IMA Journal of Numerical Analysis (2019) **00**, 1–37
doi:10.1093/imanum/dry080

Sampling inequalities for anisotropic tensor product grids

Christian Rieger*
Institute for Numerical Simulation, University of Bonn, Bonn 53115, Germany and Department of Mathematics, RWTH Aachen University, Schinkelstr. 2, Aachen D-52062, Germany
*Corresponding author: rieger@ins.uni-bonn.de

and

Holger Wendland
Department of Mathematics, University of Bayreuth, Bayreuth 95440, Germany
holger.wendland@uni-bayreuth.de

[Received on 12 April 2018; revised on 5 October 2018]

We derive sampling inequalities for discrete point sets that are of anisotropic tensor product form. Such sampling inequalities can be used to prove convergence for arbitrary stable reconstruction processes. As usual in the context of high-dimensional problems, our sampling inequalities are expressed in terms of the number of data sites, i.e., the number of points in the sparse grid. To this end, new bounds on specific monotone sets and on the number of points in an anisotropic sparse grid are derived.

Keywords: high-dimensional approximation; anisotropic sparse grids; sampling inequalities; kernel-based reconstructions.

1. Introduction

In high-dimensional approximation problems the so-called curse of dimensionality is the main obstruction to the simple upscaling of numerical algorithms that have proven to work well in low dimensions. In fact, if there is no additional structure in a high-dimensional problem then basically no method can yield satisfactory approximation qualities. In this paper, we focus on anisotropic smoothness as the main additional structure. This basically means that the functions we consider do not possess equal numbers of weak derivatives in each coordinate direction. Such problems naturally arise, for example, when dealing with simultaneous space-time discretizations of parabolic problems, see for example Griebel *et al.* (2006) and Griebel & Oeltz (2007). Other applications comprise the solution of stochastic partial differential equations, see for example Nobile *et al.* (2008a,b).

We will focus on situations in which this anisotropy is known *a priori*, as it is the case in the above-mentioned applications and hence can be incorporated into all considerations. This is in contrast to applications where the anisotropy is not known and hence dimension adaptive methods have to be developed. Such an *a priori* knowledge is in particular important in this paper, since we are not designing any new algorithm, but focus on the deterministic *a priori* error analysis by means of *sampling inequalities*.

Such sampling inequalities have been introduced in Wendland & Rieger (2005), and successively been investigated in Arcangéli *et al.* (2007) and Madych (2006) for low-dimensional domains. In Rieger & Wendland (2017) first results for high-dimensional domains in the context of classical sparse grids have been derived. It is the goal of this paper to generalize the results of Rieger & Wendland (2017) to the situation of anisotropic and more general sparse grids.

The advantage of such sampling inequalities is that they provide means to easily derive error estimates for every *stable* reconstruction method. This includes in particular kernel-based reconstructions as they are, for example, popular in radial basis function approximation, see Wendland (2005), or machine learning and support vector machines, see Steinwart & Christmann (2008). Further details can also be found in the above mentioned papers.

2. The problem: sampling inequalities for anisotropic sparse grids

We will now describe the sampling inequalities that we will prove in this paper in more detail. To this end, we first have to introduce some notation. As usual, let $\mathbb{N} := \{1, 2, \dots\}$ denote the natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We also use the notation $\mathbb{R}^d_+ = \{\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d : x_j > 0, 1 \leqslant j \leqslant d\}$ and $\mathbb{R}^d_{\geqslant 0} = \{\boldsymbol{x} \in \mathbb{R}^d : x_j \geqslant 0, 1 \leqslant j \leqslant d\}$. For $\boldsymbol{x} \in \mathbb{R}^d$ we use $\|\boldsymbol{x}\|_1 := \sum_{j=1}^d |x_j|$. For $\boldsymbol{\nu} \in \mathbb{N}_0^d$ we also use the alternative notation $|\boldsymbol{\nu}| := \|\boldsymbol{\nu}\|_1$ and $\boldsymbol{\nu}! := \prod_{j=1}^d \nu_j!$. Furthermore, for $\boldsymbol{\nu}, \boldsymbol{\mu} \in \mathbb{N}_0^d$, we write $\boldsymbol{\nu} \leqslant \boldsymbol{\mu}$ if and only if $\nu_j \leqslant \mu_j$ for $1 \leqslant j \leqslant d$ and

$$\binom{\boldsymbol{\nu}}{\boldsymbol{\mu}} := \prod_{j=1}^d \binom{\nu_j}{\mu_j}.$$

In the same way we refer to multivariate monomials and multivariate derivatives by $\boldsymbol{x}^{\boldsymbol{\nu}}$ and $D^{\boldsymbol{\nu}}$, respectively.

