0\}} \prod_{j=1}^{2} (1+\nu_j) = 4 \sum_{\{\boldsymbol{\nu}\in\Lambda_2\,:\,\varsigma_{\Lambda_2,\boldsymbol{\nu}}\neq 0\}} \prod_{j=1}^{2} \frac{1+\nu_j}{2} = 4A(n) + 4A(n-1) = \frac{8}{6}n^3 + O(n^2).$$

Hence, even though $\Lambda_1 \subsetneq \Lambda_2$, in case of non-nested quadrature points, the number of function evaluations required to compute Q_{Λ_1} is about twice the number of function evaluations required to compute Q_{Λ_2} (*cp.* (2.7) and recall that there holds an equality in (2.8) if (2.9) is satisfied). This is a consequence of the specific structure of Λ_2 which implies (in particular) $\varsigma_{\Lambda_2,\boldsymbol{\nu}} = 0$ whenever ν_j is an even number for at least one $j \in \{1,2\}$. Let us also stress that in this specific example, for $\Lambda_3 := \{(\nu_1,\nu_2) \in \mathbb{N}_0^2 \,:\, \max\{\nu_1,\nu_2\} < 2n\}$ we even have

$$\sum_{\{\boldsymbol{\nu}\in\Lambda_3\,:\,\varsigma_{\Lambda_3,\boldsymbol{\nu}}\neq 0\}} \prod_{j=1}^{2} (1+\nu_j) = (1+2n-1)(1+2n-1) = O(n^2),$$

since $\boldsymbol{\nu} = (2n-1, 2n-1)$ is the only multiindex in Λ_3 for which $\varsigma_{\Lambda_3,\boldsymbol{\nu}} \neq 0$.

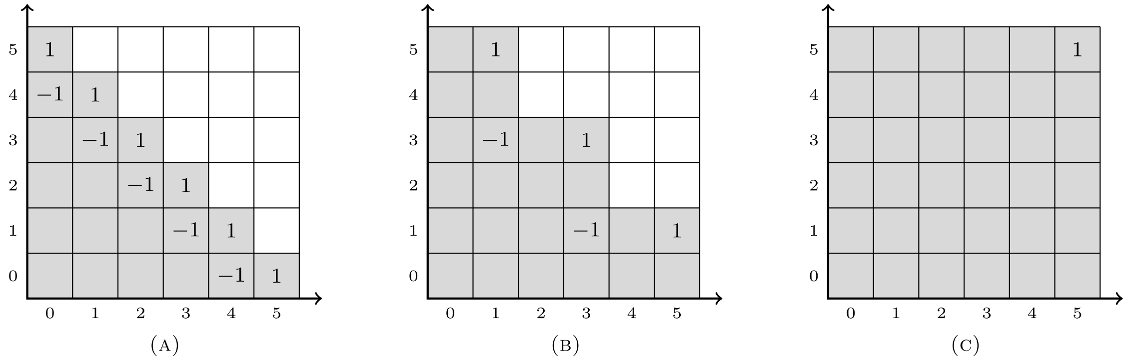


FIGURE 1. The multiindex sets $\Lambda_1(3) \subsetneq \Lambda_2(3) \subsetneq \Lambda_3(3)$ from Example 2.3. The numbers in the squares show the values of $\varsigma_{\Lambda_j,\boldsymbol{\nu}}$ in (2.5) for each $\boldsymbol{\nu} \in \Lambda_j$ with $\varsigma_{\Lambda_j,\boldsymbol{\nu}} \neq 0$. (A) Λ_1. (B) Λ_2. (C) Λ_3.

2.3.2. Admissible indices

To formalize the observation that the structure of Λ can imply $\varsigma_{\Lambda,\boldsymbol{\nu}} = 0$ for certain $\boldsymbol{\nu} \in \Lambda$, we will work with a set $\mathfrak{I} = \{\mathfrak{i}_j : j \in \mathbb{N}_0\} \subseteq \mathbb{N}_0$ of the so-called *admissible indices*. The interpretation of $\mathfrak{I}$ is as follows: We shall build Smolyak quadrature rules based on (2.5). They will have the property that $\varsigma_{\Lambda,\boldsymbol{\nu}} = 0$ for all $\boldsymbol{\nu} \in \Lambda$ for which there exists at least one $j \in \mathbb{N}$ such that $\nu_j \notin \mathfrak{I}$. In other words, Q_Λ in (2.5) will be a linear combination of tensorized quadrature rules $Q_{\boldsymbol{\nu}}$ for multiindices $\boldsymbol{\nu}$ satisfying $\nu_j \in \mathfrak{I}$ for all $j \in \mathbb{N}$, *i.e.* each ν_j must be an admissible index. This allows to control the number of function evaluations required for the computation of Q_Λ, as we show subsequently. In order to do so, in certain cases (see Rem. 5.5 ahead) it will be crucial that the set of admissible indices consists of an exponentially increasing sequence, as stated in the following assumption on $\mathfrak{I}$.

Assumption 2.4 (Admissible indices)**.** *The set $\mathfrak{I} = \{\mathfrak{i}_j : j \in \mathbb{N}_0\} \subseteq \mathbb{N}_0$ consists of the set of the strictly monotonically increasing, nonnegative sequence $(\mathfrak{i}_j)_{j\in\mathbb{N}_0}$ where $\mathfrak{i}_0 = 0$. There exists a constant $K_\mathfrak{I} \geq 1$ such that*

(i) $\mathfrak{i}_{j+1} + 1 \leq K_\mathfrak{I}(\mathfrak{i}_j + 1)$ *for all* $j \in \mathbb{N}_0$,
(ii) $\sum_{j=1}^{m}(\mathfrak{i}_j + 1) \leq K_\mathfrak{I}\mathfrak{i}_m$ *for all* $m \in \mathbb{N}$.

Remark 2.5. The concrete choice of $\mathfrak{I}$ will only influence constants (but not the convergence rates) of the convergence results presented in the following. A natural choice satisfying Assumption 2.4 is $\mathfrak{i}_{j+1} = 2^j$ for $j \in \mathbb{N}_0$, *i.e.* $\mathfrak{I} = \{0\} \cup \{2^j : j \in \mathbb{N}_0\}$.

For $x \geq 0$ denote in the following

$$\lfloor x \rfloor_\mathfrak{I} := \max\{a \in \mathfrak{I} : a \leq x\} \qquad \text{and} \qquad \lceil x \rceil_\mathfrak{I} := \min\{a \in \mathfrak{I} : a \geq x\}. \tag{2.11}$$

Application of these rounding operators to sequences is understood componentwise.

Remark 2.6. With $\mathfrak{I} = \{\mathfrak{i}_j : j \in \mathbb{N}_0\}$ as in Assumption 2.4, define

$$\mathfrak{I}_+ := \{0\} \cup \{\mathfrak{i}_j + 1 : j \in \mathbb{N}_0\}. \tag{2.12}$$

For k, $n \in \mathbb{N}_0$ it holds $\lceil k \rceil_\mathfrak{I} = \lceil n \rceil_\mathfrak{I}$ iff either $k = n = 0$ or there exists $j \in \mathbb{N}_0$ such that k, $n \in (\mathfrak{i}_j, \mathfrak{i}_{j+1}] \cap \mathbb{N}$. The latter is equivalent to k, $n \in [\mathfrak{i}_j + 1, \mathfrak{i}_{j+1} + 1) \cap \mathbb{N}$. Hence, for any $\boldsymbol{\nu}$, $\boldsymbol{\mu} \in \mathcal{F}$

$$\lceil \boldsymbol{\nu} \rceil_\mathfrak{I} = \lceil \boldsymbol{\mu} \rceil_\mathfrak{I} \quad \Leftrightarrow \quad \lfloor \boldsymbol{\nu} \rfloor_{\mathfrak{I}_+} = \lfloor \boldsymbol{\mu} \rfloor_{\mathfrak{I}_+}.$$

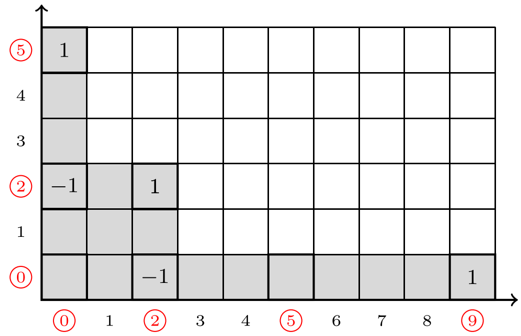


FIGURE 2. The sketch shows a set $\Lambda \subseteq \mathbb{N}_0^2$ of multiindices corresponding to the grey squares. Equation (2.13) is satisfied for some set $\mathfrak{I} = \{0, 2, 5, 9, \dots\}$. By Lemma 2.8, $\varsigma_{\Lambda,\boldsymbol{\nu}} \neq 0$ can only be true if $\nu_j \in \mathfrak{I}$ for all $j \in \mathbb{N}$. The numbers in the squares show the values of $\varsigma_{\Lambda,\boldsymbol{\nu}}$ for each $\boldsymbol{\nu} \in \Lambda$ with $\varsigma_{\Lambda,\boldsymbol{\nu}} \neq 0$.

Remark 2.7. From Assumption 2.4 (i) we infer that for every $n \in \mathbb{N}_0$ and with $\mathfrak{I}_+$ as in (2.12)

$$n \le K_{\mathfrak{I}} \lfloor n \rfloor_{\mathfrak{I}_+} \qquad \text{and} \qquad \lceil n \rceil_{\mathfrak{I}_+} \le K_{\mathfrak{I}} n.$$

We will consider sets of multiindices satisfying

$$(\boldsymbol{\nu} \in \Lambda \quad \text{and} \quad \lceil \boldsymbol{\mu} \rceil_{\mathfrak{I}} = \lceil \boldsymbol{\nu} \rceil_{\mathfrak{I}}) \quad \Rightarrow \quad \boldsymbol{\mu} \in \Lambda. \tag{2.13}$$

The following lemma in conjunction with (2.7) elucidate the significance of this property. The statement of the lemma is visualized in Figure 2. In the following we write

$$\Lambda|_{\mathfrak{I}} := \{\boldsymbol{\nu} \in \Lambda \,:\, \nu_j \in \mathfrak{I} \;\forall j \in \mathbb{N}\}.$$

Lemma 2.8. *Let $\mathfrak{I} \subseteq \mathbb{N}_0$. Let Λ be finite and downward closed with the property (2.13). Then for all $\boldsymbol{\nu} \in \Lambda \backslash \Lambda|_{\mathfrak{I}}$*

$$\varsigma_{\Lambda,\boldsymbol{\nu}} = \sum_{\{\mathbf{e} \in \{0,1\}^{\mathbb{N}} \,:\, \boldsymbol{\nu}+\mathbf{e} \in \Lambda\}} (-1)^{|\mathbf{e}|} = 0.$$

Proof. Fix $\boldsymbol{\nu} \in \Lambda \backslash (\Lambda|_{\mathfrak{I}})$. Since $\boldsymbol{\nu} \notin \Lambda|_{\mathfrak{I}}$, there exists $j \in \mathbb{N}$ with $\nu_j \notin \mathfrak{I}$. Set $A_j := \{\mathbf{e} = (e_i)_{i \in \mathbb{N}} \in \{0,1\}^{\mathbb{N}} : \boldsymbol{\nu} + \mathbf{e} \in \Lambda,\ e_j = 0\}$, and let $\mathbf{e} \in A_j$ arbitrary. By (2.13) it holds $\lceil \boldsymbol{\nu} + \mathbf{e} \rceil_{\mathfrak{I}} \in \Lambda$ since $\boldsymbol{\nu} + \mathbf{e} \in \Lambda$. Furthermore, with $\mathbf{e}_j = (\delta_{ij})_{i \in \mathbb{N}}$ we get $\lceil \boldsymbol{\nu} + \mathbf{e} + \mathbf{e}_j \rceil_{\mathfrak{I}} = \lceil \boldsymbol{\nu} + \mathbf{e} \rceil_{\mathfrak{I}}$ since $\nu_j \notin \mathfrak{I}$, and thus

$$A_j \cup \{\mathbf{e} + \mathbf{e}_j \,:\, \mathbf{e} \in A_j\} \subseteq \{\mathbf{e} \in \{0,1\}^{\mathbb{N}} \,:\, \boldsymbol{\nu} + \mathbf{e} \in \Lambda\}.$$

On the other hand, if $\boldsymbol{\delta} = (\delta_i)_{i \in \mathbb{N}} \in \{\mathbf{e} \in \{0,1\}^{\mathbb{N}} : \boldsymbol{\nu} + \mathbf{e} \in \Lambda\}$ and $\delta_j = 1$, then due to the downward closedness of Λ also $\boldsymbol{\nu} + \boldsymbol{\delta} - \mathbf{e}_j \in \Lambda$ which implies $\boldsymbol{\delta} - \mathbf{e}_j \in A_j$ and consequently

$$A_j \cup \{\mathbf{e} + \mathbf{e}_j \,:\, \mathbf{e} \in A_j\} \supseteq \{\mathbf{e} \in \{0,1\}^{\mathbb{N}} \,:\, \boldsymbol{\nu} + \mathbf{e} \in \Lambda\}.$$

Thus

$$\sum_{\{\mathbf{e} \in \{0,1\}^{\mathbb{N}} \,:\, \boldsymbol{\nu}+\mathbf{e} \in \Lambda\}} (-1)^{|\mathbf{e}|} = \sum_{\mathbf{e} \in A_j} (-1)^{|\mathbf{e}|} + \sum_{\mathbf{e} \in A_j} (-1)^{|\mathbf{e}+\mathbf{e}_j|} = \sum_{\mathbf{e} \in A_j} (-1)^{|\mathbf{e}|} - \sum_{\mathbf{e} \in A_j} (-1)^{|\mathbf{e}|} = 0.$$

$\square$

For a finite set $\Lambda \subseteq \mathcal{F}$ of multi-indices, the *effective dimension* $d(\Lambda)$ is defined as

$$d(\Lambda) := \sup_{\boldsymbol{\nu} \in \Lambda} |\operatorname{supp} \boldsymbol{\nu}|. \tag{2.14}$$

The proof of the following lemma is given in Appendix A.

Lemma 2.9. *Let $\mathfrak{I} \subseteq \mathbb{N}_0$ satisfy Assumption 2.4. Let $\Lambda \subseteq \mathcal{F}$ be finite and downward closed. Then*

$$\sum_{\boldsymbol{\nu} \in \Lambda} \prod_{j \in \mathbb{N}} (\nu_j + 1) \leq |\Lambda|^2 \qquad \text{and} \qquad \sum_{\boldsymbol{\nu} \in \Lambda|_{\mathfrak{I}}} \prod_{j \in \mathbb{N}} (\nu_j + 1) \leq K_{\mathfrak{I}}^{d(\Lambda)} |\Lambda|. \tag{2.15}$$

A key element of the present paper is the *a priori construction* of (sequences of) finite index sets $\Lambda \subseteq \mathcal{F}$ which capture provably the dominating part of gpc expansions of $(\boldsymbol{b}, \varepsilon)$-holomorphic maps. The index sets constructed in the following will satisfy

$$d(\Lambda) = o(\log(|\Lambda|)) \qquad \text{as } |\Lambda| \to \infty. \tag{2.16}$$

In this case, the number of quadrature points (also for non-nested points $\boldsymbol{\chi}$ in the sense of Def. 2.1) grows only slightly faster than linear in terms of the cardinality of the multiindex sets as the next lemma shows. Thus the properties (2.13) and (2.16) allow us to obtain good bounds on the number of required function evaluations also for non-nested quadrature points.

Lemma 2.10. *Fix $\delta > 0$. Let $\mathfrak{I}$ satisfy Assumption 2.4. Let $(\Lambda_\epsilon)_{\epsilon > 0}$ be a family of finite downward closed index sets satisfying (2.13) and (2.16). Let the quadrature points $\boldsymbol{\chi}$ be non-nested. Then*

$$|\mathrm{pts}(\Lambda_\epsilon, \boldsymbol{\chi})| \leq \sum_{\boldsymbol{\nu} \in \Lambda|_{\mathfrak{I}}} \prod_{j \in \mathbb{N}} (1 + \nu_j) = O(|\Lambda_\epsilon|^{1+\delta}) \qquad \text{as } |\Lambda_\epsilon| \to \infty.$$

Proof. W.l.o.g. we assume $|\Lambda_\epsilon| > 1$ for all $\epsilon > 0$ in the following, since the statement of the lemma only concerns the limit $|\Lambda_\epsilon| \to \infty$. Equation (2.16) is equivalently stated as: there exists a constant $C > 0$ and numbers $c_{|\Lambda_\epsilon|} > 0$ such that for all $\epsilon > 0$

$$d(\Lambda_\epsilon) \leq C \log(|\Lambda_\epsilon|) c_{|\Lambda_\epsilon|}$$

and $c_{|\Lambda_\epsilon|} \to 0$ as $|\Lambda_\epsilon| \to \infty$. We conclude

$$K_{\mathfrak{I}}^{d(\Lambda_\epsilon)} \leq K_{\mathfrak{I}}^{C \log(|\Lambda_\epsilon|) c_{|\Lambda_\epsilon|}} = e^{C \log(K_{\mathfrak{I}}) \log(|\Lambda_\epsilon|) c_{|\Lambda_\epsilon|}} = |\Lambda_\epsilon|^{C \log(K_{\mathfrak{I}}) c_{|\Lambda_\epsilon|}}.$$

Due to $c_{|\Lambda_\epsilon|} \to 0$, for any $\delta > 0$ the last term behaves like $O(|\Lambda_\epsilon|^\delta)$ as $|\Lambda_\epsilon| \to \infty$. The statement of the lemma now follows by (2.8), Lemmas 2.8 and 2.9. ☐

Remark 2.11. The bounds (2.15) are sharp in the following sense: Let $\Lambda = \{\boldsymbol{\nu} \in \mathcal{F} : \operatorname{supp} \boldsymbol{\nu} \subseteq \{1, \ldots, d\},\ \nu_j \leq N\ \forall j\}$ and set $\mathfrak{I} := \{0\} \cup \{2^j : j \in \mathbb{N}_0\}$. Then, with $N = 2^m$ for some $m \in \mathbb{N}$, we have $|\Lambda| = (N+1)^d$ and

$$\sum_{\boldsymbol{\nu} \in \Lambda} \prod_{j \in \mathbb{N}} (\nu_j + 1) = \prod_{j=1}^{d} \sum_{i=1}^{N+1} i = \left(\frac{(N+1)(N+2)}{2} \right)^d \geq 2^{-d} ((N+1)^d)^2 = 2^{-d} |\Lambda|^2, \tag{2.17}$$

as well as

$$\sum_{\boldsymbol{\nu} \in \Lambda|_{\mathfrak{I}}} \prod_{j \in \mathbb{N}} (\nu_j + 1) = \prod_{j=1}^{d} \left(1 + \sum_{i=0}^{m} (2^i + 1) \right) \geq \prod_{j=1}^{d} (1 + 2^{m+1} - 1 + m + 1) \geq (2(2^m + 1))^d \geq 2^d (N+1)^d = 2^d |\Lambda|. \tag{2.18}$$

Letting $N \to \infty$ in (2.17) and $d \to \infty$ in (2.18), a better asymptotic behaviour than quadratic in $|\Lambda|$ in the first case, and linear in $|\Lambda|$ with a constant depending exponentially on $d(\Lambda)$ in the second case cannot be expected in general.

However, these estimates may not accurately measure the actual number of function evaluations required in (2.5), since they do not take into account the fact that some (further) combination coefficients in (2.5) might vanish. Indeed, for the above example Q_Λ is the tensor product quadrature $Q_{\boldsymbol{\nu}}$ with $\nu_j = N$ if $j \le d$ and $\nu_j = 0$ otherwise. The number of function evaluations is then equal to $|\Lambda| = (N+1)^d$.

2.4. Computational cost

In the following let $u : U \to X$ be a pointwise defined function and let $\Lambda \subseteq \mathcal{F}$ be a finite downward closed index set. While the number of function evaluations is in practice a good indicator of the computational cost (in particular for PDEs where evaluating u is computationally intensive), we also analyse the error of the Smolyak quadrature in terms of the number of floating point operations required to compute $Q_\Lambda u$.

Remark 2.12. We stress that the term "computational cost" in the following merely refers to the computational complexity of evaluating $Q_\Lambda u$ in (2.5), essentially under the assumption that each evaluation of u at a point $\boldsymbol{y} \in U$ has computational complexity $O(1)$ (this will be slightly relaxed in Assumption 2.13). In particular, our present analysis does *not take into account the cost of approximating the integrand* $u(\boldsymbol{y})$, in case $u(\boldsymbol{y})$ cannot be evaluated exactly. For UQ problems, this is usually the case however, as $u(\boldsymbol{y})$ typically denotes the solution to a PDE whose coefficients depend on $\boldsymbol{y} \in U$ (*cp.* Example 2.18). While such a discussion is outside the scope of this manuscript, in [37,38] we provide an analysis of the full computational complexity (taking into account the error and the computational work stemming from the approximation of $u(\boldsymbol{y})$) of a multilevel version of the here analysed Smolyak algorithm.

We now make an assumption regarding the computational complexity of evaluating u.

Assumption 2.13. *There exists a constant $C > 0$ such that for every $\boldsymbol{\nu} \in \mathcal{F}$, u can be evaluated at each $(\chi_{\nu_j;\mu_j})_{j\in\mathbb{N}}$ for $\boldsymbol{\mu} \le \boldsymbol{\nu}$ with a number of floating point operations that is bounded by $C|\operatorname{supp}\boldsymbol{\nu}|$.*

Remark 2.14. Consider a function $u(\boldsymbol{y}) = \mathfrak{u}(\sum_{j\in\mathbb{N}} y_j\psi_j)$ as in (1.2) where $\mathfrak{u} : \mathbb{C} \to \mathbb{C}$. If $\chi_{0;0} = 0$, then the computation of $\sum_{j\in\mathbb{N}} \chi_{\nu_j;\mu_j}\psi_j = \sum_{j\in\operatorname{supp}\boldsymbol{\nu}} \chi_{\nu_j;\mu_j}\psi_j$ requires $|\operatorname{supp}\boldsymbol{\nu}|$ multiplications and $|\operatorname{supp}\boldsymbol{\nu}|-1$ additions. If $\mathfrak{u}$ can be evaluated with $O(1)$ floating point operations, then Assumption 2.13 is satisfied.

Less generally, if $u(\boldsymbol{y})$ can be evaluated with complexity $O(1)$ at every $\boldsymbol{y} \in U$, then clearly Assumption 2.13 is also fulfilled.

We point out again, that for parametric PDEs, *i.e.* where $u(\boldsymbol{y}) \in X$ denotes the solution of a PDE in a Sobolev space X, Assumption 2.13 is usually *not satisfied*. This is because $u(\boldsymbol{y})$ is typically unknown and has to be *approximated* by a numerical method such as the finite element method. The cost of evaluating an approximation of u is then linked to the discretization in space of the numerical PDE solver.

Additional to the effective dimension $d(\Lambda)$ in (2.14), the *maximal total order*

$$m(\Lambda) := \max_{\boldsymbol{\nu}\in\Lambda} |\boldsymbol{\nu}| \tag{2.19}$$

has a certain significance when analysing the computational complexity.

To bound the cost of evaluating the Smolyak quadrature $Q_\Lambda u$, we use the representations (2.4) and (2.5).

- The coefficients $(\varsigma_{\Lambda,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda} = \sum_{\{\mathbf{e}\in\{0,1\}^{\mathbb{N}}\,:\,\boldsymbol{\nu}+\mathbf{e}\in\Lambda\}} (-1)^{|\mathbf{e}|}$ can be computed with a number of floating point operations bounded by $Cd(\Lambda)|\Lambda|2^{d(\Lambda)}$: this is achieved by looping over all $\boldsymbol{\nu} \in \Lambda$, and updating the coefficient of all (at most $2^{d(\Lambda)}$) neighbours in Λ of the type $\boldsymbol{\nu} - \mathbf{e}$ for some $\mathbf{e} \in \{0,1\}^{\mathbb{N}}$ (this implies $\operatorname{supp}\mathbf{e} \subseteq \operatorname{supp}\boldsymbol{\nu}$). The computation of $|\mathbf{e}| = \sum_{j\in\operatorname{supp}\mathbf{e}} 1$ requires at most $|\operatorname{supp}\boldsymbol{\nu}| - 1 \le d(\Lambda)$ additions.

- Evaluating $Q_{\boldsymbol{\nu}}u$ in (2.4) requires knowledge of the quadrature weights $(\alpha_{n;j})$ for $j=0,\dots,n$ all $0\le n\le \max_{\boldsymbol{\nu}\in\Lambda}\nu_j\le m(\Lambda)$. These weights can be computed by solving a linear system of dimension $n\times n$. Hence this part contributes at most $C\sum_{n=0}^{m(\Lambda)} n^3\le Cm(\Lambda)^4$ floating point operations.
- To compute $Q_{\boldsymbol{\nu}}u$ in (2.4) we need to evaluate u at all points in $\{(\chi_{\nu_j;\mu_j})_{j\in\mathbb{N}} : \boldsymbol{\mu}\le\boldsymbol{\nu}\}$. Under Assumption 2.13 this requires at most $Cd(\Lambda)\prod_{j\in\mathbb{N}}(1+\nu_j)$ floating point operations, since $|\{\boldsymbol{\mu}\in\mathcal{F} : \boldsymbol{\mu}\le\boldsymbol{\nu}\}|=\prod_{j\in\mathbb{N}}(1+\nu_j)$. The computation of the quadrature weight $\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\alpha_{\nu_j;\mu_j}$ for all $\boldsymbol{\mu}\le\boldsymbol{\nu}$ requires at most $d(\Lambda)\prod_{j\in\mathbb{N}}(1+\nu_j)$ floating point operations. The summation over all $\boldsymbol{\mu}\le\boldsymbol{\nu}$ is again of complexity $\prod_{j\in\mathbb{N}}(1+\nu_j)$.

In all, we introduce

$$\operatorname{cost}(\Lambda):=\underbrace{m(\Lambda)^4}_{\text{comp. of }((\alpha_{n,j})_{j=0}^n)_{n=0}^{m(\Lambda)}}+\underbrace{d(\Lambda)2^{d(\Lambda)}|\Lambda|}_{\text{comp. of }(\varsigma_{\Lambda,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda}}+\sum_{\{\boldsymbol{\nu}\in\Lambda\,:\,\varsigma_{\Lambda,\boldsymbol{\nu}}\neq0\}}\underbrace{d(\Lambda)\prod_{j\in\mathbb{N}}(\nu_j+1)}_{\text{evaluation of }Q_{\boldsymbol{\nu}}u},\tag{2.20}$$

as a measure for the cost of evaluating the Smolyak quadrature $Q_\Lambda u$. As a consequence of Lemma 2.10 we obtain an asymptotic bound on the cost term *defined* in (2.20).

Lemma 2.15. *Fix $\delta>0$. Let $\mathfrak{I}$ satisfy Assumption 2.4. Let $(\Lambda_\epsilon)_{\epsilon>0}$ be a family of finite, downward closed index sets satisfying* (2.13)*. Let further*

$$d(\Lambda_\epsilon)=o(\log|\Lambda_\epsilon|)\qquad\text{and}\qquad m(\Lambda_\epsilon)=O(\log|\Lambda_\epsilon|)\qquad\text{as }|\Lambda_\epsilon|\to\infty.\tag{2.21}$$

Then with $\operatorname{cost}(\Lambda_\epsilon)$ *as in* (2.20)

$$\operatorname{cost}(\Lambda_\epsilon)=O(|\Lambda_\epsilon|^{1+\delta})\qquad\text{as }|\Lambda_\epsilon|\to\infty.$$

2.5. Main result

Let Z and X be two complex Banach spaces. Recall that $B_r^Z=\{\phi\in Z : \|\phi\|_Z<r\}$. A function $\mathfrak{u}:B_r^Z\to X$ is called holomorphic, if it is Fréchet differentiable. The following theorem is our main result. In the subsequent sections, we prove a slight generalization of this statement, and also provide details on the explicit construction of the index sets (see Thm. 4.3). The cost term in the formulation of the theorem was *defined* in (2.20), and we mention again that it can be interpreted as a measure of the computational cost of evaluating (2.5) *under Assumption 2.13*, also *cp.* Remark 2.12.

Theorem 2.16. *Let $(\psi_j)_{j\in\mathbb{N}}\subseteq Z$, $r>0$ and $p\in(0,1)$. Fix $\delta>0$ arbitrarily small. Assume that*

(i) $\sum_{j\in\mathbb{N}}\|\psi_j\|_Z<r$ *and* $(\|\psi_j\|_Z)_{j\in\mathbb{N}}\in\ell^p(\mathbb{N})\hookrightarrow\ell^1(\mathbb{N})$,
(ii) $\mathfrak{u}:B_r^Z\to X$ *is holomorphic and bounded,*
(iii) the quadrature points $\boldsymbol{\chi}$ *(either nested or non-nested) satisfy* (2.3).

For $\boldsymbol{y}\in U=[-1,1]^{\mathbb{N}}$ set $u(\boldsymbol{y}):=\mathfrak{u}(\sum_{j\in\mathbb{N}}y_j\psi_j)$. Then, there exists a constant $C>0$ such that for every $\epsilon>0$ there exists a finite downward closed multiindex set $\Lambda_\epsilon\subseteq\mathcal{F}$ with $|\Lambda_\epsilon|\to\infty$ as $\epsilon\to0$ and such that

$$\left\|\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y})-Q_{\Lambda_\epsilon}u\right\|_X\le C|\operatorname{pts}(\Lambda_\epsilon,\boldsymbol{\chi})|^{-\frac{2}{p}+1+\delta},\tag{2.22a}$$

and additionally

$$\left\|\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y})-Q_{\Lambda_\epsilon}u\right\|_X\le C\operatorname{cost}(\Lambda_\epsilon)^{-\frac{2}{p}+1+\delta}.\tag{2.22b}$$

Remark 2.17. More generally, in [37] we prove the following variant of Theorem 2.16, which merely assumes $\mathfrak{u}$ to be holomorphic on some open set containing all inputs rather than a ball (see (ii) below):

Let $(\psi_j)_{j\in\mathbb{N}}\subseteq Z$, $r>0$ and $p\in(0,1)$. Fix $\delta>0$ arbitrarily small. Assume that

(i) $(\|\psi_j\|_Z)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N}) \hookrightarrow \ell^1(\mathbb{N})$,
(ii) there is an open set $O \subseteq Z$ such that $\{\sum_{j\in\mathbb{N}} y_j\psi_j \ : \ \boldsymbol{y} \in U\} \subseteq O$ and $\mathfrak{u} : O \to X$ is holomorphic and bounded,
(iii) the quadrature points $\boldsymbol{\chi}$ (either nested or non-nested) satisfy (2.3).

For $\boldsymbol{y} \in U = [-1,1]^{\mathbb{N}}$ set $u(\boldsymbol{y}) := \mathfrak{u}(\sum_{j\in\mathbb{N}} y_j\psi_j)$. Then, there exists $C > 0$ such that for every $\epsilon > 0$ exists a finite downward closed multiindex set $\Lambda_\epsilon \subseteq \mathcal{F}$ such that $|\Lambda_\epsilon| \to \infty$ as $\epsilon \to 0$ and (2.22) holds.

The proof in [37] also covers general Jacobi (probability) measures whose density on $[-1,1]$ is given by $(1-x)^\alpha(1+x)^\beta C_{\alpha,\beta}$ where α, $\beta > -1$ and $C_{\alpha,\beta} = \Gamma(\alpha+\beta+2)/(2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\beta+1))$. For brevity, we provide here a proof of Theorem 2.16 corresponding to $\alpha = \beta = 0$, under stronger assumptions on the domain of holomorphy of $\mathfrak{u}$. This allows to avoid certain technicalities.

In view of Lemmas 2.10 and 2.15, it suffices to prove the asymptotic bounds (2.22) in terms of the cardinality $|\Lambda_\epsilon|$ of the multiindex sets, and to verify that Λ_ϵ complies with the assumptions of Lemmas 2.10 and 2.15. Furthermore we shall see that in case of nested points (2.22a) also holds with $\delta = 0$ (as a consequence of Lem. 2.2). We now give an example of a holomorphic function $\mathfrak{u}$ as in Theorem 2.16.

Example 2.18. Let $d \in \mathbb{N}$. Let $\mathrm{D} \subseteq \mathbb{R}^d$ be a bounded (nonempty) Lipschitz domain and set $X := H_0^1(\mathrm{D};\mathbb{C})$ so that $X' = H^{-1}(\mathrm{D};\mathbb{C})$. For $\psi \in Z := L^\infty(\mathrm{D};\mathbb{C})$ define the bounded linear operator $A(\psi) \in L(X,X')$ by

$$\langle A(\psi)u, v\rangle = \int_{\mathrm{D}} \psi \nabla u^\top \nabla v \,\mathrm{d}x.$$

Then $A \in L(Z, L(X,X'))$, and with the norm $\|u\|_X^2 := \int_{\mathrm{D}} \nabla u^\top \overline{\nabla u}\,\mathrm{d}x$ on X (here $\overline{\nabla u}$ is the complex conjugate) it holds

$$\|A\|_{L(Z,L(X,X'))} = \sup_{\|\psi\|_Z=1} \sup_{\|u\|_X=1} \sup_{\|v\|_X=1} |\langle A(\psi)u, v\rangle| = 1.$$

Suppose that $\psi_0 \in L^\infty(\mathrm{D};\mathbb{R})$ satisfies $0 < \varrho \le \psi_0(x)$ a.e. in D. Then by the (complex) Lax–Milgram Lemma, $A(\psi_0) : X \to X'$ is an isomorphism and $\|A(\psi_0)^{-1}\|_{L(X',X)} \le \varrho^{-1}$. For any $\psi \in Z$ it holds

$$\|A(\psi) - A(\psi_0)\|_{L(X,X')} = \|A(\psi - \psi_0)\|_{L(X,X')} \le \|\psi - \psi_0\|_Z.$$

Using a Neumann series, if $\|\psi - \psi_0\|_Z < \|A(\psi_0)^{-1}\|_{L(X',X)}^{-1}$, then $A(\psi) : X \to X'$ is also an isomorphism and

$$A(\psi)^{-1} = (A(\psi_0) - A(\psi_0 - \psi))^{-1} = (I - A(\psi_0)^{-1}A(\psi_0-\psi))^{-1}A(\psi_0)^{-1} = \sum_{n\in\mathbb{N}_0}(A(\psi_0)^{-1}A(\psi_0-\psi))^n A(\psi_0)^{-1}.$$

Since $(A(\psi_0)^{-1}A(h))^n A(\psi_0)^{-1} \in L(X,X)$ can be interpreted as an n-linear function of $h^n \in Z^n$, this constitutes a power series expansion (in Banach spaces) of $\psi \mapsto A(\psi)^{-1} \in L(X',X)$ around ψ_0. Due to

$$\|(A(\psi_0)^{-1}A(h))^n A(\psi_0)^{-1}\|_{L(X',X)} \le C\|h\|_Z^n \|A(\psi_0)^{-1}\|_{L(X',X)}^n$$

the power series converges to a uniformly bounded function for all elements of $\{h \in Z \ : \ \|h - \psi_0\|_Z < \|A(\psi_0)^{-1}\|_{L(X',X)}^{-1}\}$, and it is Fréchet differentiable (*i.e.* holomorphic) as a function of $\psi \in Z$ there, which is classical (see *e.g.* [7], 14.13).

Fix $F \in X'$. We showed that the solution operator $\mathfrak{u}$ mapping a diffusion coefficient $\psi \in Z$ to the unique solution $\mathfrak{u}(\psi) \in X$ of

$$\int_{\mathrm{D}} \psi \nabla \mathfrak{u}(\psi)^\top \nabla v \,\mathrm{d}x = F(v)$$

is locally a well-defined holomorphic map around $\psi_0 \in Z$, since it is given by $\mathfrak{u}(\psi) = A(\psi)^{-1}F$ and $\psi \mapsto A(\psi)^{-1}$ is holomorphic (for more details see [37], Chap. 1).

Assume that $(\psi_j)_{j\in\mathbb{N}} \subseteq Z$ and $p \in (0,1)$ are such that

$$\sum_{j\in\mathbb{N}} \|\psi_j\|_Z < \|A(\psi_0)^{-1}\|_{L(X',X)}^{-1} \qquad \text{and} \qquad (\|\psi_j\|_Z)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N}).$$

By Theorem 2.16, the Smolyak quadrature allows to approximate the Bochner integral $\int_U \mathfrak{u}(\psi_0 + \sum_{j\in\mathbb{N}} y_j\psi_j)\,\mathrm{d}\mu(\boldsymbol{y}) \in X$ with (essentially) the convergence rate $2/p - 1$.

The argument in the above example was completely independent of the concrete differential operator. The same calculation holds for any linear (differential) operator $A(\psi_0) \in L(X, X')$ which is an isomorphism and depends linearly on the data ψ_0 in some Banach space Z.

3. Summability of Taylor GPC coefficients

With $U := [-1,1]^{\mathbb{N}}$, consider $u : U \to X$, for some fixed Banach space X over $\mathbb{C}$. In this section we are concerned with the Taylor expansion

$$u(\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}} \tag{3.1}$$

of u and the summability properties of the Taylor gpc coefficients $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}}$.

3.1. $(\mathbf{b}, \boldsymbol{\varepsilon})$-holomorphy and GPC expansions

In the following Z and X are two complex Banach spaces. We now characterize the functions in Theorem 2.16 in terms of their domains of holomorphic extension. We show that they satisfy the conditions summarized in the notion of $(\boldsymbol{b}, \varepsilon)$-holomorphy, which is introduced next. This definition has similarly been used for example in [12, 13, 25].

Definition 3.1. Let $\varepsilon > 0$, $p \in (0,1)$ and $M_u > 0$. For a given sequence $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$, we say that $u : U \to X$ is $(\boldsymbol{b}, \varepsilon)$-*holomorphic*, if

(i) $u : U \to X$ is continuous,
(ii) for every sequence $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}} \subseteq (1,\infty)$ which is $(\boldsymbol{b}, \varepsilon)$-*admissible*, *i.e.* satisfies

$$\sum_{j\in\mathbb{N}} b_j(\rho_j - 1) \le \varepsilon, \tag{3.2}$$

u allows a separately holomorphic extension onto the polydisc $B_{\boldsymbol{\rho}}^{\mathbb{C}} = \times_{j\in\mathbb{N}} B_{\rho_j}^{\mathbb{C}}$ (this extension is denoted by the same symbol u in the following),
(iii) for every $(\boldsymbol{b}, \varepsilon)$-admissible sequence the extension from (ii) satisfies

$$\sup_{\boldsymbol{z}\in B_{\boldsymbol{\rho}}^{\mathbb{C}}} \|u(\boldsymbol{z})\|_X \le M_u < \infty, \tag{3.3}$$

and for two $(\boldsymbol{b}, \varepsilon)$-admissible sequences $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ the extensions from (ii) coincide on $B_{\boldsymbol{\rho}_1}^{\mathbb{C}} \cap B_{\boldsymbol{\rho}_2}^{\mathbb{C}}$.

We start with a statement about continuity, and recall that any subset of $S \subseteq \mathbb{C}^{\mathbb{N}}$ (such as $U = [-1,1]^{\mathbb{N}}$) is considered with the product topology. Hence

$$\left\{ S \cap \left(\bigtimes_{j=1}^{N} O_j \times \bigtimes_{j>N} \mathbb{C} \right) \,:\, N \in \mathbb{N},\ O_j \subseteq \mathbb{C} \quad \text{is open} \quad \forall j \in \{1,\ldots,N\} \right\}$$

is a basis of the topology on S.

Lemma 3.2. *Let $(\psi_j)_{j\in\mathbb{N}} \subseteq Z$ satisfy $(\|\psi_j\|_Z)_{j\in\mathbb{N}} \in \ell^1(\mathbb{N})$. Then $\boldsymbol{y} \mapsto \sum_{j\in\mathbb{N}} y_j\psi_j$ is continuous from U to Z.*

Proof. Fix $\epsilon > 0$ and $\boldsymbol{y} \in U$. We need to find an open set $O \subseteq U$ (open w.r.t. the topology on U) such that $\|\sum_{j\in\mathbb{N}} y_j\psi_j - \sum_{j\in\mathbb{N}} z_j\psi_j\|_Z < \epsilon$ for all $\boldsymbol{z} = (z_j)_{j\in\mathbb{N}} \in O$. Let $J \in \mathbb{N}$ be so large that $\sum_{j>J} \|\psi_j\|_Z < \epsilon/4$. Let $\delta := \epsilon/(2J)$. Then for every $\boldsymbol{z} \in O := \bigtimes_{j=1}^J \{z \in [-1,1] : |z - y_j| < \delta\} \times \bigtimes_{j>J} [-1,1]$

$$\left\|\sum_{j\in\mathbb{N}} y_j\psi_j - \sum_{j\in\mathbb{N}} z_j\psi_j\right\|_Z < \sum_{j=1}^J \delta + \sum_{j>J} 2\|\psi_j\|_Z < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Lemma 3.3. *Let $\varepsilon > 0$, $p \in (0,1)$ and $M_u > 0$. For a sequence $(\psi_j)_{j\in\mathbb{N}} \subseteq Z$ and a sequence $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}}$ assume that $\|\psi_j\|_Z \le b_j$ for all $j \in \mathbb{N}$, and $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$. With $r := \|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} + \varepsilon$ assume that $\mathfrak{u} : B_r^Z \to X$ is holomorphic (i.e. Fréchet differentiable) and $\sup_{\phi\in B_r^Z} \|\mathfrak{u}(\phi)\|_X \le M_u$. For $\boldsymbol{y} \in U$ define $u(\boldsymbol{y}) = \mathfrak{u}(\sum_{j\in\mathbb{N}} y_j\psi_j)$. Then u is $(\boldsymbol{b},\varepsilon)$-holomorphic.*

Proof. The map $u : U \to X$ defined as $u(\boldsymbol{y}) = \mathfrak{u}(\sum_{j\in\mathbb{N}} y_j\psi_j)$ is continuous, since $\mathfrak{u} : B_r^Z \to X$ is continuous (even holomorphic) and $\boldsymbol{y} \mapsto \sum_{j\in\mathbb{N}} y_j\psi_j$ is continuous from U to Z by Lemma 3.2.

Let $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}} \subseteq (1,\infty)$ be $(\boldsymbol{b},\varepsilon)$-admissible, *i.e.* $\boldsymbol{\rho}$ satisfies (3.2). Fix $\boldsymbol{z} \in B_{\boldsymbol{\rho}}^{\mathbb{C}} \subseteq \mathbb{C}^{\mathbb{N}}$. Then

$$\sum_{j\in\mathbb{N}} |z_j|\|\psi_j\|_Z \le \sum_{j\in\mathbb{N}} \|\psi_j\|_Z + \sum_{j\in\mathbb{N}} (\rho_j - 1)\|\psi_j\|_Z \le \sum_{j\in\mathbb{N}} b_j + \sum_{j\in\mathbb{N}} (\rho_j - 1)b_j \le \|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} + \varepsilon \le r. \tag{3.4}$$

Therefore $\sum_{j\in\mathbb{N}} z_j\psi_j \in Z$ is well-defined. Moreover, $\sum_{j\in\mathbb{N}} z_j\psi_j \in B_r^Z$.

Now fix $j \in \mathbb{N}$ and $(z_i)_{i\neq j} \in \bigtimes_{i\neq j} B_{\rho_i}^{\mathbb{C}}$. Then $z_j \mapsto \sum_{j\in\mathbb{N}} z_j\psi_j$ is an affine bounded (and thus holomorphic) map from $B_{\rho_j}^{\mathbb{C}} \to B_r^Z \subseteq Z$. Due to the holomorphy of $\mathfrak{u} : B_r^Z \to X$, we obtain that $u(\boldsymbol{z}) = \mathfrak{u}(\sum_{j\in\mathbb{N}} z_j\psi_j)$ is holomorphic as a function of $z_j \in B_{\rho_j}^{\mathbb{C}}$, which shows (ii).

For two $(\boldsymbol{b},\varepsilon)$-admissible sequences $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$, by definition their corresponding extensions agree on $B_{\boldsymbol{\rho}_1}^{\mathbb{C}} \cap B_{\boldsymbol{\rho}_1}^{\mathbb{C}}$. Finally, (3.3) follows by $\{\sum_{j\in\mathbb{N}} z_j\psi_j : \boldsymbol{z} \in B_{\boldsymbol{\rho}}^{\mathbb{C}}\} \subseteq B_r^Z \subseteq Z$ whenever $\boldsymbol{\rho}$ is $(\boldsymbol{b},\varepsilon)$-holomorphic, and the assumption $\sup_{\phi\in B_r^Z} \|\mathfrak{u}(\phi)\|_X \le M_u$. □

Next, we recall bounds on the norms of the Taylor coefficients.

The next lemma is essentially a consequence of the Cauchy integral theorem ([24], Thm. 2.1.2), see the proof of Lemma 2.4 from [10].

Lemma 3.4. *Let $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}} \subseteq (1,\infty)$ and assume that $u : B_{\boldsymbol{\rho}}^{\mathbb{C}} \to X$ is separately holomorphic (i.e. holomorphic in each variable), such that $\sup_{\boldsymbol{y}\in B_{\boldsymbol{\rho}}^{\mathbb{C}}} \|u(\boldsymbol{y})\|_X \le M_u < \infty$. Then for every $\boldsymbol{\nu} \in \mathcal{F}$ the Taylor gpc coefficient*

$$t_{\boldsymbol{\nu}} := \frac{\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{y})}{\boldsymbol{\nu}!}|_{\boldsymbol{y}=\boldsymbol{0}} \in X \tag{3.5}$$

satisfies the bound

$$\|t_{\boldsymbol{\nu}}\|_X \le M_u \boldsymbol{\rho}^{-\boldsymbol{\nu}}. \tag{3.6}$$

In Section 3.3 we will show that $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^1(\mathcal{F})$ for $(\boldsymbol{b},\varepsilon)$-holomorphic functions. This implies that the series $\sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}} \in X$ is pointwise well-defined for every $\boldsymbol{y} \in U$. In this case the expansion converges to $u(\boldsymbol{y})$, as recalled in the next Lemma. For a proof see, *e.g.*, Proposition 2.1.4 of [37]. Absolute convergence of a series $\sum_{j\in\mathbb{N}} x_j$ in a Banach space X means $\sum_{j\in\mathbb{N}} \|x_j\|_X < \infty$.

Lemma 3.5. *Let $p \in (0,1)$, $\varepsilon > 0$ and $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$. Let $u : U \to X$ be $(\boldsymbol{b},\varepsilon)$-holomorphic and assume that $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^1(\mathcal{F})$. Then $u(\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}}$ with uniform and absolute convergence for all $\boldsymbol{y} \in U$.*

3.2. Multiindex sets

Lemma 3.5 states that $(\boldsymbol{b},\varepsilon)$-holomorphic functions $u : U \to X$ allow representations as Taylor expansions $u(\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}$ in infinitely many variables. For a finite subset $\Lambda \subseteq \mathcal{F}$, the function $\tilde{u}(\boldsymbol{y}) := \sum_{\boldsymbol{\nu}\in\Lambda} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}$ defines an approximation to u, and for every $\boldsymbol{y} \in U$ the error can be bounded by $\|u(\boldsymbol{y}) - \tilde{u}(\boldsymbol{y})\|_X \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}\backslash\Lambda} \|t_{\boldsymbol{\nu}}\|_X$. This line of argument leads to best N-term rates, and determining suitable index sets Λ (possibly minimizing $\sum_{\boldsymbol{\nu}\in\mathcal{F}\backslash\Lambda} \|t_{\boldsymbol{\nu}}\|_X$) is typically the first step required to prove convergence rates for numerical algorithms. In order to obtain good bounds of the computational complexity, we aim to devise Λ in such a way that the asymptotics (2.21) as well as (2.13) are satisfied. This is the topic of the current subsection.

Definition 3.6. We say that $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq [0,\infty)$ is a *monotonically decreasing* sequence if $\boldsymbol{\nu} \leq \boldsymbol{\mu}$ implies $a_{\boldsymbol{\nu}} \geq a_{\boldsymbol{\mu}}$ for all $\boldsymbol{\nu}$, $\boldsymbol{\mu} \in \mathcal{F}$.

The following assumption gathers all properties required of $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, such that the set

$$\Lambda_\epsilon((a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, a_{\boldsymbol{\nu}} \geq \epsilon\}$$

satisfies the assumptions of Lemmas 2.10 and 2.15. This is shown subsequently.

Assumption 3.7. *There exist constants $C_0 > 0$, $C_\varkappa > 0$, $\varkappa > 0$, $\delta > 1$, a sequence $(f_d)_{d\in\mathbb{N}} \subseteq (0,\infty)$ with $f_d \to \infty$ as $d \to \infty$ and a set $\mathfrak{I} \subseteq \mathbb{N}_0$ satisfying Assumption 2.4 (i), such that the sequence $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq [0,\infty)$ satisfies*

(i) $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ *is monotonically decreasing (see Def. 3.6),*

(ii) $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ *has the property*

$$\lceil \boldsymbol{\nu} \rceil_{\mathfrak{I}} = \lceil \boldsymbol{\mu} \rceil_{\mathfrak{I}} \quad \Rightarrow \quad a_{\boldsymbol{\nu}} = a_{\boldsymbol{\mu}}, \tag{3.7}$$

(iii) with a decreasing rearrangement $(a_j^*)_{j\in\mathbb{N}}$ of $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ it holds

$$\begin{aligned} a_j^* &\geq C_\varkappa j^{-\varkappa} && \forall j \in \mathbb{N}, \\ \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} a_{\boldsymbol{\nu}} &\leq C_0 \delta^{-d} && \forall d \in \mathbb{N}, \\ \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} a_{\boldsymbol{\nu}} &\leq C_0\, \mathrm{e}^{-d f_d} && \forall d \in \mathbb{N}. \end{aligned} \tag{3.8}$$

Lemma 3.8. *Let $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq [0,\infty)$ satisfy Assumption 3.7 and assume that $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^q(\mathcal{F})$ for some $q > 0$. Then, for every $\epsilon > 0$ the set $\Lambda_\epsilon = \Lambda_\epsilon((a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) := \{\boldsymbol{\nu} \in \mathcal{F} \,:\, a_{\boldsymbol{\nu}} \geq \epsilon\}$ satisfies*

(i) Λ_ϵ *is finite and downward closed,*

(ii) it holds

$$(\boldsymbol{\nu} \in \Lambda_\epsilon \quad \text{and} \quad \lceil \boldsymbol{\mu} \rceil_{\mathfrak{I}} = \lceil \boldsymbol{\nu} \rceil_{\mathfrak{I}}) \quad \Rightarrow \quad \boldsymbol{\mu} \in \Lambda_\epsilon,$$

(iii) it holds

$$d(\Lambda_\epsilon) = o(\log(|\Lambda_\epsilon|)) \qquad \text{and} \qquad m(\Lambda_\epsilon) = O(\log(|\Lambda_\epsilon|)) \qquad \text{as } \epsilon \to 0. \tag{3.9}$$

Proof. Fix $\epsilon > 0$. Assume that $\boldsymbol{\nu} \leq \boldsymbol{\mu}$ and $\boldsymbol{\mu} \in \Lambda_\epsilon$. Then $a_{\boldsymbol{\mu}} \geq \epsilon$ and due to monotonicity $a_{\boldsymbol{\nu}} \geq a_{\boldsymbol{\mu}} \geq \epsilon$ so that $\boldsymbol{\nu} \in \Lambda_\epsilon$. This and the fact that $\sum_{\boldsymbol{\nu}\in\mathcal{F}} a_{\boldsymbol{\nu}}^q < \infty$ show (i). Item (ii) is an immediate consequence of (3.7) and the definition of Λ_ϵ.

To show the first statement in (3.9), note that with $\varkappa$ and $C_\varkappa$ as in Assumption 3.7 (iii), for any $d_0 \in \mathbb{N}$ and any $\epsilon > 0$ (such that $|\Lambda_\epsilon| > 0$) it holds

$$d(\Lambda_\epsilon) \geq d_0 \quad \Rightarrow \quad \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d_0\}} a_{\boldsymbol{\nu}} \geq \min_{\boldsymbol{\nu}\in\Lambda_\epsilon} a_{\boldsymbol{\nu}} \geq C_\varkappa |\Lambda_\epsilon|^{-\varkappa}.$$

Moreover, we may write $\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\ge d_0\}} a_{\boldsymbol{\nu}} \le C_0\exp(-d_0 f_{d_0})$ for the sequence $(f_d)_{d\in\mathbb{N}}$ and $C_0>0$ as in Assumption 3.7 (*i.e.* $f_d\to\infty$ as $d\to\infty$). Hence

$$\begin{aligned} d(\Lambda_\epsilon) &= \max\{d_0\in\mathbb{N}\,:\,d(\Lambda_\epsilon)\ge d_0\}\\ &\le \max\{d_0\in\mathbb{N}\,:\,C_0\exp(-d_0 f_{d_0})\ge C_\varkappa|\Lambda_\epsilon|^{-\varkappa}\}\\ &= \max\{d_0\in\mathbb{N}\,:\,d_0 f_{d_0}\le -\log(C_\varkappa/C_0)+r\log(|\Lambda_\epsilon|)\}. \end{aligned}$$

For $x\ge\inf_{d\in\mathbb{N}} df_d$ set $g(x):=\max\{d_0\in\mathbb{N}\,:\,d_0f_{d_0}\le x\}$. We claim that $g(x)=o(x)$ as $x\to\infty$. Assume on the contrary that $\limsup_{x\to\infty} g(x)/x\ne 0$. Then there exists a sequence $(x_j)_{j\in\mathbb{N}}$ with $x_j\to\infty$ and a positive constant C such that $g(x_j)\ge Cx_j$ for all $j\in\mathbb{N}$. For every $j\in\mathbb{N}$, let $d_j:=g(x_j)$. Then

$$Cx_jf_{d_j}\le g(x_j)f_{d_j}=d_jf_{d_j}\le x_j\qquad\forall j\in\mathbb{N},$$

which is a contradiction since $f_{d_j}\to\infty$ as $d_j\to\infty$. Hence $g(x)=o(x)$ as $x\to\infty$. This shows $d(\Lambda_\epsilon)=o(\log(|\Lambda_\epsilon|))$ as $|\Lambda_\epsilon|\to\infty$ or equivalently as $\epsilon\to 0$.

For $m(\Lambda_\epsilon)$ we proceed similarly. It holds for any $d_0\in\mathbb{N}$

$$m(\Lambda_\epsilon)\ge d_0\quad\Rightarrow\quad\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\ge d_0\}} a_{\boldsymbol{\nu}}\ge\min_{\boldsymbol{\nu}\in\Lambda_\epsilon} a_{\boldsymbol{\nu}}\ge C_\varkappa|\Lambda_\epsilon|^{-\varkappa}.$$

By assumption $\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\ge d_0\}} a_{\boldsymbol{\nu}}\le C_0\delta^{-d_0}$ for some $\delta>1$ and some $C_0>0$. Hence

$$m(\Lambda_\epsilon)\le\max\{d_0\in\mathbb{N}\,:\,C_0\delta^{-d_0}\ge C_\varkappa|\Lambda_\epsilon|^{-\varkappa}\}\le\frac{-\log(C_\varkappa/C_0)+\varkappa\log(|\Lambda_\epsilon|)}{\log(\delta)}=O(\log(|\Lambda_\epsilon|)),$$

which concludes the proof. $\square$

The next lemma facilitates the construction of sequences satisfying (3.7) (while leaving the asymptotic decay properties of the sequence unchanged). For its formulation recall the set $\mathfrak{I}_+=\{0\}\cup\{\mathfrak{i}_j+1\,:\,j\in\mathbb{N}_0\}$ introduced in Remark 2.6.

Lemma 3.9. *Let $k\in\mathbb{N}$ and $s>0$, let $\mathfrak{I}$ satisfy Assumption 2.4 (i) and let $\mathfrak{I}_+$ be as in (2.12). Let $(a_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}\subseteq[0,\infty)$. Define*

$$\hat{\boldsymbol{\nu}}:=(\hat\nu_j)_{j\in\mathbb{N}}\qquad\text{where}\qquad\hat\nu_j:=\begin{cases}k & \text{if}\quad 1\le\lfloor\nu_j\rfloor_{\mathfrak{I}_+}<k\\ \lfloor\nu_j\rfloor_{\mathfrak{I}_+} & \text{otherwise.}\end{cases}\tag{3.10}$$

Then there exists $C_{K_\mathfrak{I},k}>0$ depending on k and $K_\mathfrak{I}$ such that with $\hat a_{\boldsymbol{\nu}}:=a_{\hat{\boldsymbol{\nu}}}$ for all $\boldsymbol{\nu}\in\mathcal{F}$,

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}}\hat a_{\boldsymbol{\nu}}^s\le\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}a_{\boldsymbol{\nu}}^sC_{K_\mathfrak{I},k}^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}(1+\nu_j).\tag{3.11}$$

Proof. First note that $\hat{\boldsymbol{\nu}}\in\mathcal{F}_k$ for every $\boldsymbol{\nu}\in\mathcal{F}$ (*cp.* (2.1)). By Remark 2.7 it holds $\lceil 1+n\rceil_{\mathfrak{I}_+}\le K_\mathfrak{I}(1+n)$ for all $n\in\mathbb{N}_0$. Fix $\boldsymbol{\mu}\in\mathcal{F}$. Then for any $\boldsymbol{\nu}\in\mathcal{F}$

$$\hat\mu_j=\hat\nu_j\qquad\Leftrightarrow\qquad\begin{cases}\nu_j\in\{1,\dots,\min\{\mathfrak{i}_j+1\,:\,\mathfrak{i}_j+1>k\}-1\} & \text{if}\quad\hat\mu_j=k\\ \hat\mu_j\le\nu_j<\lceil 1+\hat\mu_j\rceil_{\mathfrak{I}_+} & \text{otherwise.}\end{cases}$$

Therefore, there exists a constant $C_{K_\mathfrak{I},k}$ such that for every $\boldsymbol{\mu}\in\mathcal{F}_k$

$$|\{\boldsymbol{\nu}\in\mathcal{F}\,:\,\hat{\boldsymbol{\nu}}=\boldsymbol{\mu}\}|\le\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}C_{K_\mathfrak{I},k}(1+\nu_j).$$

This implies the lemma. $\square$

The next two lemmata will be crucial to prove summability of the Taylor gpc coefficients in Section 3.3. They are a generalization of Lemma 7.1 from [10] and Theorem 7.2 from [10], in that they consider (improved) summability over $\mathcal{F}_k$ for general $k \in \mathbb{N}$ instead of just $\mathcal{F}_1$. The proofs are provided in Appendix B.

Lemma 3.10. *Let $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$, $\vartheta \geq 0$ and $R \geq 1$. Set $w_{\boldsymbol{\nu}} := R^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(1+\nu_j)^{\vartheta}$. Let $p \in (0,\infty)$ and $k \in \mathbb{N}$. The sequence $(w_{\boldsymbol{\nu}}\boldsymbol{b}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}$ belongs to $\ell^{p/k}(\mathcal{F}_k)$, iff $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$.*

Lemma 3.11. *Let $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$, $\vartheta \geq 0$ and $R \geq 1$. Set $w_{\boldsymbol{\nu}} := R^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(1+\nu_j)^{\vartheta}$. Let $p \in (0,1]$ and $k \in \mathbb{N}$. The sequence $(w_{\boldsymbol{\nu}}\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}_k}$ belongs to $\ell^{p/k}(\mathcal{F}_k)$ iff $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$.*

We now provide an example of a sequence satisfying Assumption 3.7.