Let $\Omega^{(j)} \subseteq \mathbb{R}^{n_j}$, $1 \leqslant j \leqslant d$, be a bounded domain with a sufficiently smooth boundary. We are particularly interested in the cases where $\Omega^{(j)} = [-1, 1] \subseteq \mathbb{R}$ is an interval or where n_j is small, i.e., $n_j = 2, 3$. Then, the spatial domain we are interested in is simply the Cartesian product

$$\Omega^{\otimes} := \Omega^{(1)} \times \Omega^{(2)} \times \cdots \times \Omega^{(d)}.$$

To describe the functions on $\Omega^{\otimes}$ we are interested in let $\boldsymbol{\beta} \in \mathbb{N}_0^d$ and $1 \leqslant p < \infty$ be given. Then we are interested in estimates on functions from the anisotropic Sobolev space

$$W_p^{\boldsymbol{\beta}}(\Omega^{\otimes}) = \{f \in L_p(\Omega^{\otimes}) : D^{\boldsymbol{\alpha}} f \in L_p(\Omega^{\otimes}) \text{ for } \boldsymbol{\alpha} \in \mathbb{N}_0^d \text{ with } \boldsymbol{\alpha} \leqslant \boldsymbol{\beta}\},$$

equipped with the norm

$$\|f\|^p_{W_p^{\boldsymbol{\beta}}(\Omega^{\otimes})} := \sum_{\boldsymbol{\alpha} \leqslant \boldsymbol{\beta}} \|D^{\boldsymbol{\alpha}} f\|^p_{L_p(\Omega^{\otimes})}.$$

In several situations it will not be necessary to restrict ourselves to integer orders. In the case of $p = \infty$ we will consider the classical smoothness spaces of bounded functions

$$C_b^{\boldsymbol{\beta}}(\Omega^{\otimes}) = \{f \in C_b(\Omega^{\otimes}) : D^{\boldsymbol{\alpha}} f \in C_b(\Omega^{\otimes}) \text{ for } \boldsymbol{\alpha} \in \mathbb{N}_0^d \text{ with } \boldsymbol{\alpha} \leqslant \boldsymbol{\beta}\},$$

equipped with the standard maximum norms

$$\|f\|_{W_\infty^{\boldsymbol{\beta}}(\Omega^{\otimes})} = \max_{\boldsymbol{\alpha} \leqslant \boldsymbol{\beta}} \sup_{\boldsymbol{x} \in \Omega^{\otimes}} |D^{\boldsymbol{\alpha}} f(\boldsymbol{x})|.$$

Both spaces, $W_p^{\boldsymbol{\beta}}(\Omega^{\otimes})$ and $C_b^{\boldsymbol{\beta}}(\Omega^{\otimes})$, can be considered as tensor product spaces generated by the corresponding univariate spaces, see for example Aubin (2000), Cheney & Light (1985) and Schreiber (2000). We are particularly interested in giving estimates on $f \in W_p^{\boldsymbol{\beta}}(\Omega^{\otimes})$ or $C_b^{\boldsymbol{\beta}}(\Omega)$ provided that f is only known at a discrete subset of $\Omega^{\otimes}$. To make this possible we need to assume, in the first case, that $\beta_j \geqslant n_j/p$ for $1 \leqslant j \leqslant d$ such that the Sobolev embedding theorem allows us to form point evaluations. Moreover, we have to describe the discrete point set on which f is assumed to be known. Here, motivated by applications from anisotropic sparse grids, we use the following construction. For each $1 \leqslant j \leqslant d$ let sequences of discrete, nested sets

$$\emptyset = \Xi_0^{(j)} \subseteq \Xi_1^{(j)} \subseteq \Xi_2^{(j)} \subseteq \cdots \subseteq \Omega^{(j)}$$

be given. With these univariate sets we can create sequences of multivariate sets by simply forming the Cartesian products

$$\Xi_{\boldsymbol{\lambda}}^{\otimes} = \Xi_{\lambda_1}^{(1)} \times \cdots \times \Xi_{\lambda_d}^{(d)}, \qquad \boldsymbol{\lambda} \in \mathbb{N}_0^d.$$