Lemma 3.12. *Fix $k \in \mathbb{N}$, let $\mathfrak{I}$ satisfy Assumption 2.4 and let $\mathfrak{I}_+$ be as in (2.12). Let $\boldsymbol{\varrho} = (\varrho_j)_{j\in\mathbb{N}} \subseteq (1,\infty)$ be such that $(\varrho_j^{-1})_{j\in\mathbb{N}} \in \ell^q(\mathbb{N})$ for some $q > 0$ and additionally $\varrho_j \leq C_\varkappa j^\varkappa$ for some fixed constants $\varkappa > 0$, $C_\varkappa > 0$ and all $j \in \mathbb{N}$. For all $\boldsymbol{\nu} \in \mathcal{F}$ define*

$$c_{k,\boldsymbol{\nu}} := \boldsymbol{\varrho}^{-\hat{\boldsymbol{\nu}}} \qquad \text{where} \qquad \hat{\nu}_j := \begin{cases} k & \text{if} \quad 1 \leq \lfloor \nu_j \rfloor_{\mathfrak{I}_+} < k \\ \lfloor \nu_j \rfloor_{\mathfrak{I}_+} & \text{otherwise.} \end{cases} \tag{3.12}$$

Then $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{q/k}(\mathcal{F})$ and the sequence satisfies Assumption 3.7.

Proof. First we show $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{q/k}(\mathcal{F})$. By Lemma 3.9

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} c_{k,\boldsymbol{\nu}}^{\frac{q}{k}} = \sum_{\boldsymbol{\nu}\in\mathcal{F}} (\boldsymbol{\varrho}^{-\hat{\boldsymbol{\nu}}})^{\frac{q}{k}} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} (\boldsymbol{\varrho}^{-\boldsymbol{\nu}})^{\frac{q}{k}} C_{K_{\mathfrak{I}},k}^{|\operatorname{supp}\boldsymbol{\nu}|} \prod_{j\in\mathbb{N}}(1+\nu_j).$$

Since $(\varrho_j^{-1})_{j\in\mathbb{N}} \subseteq (0,1)$ and $(\varrho_j^{-1})_{j\in\mathbb{N}} \in \ell^q(\mathbb{N})$, Lemma 3.10 implies $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{q/k}(\mathcal{F})$.

Next we check Assumption 3.7. Items (i) and (ii) are immediate consequences of Remark 2.6 and (3.12). To verify (iii) we first note that $(\varrho_j^{-k})_{j\in\mathbb{N}}$ is a subsequence of $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ and $\varrho_j^{-k} \geq C_\varkappa^{-k} j^{-\varkappa k}$ for all $j \in \mathbb{N}$, which shows the first inequality in (3.8). For the third inequality in (3.8), we use Lemma 3.13 to obtain a constant C_0 such that $\varrho_j^{-1} \leq C_0 j^{-1/q}$. If $\boldsymbol{\nu} \in \mathcal{F}$ then $\hat{\nu}_j = 0$ or $\hat{\nu}_j \geq k$ for all $j \in \mathbb{N}$. Therefore

$$\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} c_{k,\boldsymbol{\nu}} = \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} \prod_{j\in\mathbb{N}} \varrho_j^{-\hat{\nu}_j} \leq \prod_{j=1}^{d} C_0 j^{-\frac{k}{q}} = C_0^d (d!)^{-\frac{k}{q}} \leq C_0^d\, \mathrm{e}^{\frac{dk}{q}}\, d^{-\frac{dk}{q}}$$

due to $d! \geq \mathrm{e}^{-d}\, d^d$. This implies that there exists a sequence $(f_d)_{d\in\mathbb{N}}$ as stated in Assumption 3.7. Finally, for the second inequality in (3.8) we use that for all $n \in \mathbb{N}_0$ it holds $\lfloor n \rfloor_{\mathfrak{I}_+} \geq n/K_{\mathfrak{I}}$ by Remark 2.7. Thus with $\delta := \inf_{j\in\mathbb{N}} \varrho_j > 1$

$$\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} c_{k,\boldsymbol{\nu}} = \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} \prod_{j\in\mathbb{N}} \varrho_j^{-\hat{\nu}_j} \leq \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} \prod_{j\in\mathbb{N}} \varrho_j^{-\lfloor \nu_j \rfloor_{\mathfrak{I}_+}} \leq \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} \prod_{j\in\operatorname{supp}\boldsymbol{\nu}} \delta^{-\frac{\nu_j}{K_{\mathfrak{I}}}},$$

which equals $(\delta^{1/K_{\mathfrak{I}}})^{-d}$. This verifies (3.8) and Assumption 3.7. ∎

3.3. ℓ^p-summability of Taylor GPC coefficients

We now show that for $(\boldsymbol{b},\varepsilon)$-holomorphic functions with a sequence $\boldsymbol{b} \in \ell^p(\mathbb{N})$ for some $0 < p < 1$, the norms of the Taylor gpc coefficients of u belong to $\ell^{p/k}(\mathcal{F}_k)$ for every $k \in \mathbb{N}$, with $\mathcal{F}_k$ defined in (2.1). This summability is the essential property in order to verify the improved, dimension-independent algebraic convergence rate for suitably adapted Smolyak quadratures, see Section 4. N-term approximation rate bounds for Taylor and other gpc expansions have previously been established by several authors, we only mention [10–12] and the references therein. Our new contribution here is twofold: first, instead of $\mathcal{F}$ we consider the smaller sets $\mathcal{F}_k$ and in particular $\mathcal{F}_2$. As we shall see in Section 4, the set $\mathcal{F}_2$ is better suited for analyzing Smolyak-style quadrature algorithms, as it quantifies increased sparsity due to cancellation by symmetry (in the Smolyak quadratures). Our second contribution concerns a computable estimator bounding the norm of the Taylor gpc coefficients. We show that, without loss of convergence order, it can be chosen constant on certain subsets of $\mathcal{F}$. This is to be contrasted with greedy computational schemes based on numerical solutions of knapsack problems as, for example, in [3,4]. Our new, *a priori construction* allows to localize the multiindex set for the Smolyak quadrature in near linear complexity (work and memory), as explained in Section 3.1.3 of [37]. Before presenting the result we state three lemmata required in the proof.

Lemma 3.13. *Let $p \in (0,\infty)$ and let $(t_j)_{j\in\mathbb{N}}$ be nonnegative and monotonically decreasing. Then, for all $N \in \mathbb{N}$*

$$t_N \le \left(\sum_{j=1}^N t_j^p\right)^{\frac{1}{p}} N^{-\frac{1}{p}}.$$

Proof. Due to the monotonicity of $(t_j^p)_{j\in\mathbb{N}}$ it holds $t_N^p \le N^{-1}\sum_{j=1}^N t_j^p$ which implies the lemma. □

The following theorem is an extension of results in [10,12], in particular of Theorem 1.3 from [10], Theorem 2.2 from [12]. Items four and five will provide explicit constructions of multiindex sets.

Theorem 3.14. *Let $k \in \mathbb{N}$, $0 \le \vartheta < \infty$, $p \in (0,1)$ and let the set of admissible indices $\mathfrak{I} \subseteq \mathbb{N}_0$ satisfy Assumption 2.4 (i). Let $u : U \to X$ be $(\boldsymbol{b},\varepsilon)$-holomorphic for some $\boldsymbol{b} \in \ell^p(\mathbb{N})$ (see Def. 3.1). For $\boldsymbol{\nu} \in \mathcal{F}$ define $w_{\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}}(1+\nu_j)^\vartheta$.*

Then there exists $C > 0$, $C_0 > 0$ and a sequence $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ solely depending on $\mathfrak{I}$, $\boldsymbol{b}$, ϵ and ϑ such that

(i) $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ *satisfies Assumption 3.7 (with the set of admissible indices $\mathfrak{I}$),*
(ii) $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/k}(\mathcal{F})$,
(iii) the Taylor gpc coefficients $t_{\boldsymbol{\nu}}$ of u in (3.5) satisfy

$$w_{\boldsymbol{\nu}}\|t_{\boldsymbol{\nu}}\|_X \le C M_u a_{k,\boldsymbol{\nu}} \qquad \forall \boldsymbol{\nu} \in \mathcal{F}_k \tag{3.13}$$

so that in particular $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$.

Moreover

(iv) there exist $T > 1$ and $\tau_0 > 0$ such that with

$$c_{k,\boldsymbol{\nu}} := \boldsymbol{\varrho}^{-\hat{\boldsymbol{\nu}}}, \qquad \varrho_j := \max\{T, \tau_0 \min\{b_j^{-1}, j^{2/p}\}\}^{1-p}, \qquad \hat{\nu}_j := \begin{cases} k & \text{if} \quad 1 \le \lfloor \nu_j \rfloor_{\mathfrak{I}_+} < k \\ \lfloor \nu_j \rfloor_{\mathfrak{I}_+} & \text{otherwise} \end{cases} \tag{3.14}$$

it holds $(a_{k,\boldsymbol{\nu}} c_{k,\boldsymbol{\nu}}^{-1})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^1(\mathcal{F})$ and $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/(2(1-p))}(\mathcal{F})$,
(v) in case $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < C_0$, there exist τ_1, $\tau_2 > 0$ such that we have the explicit representation

$$a_{k,\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}} \max\left\{\mathrm{e}, \frac{\tau_2 \hat{\nu}_j}{|\hat{\boldsymbol{\nu}}| \max\{b_j, \tau_1 j^{-2/p}\}}\right\}^{-\hat{\nu}_j}, \qquad \hat{\nu}_j := \begin{cases} k & \text{if} \quad 1 \le \lfloor \nu_j \rfloor_{\mathfrak{I}_+} < k \\ \lfloor \nu_j \rfloor_{\mathfrak{I}_+} & \text{otherwise.} \end{cases} \tag{3.15}$$

Proof. We proceed in four steps. In the first two steps $a_{k,\boldsymbol{\nu}}$ as stated in the theorem is constructed. In the third step item (i) is shown, and finally we show (iv) in Step 4. For a constant $\tau_1 \in (0,1]$ (chosen subsequently in Step 1) throughout this proof set

$$\tilde{b}_j := \max\{b_j, \tau_1 j^{-2/p}\} \tag{3.16}$$

and $\tilde{\boldsymbol{b}} = (\tilde{b}_j)_{j\in\mathbb{N}}$. Then $\tilde{b}_j \geq b_j$ for each $j \in \mathbb{N}$, and thus the $(\boldsymbol{b},\varepsilon)$-holomorphic function u is also $(\tilde{\boldsymbol{b}},\varepsilon)$-holomorphic (*cp.* Def. 3.1). Furthermore, w.l.o.g. we assume $M_u > 0$ in Definition 3.1 (if $M_u = 0$ then $u \equiv 0$, in which case (iii) becomes trivial).

Step 1. We introduce $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ and show that the sequence is monotonically decreasing (*cp.* Def. 3.6) and that it holds (3.13). Let the constant $C_{K_\jmath,k} > 0$ be as in Lemma 3.9. Observe that with $\tilde{\boldsymbol{b}}$ as in (3.16) (where τ_1 is to be chosen), it is possible to find constants $\tau_1 \in (0,1]$, $\kappa_0 > 1$, $C_\vartheta \geq 1$ and $J \in \mathbb{N}$ with the properties

$$(1+n)^\vartheta \leq C_\vartheta \kappa_0^n \qquad \forall n \in \mathbb{N}, \tag{3.17a}$$

and with $\delta := \varepsilon/3$

$$(\kappa_0^2 - 1)\sum_{j=1}^{J-1} \tilde{b}_j + \kappa_1 \sum_{j\geq J} \tilde{b}_j < \varepsilon - \delta, \qquad \sum_{j\geq J} \tilde{b}_j < \frac{\delta}{C_\vartheta C_{K_\jmath,k}^{k/p} \kappa_0\,\mathrm{e}}, \qquad \sum_{j\geq J} \tilde{b}_j^p < \frac{\delta}{C_\vartheta C_{K_\jmath,k}\kappa_0\,\mathrm{e}}, \qquad \tilde{b}_j \leq \frac{1}{2}\ \forall j \geq J \tag{3.17b}$$

where $\mathrm{e} = \exp(1)$ and

$$\kappa_1 := C_\vartheta \kappa_0 \mathrm{e}.$$

In the general case they are obtained as follows: first set $\tau_1 = 1$. Employing $\|\tilde{\boldsymbol{b}}\|_{\ell^1(\mathbb{N})} < \infty$ we choose $\kappa_0 > 1$ with $(\kappa_0^2 - 1)\sum_{j\in\mathbb{N}} \tilde{b}_j < \varepsilon - 2\delta$ where $\delta := \varepsilon/3$, then choose C_ϑ such that $(1+n)^\vartheta \leq C_\vartheta\kappa_0^n$ for all $n \in \mathbb{N}$, and afterwards choose $J \in \mathbb{N}$ large enough such that $\kappa_1 \sum_{j\geq J} \tilde{b}_j < \delta$ and the last three conditions in (3.17b) hold. At this point we note that if

$$\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \min\left\{\frac{2\varepsilon}{3}, \frac{\delta}{C_\vartheta C_{K_\jmath,k}^{k/p}\kappa_0\,\mathrm{e}}, \left(\frac{\delta}{C_\vartheta C_{K_\jmath,k}\kappa_0\,\mathrm{e}}\right)^{1/p}, \frac{1}{2}\right\} =: C_0, \tag{3.18}$$

then we may choose $J = 1$ and fix $\tau_1 > 0$ so small that with $\tilde{b}_j = \max\{b_j, \tau_1 j^{-2/p}\}$ it also holds $\|\tilde{\boldsymbol{b}}\|_{\ell^p(\mathbb{N})} < C_0$. In this case the conditions in (3.17b) are satisfied with $J = 1$. We will use this below to show (v).

For $\boldsymbol{\nu} \in \mathcal{F}$, in the following $\boldsymbol{\nu}_E$ denotes the multiindex which coincides with $\boldsymbol{\nu}$ in the first J components and is zero otherwise, and $\boldsymbol{\nu}_F := \boldsymbol{\nu} - \boldsymbol{\nu}_E$. Set

$$\rho_{\boldsymbol{\nu};j} := \begin{cases} \kappa_0^2 & \text{if} \quad j < J, \\ \max\left\{\kappa_1, \frac{\delta\nu_j}{|\boldsymbol{\nu}_F|\tilde{b}_j}\right\} & \text{if} \quad j \geq J. \end{cases}$$

Here and in the following we adhere to the notational convention $\nu_j/|\boldsymbol{\nu}_F| = 0$ in case $|\boldsymbol{\nu}_F| = 0$. Then, with (3.17),

$$\sum_{j\in\mathbb{N}} (\rho_{\boldsymbol{\nu};j} - 1)\tilde{b}_j \leq (\kappa_0^2 - 1)\sum_{j=1}^{J-1}\tilde{b}_j + \sum_{j\geq J}\rho_{\boldsymbol{\nu};j}\tilde{b}_j \leq (\kappa_0^2 - 1)\sum_{j=1}^{J-1}\tilde{b}_j + \kappa_1\sum_{j\geq J}\tilde{b}_j + \delta\sum_{j\geq J}\frac{\nu_j}{|\boldsymbol{\nu}_F|} < \varepsilon.$$

Therefore $\boldsymbol{\rho}_{\boldsymbol{\nu}} = (\rho_{\boldsymbol{\nu};j})_{j\in\mathbb{N}}$ is $(\tilde{\boldsymbol{b}},\varepsilon)$-admissible (in the sense of Def. 3.1). Hence, with M_u as in Definition 3.1 and C_ϑ as in (3.17a), we obtain from (3.6)

$$\begin{aligned} \|u_{\boldsymbol{\nu}}\|_X \prod_{j\in\mathbb{N}} (1+\nu_j)^\vartheta &\leq M_u \left(C_\vartheta^{|\operatorname{supp}\boldsymbol{\nu}|} \prod_{j\in\operatorname{supp}\boldsymbol{\nu}} \kappa_0^{\nu_j} \right) \prod_{j\in\mathbb{N}} \rho_{\boldsymbol{\nu};j}^{-\nu_j} \\ &\leq M_u C_\vartheta^{|\operatorname{supp}\boldsymbol{\nu}|} \kappa_0^{|\boldsymbol{\nu}|} \prod_{j=1}^{J-1} \kappa_0^{-2\nu_j} \prod_{j\geq J} \max\left\{\kappa_1, \frac{\delta\nu_j}{|\boldsymbol{\nu}_F|\tilde{b}_j}\right\}^{-\nu_j} \end{aligned}$$

$$\leq M_u C_\vartheta^{J-1} \underbrace{\prod_{j=1}^{J-1} \kappa_0^{-\nu_j} \prod_{j\geq J} \max\left\{\frac{\kappa_1}{C_\vartheta\kappa_0}, \frac{\delta\nu_j}{C_\vartheta\kappa_0|\boldsymbol{\nu}_F|\tilde{b}_j}\right\}^{-\nu_j}}_{=:f_{\boldsymbol{\nu}}}. \tag{3.19}$$

We point out that $\kappa_1/(C_\vartheta\kappa_0) = \mathrm{e}$ by definition of κ_1.

We now prove that $f_{\boldsymbol{\nu}}$ is monotonically decreasing in $\boldsymbol{\nu}$. For $j < J$ and with $\mathbf{e}_j := (\delta_{ji})_{i\in\mathbb{N}}$, since $\kappa_0 > 1$ we have $f_{\boldsymbol{\nu}+\mathbf{e}_j} \leq \kappa_0^{-1} f_{\boldsymbol{\nu}} \leq f_{\boldsymbol{\nu}}$. Next, fix $j \geq J$. Note that

$$\max\left\{\mathrm{e}, \frac{\delta\nu_j}{C_\vartheta\kappa_0|\boldsymbol{\nu}_F|\tilde{b}_j}\right\} = \max\left\{\mathrm{e}, \frac{\delta\nu_j}{C_\vartheta\kappa_0\tilde{b}_j(\nu_j + \sum_{\{i\geq J\,:\,i\neq j\}}\nu_i)}\right\} \tag{3.20}$$

is monotonically increasing as a function of ν_j, and is always larger or equal to e. Therefore

$$\frac{f_{\boldsymbol{\nu}+\mathbf{e}_j}}{f_{\boldsymbol{\nu}}} \leq \max\left\{\mathrm{e}, \frac{\delta\nu_j}{C_\vartheta\kappa_0(|\boldsymbol{\nu}_F|+1)\tilde{b}_j}\right\}^{-1} \prod_{i\geq J} \frac{\max\left\{\mathrm{e}, \frac{\delta\nu_i}{C_\vartheta\kappa_0|\boldsymbol{\nu}_F|\tilde{b}_i}\right\}^{\nu_i}}{\max\left\{\mathrm{e}, \frac{\delta\nu_i}{C_\vartheta\kappa_0(|\boldsymbol{\nu}_F|+1)\tilde{b}_i}\right\}^{\nu_i}} \leq \mathrm{e}^{-1}\left(1+\frac{1}{|\boldsymbol{\nu}_F|}\right)^{|\boldsymbol{\nu}_F|} \leq 1.$$

For all $\boldsymbol{\nu} \in \mathcal{F}$ define $a_{k,\boldsymbol{\nu}} := f_{\hat{\boldsymbol{\nu}}}$ with $\hat{\boldsymbol{\nu}}$ as in (3.15). Note that $\hat{\boldsymbol{\nu}} \leq \boldsymbol{\nu}$ for all $\boldsymbol{\nu} \in \mathcal{F}_k$. Due to the monotonicity of $(f_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ it thus holds $a_{k,\boldsymbol{\nu}} \geq f_{\boldsymbol{\nu}}$ for all $\boldsymbol{\nu} \in \mathcal{F}_k$. Together with (3.19) this shows (3.13).

Finally we point out that if $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < C_0$, then as explained after (3.18), we can choose $J = 1$ so that

$$a_{k,\boldsymbol{\nu}} = f_{\hat{\boldsymbol{\nu}}} = \prod_{j\in\mathbb{N}} \max\left\{\mathrm{e}, \frac{\delta\hat{\nu}_j}{C_\vartheta\kappa_0|\hat{\boldsymbol{\nu}}|\max\{b_j, \tau_1 j^{-2/p}\}}\right\}^{-\hat{\nu}_j} \tag{3.21}$$

is of the type described in (v).

Step 2. We now show $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/k}(\mathcal{F})$. By Lemma 3.9 it holds

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} a_{k,\boldsymbol{\nu}}^{p/k} = \sum_{\boldsymbol{\nu}\in\mathcal{F}} f_{\hat{\boldsymbol{\nu}}}^{p/k} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} f_{\boldsymbol{\nu}}^{p/k} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}|} \prod_{j\in\mathbb{N}}(1+\nu_j).$$

In the following we use that by Stirling's inequalities $n^n \leq \mathrm{e}^n\, n!$ and thus

$$\frac{|\boldsymbol{\nu}|^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}^{\boldsymbol{\nu}}} \leq \mathrm{e}^{|\boldsymbol{\nu}|}\,\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \qquad \forall \boldsymbol{\nu} \in \mathcal{F}.$$

Set $\mathcal{F}_G := \{\boldsymbol{\nu}_G \,:\, \boldsymbol{\nu} \in \mathcal{F}\}$, $G \in \{E, F\}$. Employing the definition of $f_{\boldsymbol{\nu}}$ in (3.19), $d_j := C_{K_\jmath,k}^{k/p} C_\vartheta\kappa_0\mathrm{e}\tilde{b}_j/\delta$ and $\tilde{d}_j := d_{j+J-1}$, $j \in \mathbb{N}$, we get

$$\begin{aligned}
\sum_{\boldsymbol{\nu}\in\mathcal{F}_k} a_{k,\boldsymbol{\nu}} &\leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}_E|} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}_F|} f_{\boldsymbol{\nu}}^{p/k} \prod_{j\in\mathbb{N}}(1+\nu_j)\\
&\leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} C_{K_\jmath,k}^{J-1} C_{K_\jmath,k}^{|\boldsymbol{\nu}_F|} f_{\boldsymbol{\nu}}^{p/k} \prod_{j\in\mathbb{N}}(1+\nu_j)\\
&\leq C_{K_\jmath,k}^{J-1} \sum_{\boldsymbol{\mu}\in\mathcal{F}_E\cap\mathcal{F}_k} \kappa_0^{-|\boldsymbol{\mu}|p/k} \left(\prod_{i=1}^{J-1}(1+\mu_i)\right) \sum_{\boldsymbol{\nu}\in\mathcal{F}_F\cap\mathcal{F}_k} \left(\frac{|\boldsymbol{\nu}|^{|\boldsymbol{\nu}|}}{\boldsymbol{\nu}^{\boldsymbol{\nu}}} \prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\right.\\
&\qquad \times \left.\left(\frac{C_{K_\jmath,k}^{k/p} C_\vartheta\kappa_0\tilde{b}_j}{\delta}\right)^{\nu_j}\right)^{p/k} \left(\prod_{j\geq J}(1+\nu_j)\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq C_{K_{\mathfrak{I}},k}^{J-1} \sum_{\boldsymbol{\mu}\in\mathcal{F}_E\cap\mathcal{F}_k} \kappa_0^{-|\boldsymbol{\mu}|p/k} \left(\prod_{i=1}^{J-1}(1+\mu_i)\right) \sum_{\boldsymbol{\nu}\in\mathcal{F}_F\cap\mathcal{F}_k} \left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\boldsymbol{d}^{\boldsymbol{\nu}}\right)^{p/k} \left(\prod_{j\geq J}(1+\nu_j)\right)\\
&= C_{K_{\mathfrak{I}},k}^{J-1} \sum_{\boldsymbol{\mu}\in\mathcal{F}_E} \kappa_0^{-|\boldsymbol{\mu}|p/k} \left(\prod_{i=1}^{J-1}(1+\mu_i)\right) \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} \left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\tilde{\boldsymbol{d}}^{\boldsymbol{\nu}}\right)^{p/k} \left(\prod_{j\in\mathbb{N}}(1+\nu_j)\right). \qquad (3.22)
\end{aligned}
$$

We have $\|(\tilde{d}_j)_{j\in\mathbb{N}}\|_{\ell^p(\mathbb{N})} \leq C\|\tilde{\boldsymbol{b}}\|_{\ell^p(\mathbb{N})} < \infty$. Furthermore, due to (3.17b) it holds

$$
\|(\tilde{d}_j)_{j\in\mathbb{N}}\|_{\ell^1(\mathbb{N})} = \frac{C_{K_{\mathfrak{I}},k}^{k/p} C_\vartheta \kappa_0 \mathrm{e}}{\delta} \sum_{j\geq J} \tilde{b}_j < 1.
$$

Therefore, both sums on the right-hand side of (3.22) are finite according to Lemmas 3.10 and 3.11.

Step 3. We prove that $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ satisfies Assumption 3.7.

Since $(f_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is monotonically decreasing, and since $\boldsymbol{\nu} \leq \boldsymbol{\mu}$ implies $\hat{\boldsymbol{\nu}} \leq \hat{\boldsymbol{\mu}}$ (*cp.* (3.10)), also $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is monotonically decreasing.

To see (3.7) assume that $\lceil\boldsymbol{\nu}\rceil_{\mathfrak{I}} = \lceil\boldsymbol{\mu}\rceil_{\mathfrak{I}}$. By Remark 2.6 we then have $\lfloor\boldsymbol{\nu}\rfloor_{\mathfrak{I}_+} = \lfloor\boldsymbol{\mu}\rfloor_{\mathfrak{I}_+}$. By definition of $\hat{\boldsymbol{\nu}}$, this implies $\hat{\boldsymbol{\nu}} = \hat{\boldsymbol{\mu}}$ and therefore $a_{k,\boldsymbol{\nu}} = f_{\hat{\boldsymbol{\nu}}} = f_{\hat{\boldsymbol{\mu}}} = a_{k,\boldsymbol{\mu}}$.

It remains to show Assumption 3.7 (iii). Denote $\mathbf{e}_j = (\delta_{ij})_{i\in\mathbb{N}}$. The sequence $(a_{k,\mathbf{e}_j})_{j\geq J}$ is a subsequence of $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$. By (3.19) and (3.16), it holds (since $\tilde{b}_j \to 0$ as $j\to\infty$)

$$
a_{k,\mathbf{e}_j} = f_{\hat{\mathbf{e}}_j} = f_{k\mathbf{e}_j} = \max\left\{\mathrm{e}, \frac{\delta}{C_\vartheta\kappa_0\tilde{b}_j}\right\}^{-k} \geq C\tilde{b}_j^k \geq Cj^{-2k/p}.
$$

This shows the first inequality in (3.8) with $\varkappa = 2k/p > 0$.

For the third property in (3.8) we use $\tilde{\boldsymbol{b}} \in \ell^p(\mathbb{N})$, so that by Lemma 3.13 we have $\tilde{b}_j \leq C_{\tilde{\boldsymbol{b}}} j^{-1/p}$ for some $C_{\tilde{\boldsymbol{b}}} < \infty$. Then for $d > J$ with (3.19) and due to the monotonicity of $(f_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$

$$
\begin{aligned}
\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} a_{k,\boldsymbol{\nu}} &\leq \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} f_{\boldsymbol{\nu}}\\
&\leq \prod_{j=J}^{d}\left(\frac{C_\vartheta\kappa_0(d-J)}{\delta}\right)\tilde{b}_j \leq \prod_{j=J}^{d}\left(\frac{C_\vartheta(d-J)}{\delta}\right)C_{\tilde{\boldsymbol{b}}}j^{-1/p}\\
&\leq \left(\frac{C_{\tilde{\boldsymbol{b}}}C_\vartheta\kappa_0(d-J)}{\delta}\right)^{d-J+1} d^d \prod_{j=J}^{d} j^{-1/p} \leq ((J-1)!)^{1/p} C^d d^d (d!)^{-1/p},
\end{aligned}
$$

where $C = (C_{\tilde{\boldsymbol{b}}}C_\vartheta)/\delta$. By Stirling's inequality, $d! \geq d^d\,\mathrm{e}^{-d}$ for all $d\in\mathbb{N}$. Therefore, there exists a constant $C > 0$ such that for every $d\in\mathbb{N}$ holds with $c = 1/p - 1 > 0$

$$
\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\operatorname{supp}\boldsymbol{\nu}|\geq d\}} a_{k,\boldsymbol{\nu}} \leq C^d d^{-cd}.
$$

This shows the third property in (3.8).

Finally, we show the second property in (3.8). By Remark 2.7 it holds $\lfloor n\rfloor_{\mathfrak{I}_+} \geq n/K_{\mathfrak{I}}$ for all $n\in\mathbb{N}_0$. Using that $\hat{\boldsymbol{\nu}} \geq \lfloor\boldsymbol{\nu}\rfloor_{\mathfrak{I}_+}$ and that $(f_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}$ is monotonically decreasing we get

$$
\begin{aligned}
\sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} a_{k,\boldsymbol{\nu}} = \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} f_{\hat{\boldsymbol{\nu}}} &\leq \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} f_{\lfloor\boldsymbol{\nu}\rfloor_{\mathfrak{I}_+}} \leq \sup_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|\geq d\}} \prod_{j=1}^{J-1}\kappa_0^{-\nu_j/K_{\mathfrak{I}}}\prod_{j\geq J}\mathrm{e}^{-\nu_j/K_{\mathfrak{I}}}\\
&\leq \left(\min\{\kappa_0,\mathrm{e}\}^{1/K_{\mathfrak{I}}}\right)^{-d},
\end{aligned}
$$

which shows the second property in (3.8).

Step 4. We show (iv). By definition $a_{k,\boldsymbol{\nu}} = f_{\hat{\boldsymbol{\nu}}}$ and $c_{k,\boldsymbol{\nu}} = \varrho^{-\hat{\boldsymbol{\nu}}}$ where

$$\varrho_j = \max\{T, \tau_0 \min\{b_j^{-1}, j^{2/p}\}\}^{1-p}$$

and the constants $T > 1$, $\tau_0 > 0$ are still at our disposal. Lemma 3.9 gives

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} a_{k,\boldsymbol{\nu}} c_{k,\boldsymbol{\nu}}^{-1} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} f_{\boldsymbol{\nu}} \varrho^{\boldsymbol{\nu}} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}|} \prod_{j\in\mathbb{N}} (1+\nu_j).$$

Fix $T \in (1, \min\{\kappa_0^{1/(1-p)}, 2\})$. Let $\tau_0 \in (0,1]$ be so small that $\max\{T, \tau_0 \tilde{b}_j^{-1}\} \leq T$ for all $j < J$. By (3.17b) we have $\tilde{b}_j = \max\{b_j, \tau_1 j^{-2/p}\} \leq 1/2$ and thus $\tilde{b}_j^{-1} \geq 2 \geq T$ for all $j \geq J$. Due to τ_0, $\tau_1 \in (0,1]$ we get

$$\varrho_j = \max\{T, \tau_0 \min\{b_j^{-1}, j^{2/p}\}\}^{1-p} \leq \min\{b_j^{-1}, \tau_1^{-1} j^{2/p}\}^{1-p} = \tilde{b}_j^{p-1} \qquad \forall j \geq J.$$

Then by definition of $f_{\boldsymbol{\nu}}$ in (3.19)

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}} a_{k,\boldsymbol{\nu}} c_{k,\boldsymbol{\nu}}^{-1} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}|} \left(\prod_{i\in\operatorname{supp}\boldsymbol{\nu}} (1+\nu_i)\right) \left(\prod_{j=1}^{J-1} \left(\frac{T^{1-p}}{\kappa_0}\right)^{\nu_j}\right) \left(\prod_{j\geq J} \left(\frac{C_\vartheta \kappa_0}{\delta} \frac{|\boldsymbol{\nu}_F|}{\nu_j} \tilde{b}_j^p\right)^{\nu_j}\right).$$

Using once more $n! \geq n^n \mathrm{e}^{-n}$, similar as before we get with $d_j := (C_\vartheta \kappa_0 \mathrm{e}/\delta)\tilde{b}_{j+J-1}^p$ for $j \in \mathbb{N}$ and $\boldsymbol{d} = (d_j)_{j\in\mathbb{N}}$

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\mathcal{F}} a_{k,\boldsymbol{\nu}} c_{k,\boldsymbol{\nu}}^{-1} &\leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}|} \left(\prod_{i\in\operatorname{supp}\boldsymbol{\nu}} (1+\nu_i)\right) \frac{|\boldsymbol{\nu}_F|!}{\boldsymbol{\nu}_F!} \left(\prod_{j=1}^{J-1} \left(\frac{T^{1-p}}{\kappa_0}\right)^{\nu_j}\right) \left(\prod_{j\geq J} \left(\frac{C_\vartheta \kappa_0 \mathrm{e}}{\delta} \tilde{b}_j^p\right)^{\nu_j}\right) \\ &\leq \left(\sum_{\boldsymbol{\mu}\in\mathbb{N}_0^{J-1}} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\mu}|} \prod_{j=1}^{J-1} (1+\mu_j) \left(\frac{T^{1-p}}{\kappa_0}\right)^{|\boldsymbol{\mu}|}\right) \left(\sum_{\boldsymbol{\nu}\in\mathcal{F}_k} \frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \boldsymbol{d}^{\boldsymbol{\nu}} C_{K_\jmath,k}^{|\operatorname{supp}\boldsymbol{\nu}|} \prod_{j\in\mathbb{N}} (1+\nu_j)\right). \end{aligned} \tag{3.23}$$

By (3.17b) we have

$$\sum_{j\in\mathbb{N}} d_j = \sum_{j\geq J} \frac{C_{K_\jmath,k} C_\vartheta \kappa_0 \mathrm{e}}{\delta} \tilde{b}_j^p < 1.$$

Therefore both sums in (3.23) are finite by Lemmas 3.10 and 3.11.

Finally, since $(\tilde{b}_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$, with $\varrho_j^{-1} = \max\{T, \tau_0 \min\{b_j^{-1}, j^{2/p}\}\}^{1-p}$, we have $(\varrho_j^{-1})_{j\in\mathbb{N}} \in \ell^{p/(1-p)}(\mathbb{N})$ and $\inf_{j\in\mathbb{N}} \varrho_j > 1$. Therefore $(c_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/(2(1-p))}(\mathcal{F})$ by Lemma 3.12. $\square$

Remark 3.15. Whenever $\boldsymbol{b} \in \ell^p(\mathbb{N})$ is a positive sequence, and τ_1, $\tau_2 > 0$, then the sequence $(a_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ defined in (3.15) belongs to $\ell^{p/k}(\mathcal{F})$. This follows by similar arguments as used in the proof of Theorem 3.14.

4. Smolyak convergence rates

Hereafter the main results of this paper are established. First, we show some elementary properties of the Smolyak quadrature operator. In particular it will be verified that any multivariate monomial $\boldsymbol{y}^{\boldsymbol{\nu}}$ with $\boldsymbol{\nu} \in \mathcal{F}\backslash\mathcal{F}_2$ is integrated exactly. Subsequently the dimension-independent convergence rate of $2/p - 1$ for the Smolyak quadrature with nested quadrature rules in terms of number of number of quadrature points is given for $(\boldsymbol{b}, \varepsilon)$-holomorphic functions with $\boldsymbol{b} \in \ell^p(\mathbb{N})$ for some $0 < p < 1$. For non-nested quadrature points, nearly the same convergence rate is obtained. Similarly, we obtain the same algebraic convergence in terms of the cost measure (which counts the number of required floating operations) introduced in Section 2.4.

4.1. Properties of the Smolyak quadrature

Lemma 4.1. *Let $\Lambda \subseteq \mathcal{F}$ be finite and downward closed. Then*

(i) for $\boldsymbol{\nu} \in \mathcal{F}$ it holds $Q_\Lambda \boldsymbol{y}^{\boldsymbol{\nu}} = Q_{\{\boldsymbol{\mu}\in\Lambda\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}}\boldsymbol{y}^{\boldsymbol{\nu}}$,
(ii) $Q_\Lambda P = \int_U P(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y})$ for all $P \in \operatorname{span}\{\boldsymbol{y}^{\boldsymbol{\nu}} : \boldsymbol{\nu}\in\Lambda\}$,
(iii) if $\chi_{0;0}=0$, then $Q_\Lambda P = \int_U P(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) = 0$ for all $P \in \operatorname{span}\{\boldsymbol{y}^{\boldsymbol{\nu}} : \boldsymbol{\nu}\in\mathcal{F}\backslash\mathcal{F}_2\}$,
(iv) if (2.3) holds for some $\vartheta\ge 1$, then for all $\boldsymbol{\nu}\in\mathcal{F}$

$$|Q_\Lambda \boldsymbol{y}^{\boldsymbol{\nu}}| \le \prod_{j\in\mathbb{N}}(1+\nu_j)^{\vartheta+1}.$$

Proof. Fix $\boldsymbol{\nu}\in\mathcal{F}$. Due to $Q_n y^k = \int_{-1}^1 y^k\,\mathrm{d}y/2$ for all $n\ge k$ we have $(\bigotimes_{j\in\mathbb{N}}(Q_{\mu_j}-Q_{\mu_{j-1}}))(\boldsymbol{y}^{\boldsymbol{\nu}})=0$ whenever there exists $j\in\mathbb{N}$ such that $\mu_j>\nu_j$. Thus

$$Q_\Lambda \boldsymbol{y}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{\mu}\in\Lambda}\left(\bigotimes_{j\in\mathbb{N}}(Q_{\mu_j}-Q_{\mu_{j-1}})\right)\boldsymbol{y}^{\boldsymbol{\nu}} = \sum_{\{\boldsymbol{\mu}\in\Lambda\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}}\left(\bigotimes_{j\in\mathbb{N}}(Q_{\mu_j}-Q_{\mu_{j-1}})\right)\boldsymbol{y}^{\boldsymbol{\nu}},$$

which shows (i). Next observe that due to the convention $Q_{-1}\equiv 0$

$$\sum_{\{\boldsymbol{\mu}\in\mathcal{F}\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}}\left(\bigotimes_{j\in\mathbb{N}}(Q_{\mu_j}-Q_{\mu_{j-1}})\right) = \bigotimes_{j\in\mathbb{N}}\sum_{i=0}^{\nu_j}(Q_i-Q_{i-1}) = \bigotimes_{j\in\mathbb{N}}Q_{\nu_j} = Q_{\boldsymbol{\nu}}.$$

Therefore, if $\boldsymbol{\nu}\in\Lambda$ then by (i) it holds $Q_\Lambda\boldsymbol{y}^{\boldsymbol{\nu}} = Q_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}} = \prod_{j\in\mathbb{N}}Q_{\nu_j}y_j^{\nu_j} = \int_U \boldsymbol{y}^{\boldsymbol{\nu}}\,\mathrm{d}\mu(\boldsymbol{y})$.

For (iii) consider the univariate quadrature operator $Q_n : C_0([-1,1])\to\mathbb{R}$, employing $n+1$ distinct quadrature points in $[-1,1]$. The monomial $y\mapsto y$ satisfies $Q_n y = \int_{-1}^1 y\,\mathrm{d}y/2 = 0$ for all $n\in\mathbb{N}_0$: this is true for $n\ge1$, as stated at the beginning of the proof. It is true for $n=0$, because $Q_0 y = \chi_{0;0}=0$. For $\boldsymbol{\nu}\in\mathcal{F}$ and $\boldsymbol{\mu}\in\mathcal{F}\backslash\mathcal{F}_2$ arbitrary there exists j with $\mu_j=1$ and thus

$$Q_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\mu}} = \left(\bigotimes_{j\in\mathbb{N}}Q_{\nu_j}\right)\boldsymbol{y}^{\boldsymbol{\mu}} = \prod_{j\in\mathbb{N}}Q_{\nu_j}y_j^{\mu_j} = 0 = \int_U \boldsymbol{y}^{\boldsymbol{\mu}}\,\mathrm{d}\mu(\boldsymbol{y}),$$

which by (2.5) gives $Q_\Lambda\boldsymbol{y}^{\boldsymbol{\mu}} = 0 = \int_U \boldsymbol{y}^{\boldsymbol{\nu}}\,\mathrm{d}\mu(\boldsymbol{y})$ for all $\boldsymbol{\mu}\in\mathcal{F}\backslash\mathcal{F}_2$.

For item (iv), fix $\boldsymbol{\nu}\in\mathcal{F}$. By (i) and (2.3) we can bound $|Q_\Lambda\boldsymbol{y}^{\boldsymbol{\nu}}|$ by

$$\left|\sum_{\boldsymbol{\mu}\le\boldsymbol{\nu}}\prod_{j\in\mathbb{N}}(Q_{\mu_j}-Q_{\mu_j-1})y_j^{\nu_j}\right| \le \sum_{\boldsymbol{\mu}\le\boldsymbol{\nu}}\prod_{j\in\mathbb{N}}\left((1+\mu_j)^\vartheta+\mu_j^\vartheta\right) = \prod_{j\in\mathbb{N}}\sum_{i=0}^{\nu_j}\left((1+i)^\vartheta+i^\vartheta\right).$$

So we need to show $\sum_{i=0}^m((1+i)^\vartheta+i^\vartheta)\le(1+m)^{\vartheta+1}$. The statement is true for $m=0$. For the induction step we get $\sum_{i=0}^{m+1}((1+i)^\vartheta+i^\vartheta)\le(1+m)^{\vartheta+1}+(2+m)^\vartheta+(1+m)^\vartheta$. It suffices to show that $((1+m)^{\vartheta+1}+(2+m)^\vartheta+(1+m)^\vartheta)/(2+m)^\vartheta\le 2+m$. The latter is equivalent to $((1+m)/(2+m))^\vartheta(2+m)\le 1+m$. This is satisfied because $\vartheta\ge1$. ☐

Remark 4.2. Let $-\infty\le a<b\le\infty$ and let η be a probability measure on (a,b) equipped with the Borel σ-Algebra. The idea of Lemma 4.1 (iii) is generalized as follows. Set $\chi_{0,0} := \int_a^b y\,\mathrm{d}\eta(y)$. Then the one point quadrature rule $Q_0 : f\mapsto f(\chi_{0,0})$ w.r.t. the measure η is exact on $\operatorname{span}\{1,y\}$: it holds $Q_0 1 = 1 = \int_a^b 1\,\mathrm{d}\eta(y)$ and $Q_0 y = \chi_{0,0} = \int_a^b y\,\mathrm{d}\eta(y)$.

4.2. Convergence rates

We now turn to the proof of Theorem 2.16. Due to Lemma 3.3, Theorem 2.16 is implied by the following, stronger statement.

Theorem 4.3. *Let X be a Banach space, $U = [-1,1]^{\mathbb{N}}$ and let $u : U \to X$ be $(\boldsymbol{b}, \varepsilon)$-holomorphic (see Def. 3.1) for a sequence $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ and some $p \in (0,1)$. Let the quadrature points $\boldsymbol{\chi}$ in (2.6) satisfy the bound (2.3) for some $\vartheta \geq 0$, and let the set of admissible indices $\mathfrak{I} \subseteq \mathbb{N}_0$ satisfy Assumption 2.4.*

Then for any $\delta > 0$ there exists a constant C such that

(i) with $(a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ as in Theorem 3.14 for $\tilde{\vartheta} := \vartheta + 1$, for every $\epsilon > 0$ the set $\Lambda_\epsilon := \{\boldsymbol{\nu} \in \mathcal{F} \,:\, a_{2,\boldsymbol{\nu}} \geq \epsilon\}$ is finite and downward closed and (cp. (2.7))

$$\left\| \int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u \right\|_X \leq C|\mathrm{pts}(\Lambda_\epsilon, \boldsymbol{\chi})|^{-(\frac{2}{p}-1)+\delta} \tag{4.1}$$

as well (cp. (2.20))

$$\left\| \int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u \right\|_X \leq C\mathrm{cost}(\Lambda_\epsilon)^{-(\frac{2}{p}-1)+\delta}, \tag{4.2}$$

(ii) with $(c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ as in Theorem 3.14 for $\tilde{\vartheta} := \vartheta + 1$, for every $\epsilon > 0$ the set $\Lambda_\epsilon := \{\boldsymbol{\nu} \in \mathcal{F} \,:\, c_{2,\boldsymbol{\nu}} \geq \epsilon\}$ is finite and downward closed, and (cp. (2.7))

$$\left\| \int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u \right\|_X \leq C|\mathrm{pts}(\Lambda_\epsilon, \boldsymbol{\chi})|^{-(\frac{2}{p}-2)+\delta} \tag{4.3}$$

as well (cp. (2.20))

$$\left\| \int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u \right\|_X \leq C\mathrm{cost}(\Lambda_\epsilon)^{-(\frac{2}{p}-2)+\delta}, \tag{4.4}$$

(iii) if the points $\boldsymbol{\chi}$ are nested, then (4.1) and (4.3) remain true for $\delta = 0$, and Assumption 2.4 (ii) (exponential increase of the admissible indices) on $\mathfrak{I}$ can be dropped.

We refer to Remark 5.5 for more details on the concrete choice of the set $\mathfrak{I}$.

Remark 4.4. The convergence rate for $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ in Theorem 4.3 (ii) is off by a factor 1 compared to the index sets $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ in Theorem 4.3 (i). In Lemma 1.4.19 of [37] we give an example which shows that this is not due to a rough estimate, but the index sets $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ are in fact suboptimal in general. However, in our numerical experiments we shall see that the index sets $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ seem to perform better in practice than $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, see Figure 9.

Proof of Theorem 4.3. We start with (i) and let $\Lambda_\epsilon = \Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, where $(a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is as in Theorem 3.14

By Theorem 3.14 (iii), the Taylor gpc coefficients $(t_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq X$ of u satisfy $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^p(\mathcal{F}) \hookrightarrow \ell^1(\mathcal{F})$. By Lemma 3.5, $u(\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}$ converges absolutely in $C^0(U,X)$. Fix $\epsilon > 0$. As $Q_{\Lambda_\epsilon} : C^0(U) \to X$ is a bounded linear operator, by Lemma 4.1 (ii) and (iii)

$$Q_{\Lambda_\epsilon} u = Q_{\Lambda_\epsilon} \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}} = \sum_{\boldsymbol{\nu}\in\mathcal{F}} t_{\boldsymbol{\nu}} Q_{\Lambda_\epsilon} \boldsymbol{y}^{\boldsymbol{\nu}} = \int_U \sum_{\boldsymbol{\nu}\in\Lambda_\epsilon} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}\,\mathrm{d}\mu(\boldsymbol{y}) + \sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon} t_{\boldsymbol{\nu}} Q_{\Lambda_\epsilon} \boldsymbol{y}^{\boldsymbol{\nu}}, \tag{4.5}$$

where the latter sum is absolutely convergent in X. Lemma 4.1 (iii) also implies $\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) = \int_U \sum_{\boldsymbol{\nu}\in\mathcal{F}_2} t_{\boldsymbol{\nu}} \boldsymbol{y}^{\boldsymbol{\nu}}\,\mathrm{d}\mu(\boldsymbol{y})$. Using Theorem 3.14 (iii) and Lemma 4.1 (iv) we get that there exists a constant $C > 0$

such that for every $\epsilon > 0$

$$\begin{aligned}\left\|\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon}u\right\|_X &\le \left\|\int_U \sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon} t_{\boldsymbol{\nu}}\boldsymbol{y}^{\boldsymbol{\nu}}\,\mathrm{d}\mu(\boldsymbol{y})\right\|_X + \sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon}\|t_{\boldsymbol{\nu}}\|_X|Q_{\Lambda_\epsilon}\boldsymbol{y}^{\boldsymbol{\nu}}| \\ &\le \sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon}\|t_{\boldsymbol{\nu}}\|_X\|\boldsymbol{y}^{\boldsymbol{\nu}}\|_{C^0(U,\mathbb{R})}\left(1+\prod_{j\in\mathbb{N}}(\nu_j+1)^{\vartheta+1}\right) \\ &\le C\sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon} a_{2,\boldsymbol{\nu}} \le C\sum_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,a_{2,\boldsymbol{\nu}}<\epsilon\}} a_{2,\boldsymbol{\nu}}. \end{aligned} \tag{4.6}$$

Exploiting $(a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/2}(\mathcal{F})$ allows to bound the last sum by $C|\Lambda_\epsilon|^{1-2/p}$. This follows by rearranging the sequence $(a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ as a monotonically decreasing sequence $(a_j^*)_{j\in\mathbb{N}}$, so that Lemma 3.13 gives $a_j^* \le Cj^{-2/p}$ and consequently $\sum_{j>N} a_j^* \le C\int_N^\infty x^{-2/p}\,\mathrm{d}x \le CN^{1-2/p}$.

In case the points are nested we have $|\mathrm{pts}(\Lambda_\epsilon,\boldsymbol{\chi})| = |\Lambda_\epsilon|$ by Lemma 2.2, which shows (4.1) for $\delta = 0$, and thus the statement in (iii) in this case. If the points are non-nested, then we use that for any $\delta > 0$ it holds $|\mathrm{pts}(\Lambda_\epsilon,\boldsymbol{\chi})| = O(|\Lambda_\epsilon|^{1+\delta})$ as $\epsilon \to 0$. This is an immediate consequence of Theorem 3.14 (i), Lemmas 3.8 and 2.10. This shows (4.1) also for non-nested points.

For (4.2) we argue similarly by invoking Theorem 3.14 (i), Lemmas 3.8 and 2.10.

Next we prove (ii), *i.e.* in the following $\Lambda_\epsilon = \Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu}\in\mathcal{F} : c_{2,\boldsymbol{\nu}} \ge \epsilon\}$, where $(c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is as in Theorem 3.14 (iv). As in (4.6) we obtain

$$\left\|\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon}u\right\|_X \le C\sum_{\boldsymbol{\nu}\in\mathcal{F}_2\setminus\Lambda_\epsilon} a_{2,\boldsymbol{\nu}} \le C\left(\sup_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_\epsilon} c_{2,\boldsymbol{\nu}}\right)\left(\sum_{\boldsymbol{\mu}\in\mathcal{F}_2\setminus\Lambda_\epsilon} a_{2,\boldsymbol{\mu}}c_{2,\boldsymbol{\mu}}^{-1}\right).$$

Since $(c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/(2(1-p))}(\mathcal{F})$ and $(a_{2,\boldsymbol{\nu}}c_{2,\boldsymbol{\nu}}^{-1})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^1(\mathcal{F})$ by Theorem 3.14 (iv), Lemma 3.13 implies

$$\left\|\int_U u(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon}u\right\|_X \le C\sup_{\boldsymbol{\nu}\in\mathcal{F}\setminus\Lambda_\epsilon} c_{2,\boldsymbol{\nu}} \le C|\Lambda_\epsilon|^{-2/p-2}.$$

For nested points, Lemma 2.2 then implies (4.3), which also shows (iii) in this case. In order to prove (4.4) as well as the estimate (4.3) for non-nested points, we use the fact that $(c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ satisfies Assumption 3.7 by Lemma 3.12, so that we can employ Lemmas 3.8, 2.10 and 2.15 as above. ☐

Remark 4.5. In the papers [20, 22], rather than $(\boldsymbol{b},\varepsilon)$-holomorphy, a requirement of the following type is presumed:

$$\begin{array}{l} u \text{ is separately holomorphic and uniformly bounded on some polydisc} \\ B_{\boldsymbol{\rho}}^{\mathbb{C}} \subseteq \mathbb{C}^{\mathbb{N}}, \text{ where } \rho_j > 1 \text{ for all } j\in\mathbb{N} \text{ and } (\rho_j^{-1})_{j\in\mathbb{N}} \in \ell^p(\mathbb{N}),\ p\in(0,1). \end{array} \tag{4.7}$$

In these references, under assumptions similar to (4.7), dimension-independent convergence rates $(1/p-1)$ and $(1/p-1)/2$, respectively, are established (see [20], Cor. 5.9, [22], Assumption 4.2, Thm. 5.5 for the precise assumptions and statements).

Let u be $(\boldsymbol{b},\varepsilon)$-holomorphic for some $\boldsymbol{b} \in \ell^p(\mathbb{N})$ and some $p\in(0,1)$, $\varepsilon > 0$. Let $\kappa > 1$ be so small and $J\in\mathbb{N}$ be so large that $(\kappa-1)\sum_{j\in\mathbb{N}} b_j + \sum_{j>J} b_j^p < \varepsilon$. This is possible because $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})}, \|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$. Set $\rho_j := \kappa$ for $j \le J$ and $\rho_j := \max\{\kappa, b_j^{p-1}\}$ for $j > J$. Then $\sum_{j\in\mathbb{N}} b_j(\rho_j - 1) \le \sum_{j\in\mathbb{N}}(\kappa-1)b_j + \sum_{j>J} b_j^p \le \varepsilon$. Thus $(\boldsymbol{b},\varepsilon)$-holomorphy implies (4.7) with this $\boldsymbol{\rho}$. Note that $(\rho_j^{-1})_{j\in\mathbb{N}} \in \ell^{p/(1-p)}(\mathbb{N})$ and $p/(1-p) > p$. On the other hand, (4.7) implies $(\tilde{\boldsymbol{b}},1)$-holomorphy, with $\tilde{b}_j := (\rho_j-1)^{-1}$ and $(\tilde{b}_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$: if $\tilde{\boldsymbol{\rho}}$ is arbitrary with $\sum_{j\in\mathbb{N}} \tilde{b}_j(\tilde{\rho}_j - 1) < 1$, then $\tilde{b}_j(\tilde{\rho}_j-1) < 1$, and thus $(\tilde{\rho}_j-1)/(\rho_j-1) < 1$ implying $\tilde{\rho}_j < \rho_j$ for each $j\in\mathbb{N}$. Since

u allows a bounded holomorphic extension to $B^{\mathbb{C}}_{\boldsymbol{\rho}}$ by (4.7), it also allows a bounded holomorphic extension to $B^{\mathbb{C}}_{\tilde{\boldsymbol{\rho}}} \subseteq B^{\mathbb{C}}_{\boldsymbol{\rho}}$. Hence $(\boldsymbol{b},\varepsilon)$-holomorphy is more general than (4.7).

In summary, Theorem 4.3 improves the dimension-independent convergence rates $1/p-1$, $(1/p-1)/2$ for the anisotropic Smolyak quadrature proved in [20,22] to $2/p-1$, *i.e.* by more than a factor 2 and 4, respectively, and under weaker assumptions regarding the domain of holomorphy (namely $(\boldsymbol{b},\varepsilon)$-holomorphy rather than (4.7)). We explain this in more detail in Examples 5.2 and 5.3 ahead.

5. Numerical experiments

This section reports on the numerical testing, which we have performed for the presented algorithm. More details on the construction of the index sets will be given in Section 5.1. We shall see, that there is a large preasymptotic range, which is addressed in Section 5.3. Afterwards, in Section 5.4 we consider the integration of two real valued test functions.

We now introduce the two test integrands and discuss the proven convergence rate of the Smolyak quadrature implied by Theorem 4.3. Additionally, we compare it with the results of [20, 22].

Remark 5.1. Some of the convergence rates presented in Theorem 4.3 only hold up to some (arbitrarily small) $\delta > 0$. Throughout what follows, the mentioned convergence rates are usually understood up to $\delta > 0$. We omit to state this at every instance.

Example 5.2. Let $p \in (0,1)$ and assume that $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$ satisfies $\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$ and $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$. Define

$$u_1(\boldsymbol{y}) := \prod_{j\in\mathbb{N}} (1+b_j y_j)^{-1} \qquad \boldsymbol{y} \in U. \tag{5.1}$$

(i) Fix $\varepsilon \in (0, 1-\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})})$ and let $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}} \subseteq (1,\infty)$ be $(\boldsymbol{b},\varepsilon)$-admissible, *i.e.* $\sum_{j\in\mathbb{N}} b_j(\rho_j - 1) < \varepsilon$ (*cp.* Def. 3.1). Fix $\boldsymbol{z} \in B^{\mathbb{C}}_{\boldsymbol{\rho}} \subseteq \mathbb{C}^{\mathbb{N}}$ and set $\delta := \varepsilon + \|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$. We can find a constant C_δ such that for $0 \le x \le \delta$ it holds $\log(1/(1-x)) \le C_\delta x$. Since $b_j\rho_j = b_j(\rho_j - 1) + b_j \le \delta < 1$, we get

$$|u_1(\boldsymbol{z})| = \left|\prod_{j\in\mathbb{N}} (1+b_j z_j)^{-1}\right| \le \prod_{j\in\mathbb{N}} (1-b_j\rho_j)^{-1} \le \exp\left(C_\delta \sum_{j\in\mathbb{N}} b_j\rho_j\right).$$

The last term is finite (independent of $\boldsymbol{\rho}$) because $\sum_{j\in\mathbb{N}} b_j\rho_j = \sum_{j\in\mathbb{N}} b_j(\rho_j-1)+\sum_{j\in\mathbb{N}} b_j \le \varepsilon+\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < \infty$. Therefore u allows a well-defined uniformly bounded extension to $B^{\mathbb{C}}_{\boldsymbol{\rho}}$. Clearly $u(\boldsymbol{z})$ is holomorphic in each $z_j \in B^{\mathbb{C}}_{\rho_j}$. Continuity of $U \ni \boldsymbol{y} \mapsto u_1(\boldsymbol{y})$ is easily checked, and thus u is $(\boldsymbol{b},\varepsilon)$-holomorphic. By Theorem 4.3, the asymptotic convergence rate of the Smolyak quadrature is at least $2/p-1$.

(ii) Consider now assumption (4.7), *i.e.* the requirement which was similarly presumed in [20, 22]. We wish to find $\boldsymbol{\rho} = (\rho_j)_{j\in\mathbb{N}}$ such that u allows a uniformly bounded holomorphic extension onto the polydisc $B^{\mathbb{C}}_{\boldsymbol{\rho}}$. In view of Remark 4.5, the sequence $\boldsymbol{\rho}$ should be chosen such that $(\rho_j^{-1})_{j\in\mathbb{N}} \in \ell^{\tilde{p}}(\mathbb{N})$ for some possibly small $\tilde{p} > 0$.
For $0 \le x < 1$ we have $1/(1-x) \ge 1+x$ and furthermore $\log(1+x) \ge x/2$, which gives $-\log(1-x) \ge x/2$. Thus for $\boldsymbol{z} := (-\rho_j/2)_{j\in\mathbb{N}} \in B^{\mathbb{C}}_{\boldsymbol{\rho}}$

$$|u_1(\boldsymbol{z})| = \prod_{j\in\mathbb{N}} (1-b_j\rho_j/2)^{-1} = \exp\left(-\sum_{j\in\mathbb{N}} \log(1-b_j\rho_j/2)\right) \ge \exp\left(\frac{1}{4}\sum_{j\in\mathbb{N}} b_j\rho_j\right).$$

Hence $\boldsymbol{\rho}$ must satisfy $\sum_{j\in\mathbb{N}} \rho_j b_j < \infty$. This implies $\rho_j^{-1} = b_j/c_j$ for some sequence $(c_j)_{j\in\mathbb{N}} \in \ell^1(\mathbb{N})$. Suppose that $(\rho_j^{-1})_{j\in\mathbb{N}} \in \ell^{\tilde p}(\mathbb{N})$ for some $0 < \tilde p < 1$. Then with $\hat p := \tilde p/(1+\tilde p) < 1$

$$\sum_{j\in\mathbb{N}} b_j^{\hat p} = \sum_{j\in\mathbb{N}} \left(\frac{b_j}{c_j}\right)^{\hat p} c_j^{\hat p} \le \left(\sum_{j\in\mathbb{N}} \left(\frac{b_j}{c_j}\right)^{\frac{\hat p}{1-\hat p}}\right)^{1-\hat p} \left(\sum_{j\in\mathbb{N}} c_j\right)^{\hat p} = \left(\sum_{j\in\mathbb{N}} \left(\frac{b_j}{c_j}\right)^{\tilde p}\right)^{1-\hat p} \left(\sum_{j\in\mathbb{N}} c_j\right)^{\hat p}$$

and we obtain $\boldsymbol{b} \in \ell^{\hat p}(\mathbb{N})$. Assuming that $p > 0$ was an optimal choice, in the sense that $\boldsymbol{b} \in \ell^p(\mathbb{N})$ but $\boldsymbol{b} \notin \ell^q(\mathbb{N})$ with $q < p$, it must hold $\hat p = \tilde p/(1+\tilde p) \ge p$, and therefore $\tilde p \ge p/(1-p)$. Hence $(\rho_j^{-1})_{j\in\mathbb{N}}$, can at best be in $\ell^{p/(1-p)}(\mathbb{N})$. One possible choice achieving this is $\rho_j := \max\{\kappa, b_j^{p-1}\}$, with $\kappa > 1$ fulfilling $\kappa\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$. One checks that u then allows a uniformly bounded extension onto $B_{\boldsymbol{\rho}}^{\mathbb{C}}$ and it holds $(\rho_j^{-1}) \in \ell^{\tilde p}(\mathbb{N})$ with $\tilde p := p/(1-p)$. The statements in Corollary 5.9 of [20] and Assumption 4.2, Theorem 5.5 of [22], then essentially give the convergence rates $s_1 := \tilde p^{-1} - 1 = 1/p - 2$ and $s_2 := (\tilde p^{-1} - 1)/2 = 1/(2p) - 1$. In comparison, Theorem 4.3 gives the convergence rate $2/p - 1 = 2s_1 + 3 = 4s_2 + 3$.

Example 5.3. Let $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}} \subseteq (0,\infty)$ satisfy $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$, and define

$$u_2(\boldsymbol{y}) := \left(1 + \sum_{j\in\mathbb{N}} b_j y_j\right)^{-1} \qquad \boldsymbol{y} \in U. \tag{5.2}$$

With $\mathfrak{u}(z) := 1/(1+z)$ we have $u_2(\boldsymbol{y}) = \mathfrak{u}(\sum_{j\in\mathbb{N}} y_j b_j)$. Hence, Lemma 3.3 implies u to be $(\boldsymbol{b},\varepsilon)$-holomorphic for any fixed $\varepsilon \in (0, 1 - \|\boldsymbol{b}\|_{\ell^1(\mathbb{N})})$.

Similar as in Example 5.2, the corresponding results in Corollary 5.9 of [20], Assumption 4.2, Theorem 5.5 of [22] give the convergence rates $s_1 = 1/p - 2$ and $s_2 = 1/(2p) - 1$, while Theorem 4.3 implies the convergence rate $2/p - 1 = 2s_1 + 3 = 4s_2 + 3$ in terms of the number of quadrature points.

Remark 5.4. Differentiating u_1, u_2 in (5.1), (5.2) for some $\boldsymbol{\nu} \in \mathcal{F}$ we find

$$\frac{1}{\boldsymbol{\nu}!}\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_1(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{0}} = (-1)^{|\boldsymbol{\nu}|}\boldsymbol{b}^{\boldsymbol{\nu}} \qquad \text{and} \qquad \frac{1}{\boldsymbol{\nu}!}\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_2(\boldsymbol{y})|_{\boldsymbol{y}=\boldsymbol{0}} = (-1)^{|\boldsymbol{\nu}|}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\boldsymbol{b}^{\boldsymbol{\nu}}.$$

Thus the modulus of the Taylor gpc coefficients of u_1, u_2 agree with the sequences in Lemmas 3.10 and 3.11 (for $\vartheta = 0$ and $R = 1$).

5.1. *A priori* construction of quadrature rules

We consider two different types of quadrature points: sections of a Leja sequence serve as an example of nested quadrature points, and the Gauss-Legendre points will be used as an example of non-nested quadrature points. To construct a quadrature rule for $(\boldsymbol{b},\varepsilon)$-holomorphic functions, throughout Section 5.1 the sequence $\boldsymbol{b} = (b_j)_{j\in\mathbb{N}}$ in Definition 3.1 is assumed to satisfy

$$b_j = \theta j^{-r} \qquad \forall j \in \mathbb{N}, \tag{5.3}$$

for some fixed values of $\theta \in (0,1)$, $r > 1$ and a constant C. Then $\boldsymbol{b} = (\theta j^{-r})_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ for any $p > 1/r$.

5.1.1. *Leja quadrature*

So called Leja sequences provide nested quadrature points which possess polynomial bounds on the growth of the Lebesgue constant. We use the following construction given in Section 3 of [8]. Set $\varphi_0 := 0$, $\varphi_1 := \pi$, $\varphi_2 := \pi/2$ and

$$\varphi_{2n+1} := \frac{\varphi_{n+1}}{2}, \qquad \varphi_{2n+2} := \varphi_{2n+1} + \pi \qquad \forall n \ge 1.$$

Now let $\chi_n := \cos(\varphi_n)$ for all $n \in \mathbb{N}_0$. For every $n \in \mathbb{N}_0$ and $j \in \{0, \dots, n\}$ we define $\chi^{\rm leja}_{n;0} := 0$, $\chi^{\rm leja}_{n;1} := 1$, $\chi^{\rm leja}_{n;2} := -1$ and $\chi^{\rm leja}_{n;j} := \chi_n$ for $j \geq 3$. As shown in Theorem 3.1 of [8] there holds a bound of the type (2.3), also see [5,6]. This yields nested one dimensional quadrature points (*cp.* Def. 2.1).

Theorem 4.3 proposes two strategies to determine sets of multiindices Λ_ϵ providing proven asymptotic convergence of the Smolyak quadrature. First, let $(\tilde{c}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ be as in (3.12) with $\varrho_j = \max\{T, \tau_0 \min\{b_j^{-1}, j^{2/p}\}\}^{1-p}$ as in Theorem 3.14 (iv). Here the constants $T > 1$ and $\tau_0 > 0$ are in practice unknown. We simplify this by setting $\varrho_j = b_j^{p-1}$. With $\mathfrak{I} = \mathbb{N}_0$ in (3.12) and with (5.3) we arrive at

$$\tilde{c}_{2,\boldsymbol{\nu}} = \prod_{j\in\mathbb{N}} (\theta j^{-r})^{(1-p)\hat{\nu}_j} \qquad \text{where} \qquad \hat{\nu}_j = \begin{cases} 2 & \text{if} \quad \nu_j = 1 \\ \nu_j & \text{otherwise.} \end{cases} \tag{5.4}$$

Note that $\mathfrak{I} = \mathbb{N}_0$ satisfies Assumption 2.4 (i), but not Assumption 2.4 (ii). Due to the nestedness of the univariate points $\boldsymbol{\chi}$, Theorem 4.3, item (iii) is applicable. With $\Lambda_\epsilon((\tilde{c}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, \tilde{c}_{2,\boldsymbol{\nu}} \geq \epsilon\}$, Theorem 4.3 suggests the convergence rate $2r - 2$ for $(\boldsymbol{b}, \varepsilon)$-holomorphic functions, where $\boldsymbol{b}$ is as in (5.3). Due to

$$\{\boldsymbol{\nu} \in \mathcal{F} \,:\, \tilde{c}_{2,\boldsymbol{\nu}} \geq \epsilon\} = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, \tilde{c}_{2,\boldsymbol{\nu}}^s \geq \epsilon^s\}$$

for any $s > 0$, the choice of exponent $1 - p$ in (5.4) is irrelevant for the definition of the index sets Λ_ϵ. Thus we set

$$c^{\rm leja}_{2,\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}} (\theta j^{-r})^{\hat{\nu}_j} \qquad \text{where} \qquad \hat{\nu}_j = \begin{cases} 2 & \text{if} \quad \nu_j = 1 \\ \nu_j & \text{otherwise} \end{cases} \tag{5.5a}$$

and

$$\Lambda_\epsilon((c^{\rm leja}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, (c^{\rm leja}_{2,\boldsymbol{\nu}}) \geq \epsilon\}. \tag{5.5b}$$

Next we employ Theorem 3.14 (v) to construct a second choice of indexsets. Simplifying (3.15) by choosing $\tau_1 = \tau_2 = 1$, we get

$$a^{\rm leja}_{2,\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}} \max\left\{ \mathrm{e}, \frac{\hat{\nu}_j}{|\hat{\boldsymbol{\nu}}|\theta j^{-r}} \right\}^{-\hat{\nu}_j} \text{where} \qquad \hat{\nu}_j = \begin{cases} 2 & \text{if} \quad \nu_j = 1 \\ \nu_j & \text{otherwise} \end{cases} \tag{5.6a}$$

and

$$\Lambda_\epsilon((a^{\rm leja}_{k,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, a^{\rm leja}_{2,\boldsymbol{\nu}} \geq \epsilon\}. \tag{5.6b}$$

In this case Theorems 4.3 and 3.14 (v) imply the convergence rate $2r - 1$ for the Smolyak quadrature, provided that θ is small enough depending on u (and provided that the above choice of $\tau_1 = \tau_2 = 1$ was viable according to Thm. 3.14 (v)).

5.1.2. *Gauss-Legendre quadrature*

For every $n \in \mathbb{N}_0$ denote by $(\chi^{\rm gauss}_{n;j})_{j=0}^n$ the $n+1$ unique roots of the nth Legendre polynomial in the interval $[-1, 1]$. The one dimensional quadrature Q_n in (2.2) then integrates exactly all polynomials of degree $2n+1$ as is well-known. With $\mathfrak{I} = \{2^j - 1 \,:\, j \in \mathbb{N}_0\}$ and $\mathfrak{I}_+ = \{0\} \cup \{2^j \,:\, j \in \mathbb{N}_0\}$ (*cp.* Rem. 2.6 and note that $\mathfrak{I}$ satisfies Assumption 2.4), set

$$c^{\rm gauss}_{2,\boldsymbol{\nu}} := \prod_{j\in\mathbb{N}} (\theta j)^{-2r\lfloor \nu_j \rfloor_{\mathfrak{I}_+}} \tag{5.7a}$$

and

$$\Lambda_\epsilon((c^{\rm gauss}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu} \in \mathcal{F} \,:\, (c^{\rm gauss}_{2,\boldsymbol{\nu}}) \geq \epsilon\}. \tag{5.7b}$$

This definition deviates from the formula in (3.12): the factor 2 in the exponent in (5.7a) accounts for the fact that Q_n integrates exactly polynomials of degree $2n+1$ (and not just $n+1$). The sets in (5.7) can be considered as a heuristic choice here, but we also refer to Section 5.1.1 from [37] which provides a justification for this definition.

For the second choice of indexsets suggested by Theorem 3.14 (v), we similarly define

$$a_{2,\boldsymbol{\nu}}^{\text{gauss}} := \prod_{j\in\mathbb{N}} \max\left\{\mathrm{e}, \frac{2\lfloor\nu_j\rfloor_{\mathfrak{I}_+}}{2|\lfloor\boldsymbol{\nu}\rfloor_{\mathfrak{I}_+}|\theta j^{-r}}\right\}^{-2\lfloor\nu_j\rfloor_{\mathfrak{I}_+}} \tag{5.8a}$$

and

$$\Lambda_\epsilon((a_{2,\boldsymbol{\nu}}^{\text{gauss}})_{\boldsymbol{\nu}\in\mathcal{F}}) = \{\boldsymbol{\nu}\in\mathcal{F} \,:\, a_{2,\boldsymbol{\nu}}^{\text{gauss}} \geq \epsilon\}. \tag{5.8b}$$

5.1.3. Decay of the Taylor GPC coefficients

Consider the two sequences $(c_{2,\boldsymbol{\nu}}^{\text{leja}})_{\boldsymbol{\nu}\in\mathcal{F}}$ and $(a_{2,\boldsymbol{\nu}}^{\text{leja}})_{\boldsymbol{\nu}\in\mathcal{F}}$ from Section 5.1.1. By Lemma 3.12 and Remark 3.15 it holds $(c_{2,\boldsymbol{\nu}}^{\text{leja}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/2}(\mathcal{F})$ and $(a_{2,\boldsymbol{\nu}}^{\text{leja}})_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/2}(\mathcal{F})$ for any $p > 1/r$. Denote by $(c_{2,j}^*)_{j\in\mathbb{N}}$ and $(a_{2,j}^*)_{j\in\mathbb{N}}$ two monotonically decreasing rearrangements. By Lemma 3.13, for any $\delta > 0$ there exists a constant C such that for all $j\in\mathbb{N}$

$$c_{2,j}^* \leq Cj^{-2r+\delta} \qquad \text{and} \qquad a_{2,j}^* \leq Cj^{-2r+\delta}. \tag{5.9}$$

Figure 3 depicts the decay of these sequences for different values of r and θ. The rates in (5.9) are in general not obtained in Figure 3, as there appears to be a large preasymptotic range for larger θ. Decreasing θ improves the situation in the plotted range of j. For very small θ, the rates come close to the ones predicted by (5.9).

By Remark 5.4 and by definition of $c_{2,\boldsymbol{\nu}}^{\text{leja}}$, it holds $|t_{\boldsymbol{\nu}}| = (c_{2,\boldsymbol{\nu}}^{\text{leja}})$ for all $\boldsymbol{\nu}\in\mathcal{F}_2$ for the Taylor coefficient $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u_1(\boldsymbol{y})|_{\boldsymbol{y}=0}/\boldsymbol{\nu}!$ of the function u_1 from Example 5.2. Similarly, by Theorem 3.14, it holds $\|t_{\boldsymbol{\nu}}\|_X \leq Ca_{2,\boldsymbol{\nu}}^{\text{leja}}$ for the Taylor gpc coefficients $(t_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \subseteq X$ of any $(\boldsymbol{b},\varepsilon)$-holomorphic function, *provided that* $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})}$ *is small enough* as stated in Theorem 3.14 (v). Figure 3 suggests that there is a preasymptotic range, where the norms of the Taylor gpc coefficients decay slower than implied by Lemma 3.13 and the fact that $(\|t_{\boldsymbol{\nu}}\|_X)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$ as stated in Theorem 3.14 (iii). Since the proof of Theorem 4.3 heavily relies on this decay (for $k = 2$), we expect to have a range of preasymptotic convergence with subpar convergence of the Smolyak quadrature.

5.2. Quadrature algorithm

For the convenience of the reader we now briefly summarize our algorithm to approximate the integral of a $(\boldsymbol{b},\varepsilon)$-holomorphic function $u : U \to X$ (*cp.* Def. 3.1):

(i) Choose (univariate) **quadrature points** $\boldsymbol{\chi} = ((\chi_{n,j})_{j=0}^n)_{n\in\mathbb{N}_0} \subseteq [-1,1]$, such that the norms of the corresponding univariate quadrature rules are polynomially bounded according to (2.3).
(ii) Choose a suitable set $\mathfrak{I} \subseteq \mathbb{N}_0$ **of admissible indices**. In case the quadrature points are nested (see Def. 2.1), one can simply set $\mathfrak{I} = \mathbb{N}_0$ (*cp.* Thm. 4.3 (iii)). For non-nested quadrature points, $\mathfrak{I}$ should satisfy Assumption 2.4, we can choose for example $\mathfrak{I} = \{0\}\cup\{2^j \,:\, j\in\mathbb{N}_0\}$ (see Rem. 5.5 for more details).
(iii) Using $\mathfrak{I}$ from (ii) and the sequence $\boldsymbol{b} = (b_i)_{i\in\mathbb{N}}$ from Definition 3.1, **define a sequence** $(a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ (or $(c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$) *via* the formula provided in Theorem 3.14. For our numerical experiments we simply set the unknown constants T, τ_0, τ_1 and τ_2 in Theorem 3.14 to 1, see the formulas in Section 5.1.
(iv) Given $\epsilon > 0$ **determine** $\Lambda_\epsilon = \{\boldsymbol{\nu}\in\mathcal{F} \,:\, a_{2,\boldsymbol{\nu}} \geq 0\}$ (or $\Lambda_\epsilon = \{\boldsymbol{\nu}\in\mathcal{F} \,:\, c_{2,\boldsymbol{\nu}} \geq 0\}$). This can be achieved in almost linear complexity as explained in Section. 3.1.3 of [37] (under certain assumptions including $\boldsymbol{b}$ to be monotonically decreasing).
(v) Determine all **combination coefficients** $(\varsigma_{\Lambda_\epsilon,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ in (2.5).
(vi) **Evaluate the integrand** u at all points in $\mathrm{pts}(\Lambda_\epsilon,\boldsymbol{\chi})$, see (2.7).
(vii) **Compute** $Q_{\Lambda_\epsilon} u$ using (2.4) and (2.5).

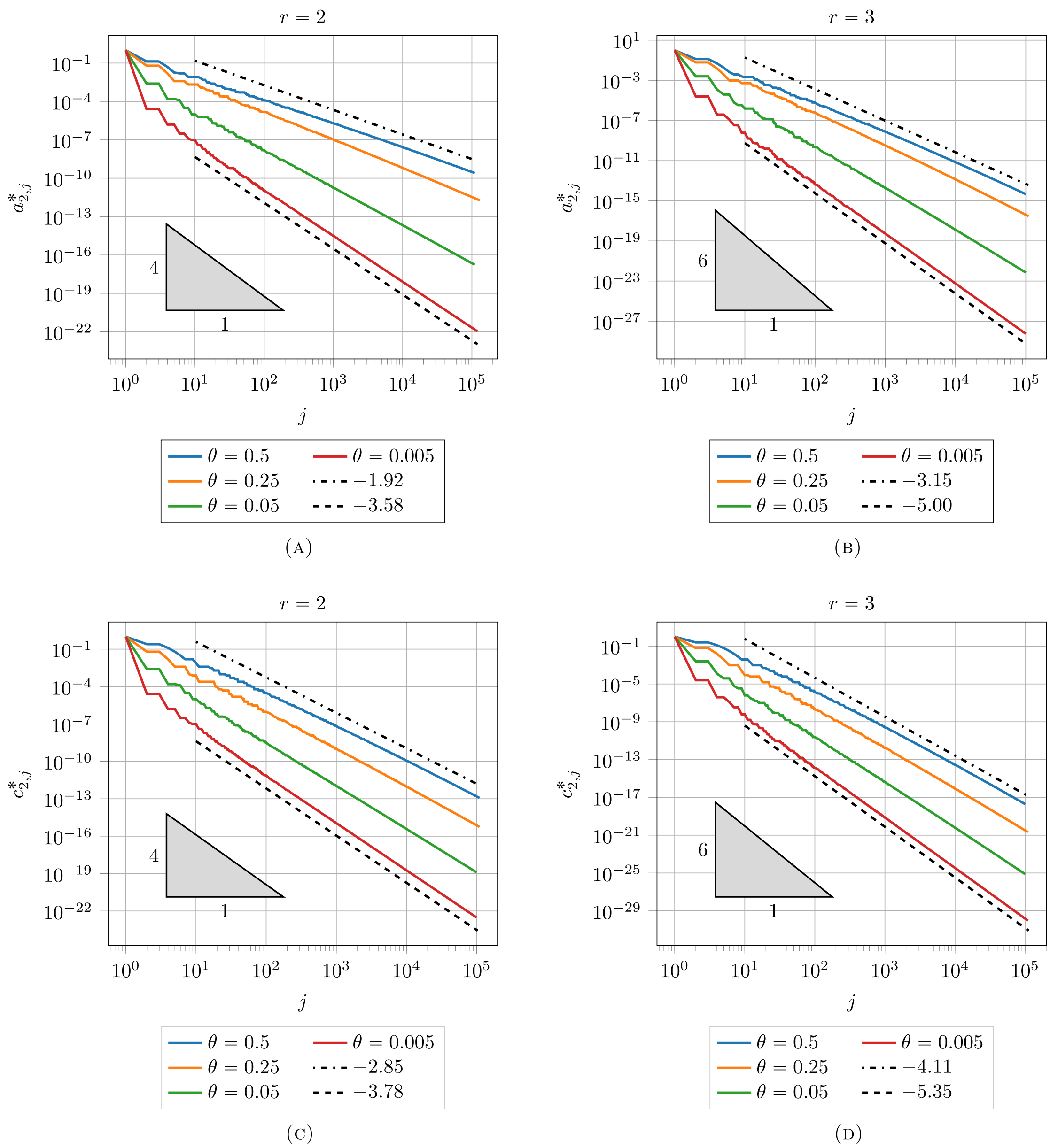


FIGURE 3. Decay of monotonically decreasing rearrangements $(a^*_{2,j})_{j\in\mathbb{N}}$ and $(c^*_{2,j})_{j\in\mathbb{N}}$ of $(a^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ and $(c^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ in (5.6), (5.5). In all cases, the asymptotic algebraic decay rate is $2r-\delta$ for any $\delta > 0$ as stated in (5.9). (A) $(a^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, $r = 2$. (B) $(a^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, $r = 3$. (C) $(c^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, $r = 2$. (D) $(c^{\mathrm{leja}}_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, $r = 3$.

Remark 5.5. Theorem 4.3 provides error bounds either in terms of the number of quadrature points (*i.e.* the number of required function evaluations) or in terms of the cost quantity defined in (2.20). The following table summarizes which parts of Assumption 2.4 the set $\mathfrak{I}$ needs to satisfy in order for our convergence theory to hold:

	Nested points	Non-nested points
nr. of points	Assumption 2.4 (i)	Assumption 2.4 (i), (ii)
cost	Assumption 2.4 (i), (ii)	Assumption 2.4 (i), (ii)

That is, Assumption 2.4 (ii) is always required, except for the case where we have nested points and measure the error in terms of the number of function evaluations. As mentioned before, a natural choice for $\mathfrak{I}$ is $\mathfrak{I} = \{0\} \cup \{2^j : j \in \mathbb{N}_0\}$ in case Assumption 2.4 (i), (ii) has to be satisfied, and $\mathfrak{I} = \mathbb{N}_0$ if only Assumption 2.4 (i) has to be satisfied.

Remark 5.6. Above we assumed given quadrature points $\boldsymbol{\chi} = ((\chi_{n,j})_{j=0}^n)_{n\in\mathbb{N}_0}$. Some quadrature rules only provide univariate quadrature points for certain but not all $n \in \mathbb{N}_0$. For example, the Clenshaw-Curtis quadrature is given through $\chi_{0,0} = 0$ and

$$\chi_{2^k,j} = \cos\left(\frac{j\pi}{2^k}\right), \qquad j \in \{0, \dots, 2^k\},\ k \in \mathbb{N}_0.$$

Thus $(\chi_{n,j})_{j=0}^n$ is only defined for $n \in \{0\} \cup \{2^k : k \in \mathbb{N}_0\}$. Such a quadrature rule still fits our setting, namely by setting $\mathfrak{I} := \{0\} \cup \{2^k : k \in \mathbb{N}_0\}$ in the above algorithm: As explained in Section 2.3 (see in particular (2.5) and Lem. 2.8), the algorithm then realizes a quadrature rule $Q_{\Lambda_\epsilon} = \sum_{\{\boldsymbol{\nu}\in\Lambda_\epsilon : \varsigma_{\Lambda_\epsilon,\boldsymbol{\nu}}\neq 0\}} \varsigma_{\Lambda_\epsilon,\boldsymbol{\nu}} Q_{\boldsymbol{\nu}}$ that is a linear combination of tensorized quadrature rules $Q_{\boldsymbol{\nu}} = \bigotimes_{j\in\mathbb{N}} Q_{\nu_j}$ for multiindices $\boldsymbol{\nu}$ with $\nu_j \in \mathfrak{I}$ for all $j \in \mathbb{N}$. All of those tensorized quadrature rules $Q_{\boldsymbol{\nu}}$ are well-defined.

To formally satisfy the requirements of our results, one can simply "fill in" the missing quadrature points by defining $(\chi_{n,j})_{j=0}^n$ for instance as the Gauss points whenever $n \in \mathbb{N}_0 \backslash \mathfrak{I}$. This has no effect on Q_{Λ_ϵ}, since it does not change $Q_{\boldsymbol{\nu}}$ for multiindices $\boldsymbol{\nu}$ with $\nu_j \in \mathfrak{I}$ for all $j \in \mathbb{N}$.

5.3. Preasymptotic behaviour

In the range shown in Figure 3, *for values of the scaling parameter* $\theta \in (0,1)$ *close to* 1, the observed convergence rates appear to contradict the predicted asymptotic rates as noted in Section 5.1.3. To understand this, we investigate in more detail the decay of the (modulus of the) Taylor gpc coefficients $(\prod_{j\in\mathbb{N}}(\theta j^{-r})^{\nu_j})_{\boldsymbol{\nu}\in\mathcal{F}}$ of the function in Example 5.2 for $b_j = \theta j^{-r}$ and some fixed values of θ and r (*cp.* Rem. 5.4). This sequence can be written as

$$(\theta^{|\boldsymbol{\nu}|}\boldsymbol{\rho}^{-r\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}} \qquad \text{where} \qquad \boldsymbol{\rho} = (j)_{j\in\mathbb{N}}. \tag{5.10}$$

We partition $\mathcal{F}_k$, $k \in \{1,2\}$, into subsets of *m-homogeneous multiindices*

$$\mathcal{F}_1^m := \{\boldsymbol{\nu} \in \mathcal{F}_1 : |\boldsymbol{\nu}| = m\} \qquad \text{and} \qquad \mathcal{F}_2^m := \{\boldsymbol{\nu} \in \mathcal{F}_2 : |\boldsymbol{\nu}| = m\}. \tag{5.11}$$

For $k \in \{1,2\}$ denote

$$(x_{k;j})_{j\in\mathbb{N}}, \text{ a decreasing rearrangement of } (\theta^{|\boldsymbol{\nu}|}\boldsymbol{\rho}^{-r\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k} \tag{5.12}$$

and for $m \in \mathbb{N}$

$$(x_{k;m;j})_{j\in\mathbb{N}}, \text{ a decreasing rearrangement of } (\theta^{m}\boldsymbol{\rho}^{-r\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k^m}. \tag{5.13}$$

The next lemma describes the asymptotic decay of these sequences.

Lemma 5.7. *Fix* k, $m \in \mathbb{N}$ *and* $\theta \in (0,1)$, $r > 0$ *in* (5.13). *For every* $\delta > 0$ *exists* $C > 0$ *(depending on* δ, k, m, θ *and* r*) such that*

$$\forall j \in \mathbb{N}: \quad x_{k;j} \leq C j^{-kr+\delta} \qquad \text{and} \qquad x_{k;m;j} \leq C j^{-kr+\delta}. \tag{5.14}$$

Proof. By Lemma 3.10, $((\theta\boldsymbol{\rho}^{-r})^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{1/(kr)+\delta}(\mathcal{F}_k)$. Lemma 3.13 implies (5.14) for $(x_{k;j})_{j\in\mathbb{N}}$. Since $(x_{k;m;j})_{j\in\mathbb{N}}$ is a subsequence of $(x_{k;j})_{j\in\mathbb{N}}$, also the second bound in (5.14) is satisfied. $\square$

In Section 5.3.1 we will show that certain logarithmic factors are involved in the decay of $(x_{1;m;j})_{j\in\mathbb{N}}$, so that the algebraic rate r in (5.14) (for $k=1$) is observed only for large values of j. The case $k=1$ is more relevant for stochastic collocation (*i.e.* interpolation rather than quadrature), but the analysis in Section 5.3.1 explains to some extent the preasymptotic behaviour of these sequences. In Section 5.3.2, we establish a formula for a lower bound of the sequence $(x_{2;j})_{j\in\mathbb{N}}$ (*i.e.* $k=2$). A plot of this lower bound (see Fig. 6) will show that (for large θ) the asymptotic regime is reached only for very large values of j.

5.3.1. Decay w.r.t. $\mathcal{F}_1^m$

Throughout the following, log denotes the natural logarithm.

Lemma 5.8. *Let* $r>0$, $\boldsymbol{\rho}=(j)_{j\in\mathbb{N}}$ *and* $m\in\mathbb{N}$. *For* $R\geq 0$ *set*

$$A_m(R) := \sum_{\{\boldsymbol{\nu}\in\mathcal{F}_1^m\,:\,\boldsymbol{\rho}^{-r\boldsymbol{\nu}}\geq R^{-r}\}} \frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}.$$

Then $A_m(R)=0$ *if* $R<1$ *and with* $c_0:=1-\log(2)\in(0,1)$ *for all* $R\geq 1$

$$c_0^m R\sum_{i=0}^{m-1}\frac{(c_0^{-1}\log(R))^i}{i!} \leq A_m(R) \leq 2^{m-1}R\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}. \tag{5.15}$$

Proof. For $R\in[0,1)$ the sum is over the empty set, so let $R\geq 1$ in the following. Then

$$A_m(R) = \sum_{\{\boldsymbol{\nu}\in\mathcal{F}\,:\,|\boldsymbol{\nu}|=m,\ \boldsymbol{\rho}^{-r\boldsymbol{\nu}}\geq R^{-r}\}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} = \left|\left\{(i_1,\ldots,i_m)\in\mathbb{N}^m \,:\, \prod_{j=1}^m i_j^{-r}\geq R^{-r}\right\}\right|,$$

since for every $\boldsymbol{\nu}\in\mathcal{F}$ with $|\boldsymbol{\nu}|=m$, there exist exactly $|\boldsymbol{\nu}|!/\boldsymbol{\nu}!$ elements $(i_1,\ldots,i_m)$ of $\mathbb{N}^m$ such that $|\{j\in\{1,\ldots,m\}\,:\,i_j=l\}|=\nu_l$ for all $l\in\mathbb{N}$. With $N:=\lfloor R\rfloor\in\mathbb{N}$ we have

$$A_{m+1}(R)=\sum_{j=1}^N\left|\{(i_1,\ldots,i_m)\,:\,j^{-r}\prod_{l=1}^m i_l^{-r}\geq R^{-r}\}\right| = \sum_{j=1}^N\left|\{(i_1,\ldots,i_m)\,:\,\prod_{l=1}^m i_l^{-r}\geq (R/j)^{-r}\}\right| = \sum_{j=1}^N A_m(R/j). \tag{5.16}$$

To prove the upper bound in (5.15), we proceed by induction over m. For $m=1$ it holds $i_1^{-r}\geq R^{-r}$ iff $i_1\leq R$, so that $A_1(R)=\lfloor R\rfloor$ and the estimate is satisfied. Next, employing (5.16) and the induction hypothesis

$$A_{m+1}\leq 2^{m-1}\sum_{j=1}^N\frac{R}{j}\sum_{i=0}^{m-1}\frac{\log(R/j)^i}{i!} = 2^{m-1}\sum_{j=1}^N\sum_{i=0}^{m-1}\frac{1}{i!}\frac{\log(R/j)^i}{j}.$$

For any $i\in\mathbb{N}$ and all $x\in[1,R]$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\log(R/x)^i}{x}\right) = \frac{-i\log(R/x)^{i-1}-\log(R/x)^i}{x^2}\leq 0.$$

Therefore $f(x):=2^{m-1}\sum_{i=0}^{m-1}\log(R/x)^i/(x\cdot i!)$ is monotonically decreasing for $x\in[1,R]$. Thus $\sum_{j=1}^N f(j)\leq f(1)+\int_1^R f(x)\,\mathrm{d}x$, giving

$$\sum_{j=1}^{N} 2^{m-1}\sum_{i=0}^{m-1}\frac{1}{i!}\frac{\log(R/j)^i}{j} \le f(1)+\int_1^R\sum_{i=0}^{m-1}\frac{2^i}{i!}\frac{\log(R/x)^i}{x}\,\mathrm{d}x = f(1)+2^{m-1}\sum_{i=0}^{m-1}\frac{1}{i!}\int_0^{\log(R)}(\log(R)-y)^i\,\mathrm{d}y$$

$$= 2^{m-1}\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}+2^{m-1}\sum_{i=0}^{m-1}\frac{1}{i!}\frac{\log(R)^{i+1}}{i+1} \le 2^m\sum_{i=0}^{m}\frac{\log(R)^i}{i!}, \tag{5.17}$$

which concludes the proof of the upper bound.

For the lower bound, the case $m=1$ follows by $Rc_0 \le \lfloor R\rfloor = A_1(R)$ where $c_0=(1-\log(2))<1/2$. With (5.16), due to the induction hypothesis

$$A_{m+1}(R)=\sum_{j=1}^{N}A_m(R/j)\ge R\sum_{j=1}^{N}\frac{R}{j}\sum_{i=0}^{m-1}\frac{c_0^{m-i}\log(R/j)^i}{i!}.$$

Note that for $\lfloor R\rfloor = N\ge 1$

$$\sum_{j=1}^{N}\frac{1}{j}\ge 1+\int_2^{N+1}\frac{1}{x}\,\mathrm{d}x \ge 1-\int_1^2\frac{1}{x}\,\mathrm{d}x+\int_1^R\frac{1}{x}\,\mathrm{d}x = c_0+\log(R).$$

Hence, using (as above) that $f(x):=\sum_{i=0}^{m-1}c_0^{m-i}\log(R/x)^i/x$ is monotonically decreasing for $x\in[1,R]$ so that $\sum_{j=1}^{N}f(j)\ge\int_1^R f(x)\,\mathrm{d}x$, similar as in (5.17) we get

$$\begin{aligned}\sum_{j=1}^{N}\sum_{i=0}^{m-1}\frac{1}{i!}\frac{c_0^{m-i}\log(R/j)^i}{j} &= \sum_{j=1}^{N}\frac{c_0^m}{j}+\sum_{j=1}^{N}\sum_{i=1}^{m-1}\frac{c_0^{m-i}}{i!}\frac{\log(R/j)^i}{j}\\ &\ge c_0^{m+1}+c_0^m\log(R)+\sum_{i=1}^{m-1}\int_1^R\frac{c_0^{m-i}}{i!}\frac{\log(R/x)^i}{x}\,\mathrm{d}x\\ &= \sum_{i=0}^{m}\frac{c_0^{m+1-i}\log(R)^i}{i!},\end{aligned}$$

which proves the lower bound in (5.15). $\square$

With Lemma 5.8 and $c_0:=1-\log(2)\in(0,1)$, we observe for $R\ge 1$

$$f_m(R):=\frac{c_0^m}{m!}R\sum_{i=0}^{m-1}\frac{(c_0^{-1}\log(R))^i}{i!}\le|\{\boldsymbol{\nu}\in\mathcal{F}_1^m\,:\,\boldsymbol{\rho}^{-r\boldsymbol{\nu}}\ge R^{-r}\}|\le 2^{m-1}R\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}=:g_m(R), \tag{5.18}$$

which immediately gives:

Lemma 5.9. *For $j\in\mathbb{N}$ let $R_j\ge 1$ and $S_j\ge 1$ be such that $f_m(R_j)=j$ and $g_m(S_j)=j$. Then with $x_{1;m;j}$ as in (5.13)*

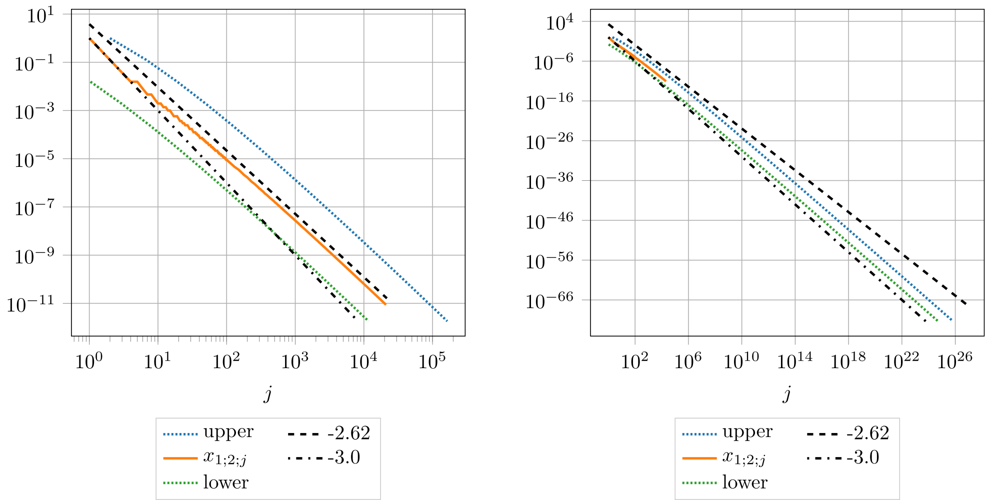
$$\theta^m R_j^{-r}\le x_{1;m;j}\le\theta^m S_j^{-r}\qquad\forall j\in\mathbb{N}. \tag{5.19}$$

Lemma 5.9, gives the parametrized curves

$$(f_m(R),\theta^m R^{-r})\qquad\text{and}\qquad(g_m(R),\theta^m R^{-r}) \tag{5.20}$$

for $R\ge 1$, which are lower and upper bounds of $(x_{1;m;j})$ at every R_j, S_j where $f_m(R_j)=j$ and $g_m(S_j)=j$.

To estimate the local algebraic decay of the upper bound for m in Lemma 5.9, we need to compute the slope of the curve $(\log(g_m(R)),\log(\theta^m R^{-r}))$. At $(\log(g_m(R)),\log(\theta^m R^{-r}))$ it equals



(A) (B)

FIGURE 4. Decay of $(x_{1;2;j})_{j\in\mathbb{N}}$ in (5.13) (*i.e.* $m = 2$ and $k = 1$), for $r = 3$, $\theta = 1$. Additionally, the lower and upper bounds of $x_{1;2;j}$ in (5.20) (*cp.* (5.19)) are depicted. For any $\delta > 0$ there exists $C > 0$ such that $x_{1;2;j} \le Cj^{-3+\delta}$ for all $j \in \mathbb{N}$. For small j, a worse, preasymptotic rate is observed. (A) $(x_{1;2;j})_{j\in\mathbb{N}}$ and lower/upper bound. (B) $(x_{1;2;j})_{j\in\mathbb{N}}$ and lower/upper bound for larger range of j.

$$\frac{\frac{d}{dR}\log(\theta^m R^{-r})}{\frac{d}{dR}\log(g_m(R))} = -\frac{r\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}}{g_m'(R)} = -\frac{r\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}}{\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}+\sum_{i=0}^{m-2}\frac{\log(R)^i}{i!}} = -\frac{r}{1+\frac{\sum_{i=0}^{m-2}\frac{\log(R)^i}{i!}}{\sum_{i=0}^{m-1}\frac{\log(R)^i}{i!}}}\cdot$$

For example, if $m = 2$, then the upper bound at position $j = g_2(S_j) = S_j(1 + \log(S_j))$ locally decreases at the algebraic rate

$$\frac{r}{1+\frac{1}{1+\log(S_j)}}\cdot \tag{5.21}$$

A similar deliberation for the lower bound in (5.19) gives the rate $r/(1 + c_0^{-1}/(1 + c_0^{-1}\log(R_j)))$ at position $j = f_2(R_j) = R_j(c_0 + \log(R_j))c_0/2$. The logarithmic term $\log(S_j)$ in (5.21) explains why a rate close to r is only observed for large j. Due to the additional (higher order) logarithmic terms in (5.18), in a given, fixed range of j, the rate of decay becomes worse as m grows.

Figure 4 shows the sequence $(x_{1;2;j})_{j\in\mathbb{N}}$ (*i.e.* $m = 2$) for $r = 3$ together with the lower and upper bounds in (5.20). For small j, the behaviour of $(x_{1;m;j})_{j\in\mathbb{N}}$ is far from j^{-r}. The plot of the bounds for larger values of j shows that the rate will eventually approach r.

5.3.2. *Decay w.r.t.* $\mathcal{F}_2$

For the convergence rate analysis of the Smolyak quadrature, we are mainly interested in the sequence $x_{2;j}$ in (5.13), *i.e.* the decreasing rearrangement of $(\theta^{|\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(j^{-r\nu_j}))_{\boldsymbol{\nu}\in\mathcal{F}_2}$. Here and in the following, we fix $\theta \in (0, 1)$ and $r > 0$.

We first discuss the decay of $(x_{2;m;j})_{j\in\mathbb{N}}$ (*cp.* (5.13)) for different $m \in \mathbb{N}$. Recall that by (5.14), for any $\delta > 0$ there exists C such that $(x_{2;m;j})_{j\in\mathbb{N}} \le Cj^{-2r}$ for all $j \in \mathbb{N}$.

- $m = 1$: Since $\mathcal{F}_2^1 = \{\boldsymbol{\nu} \in \mathcal{F}_2 : |\boldsymbol{\nu}| = 1\} = \emptyset$ this case is trivial.
- $m = 2$: With $\mathbf{e}_j = (\delta_{ij})_{i\in\mathbb{N}}$ we have $\mathcal{F}_2^2 = \{2\mathbf{e}_j\ :\ j \in \mathbb{N}\}$ and $\{\boldsymbol{\rho}^{-r\boldsymbol{\nu}}\ :\ \boldsymbol{\nu} \in \mathcal{F}_2^2\} = \{j^{-2r}\ :\ j \in \mathbb{N}\}$ so that $x_{2;2;j} = j^{-2r}$, and the decay predicted by (5.14) is apparent also for small j.

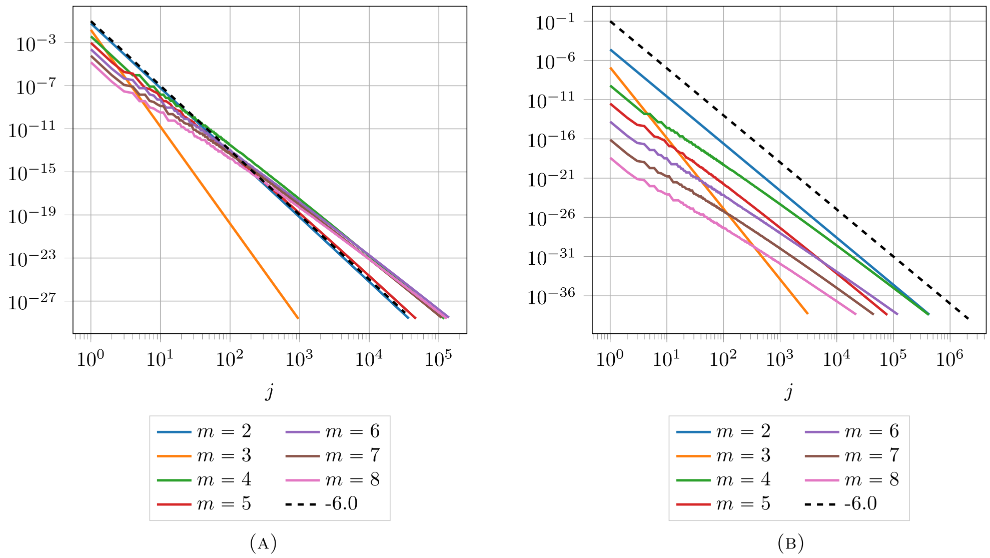


FIGURE 5. Decay of $(x_{2;m;j})_{j\in\mathbb{N}}$ in (5.13) for $r = 3$ and different values of θ. For any $\delta > 0$ and all $m \geq 2$ there exists $C > 0$ such that $x_{2;m;j} \leq Cj^{-6+\delta}$ for all $j \in \mathbb{N}$. (A) $\theta = 0.25$. (B) $\theta = 0.005$.

– $m = 3$: It holds $\mathcal{F}_2^3 = \{3\mathbf{e}_j : j \in \mathbb{N}\}$ and thus $\{\boldsymbol{\rho}^{-r\boldsymbol{\nu}} : \boldsymbol{\nu} \in \mathcal{F}_2^3\} = \{j^{-3r} : j \in \mathbb{N}\}$. Hence $m = 3$ can be considered as a special case, since $x_{2;3;j} = j^{-3r}$ and the decay is even faster than j^{-2r}, see Figure 5.
– $m = 4$: We have

$$\left|\{\boldsymbol{\nu} \in \mathcal{F}_2^4 : \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \geq R^{-r}\}\right| = \left|\{\boldsymbol{\nu} \in \mathcal{F}_1^2 : \boldsymbol{\rho}^{-r2\boldsymbol{\nu}} \geq R^{-r}\}\right| = \left|\{\boldsymbol{\nu} \in \mathcal{F}_1^2 : \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \geq R^{-r/2}\}\right|$$

and thus with (5.18)

$$f_2(R^{1/2}) \leq \left|\{\boldsymbol{\nu} \in \mathcal{F}_2^4 : \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \geq R^{-r}\}\right| \leq g_2(R^{1/2}).$$

Considering the parametrized curves $(f_2(R^{1/2}), \theta^4 R^{-r})$, $(g_2(R^{1/2}), \theta^4 R^{-r})$ for $R \geq 1$, a computation similar to the one before (5.21) implies that the decay of $(x_{2;4;j})_{j\in\mathbb{N}}$ in the preasymptotic range is worse than what (5.14) suggests, due to the logarithmic factors occurring in f_2, g_2.
– $m > 4$: Similar arguments as in the case $m = 4$ apply, and we expect the decay rate to further diminish as m grows. The precise behaviour depends on the number of possibilities to write m as a sum of integers in $\mathbb{N}\backslash\{1\}$: for example $\{x_{2;5;j} : j \in \mathbb{N}\} = \{k^{-2}l^{-3} : k \neq l \in \mathbb{N}\}$ decreases faster than $\{x_{2;4;j} : j \in \mathbb{N}\} = \{k^{-2}l^{-2} : k < l \in \mathbb{N}\}$, as Figure 5 right panel shows.

Implications for $(x_{2;j})_{j\in\mathbb{N}}$ are as follows. All terms belonging to $\mathcal{F}_2^m$, *i.e.*

$$(\theta\boldsymbol{\rho}^{-r})^{\boldsymbol{\nu}} = \theta^m \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \qquad \forall\, \boldsymbol{\nu} \in \mathcal{F}_2^m, \tag{5.22}$$

are *scaled by the common factor* θ^m: the smaller θ, the fewer multiindices of high total order m (which, in the preasymptotic range, decay slower than expected as we have noticed) will be among the N largest ones. This is depicted in Figure 5 which shows the sequences $(x_{2;m;j})_{j\in\mathbb{N}}$ for $m \in \{2, \dots, 8\}$ and two different values $\theta \in \{0.25, 0.005\}$.

If $0 < \theta < 1$ is small then, due to the factor θ^m in (5.22), only few multiindices of order $m \geq 4$ occur among the largest, and essentially $((\theta\boldsymbol{\rho}^{-r})^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_2^2\cup\mathcal{F}_2^3}$ governs the decay of x_j for small j, thus yielding the expected rate $2r - \delta$. On the other hand, as θ draws closer to 1, more higher order multiindices contribute to the largest j terms, resulting in a longer preasymptotic range with slower decay.

To numerically verify these heuristic considerations, we determine a lower bound of $x_{2;j}$. With f_m as in (5.18), for $R \geq 1$ there holds

$$f_m(R) \leq |\{\boldsymbol{\nu} \in \mathcal{F}_1^m \,:\, \boldsymbol{\rho}^{-r2\boldsymbol{\nu}} \geq R^{-2r}\}| \leq |\{\boldsymbol{\nu} \in \mathcal{F}_2^{2m} \,:\, \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \geq R^{-2r}\}|. \tag{5.23}$$

We extend f_m *via* $f_m(R) := 0$ for all $R \in [0, 1)$, and (5.23) then remains true also for $R < 1$. Then

$$\begin{aligned} F(R) &:= 1 + \sum_{m\in\mathbb{N}} f_m(\theta^{2m/2r}R) \leq |\{\mathbf{0}\}| + \sum_{m\in\mathbb{N}} |\{\boldsymbol{\nu} \in \mathcal{F}_2^{2m} \,:\, \boldsymbol{\rho}^{-r\boldsymbol{\nu}} \geq (\theta^{2m/2r}R)^{-2r}\}| \\ &= |\{\mathbf{0}\}| + \sum_{m\in\mathbb{N}} |\{\boldsymbol{\nu} \in \mathcal{F}_2^{2m} \,:\, (\theta\boldsymbol{\rho}^{-r})^{\boldsymbol{\nu}} \geq R^{-2r}\}| \leq |\{\boldsymbol{\nu} \in \mathcal{F}_2 \,:\, (\theta\boldsymbol{\rho}^{-r})^{\boldsymbol{\nu}} \geq R^{-2r}\}|, \end{aligned} \tag{5.24}$$

which gives:

Lemma 5.10. *For $j \in \mathbb{N}$ let $R_j \geq 1$ be such that $F(R_j) = j$. For the sequence $x_{2;j}$ it holds $R_j^{-2r} \leq x_{2;j}$.*

Figure 6 depicts the decay of $(x_{2;j})_{j\in\mathbb{N}}$ as well as the lower bound in Lemma 5.10 for $r = 3$ and $\theta = 0.25$. The measured rate of $(x_{2;j})_{j\in\mathbb{N}}$ in the observed range of j is merely 4.96 and not close to 6 as suggested by (5.14). For the plotted range of j in Figure 6A up to about $j = 10^6$, the lower bound from Lemma 5.10 seems to capture well the preasymptotic behaviour of $(x_{2;j})_{j\in\mathbb{N}}$. Plotting the lower bound for larger values of j up to about $j = 10^{55}$, we observe that its algebraic decay rate eventually increases to approach $2r = 6$, however only very slowly. This suggests, that if $\theta > 0$ is not small enough, then the range where the Taylor gpc coefficients of u_1 from Example 5.2 will show the predicted algebraic decay only occurs for j so large that it is not relevant in practice.

Finally, for general $(\boldsymbol{b}, \varepsilon)$-holomorphic functions, in the proof of Theorem 3.14 we derived estimates of the norms of the Taylor gpc coefficients which were of the type $\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!$ (also see Rem. 5.4). In this section we have analysed in more detail a sequence of the type $(\boldsymbol{b}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$, which corresponds to the Taylor gpc coefficients of u_1 in Example 5.2. Due to the additional term $|\boldsymbol{\nu}|!/\boldsymbol{\nu}!$, it can be expected that the preasymptotic effect is even stronger in the general case.

Remark 5.11. The case $k = 1$ is relevant for stochastic collocation algorithms (*i.e.* interpolation instead of quadrature). Similar as in (5.24), we can define $G(R) := 1 + \sum_{m\in\mathbb{N}} g_m(\theta^{m/r}R)$ and deduce that the curve $(G(R), R^{-r})$ provides an upper bound for the behaviour of $(x_{1;j})_{j\in\mathbb{N}}$ in (5.13). By Lemma 5.8 it holds $g_m(\theta^{m/r}R) \leq 2^{m-1}\theta^{2m/r}R^2 \leq (2\theta^{2/r})^m R^2$ for all $m \in \mathbb{N}$, and therefore $G(R) \leq 1 + (2\theta^{2/r})/(1 - 2\theta^{2/r})R^2$. For u_1 in Example 5.2 (*cp.* Rem. 5.4), we conclude that as long as θ is small enough such that the constant $(2\theta^{1/r})/(1-2\theta^{1/r})$ is (moderately) bounded, the preasymptotic error convergence of the interpolation error can be expected to be at worst half of the proven convergence rate, which is in this case $(r - 1)/2$.

This can be extended to general $(\boldsymbol{b}, \varepsilon)$-holomorphic functions, by constructing indexsets based on the sequence $c_{1;\boldsymbol{\nu}}$ as stated in Theorem 3.14 (iv) (for $k = 1$): if $\mathfrak{I} = \mathbb{N}_0$ in (3.12), then $c_{1;\boldsymbol{\nu}}$ is exactly of the type $\prod_{j\in\mathbb{N}} \varrho_j^{-\nu_j}$ (*i.e.* like the sequence analysed in the current section).

5.4. Real valued model parametric integrand functions

We now test the convergence of the Smolyak quadrature for the functions u_1, u_2 in Examples 5.2 and 5.3. For u_2 we also refer to [22] where computations for almost the same integrand were done with the method suggested in their paper.

5.4.1. Model integrand u_1

Let

$$u_1(\boldsymbol{y}) = \prod_{j\in\mathbb{N}} \frac{1}{1 + y_j\theta j^{-r}} \qquad \boldsymbol{y} \in U \tag{5.25}$$

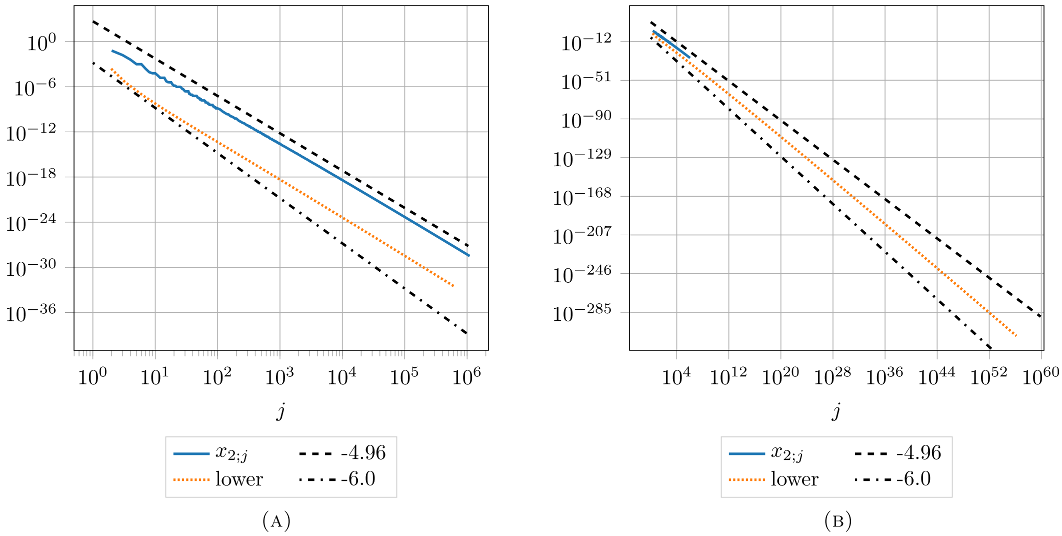


FIGURE 6. Decay of $(x_{2;j})_{j\in\mathbb{N}}$ in (5.12) for $\theta = 0.25$ and $r = 3$. The lower bound is given in Lemma 5.10. For any $\delta > 0$ there exists $C > 0$ such that $x_{2;j} \le Cj^{-2r+\delta} = Cj^{-6+\delta}$ for all $j \in \mathbb{N}$. In the preasymptotic range a worse rate is observed. (A) $(x_{2;j})_{j\in\mathbb{N}}$ and lower bound. (B) $(x_{2;j})_{j\in\mathbb{N}}$ and lower bound for larger range of j.

be as in (5.1) with $b_j := \theta j^{-r}$, $0 < \theta < 1$, $r > 1$. As explained in Example 5.2, u_1 is $(\boldsymbol{b}, \varepsilon)$-holomorphic, and by Theorem 4.3 the Smolyak quadrature can achieve the convergence rate $2r - 1$ (*cp.* Rem. 5.1) in terms of the number of quadrature points if optimal indexsets are chosen.

Figure 7 shows the absolute error $|\int_U u_1(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u_1|$ for different values of r and θ, and with Λ_ϵ as in Sections 5.1.1 and 5.1.2. Note that (up to the guessing of constants and simplifications in Sects. 5.1.1 and 5.1.2), Theorem 4.3 implies the convergence rates $2r-1$ for $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ as in (5.6) or (5.8) and $2r-2$ for $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ as in (5.5) or (5.7). The reference value for $\int_U u_1(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y})$ was computed directly as $\int_U u_1(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) = \prod_{j\in\mathbb{N}} \log((1+b_j)/(1-b_j))/(2b_j)$.

Even though the Gauss-Legendre points are *not nested*, we observe that the Leja points and the Gauss-Legendre points perform equally well in terms of the total number of function evaluations. Furthermore, the index sets $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ deliver slightly better error convergence than $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$. This is not surprising, as $(c_{2;\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}}$ is a sequence resembling the Taylor gpc coefficients of u_1, see Remark 5.4 and also Figure 9. As expected, the convergence rate (which asymptotically only depends on r), strongly depends on θ. For large θ a preasymptotic range of subpar convergence is observed. This can be explained by the preasymptotic behaviour of the decay of the Taylor gpc coefficients which we analysed in Section 5.3. For very small θ, we get close to the proven convergence rate $2r - 1$, *e.g.* for $r = 2$ and $\theta = 0.005$ we observe convergence rates of about 2.68 and 2.81 depending on the chosen index sets. The plots confirm that considerably faster convergence than the previously proved rate $r - 1$ is in principle attainable.

5.4.2. Model integrand u_2

Let

$$u_2(\boldsymbol{y}) = \frac{1}{1 + \theta \sum_{j\in\mathbb{N}} y_j j^{-r}} \tag{5.26}$$

be as in (5.2) with $b_j := \theta j^{-r}$, $r > 1$ and $\theta > 0$ small enough such that $\theta \sum_{j\in\mathbb{N}} j^{-r} < 1$. By Example 5.3, u_2 is $(\boldsymbol{b}, \varepsilon)$-holomorphic, and Theorem 4.3 implies that the Smolyak quadrature can achieve the convergence rate $2r - 1$ in terms of the number of quadrature points if optimal index sets are chosen. Figure 8 shows the

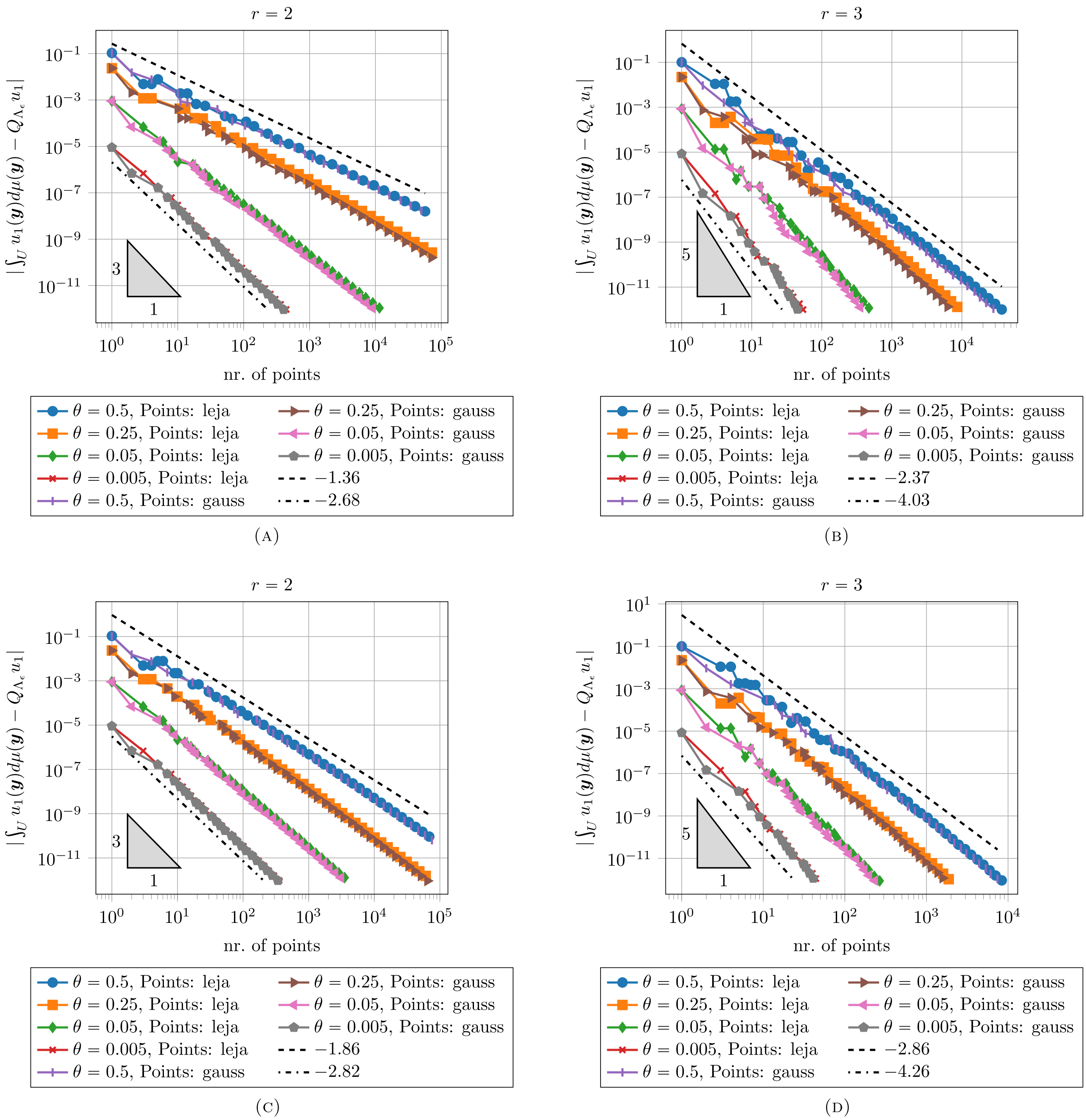


FIGURE 7. Quadrature error $|\int_U u_1(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u_1|$ for u_1 in (5.25), for different values of r and θ. The plot shows the absolute error in terms of the number of quadrature points $|\mathrm{pts}(\Lambda_\epsilon, \boldsymbol{\chi})|$ (*cp.* (2.7)). (A) $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 2$. (B) $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 3$. (C) $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 2$. (D) $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 3$.

convergence of the absolute error $|\int_U u_2(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon}u|$ for different values of r and θ. Again we compare the convergence for either nested Leja quadrature points or non-nested Gauss-Legendre quadrature points, and different *a priori* constructions of multiindices as explained in Sections 5.1.1 and 5.1.2. As before, (up to the guessing of constants and simplifications in Sects. 5.1.1 and 5.1.2), Theorem 4.3 implies the convergence rates $2r-1$ for $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ as in (5.6) or (5.8) and $2r-2$ for $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ as in (5.5) or (5.7).

The reference value for $\int_U u_2(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y})$ has been computed with a higher order quasi Monte Carlo rule (a so-called high-order, Interlaced Polynomial Lattice rule adapted to the model integrand, with suitable digit interlacing parameter, see [16] and the references there) utilizing $2^{20}\sim 10^6$ quadrature points applied to the function u restricted to the first 1024 dimensions.

The observations are similar as for u_1. The (preasymptotic) convergence rate strongly depends on the scaling parameter θ. Leja and Gauss-Legendre quadrature deliver almost the same error w.r.t. the number of function evaluations, and the index sets $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$ perform (slightly) better than $\Lambda_\epsilon((a_{2;\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$. This is observed in Figure 9 where we compare the error for both sequences directly.

5.4.3. Comparison with an adaptive method

We consider the model parametric integrand u_2 defined in (5.2), with $b_j := \theta j^{-r}$ for $r=2$ and $\theta>0$. In the following, our method is compared with a variant of the dimension adaptive algorithm described in [18] which we outline briefly for completeness. For some finite, downward closed set of multiindices $\{\mathbf{0}\}\neq\Lambda\subseteq\mathcal{F}$, following [9] we introduce the reduced set of neighbours

$$\mathcal{N}(\Lambda) := \{\boldsymbol{\nu}\in\mathcal{F} : \boldsymbol{\nu}\notin\Lambda,\ \boldsymbol{\nu}-\mathbf{e}_j\in\Lambda\ \forall j\in\operatorname{supp}\boldsymbol{\nu},\ \nu_j=0\ \forall j>\max_{\boldsymbol{\mu}\in\Lambda}\max\{i\in\mathbb{N} : \mu_i\neq 0\}+1\},$$

with the special case $\mathcal{N}(\{\mathbf{0}\}) := \{(1,0,0,\dots)\}$. Algorithm 1 shows the used adaptive method. Also recall, that $Q_{-1}:=0$ and for notational convenience also $Q_{-2}:=0$ in the following. As in (2.2), for $n\in\mathbb{N}_0$, Q_n stands for the one dimensional interpolatory quadrature employing the $n+1$ points $(\chi_j)_{j=0}^n$ in $[-1,1]$. In the following the quadrature points for the adaptive method and for the *a priori* choice of index sets consist of the Leja points introduced in Section 5.1.1.

Algorithm 1. AdaptiveSmolyak(integrand $u:[-1,1]^{\mathbb{N}}\to\mathbb{R}$, number of multiindices $M\in\mathbb{N}$).

$\Lambda_{\mathrm{act}} \leftarrow \{\mathbf{0}\}$
$\Lambda_{\mathrm{tot}} \leftarrow \{\mathbf{0}\}$
$\Delta_{\mathbf{0}} \leftarrow \bigotimes_{j\in\mathbb{N}} Q_0 u$
while $|\Lambda_{\mathrm{act}}| < M$ **do**
  $\Lambda_{\mathrm{new}} \leftarrow \mathcal{N}(\Lambda_{\mathrm{act}})\backslash\Lambda_{\mathrm{tot}}$
  $\Lambda_{\mathrm{tot}} \leftarrow \Lambda_{\mathrm{tot}}\cup\Lambda_{\mathrm{new}}$
  for $\boldsymbol{\nu}\in\Lambda_{\mathrm{new}}$ **do**
    $\Delta_{\boldsymbol{\nu}} \leftarrow \bigotimes_{j\in\mathbb{N}}(Q_{2\nu_j}-Q_{2(\nu_j-1)})u$
  end for
  $\boldsymbol{\mu} \leftarrow \operatorname{argmax}\{|\Delta_{\boldsymbol{\nu}}| : \boldsymbol{\nu}\in\Lambda_{\mathrm{tot}}\backslash\Lambda_{\mathrm{act}}\}$
  $\Lambda_{\mathrm{act}} \leftarrow \Lambda_{\mathrm{act}}\cup\{\boldsymbol{\mu}\}$
end while
$Q_{\Lambda_{\mathrm{act}}}u \leftarrow \sum_{\boldsymbol{\nu}\in\Lambda_{\mathrm{act}}}\Delta_{\boldsymbol{\nu}}u$
$Q_{\Lambda_{\mathrm{tot}}}u \leftarrow \sum_{\boldsymbol{\nu}\in\Lambda_{\mathrm{tot}}}\Delta_{\boldsymbol{\nu}}u$

Figure 10 shows a comparison of the error convergence for the adaptive Smolyak algorithm, and the Smolyak algorithm with the *a priori* index sets $\Lambda_\epsilon(((c_{2;\boldsymbol{\nu}}^{\mathrm{leja}}))_{\boldsymbol{\nu}\in\mathcal{F}})$ from Section 5.1.1. The plots show the error vs. number of quadrature points. In case of the adaptive algorithm, we plot the curve for the set of accepted indices Λ_{act} and for the set of total indices Λ_{tot}, as computed by Algorithm 1.

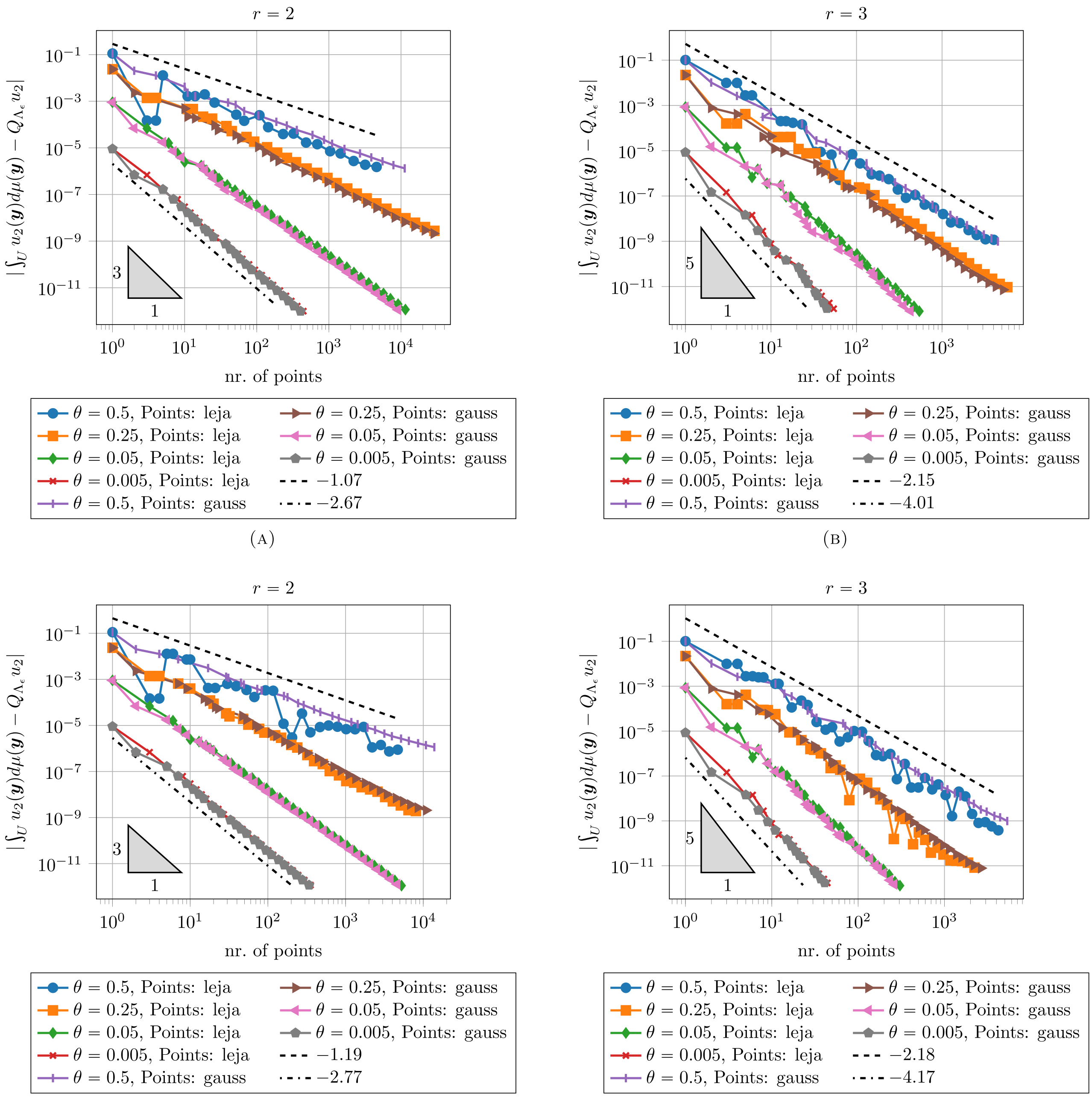


FIGURE 8. Quadrature error $|\int_U u_2(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u_2|$ for u_2 in (5.25), for different values of r and θ. The plot shows the absolute error in terms of the number of quadrature points $|\mathrm{pts}(\Lambda_\epsilon, \boldsymbol{\chi})|$ (*cp.* (2.7)). (A) $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 2$. (B) $\Lambda_\epsilon((a_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 3$. (C) $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 2$. (D) $\Lambda_\epsilon((c_{2,\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}})$, $r = 3$.

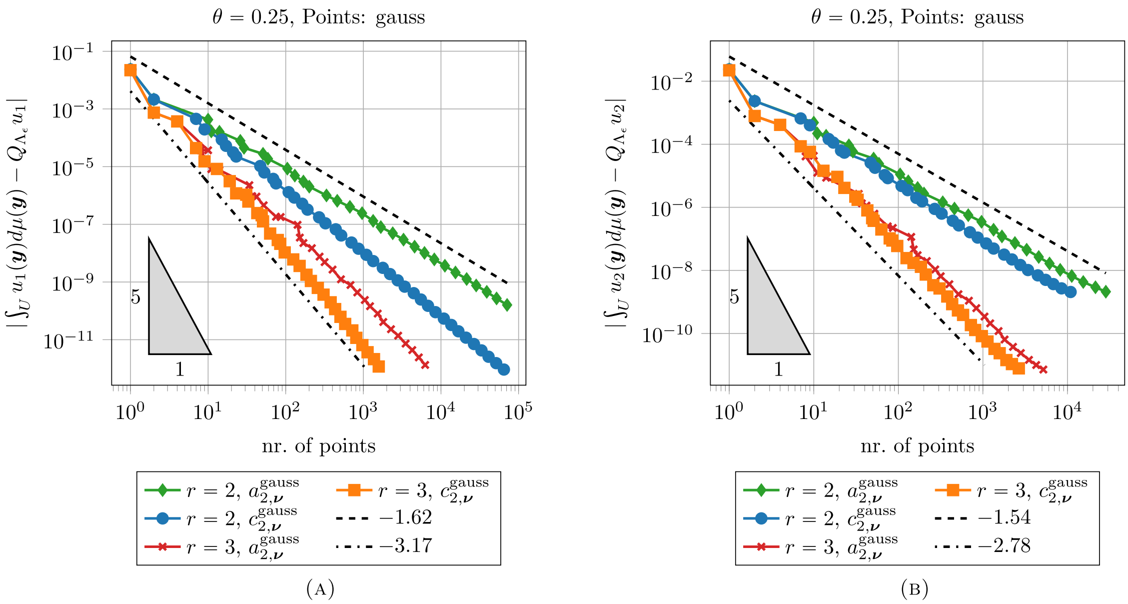


FIGURE 9. Quadrature error $|\int_U u_j(\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{y}) - Q_{\Lambda_\epsilon} u_j|$ for u_j, $j \in \{1,2\}$, in (5.25) and (5.26) for $\theta = 0.25$, $r \in \{2,3\}$ and using Gauss-Legendre quadrature points. The plot compares the error convergence for two different quadrature rules based on multiindex sets built by either using the sequence $(a_{2,\boldsymbol{\nu}}^{\text{gauss}})_{\boldsymbol{\nu}\in\mathcal{F}}$ or $(c_{2,\boldsymbol{\nu}}^{\text{gauss}})_{\boldsymbol{\nu}\in\mathcal{F}}$, see Section 5.1. (A) u_1. (B) u_2.

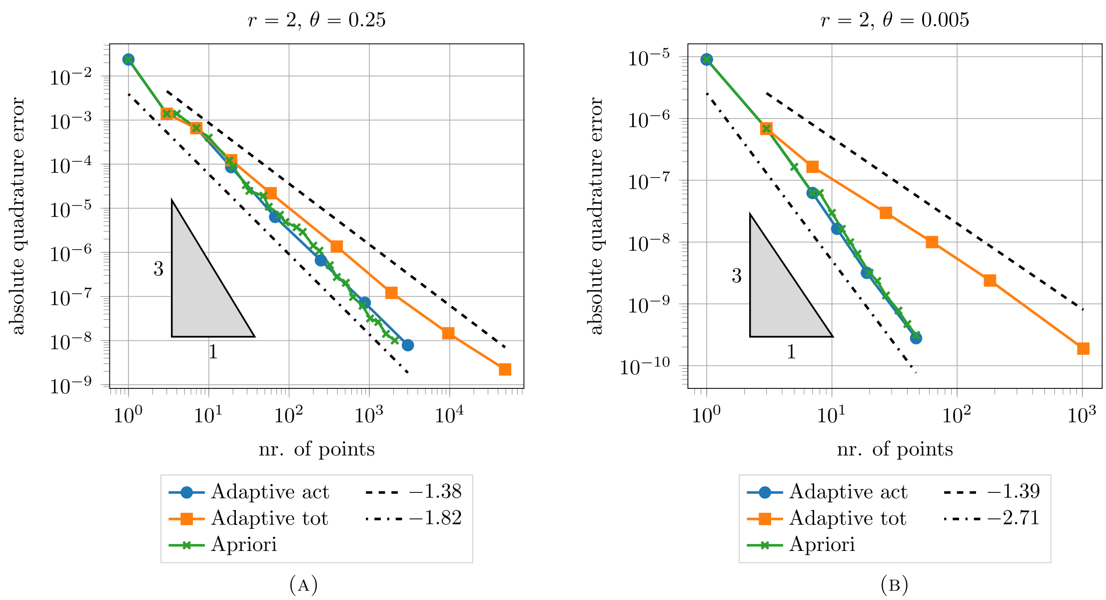


FIGURE 10. Absolute quadrature error for u_2 in (5.26), with $r = 2$ and different values of θ. We compare convergence of the adaptive algorithm in Algorithm 1 with the Smolyak quadrature based on our *a priori* choice of index sets. In both cases the same Leja quadrature points (see Sect. 5.1.1) are used. (A) $\theta = 0.25$. (B) $\theta = 0.005$.

In order to find the set Λ_{act}, Algorithm 1 also requires to evaluate the integrand at quadrature points belonging to the total set Λ_{tot}. Thus, the curve for the accepted multiindices Λ_{act} should be considered as a benchmark, whereas the curve for the total set of indices Λ_{tot} can be seen as a practically obtainable computation in terms of error vs. number of quadrature points (*i.e.* number of function evaluations). We observe, that our *a priori* chosen quadrature points are as good, as the ones obtained by the adaptive method and denoted by Λ_{act} above. This implies, that the *a priori* choice captures well the most important multiindices.

Comparing with Λ_{tot}, our method even outperforms the adaptive algorithm when θ becomes small. We mention that it was already reported earlier that *a priori* choices of index sets can perform superior to adaptive methods, see, *e.g.*, [3]. We note that the convergence for the *a priori* choice (and for the adaptive algorithm in terms of Λ_{act}) improves as θ decreases, while the convergence rate of the adaptive algorithm in terms of Λ_{tot} does *not increase* as θ decreases. For $\theta = 0.005$, the convergence rate of the adaptive algorithm w.r.t. Λ_{tot}, is only about half the convergence rate obtained with the *a priori* chosen set. This is not a coincidence, and we explain this with an example in more detail in Section 5.2.2 of [37] (see in particular [37], Example 5.2.2). We point out that one of the main advantages of determining the quadrature rule *a priori* instead of adaptively, is that it allows to compute all function values at the quadrature points in parallel, which is in general not possible for the adaptive algorithm in [18].

6. Conclusions and generalizations

We have analysed convergence rates of Smolyak quadratures for classes of smooth, Banach space valued, parametric functions with a suitable sparsity as stated in Definition 3.1. We proved that exploiting certain cancellation properties implied by the combination coefficients and the symmetry of the marginal probability measures allow for the dimension independent convergence rate $2/p - 1$ for p-summable sequences of (norms of) Taylor gpc coefficients of the parametric integrand functions. This is superior to previously known rates established, for example, in [20, 21], of N-term gpc approximation of the integrand obtained in [11], or for Higher Order Quasi-Monte Carlo integration in [14], under analogous sparsity assumptions on the parametric integrands. We also provided an *a priori* construction algorithm of integrand-adapted sparse grids whose complexity (work and memory) scales near linearly with respect to the number quadrature points. Additionally, all convergence rate bounds were shown w.r.t. the number of quadrature points, showing in particular that essentially the same convergence rates can be obtained for both nested and non-nested univariate quadrature points $\boldsymbol{\chi}$. Numerical experiments showed that the dimension-independent convergence rates are achieved with a moderate number of quadrature points provided that the scaling parameter $\theta > 0$ was small enough. For the considered test functions, this amounts to the integrand having *small deviation from their 'nominal', average, values*. We explain, by a refined analysis of the error bounds for a class of model parametric integrands, that the asymptotic range where the (dimension-independent) convergence rate $2/p - 1$ is visible could appear only for a prohibitively large number of quadrature points.

Convergence rates which are superior to N-term approximation bounds for the parametric integrands have been reported in numerical experiments for example in [32]. Concrete *a priori* estimates on gpc coefficients that may be exploited to *a priori* determine suitable index sets by *e.g.* greedy searches or by knapsack solvers were also given in these references. The presently proposed variants of the Smolyak algorithm, in particular exploiting multiindices containing a 1, appear to be new. As we prove and verify in numerical experiments, this results in an algorithm that performs comparably to the currently best (heuristic) adaptive algorithms, from [17, 18] as shown in in Figure 10.

The complexity of the Smolyak quadrature was investigated under p-summability of sequences of (X-norms of) Taylor gpc coefficients, as implied by $(\boldsymbol{b}, \varepsilon)$-holomorphy. This condition is known to hold for broad classes of holomorphic-parametric operator equations as shown in [12], and also for the corresponding Bayesian inverse problems [32,34]. We emphasize that our key findings, notably the observation that all linear terms are integrated exactly by any Smolyak quadrature, remain valid for other measures μ, presuming that the one point rule in the Smolyak construction integrates linear polynomials exactly (*cp.* Rem. 4.2). In particular, similar improvements as shown in this paper also hold in other contexts. For example, for linear, affine-parametric diffusion problems with coefficient functions $\psi_j(x)$ that exhibit localized supports (as occur for example in a wavelet expansion), improved summability of the Taylor gpc coefficients of the parametric solution was verified in Theorem 1.2 of [1]. In Chapter 3 of [37] we show that this entails a corresponding improvement of the convergence rate for Smolyak quadratures.

Another particular case in point are Gaussian measures μ. Here, for certain PDEs bounds on Hermite Chaos coefficients can be obtained by real-variable bootstrapping on the parametric PDE (see [19, 26, 29]), so that similar conclusions for the corresponding Smolyak algorithms could be expected.

In many practical settings the evaluation of the integrand is presumed to be far more costly than performing the quadrature itself. For integrands exhibiting low sparsity, using a large number of quadrature points becomes inevitable. The near linear scaling of the cost in terms of the number of quadrature points makes the algorithm feasible also for such problems.

In this paper we assumed the integrand to allow exact evaluation at each quadrature point. In general, for UQ problems the integrand is given as the solution to some PDE, which needs to be approximated by a numerical scheme. This is addressed in [38], where we perform a fully discrete error analysis taking into account the cost of approximating the function values at the quadrature points.

Appendix A. Proof of Lemma 2.9

Proof of Lemma 2.9. The first inequality follows by the downward closedness of Λ so that

$$\sum_{\boldsymbol{\nu}\in\Lambda}\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}(\nu_j+1)=\sum_{\boldsymbol{\nu}\in\Lambda}|\{\boldsymbol{\mu}\in\Lambda\,:\,\boldsymbol{\mu}\le\boldsymbol{\nu}\}|\le\sum_{\boldsymbol{\nu}\in\Lambda}|\Lambda|=|\Lambda|^2.$$

We claim that if $\Gamma\subseteq\mathcal{F}$ is finite and satisfies for some $n\in\mathbb{N}$ and $A\subseteq\mathbb{N}$ with $|A|=n$ that

$$(\operatorname{supp}\boldsymbol{\nu}=A\quad\forall\boldsymbol{\nu}\in\Gamma)\qquad\text{and}\qquad((\boldsymbol{\nu}\in\Gamma,\ \boldsymbol{\mu}\le\boldsymbol{\nu},\ \operatorname{supp}\boldsymbol{\mu}=A)\quad\Rightarrow\quad\boldsymbol{\mu}\in\Gamma)\tag{A.1}$$

then

$$\sum_{\boldsymbol{\nu}\in\Gamma|_{\mathfrak{J}}}\prod_{j\in\mathbb{N}}(1+\nu_j)\le K_{\mathfrak{J}}^n|\Gamma|.\tag{A.2}$$

Suppose that (A.2) is true. Partitioning Λ in $\{\mathbf{0}\}$ and finitely many disjoint sets Γ of the type (A.1), this immediately implies the second inequality in (2.15).

We show (A.2) by induction. For $n=1$ assume w.l.o.g. that $A=\{1\}$. Then by Assumption 2.4 (ii)

$$\sum_{\boldsymbol{\nu}\in\Gamma|_{\mathfrak{J}}}\prod_{j\in\mathbb{N}}(1+\nu_j)=\sum_{\boldsymbol{\nu}\in\Gamma|_{\mathfrak{J}}}(1+\nu_1)\le K_{\mathfrak{J}}|\Gamma|.$$

For the induction step assume that the statement is true for $n-1\ge1$, and assume w.l.o.g. that $A=\{1,\dots,n\}$. For every $i\in\mathbb{N}$ set $\Gamma_i:=\{\boldsymbol{\mu}\in\mathcal{F}\,:\,(i,\boldsymbol{\mu})\in\Gamma\}$. Then each Γ_i is of the type (A.1) for the set $\tilde{A}=\{1,\dots,n-1\}$, so that we can apply the induction hypothesis to it. Therefore

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\Gamma|_{\mathfrak{J}}}\prod_{j\in\mathbb{N}}(1+\nu_j)&=\sum_{0<i\in\mathfrak{J}}(1+i)\sum_{\boldsymbol{\mu}\in\Gamma_i|_{\mathfrak{J}}}\prod_{j\in\mathbb{N}}(1+\mu_j)\le\sum_{0<i\in\mathfrak{J}}(1+i)K_{\mathfrak{J}}^{n-1}|\Gamma_i|\\&=K_{\mathfrak{J}}^{n-1}\sum_{0<i\in\mathfrak{J}}(1+i)\sum_{\boldsymbol{\mu}\in\Gamma_i}1=K_{\mathfrak{J}}^{n-1}\sum_{\boldsymbol{\mu}\in\mathcal{F}}\sum_{\{0<i\in\mathfrak{J}\,:\,(i,\boldsymbol{\mu})\in\Gamma\}}(1+i)\\&\le K_{\mathfrak{J}}^{n-1}\sum_{\boldsymbol{\mu}\in\mathcal{F}}K_{\mathfrak{J}}|\{i\in\mathbb{N}\,:\,(i,\boldsymbol{\mu})\in\Gamma\}|=K_{\mathfrak{J}}^n|\Gamma|,\end{aligned}$$

where we used again Assumption 2.4 (ii) for the last inequality. ∎

Appendix B. Proof of Lemmas 3.10 and 3.11

Proof of Lemma 3.10. We start with $\vartheta = 0$ and $R = 1$ (*i.e.* $w_{\boldsymbol{\nu}} = 1$ for all $\boldsymbol{\nu} \in \mathcal{F}$). Fix $k \in \mathbb{N}$. Observe that $\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$ and $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ are necessary in order for $(\boldsymbol{b}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$ to hold: For every fixed $j \in \mathbb{N}$ the sequence $(b_j^{lp/k})_{l\geq k}$ is a subsequence of $(\boldsymbol{b}^{\boldsymbol{\nu}p/k})_{\boldsymbol{\nu}\in\mathcal{F}_k}$, which implies necessity of $\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$. Furthermore $(b_j^p)_{j\in\mathbb{N}}$ is a subsequence of $(\boldsymbol{b}^{\boldsymbol{\nu}p/k})_{\boldsymbol{\nu}\in\mathcal{F}_k}$ so that $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ is also a necessary condition.

On the other hand, since $\log(1 + x) \leq x$ for all $x \geq 0$ we have

$$\begin{aligned}\|(\boldsymbol{b}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k}\|_{\ell^{\frac{p}{k}}(\mathcal{F}_k)}^{\frac{p}{k}} &= \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} (\boldsymbol{b}^{\boldsymbol{\nu}})^{\frac{p}{k}} = \prod_{j\in\mathbb{N}}\left(1 + \sum_{\{l\in\mathbb{N}\,:\,l\geq k\}} b_j^{\frac{lp}{k}}\right) = \prod_{j\in\mathbb{N}}\left(1 + \frac{b_j^{\frac{pk}{k}}}{1 - b_j^{\frac{p}{k}}}\right)\\ &= \exp\left(\sum_{j\in\mathbb{N}} \log\left(1 + \frac{b_j^p}{1 - b_j^{\frac{p}{k}}}\right)\right) \leq \exp\left(\sum_{j\in\mathbb{N}} \frac{b_j^p}{1 - b_j^{\frac{p}{k}}}\right) \leq \exp\left(\frac{1}{1 - \|\boldsymbol{b}\|_{\ell^\infty}^{\frac{p}{k}}}\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})}^p\right). \end{aligned} \tag{B.1}$$

This proves the lemma for $\vartheta = 0$ and $R = 1$. To finish the proof it suffices to show that under the assumptions $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^\infty(\mathbb{N})} < 1$ it holds for any $\vartheta > 0$ that $(w_{\boldsymbol{\nu}} b^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$ where $w_{\boldsymbol{\nu}} = R^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(1 + \nu_j)^\vartheta$.

Fix $\vartheta > 0$ and $R \geq 1$. Let $\tilde\vartheta > \vartheta$ be so large that $2^{\tilde\vartheta-\vartheta} \geq R$. Then $R^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(1 + \nu_j)^\vartheta \leq \prod_{j\in\mathbb{N}}(1 + \nu_j)^{\tilde\vartheta}$. Let $\delta > 1$ be so large that $(1 + n)^{\tilde\vartheta} \leq \delta^n$ for all $n \in \mathbb{N}$, let $J \in \mathbb{N}$ be so large that $b_j < 1/(2\delta)$ for all $j > J$, and let $\kappa > 1$ be so small that $b_j\kappa < 1$ for all $j \leq J$. Define $\tilde{\boldsymbol{b}} \in \ell^p(\mathbb{N})$ by $\tilde b_j := \kappa b_j$ if $j \leq J$ and $\tilde b_j := \delta b_j$ otherwise. Then $\|\tilde{\boldsymbol{b}}\|_{\ell^\infty(\mathbb{N})} < 1$ and $\|\tilde{\boldsymbol{b}}\|_{\ell^p(\mathbb{N})} < \infty$. Moreover, with $C_0 := \sup_{n\in\mathbb{N}}(1 + n)^{\tilde\vartheta}/\kappa^n < \infty$, for all $\boldsymbol{\nu} \in \mathcal{F}$

$$w_{\boldsymbol{\nu}} = \prod_{j\in\mathbb{N}}(1 + \nu_j)^{\tilde\vartheta} \leq \prod_{j=1}^{J} C_0\kappa^{\nu_j}\prod_{i>J}\delta^{\nu_i} = C_0^J\prod_{j=1}^{J}\kappa^{\nu_j}\prod_{i>J}\delta^{\nu_i}.$$

Thus $\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}(w_{\boldsymbol{\nu}}\boldsymbol{b}^{\boldsymbol{\nu}})^{p/k} \leq C_0^{Jp/k}\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}(\tilde{\boldsymbol{b}}^{\boldsymbol{\nu}})^{p/k}$ which is finite by what we have shown above. $\square$

Lemma B.1. *Let $\rho > 1$ and fix $k \in \mathbb{N}$. Then there exists a constant $C_{k,\rho}$ depending on ρ, k, such that for all $\boldsymbol{\nu} \in \mathcal{F}\backslash\{\boldsymbol{0}\}$ the multiindex $k\boldsymbol{\nu} := (k\nu_j)_{j\in\mathbb{N}} \in \mathcal{F}_k \subseteq \mathcal{F}$ satisfies*

$$(2\pi)^{\frac{1-k}{2}}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k \leq \frac{|k\boldsymbol{\nu}|!}{(k\boldsymbol{\nu})!} \leq C_{k,\rho}^{|\operatorname{supp}\boldsymbol{\nu}|}\rho^{|\boldsymbol{\nu}|}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k. \tag{B.2}$$

Proof. We begin with the lower bound. Recall that $\sqrt{2\pi}n^{n+\frac{1}{2}}\,\mathrm{e}^{-n} \leq n! \leq n^{n+\frac{1}{2}}\,\mathrm{e}^{-n+1}$ for all $n \in \mathbb{N}$ by Stirling's approximation, see for example [31]. Thus

$$\begin{aligned}\frac{|k\boldsymbol{\nu}|!}{(k\boldsymbol{\nu})!} &\geq \frac{\sqrt{2\pi}(k|\boldsymbol{\nu}|)^{k|\boldsymbol{\nu}|+\frac{1}{2}}\exp(-k|\boldsymbol{\nu}|)}{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}(k\nu_j)^{k\nu_j+\frac{1}{2}}\exp(-k\nu_j + 1)}\\ &= \frac{\sqrt{2\pi}}{\exp(|\operatorname{supp}\boldsymbol{\nu}|)}\frac{k^{k|\boldsymbol{\nu}|+\frac{1}{2}}}{k^{k|\boldsymbol{\nu}|+\frac{|\operatorname{supp}\boldsymbol{\nu}|}{2}}}\frac{|\boldsymbol{\nu}|^{k|\boldsymbol{\nu}|+\frac{1}{2}}\exp(-k|\boldsymbol{\nu}|)}{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j^{k\nu_j+\frac{1}{2}}\exp(-k\nu_j)}\\ &= \frac{\sqrt{2\pi}k^{\frac{1-|\operatorname{supp}\boldsymbol{\nu}|}{2}}}{\exp(|\operatorname{supp}\boldsymbol{\nu}|)}\frac{(2\pi)^{-\frac{k}{2}}|\boldsymbol{\nu}|^{\frac{1-k}{2}}\left(\sqrt{2\pi}|\boldsymbol{\nu}|^{|\boldsymbol{\nu}|+\frac{1}{2}}\exp(-|\boldsymbol{\nu}|)\right)^k}{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\exp(-k)\nu_j^{\frac{1-k}{2}}\left(\nu_j^{\nu_j+\frac{1}{2}}\exp(-\nu_j + 1)\right)^k}\\ &\geq (2\pi)^{\frac{1-k}{2}}\left(\frac{\exp(k)}{k^{\frac{1}{2}}\,\mathrm{e}}\right)^{|\operatorname{supp}\boldsymbol{\nu}|}\left(\frac{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j}{|\boldsymbol{\nu}|}\right)^{\frac{k-1}{2}}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k. \end{aligned} \tag{B.3}$$

We claim that

$$f(\boldsymbol{\nu}) := \left(\frac{\exp(k-1)}{k^{\frac{1}{2}}}\right)^{|\operatorname{supp}\boldsymbol{\nu}|}\left(\frac{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j}{|\boldsymbol{\nu}|}\right)^{\frac{k-1}{2}} \geq 1, \tag{B.4}$$

for all $\mathbf{0} \neq \boldsymbol{\nu} \in \mathcal{F}$, which then gives the lower bound in (B.2). In order to see this we use induction on $n = |\boldsymbol{\nu}|$. The case $n = 1$ is trivial because $\exp(k-1)/k^{1/2} \geq 1$ for all $k \in \mathbb{N}$ and $\prod_{j\in\mathbb{N}}\nu_j = |\boldsymbol{\nu}|$ in this case. For the induction step let $\mathbf{e}_i = (\delta_{ij})_{j\in\mathbb{N}}$, fix an integer $n > 1$ and suppose that $f(\boldsymbol{\nu}) \geq 1$ for all $\boldsymbol{\nu} \in \mathcal{F}$ with $|\boldsymbol{\nu}| = n$. First assume $i \in \operatorname{supp}\boldsymbol{\nu}$ so that $|\operatorname{supp}\boldsymbol{\nu}| = |\operatorname{supp}(\boldsymbol{\nu}+\mathbf{e}_i)|$. Then

$$f(\boldsymbol{\nu}+\mathbf{e}_i) \geq f(\boldsymbol{\nu}) \qquad \Leftrightarrow \qquad \frac{(\nu_i+1)\prod_{j\neq i}\nu_j}{|\boldsymbol{\nu}|+1} \geq \frac{\prod_j \nu_j}{|\boldsymbol{\nu}|} \qquad \Leftrightarrow \qquad \frac{\nu_i+1}{|\boldsymbol{\nu}|+1} \geq \frac{\nu_i}{|\boldsymbol{\nu}|},$$

which is true so that $f(\boldsymbol{\nu}+\mathbf{e}_i) \geq f(\boldsymbol{\nu}) \geq 1$. Next let $i \notin \operatorname{supp}\boldsymbol{\nu}$. Then $\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j = \prod_{j\in\operatorname{supp}(\boldsymbol{\nu}+\mathbf{e}_i)}(\boldsymbol{\nu}+\mathbf{e}_i)_j$ and with $n = |\boldsymbol{\nu}|$

$$\frac{f(\boldsymbol{\nu}+\mathbf{e}_i)}{f(\boldsymbol{\nu})} = \frac{\exp(k-1)}{k^{\frac{1}{2}}}\left(\frac{n}{n+1}\right)^{\frac{k-1}{2}} \geq \frac{\exp(k-1)}{k^{\frac{1}{2}}}\left(\frac{1}{2}\right)^{\frac{k-1}{2}} := ng(k). \tag{B.5}$$

We have $g(1) = 1$. Moreover for $k \geq 1$

$$g'(k) = \frac{2^{-\frac{1+k}{2}}\exp(k-1)\left((2-\log(2))k-1\right)}{k^{\frac{3}{2}}} \geq 0,$$

which shows $g(k) \geq g(1) \geq 1$ for all $k \in \mathbb{N}$ and therefore $f(\boldsymbol{\nu}+\mathbf{e}_i) \geq f(\boldsymbol{\nu}) \geq 1$ by (B.5). This concludes the proof of the claim (B.4) which further implies the lower bound in (B.2).

For the upper bound, we fix $\mathbf{0} \neq \boldsymbol{\nu} \in \mathcal{F}$ and use again Stirling's inequalities to obtain

$$\begin{aligned}
\frac{|k\boldsymbol{\nu}|!}{(k\boldsymbol{\nu})!} &\leq \frac{(k|\boldsymbol{\nu}|)^{k|\boldsymbol{\nu}|+\frac{1}{2}}\exp(-k|\boldsymbol{\nu}|+1)}{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\sqrt{2\pi}(k\nu_j)^{k\nu_j+\frac{1}{2}}\exp(-k\nu_j)} \\
&= \frac{\mathrm{e}(2\pi)^{-\frac{k}{2}}|\boldsymbol{\nu}|^{\frac{1-k}{2}}\left(\sqrt{2\pi}|\boldsymbol{\nu}|^{|\boldsymbol{\nu}|+\frac{1}{2}}\exp(-|\boldsymbol{\nu}|)\right)^k}{\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\sqrt{2\pi}\exp(-k)\nu_j^{\frac{1-k}{2}}\left(\nu_j^{\nu_j+\frac{1}{2}}\exp(-\nu_j+1)\right)^k}\frac{k^{k|\boldsymbol{\nu}|+\frac{1}{2}}}{k^{k|\boldsymbol{\nu}|+\frac{1}{2}|\operatorname{supp}\boldsymbol{\nu}|}} \\
&\leq \frac{\mathrm{e}(2\pi)^{-\frac{k}{2}}|\boldsymbol{\nu}|^{\frac{1-k}{2}}}{\left(\sqrt{2\pi}\exp(-k)\right)^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j^{\frac{1-k}{2}}}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k \\
&\leq \mathrm{e}(2\pi)^{-\frac{k}{2}}\left(\frac{\exp(k)}{\sqrt{2\pi}}\right)^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j^{\frac{k-1}{2}}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k.
\end{aligned} \tag{B.6}$$

Since $\rho > 1$, there exists a constant $\tilde{C}_\rho$ such that $n^{(k-1)/2} \leq \tilde{C}_\rho\rho^n$ for all $n \in \mathbb{N}$. Thus $\prod_{j\in\operatorname{supp}\boldsymbol{\nu}}\nu_j^{(k-1)/2} \leq \tilde{C}_\rho^{|\operatorname{supp}\boldsymbol{\nu}|}\rho^{|\boldsymbol{\nu}|}$. The upper bound in (B.2) then follows *via* (B.6), for instance with $C_{k,\rho} := \tilde{C}_\rho\exp(k+1)(2\pi)^{-\frac{1}{2}}$. $\square$

Proof of Lemma 3.11. We start again with the case $\vartheta = 0$. W.l.o.g. we assume throughout $b_j > 0$ for all $j \in \mathbb{N}$.

Step 1. For $k = 1$, $p = 1$ we have

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\boldsymbol{b}^{\boldsymbol{\nu}} = \sum_{l\in\mathbb{N}_0}\left(\sum_{j\in\mathbb{N}}b_j\right)^l = \frac{1}{1-\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})}} < \infty, \tag{B.7}$$

which, due to $\mathcal{F}_1 = \mathcal{F}$, gives $(\boldsymbol{b}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_1} \in \ell^1(\mathcal{F}_1)$ iff $\|\boldsymbol{b}\|_{\ell^1} < 1$.

Step 2. We show that for any $p \in (0,1]$ and $k \in \mathbb{N}$ the conditions $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$ are necessary in order for $(\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$ to hold. It is clear that $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ must be satisfied, since $(b_j^p)_{j\in\mathbb{N}}$ is a subsequence of $((\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)^{p/k})_{\boldsymbol{\nu}\in\mathcal{F}_k}$. Next, it suffices to verify necessity of $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$ for $p = 1$. Let $k \in \mathbb{N}$. With Lemma B.1 it holds

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{1}{k}} \geq \sum_{\boldsymbol{\nu}\in\mathcal{F}}\left(\boldsymbol{b}^{k\boldsymbol{\nu}}\frac{|k\boldsymbol{\nu}|!}{k\boldsymbol{\nu}!}\right)^{\frac{1}{k}} \geq C\sum_{\boldsymbol{\nu}\in\mathcal{F}}\left(\boldsymbol{b}^{k\boldsymbol{\nu}}\left(\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^k\right)^{\frac{1}{k}} = C\sum_{\boldsymbol{\nu}\in\mathcal{F}}\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}.$$

According to (B.7), the last sum is finite iff $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$. This shows that for any value of $p \in (0,1]$ and $k \in \mathbb{N}$, the stated conditions are necessary.

Step 3. Fix an integer $k > 1$. We claim that for every $\boldsymbol{\nu} \in \mathcal{F}_k$, there exists $\boldsymbol{\mu} \in \mathcal{F}$ such that

$$(\mu_j \in \{ki \,:\, i \in \mathbb{N}_0\} \quad \text{and} \quad |\nu_j - \mu_j| < k)\ \forall j \in \mathbb{N}, \qquad \boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \leq k^{k|\operatorname{supp}\boldsymbol{\nu}|}\boldsymbol{b}^{\boldsymbol{\mu}}\frac{|\boldsymbol{\mu}|!}{\boldsymbol{\mu}!}. \tag{B.8}$$

To show this claim fix $\boldsymbol{\nu} \in \mathcal{F}_k$ and assume for the moment that there exists $j_0 \in \mathbb{N}$ such that $\nu_{j_0} \notin \{ki \,:\, i \in \mathbb{N}_0\}$ and $\nu_i \in \{ki \,:\, i \in \mathbb{N}_0\}$ for all $i \neq j_0$. By definition of $\mathcal{F}_k$, this implies $\nu_{j_0} > k$. Assume first that

$$b_{j_0}^{-1}\frac{\nu_{j_0}}{|\boldsymbol{\nu}|} \geq 1. \tag{B.9}$$

Then for $r \in \{1, \ldots, k-1\}$

$$b_{j_0}^{-1}\frac{\nu_{j_0} - r}{|\boldsymbol{\nu}| - r} = b_{j_0}^{-1}\frac{\nu_{j_0}}{|\boldsymbol{\nu}|}\frac{|\boldsymbol{\nu}|}{|\boldsymbol{\nu}| - r}\frac{\nu_{j_0} - r}{\nu_{j_0}} \geq \frac{\nu_{j_0} - r}{\nu_{j_0}} \geq \frac{1}{k}, \tag{B.10}$$

because $\nu_{j_0} > k$ and $r < k$. Define $\boldsymbol{\mu} = (\mu_{j_0})_{j\in\mathbb{N}} \in \mathcal{F}$ by

$$\mu_i := \begin{cases} \nu_i & \text{if} \quad i \neq j \\ \max\{nk \,:\, n \in \mathbb{N},\ nk \leq \nu_{j_0}\} & \text{if} \quad i = j \end{cases}$$

for all $i \in \mathbb{N}$. Then $|\nu_{j_0} - \mu_{j_0}| < k$ and by (B.10)

$$\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \leq \boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}k^{\nu_{j_0} - \mu_{j_0}}\prod_{r=0}^{\nu_{j_0} - \mu_{j_0} - 1} b_{j_0}^{-1}\frac{\nu_{j_0} - r}{|\boldsymbol{\nu}| - r} = k^{\nu_{j_0} - \mu_{j_0}}\boldsymbol{b}^{\boldsymbol{\mu}}\frac{|\boldsymbol{\mu}|!}{\boldsymbol{\mu}!} \leq k^k\boldsymbol{b}^{\boldsymbol{\mu}}\frac{|\boldsymbol{\mu}|!}{\boldsymbol{\mu}!},$$

which shows that $\boldsymbol{\mu}$ satisfies (B.8).

Next, suppose that (B.9) does not hold. Then $b_{j_0}|\boldsymbol{\nu}|/\nu_{j_0} > 1$ and therefore for $r \in \{1, \ldots, k-1\}$

$$b_{j_0}\frac{|\boldsymbol{\nu}| + r}{\nu_{j_0} + r} \geq b_{j_0}\frac{|\boldsymbol{\nu}|}{\nu_{j_0}}\frac{|\boldsymbol{\nu}| + r}{|\boldsymbol{\nu}|}\frac{\nu_{j_0}}{\nu_{j_0} + r} \geq \frac{\nu_{j_0}}{\nu_{j_0} + r} \geq \frac{1}{k}.$$

With $\boldsymbol{\mu} = (\mu_{j_0})_{j\in\mathbb{N}} \in \mathcal{F}$ defined by

$$\mu_i := \begin{cases} \nu_i & \text{if} \quad i \neq j \\ \min\{nk \,:\, n \in \mathbb{N},\ nk \geq \nu_{j_0}\} & \text{if} \quad i = j \end{cases}$$

we then have $|\mu_{j_0} - \nu_{j_0}| < k$ and similar as before

$$\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \leq \boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}k^{\mu_{j_0} - \nu_{j_0}}\prod_{r=0}^{\mu_{j_0} - \nu_{j_0} - 1} b_{j_0}\frac{|\boldsymbol{\nu}| + r}{\nu_{j_0} + r} = k^{\mu_{j_0} - \nu_{j_0}}\boldsymbol{b}^{\boldsymbol{\mu}}\frac{|\boldsymbol{\mu}|!}{\boldsymbol{\mu}!} \leq k^k\boldsymbol{b}^{\boldsymbol{\mu}}\frac{|\boldsymbol{\mu}|!}{\boldsymbol{\mu}!},$$

which again shows that $\boldsymbol{\mu}$ satisfies (B.8).

For the general case, where there might exist several indices j with $\nu_j \notin \{ki : i \in \mathbb{N}_0\}$, we repeat the above procedure for all such j to find $\boldsymbol{\mu}$ satisfying (B.8). This verifies the claim.

Step 4. In this step we prove that for $p = 1$ and $1 < k \in \mathbb{N}$, the conditions $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$ imply $(\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$.

If $\boldsymbol{\mu} \in \mathcal{F}$ with $\mu_j \in \{ki : i \in \mathbb{N}_0\}$ for all $j \in \mathbb{N}$ then

$$|\{\boldsymbol{\nu} \in \mathcal{F}_k : |\nu_j - \mu_j| < k,\ \forall\, j \in \mathbb{N}\}| \leq (2k-1)^{|\operatorname{supp}\boldsymbol{\mu}|}. \tag{B.11}$$

With $\boldsymbol{\mu}_{\boldsymbol{\nu}}$ denoting the multiindex constructed in Step 3 and satisfying (B.8), we get with (B.11)

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{1}{k}} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k} k^{|\operatorname{supp}\boldsymbol{\nu}|}\left(\boldsymbol{b}^{\boldsymbol{\mu}_{\boldsymbol{\nu}}}\frac{|\boldsymbol{\mu}_{\boldsymbol{\nu}}|!}{\boldsymbol{\mu}_{\boldsymbol{\nu}}!}\right)^{\frac{1}{k}} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}}(2k-1)^{|\operatorname{supp}\boldsymbol{\nu}|}k^{|\operatorname{supp}\boldsymbol{\nu}|}\left(\boldsymbol{b}^{k\boldsymbol{\nu}}\frac{|k\boldsymbol{\nu}|!}{(k\boldsymbol{\nu})!}\right)^{\frac{1}{k}}. \tag{B.12}$$

Now let $\rho > 1$ be so small that $\|\rho^{1/k}\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$, which is possible because $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$ by assumption. Then, employing Lemma B.1, the right-hand side of (B.12) is bounded by

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}}(k(2k-1))^{|\operatorname{supp}\boldsymbol{\nu}|}C_{k,\rho}^{\frac{|\operatorname{supp}\boldsymbol{\nu}|}{k}}\left(\rho^{\frac{1}{k}}\boldsymbol{b}\right)^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}}\tilde{C}_{k,\rho}^{|\operatorname{supp}\boldsymbol{\nu}|}\left(\rho^{\frac{1}{k}}\boldsymbol{b}\right)^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}, \tag{B.13}$$

where $\tilde{C}_{k,\rho} := k(2k-1)C_{k,\rho}^{1/k}$. Now let $J \in \mathbb{N}$ be so large that with $\tilde{b}_j := \rho^{\frac{1}{k}}b_j$ if $j \leq J$ and $\tilde{b}_j := \tilde{C}_{k,\rho}\rho^{\frac{1}{k}}b_j$ if $j > J$, it holds $\|\tilde{\boldsymbol{b}}\|_{\ell^1(\mathbb{N})} < 1$. With this choice, by (B.12) and (B.13) we arrive at

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{1}{k}} \leq \tilde{C}_{k,\rho}^{J-1}\sum_{\boldsymbol{\nu}\in\mathcal{F}}\tilde{\boldsymbol{b}}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!} < \infty, \tag{B.14}$$

where the last series is finite by (B.7) and because $\|\tilde{\boldsymbol{b}}\|_{\ell^1} < 1$. This concludes the proof for $k > 1$, $p = 1$.

Step 5. It remains to show that $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$ imply $(\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$ for $k \geq 1$ and $p \in (0,1)$. As shown in the proof of Theorem 7.2[10], with $p' := p/(1-p)$ one can construct sequences $\boldsymbol{\gamma} = (\gamma_j)_{j\in\mathbb{N}}$, $\boldsymbol{\delta} = (\delta_j)_{j\in\mathbb{N}}$ such that

$$\|\boldsymbol{\gamma}\|_{\ell^1(\mathbb{N})} < 1, \qquad \|\boldsymbol{\delta}\|_{\ell^\infty(\mathbb{N})} < 1, \qquad \|\boldsymbol{\delta}\|_{\ell^{p'}(\mathbb{N})} < \infty \qquad \text{and} \qquad b_j \leq \delta_j\gamma_j \quad \forall\, j \in \mathbb{N} \tag{B.15}$$

(essentially $\gamma_j \sim b_j^p$ and $\delta_j \sim b_j^{1-p}$). We get

$$\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{b}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{p}{k}} \leq \sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{\gamma}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{p}{k}}\boldsymbol{\delta}^{\frac{p}{k}} \leq \left(\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\left(\boldsymbol{\gamma}^{\boldsymbol{\nu}}\frac{|\boldsymbol{\nu}|!}{\boldsymbol{\nu}!}\right)^{\frac{1}{k}}\right)^{p}\left(\sum_{\boldsymbol{\nu}\in\mathcal{F}_k}\boldsymbol{\delta}^{\boldsymbol{\nu}\frac{p}{k(1-p)}}\right)^{1-p}.$$

Using (B.15), the first sum is finite by the statement of the current Lemma for $p = 1$ (already shown in Step 4), and the second sum is finite since $(\boldsymbol{\delta}^{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p'/k}(\mathcal{F}_k)$ according to Lemma 3.10. This proves $(\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}_k} \in \ell^{p/k}(\mathcal{F}_k)$.

Step 6. We have shown the lemma for $\vartheta = 0$. In order to finish the proof, it suffices to verify that under the assumptions $\|\boldsymbol{b}\|_{\ell^p(\mathbb{N})} < \infty$ and $\|\boldsymbol{b}\|_{\ell^1(\mathbb{N})} < 1$, for any fixed $k \in \mathbb{N}$ and $\vartheta > 0$ with $w_{\boldsymbol{\nu}} = R^{|\operatorname{supp}\boldsymbol{\nu}|}\prod_{j\in\mathbb{N}}(1+\nu_j)^{\vartheta}$ it holds $(w_{\boldsymbol{\nu}}\boldsymbol{b}^{\boldsymbol{\nu}}|\boldsymbol{\nu}|!/\boldsymbol{\nu}!)_{\boldsymbol{\nu}\in\mathcal{F}} \in \ell^{p/k}(\mathcal{F}_k)$. This can be shown by the same argument used at the end of the proof of Lemma 3.10. $\square$

Acknowledgements. Supported in part by SNSF Grant No. 159940. JZ is supported by SNSF fellowship 184530. The results of the paper were obtained as part of the Ph.D. studies of JZ at the Seminar for Applied Mathematics at ETH Zürich .

References

[1] M. Bachmayr, A. Cohen and G. Migliorati, Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients. *ESAIM: M2AN* **51** (2017) 321–339.

[2] V. Barthelmann, E. Novak and K. Ritter, High dimensional polynomial interpolation on sparse grids. Multivariate polynomial interpolation. *Adv. Comput. Math.* **12** (2000) 273–288.

[3] J. Beck, R. Tempone, F. Nobile and L. Tamellini, On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods. *Math. Models Methods Appl. Sci.* **22** (2012) 1250023, 33.

[4] H.-J. Bungartz and M. Griebel, Sparse grids. *Acta Numer.* **13** (2004) 147–269.

[5] J.-P. Calvi and P.V. Manh, Lagrange interpolation at real projections of Leja sequences for the unit disk. *Proc. Amer. Math. Soc.* **140** (2012) 4271–4284.

[6] J.-P. Calvi and M. Phung Van, On the Lebesgue constant of Leja sequences for the unit disk and its applications to multivariate interpolation. *J. Approx. Theory* **163** (2011) 608–622.

[7] S.B. Chae, Holomorphy and Calculus in Normed Spaces. With an appendix by Angus E. Taylor. In: Vol. 92 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York (1985).

[8] M.A. Chkifa, On the Lebesgue constant of Leja sequences for the complex unit disk and of their real projection. *J. Approx. Theory* **166** (2013) 176–200.

[9] A. Chkifa, A. Cohen, R. DeVore and C. Schwab, Adaptive algorithms for sparse polynomial approximation of parametric and stochastic elliptic pdes. *ESAIM: M2AN* **47** (2013) 253–280.

[10] A. Cohen, R. DeVore and C. Schwab, Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.* **10** (2010) 615–646.

[11] A. Cohen, R. Devore and C. Schwab, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Anal. Appl. (Singap.)* **9** (2011) 11–47.

[12] A. Cohen, A. Chkifa and C. Schwab, Breaking the curse of dimensionality in sparse polynomial approximation of parametric pdes. *J. Math. Pures Appl.* **103** (2015) 400–428.

[13] A. Cohen, C. Schwab and J. Zech, Shape holomorphy of the stationary Navier–Stokes equations. *SIAM J. Math. Anal.* **50** (2018) 1720–1752.

[14] J. Dick, F.Y. Kuo, Q.T.L. Gia, D. Nuyens and C. Schwab, Higher order QMC Petrov–Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.* **52** (2014) 2676–2702.

[15] J. Dick, F.Y. Kuo and I.H. Sloan, High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer.* **22** (2013) 133–288.

[16] R.N. Gantner and C. Schwab, Computational higher order quasi-Monte Carlo integration. In: Vol. 163 of *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014* (2016) 271–288.

[17] T. Gerstner and M. Griebel, Numerical integration using sparse grids. *Numer. Algorithms* **18** (1998) 209–232.

[18] T. Gerstner and M. Griebel, Dimension-adaptive tensor-product quadrature. *Computing* **71** (2003) 65–87.

[19] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numer. Math.* **131** (2015) 329–368.

[20] M. Griebel and J. Oettershagen, On tensor product approximation of analytic functions. *J. Approx. Theory* **207** (2016) 348–379.

[21] A.-L. Haji-Ali, F. Nobile, L. Tamellini and R. Tempone, Multi-index stochastic collocation for random PDEs. *Comput. Methods Appl. Mech. Eng.* **306** (2016) 95–122.

[22] A.-L. Haji-Ali, H. Harbrecht, M. Peters and M. Siebenmorgen, Novel results for the anisotropic sparse grid quadrature. *J. Complexity* **47** (2018) 62–85.

[23] H. Harbrecht, M. Peters and M. Siebenmorgen, Analysis of the domain mapping method for elliptic diffusion problems on random domains. *Numer. Math.* **134** (2016) 823–856.

[24] M. Hervé, Analyticity in Infinite-dimensional Spaces. In: Vol. 10 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin (1989).

[25] R. Hiptmair, L. Scarabosio, C. Schillings and C. Schwab, Large deformation shape uncertainty quantification in acoustic scattering. *Adv. Comput. Math.* **44** (2018) 1475–1518.

[26] V.H. Hoang and C. Schwab, N-term Wiener chaos approximation rate for elliptic PDEs with lognormal Gaussian random inputs. *Math. Models Methods Appl. Sci.* **24** (2014) 797–826.

[27] C. Jerez-Hanckes, C. Schwab and J. Zech, Electromagnetic wave scattering by random surfaces: shape holomorphy. *Math. Mod. Meth. Appl. Sci.* **27** (2017) 2229–2259.

[28] F.Y. Kuo and D. Nuyens, Application of quasi-Monte Carlo methods to elliptic PDEs with random diffusion coefficients: a survey of analysis and implementation. *Found. Comput. Math.* **16** (2016) 1631–1696.

[29] F. Kuo, R. Scheichl, C. Schwab, I. Sloan and E. Ullmann, Multilevel Quasi-Monte Carlo Methods for Lognormal Diffusion Problems. Technical Report 2015-22, Seminar for Applied Mathematics, ETH Zürich (2016).

[30] F. Nobile, R. Tempone and C.G. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.* **46** (2008) 2411–2442.

[31] H. Robbins, A remark on Stirling's formula. *Amer. Math. Monthly* **62** (1955) 26–29.

[32] C. Schillings and C. Schwab, Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. *Inverse Prob.* **29** (2013) 065011.

[33] C. Schwab and C. Gittelson, Sparse tensor discretizations of high-dimensional parametric and stochastic pdes. *Acta Numer.* **20** (2011) 291–467.

[34] C. Schwab and A.M. Stuart, Sparse deterministic approximation of Bayesian inverse problems. *Inverse Prob.* **28** (2012) 045003.

[35] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions. *Sov. Math. Dokl.* **4** (1963) 240–243.

[36] R.A. Todor and C. Schwab, Convergence rates of sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.* **44** (2007) 232–261.

[37] J. Zech, *Sparse-grid approximation of high-dimensional parametric PDEs*. Ph.D. thesis. Dissertation 25683, ETH Zürich (2018). DOI: 10.3929/ethz-b-000340651.

[38] J. Zech, D. Dũng and C. Schwab, Multilevel approximation of parametric and stochastic PDEs. *Math. Models Methods Appl. Sci.* **29** (2019) 1753–1817.