The union of several of these multivariate point sets then form what we will call an *anisotropic sparse grid*. To be more precise, following Nobile *et al.* (2008a), we define for $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $\ell \in \mathbb{N}$ the index sets

$$\mathscr{I}_{\boldsymbol{\omega}}(\ell, d) := \left\{ \boldsymbol{\nu} \in \mathbb{N}^d : \sum_{j=1}^{d} (\nu_j - 1)\omega_j \leqslant \ell \min_{1 \leqslant j \leqslant d} \omega_j \right\}, \tag{2.1}$$

$$\mathscr{J}_{\boldsymbol{\omega}}(\ell, d) := \mathscr{I}_{\boldsymbol{\omega}}(\ell, d) \setminus \mathscr{I}_{\boldsymbol{\omega}}\left(\ell - \frac{\|\boldsymbol{\omega}\|_1}{\omega_{\min}}, d\right), \tag{2.2}$$

where the first one is a special case of a *monotone index set*. With this we are able to define an *anisotropic sparse grid point set* as

$$\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d) := \bigcup_{\boldsymbol{\lambda} \in \mathscr{J}_{\boldsymbol{\omega}}(\ell, d)} \Xi_{\boldsymbol{\lambda}}^{\otimes} = \bigcup_{\boldsymbol{\lambda} \in \mathscr{J}_{\boldsymbol{\omega}}(\ell, d)} \Xi_{\lambda_1}^{(1)} \times \cdots \times \Xi_{\lambda_d}^{(d)}. \tag{2.3}$$

With this notation at hand we can present one of our main results that will follow from the theory we will develop in this paper. The precise conditions of the following theorem will be given later, see Theorem 4.9. Nonetheless, the theorem already shows a typical connection between the smoothness vector $\boldsymbol{\beta}$ defining the Sobolev space and the weight vector $\boldsymbol{\omega}$ defining the anisotropic sparse grid.

THEOREM 2.1 Let $\boldsymbol{\beta} = \boldsymbol{\omega} \in \mathbb{N}^d$. Under certain assumptions on the grid points $\Xi_k^{(j)}$ there is a constant $C > 0$ such that

$$\|f\|_{L_\infty(\Omega^{\otimes})} \leqslant C(\log N)^{(2d-1)+(d-1)\omega_{\min}} N^{-\omega_{\min}} \|f\|_{W_\infty^{\boldsymbol{\beta}}(\Omega^{\otimes})} + C(\log N)^d \|f\|_{\ell_\infty(\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d))}. \qquad f \in C_b^{\boldsymbol{\beta}}(\Omega^{\otimes}).$$

Here, $\omega_{\min} = \min_{1 \leqslant j \leqslant d} \omega_j$ and N denotes the number of points in the sparse grid $\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d)$. Furthermore, $\|f\|_{\ell_\infty(\Xi)} := \max_{\boldsymbol{x} \in \Xi} |f(\boldsymbol{x})|$ denotes the discrete ℓ_∞-norm.

The above theorem generalizes the findings for isotropic sparse grids from Rieger & Wendland (2017). The result shows that such sampling inequalities also hold for the specific anisotropic sparse grids mentioned above. As usual in this context we assume that the weight vector $\boldsymbol{\omega}$ is a given, fixed vector, that is, for example, determined by the smoothness of the functions involved. Our results then concentrate on finding bounds for varying $\ell \in \mathbb{N}$. As for ℓ tending to infinity the corresponding anisotropic sparse grids look more and more like isotropic sparse grids, the sampling inequalities should reflect this, as it is the case in the above theorem. However, the techniques developed in this paper will also allow us to derive more sophisticated sampling inequalities and hence error estimates when the weight vector $\boldsymbol{\omega}$ is not fixed *a priori*. The situation we have in mind is where the vector $\boldsymbol{\omega}$ can (in a certain range) be optimized. Of course, from an approximation rate perspective, and as large as possible, $\boldsymbol{\omega}$ is beneficial. However, the smoothness norms are typically growing in $\boldsymbol{\omega}$. Hence, in order to get the best possible estimates, these two influences have to be balanced. This will be the subject of a subsequent paper.

This paper is organized as follows. In the next section we describe and collect necessary results on monotone sets, including new results or new proofs for specific integrals over simplices. This is required for giving bounds on the computational cost, which means here particularly relating the number of points in the sparse grid to the number of indices in the monotone index set. We employ the approach of Griebel & Oettershagen (2016) and Nobile *et al.* (2008a) when it comes to estimating the number of indices in the index set. However, in contrast to these sources, we give new proofs based upon the Faà di Bruno formula for derivatives of concatenated multivariate functions and the Hermite–Gnocci theorem and the Peano kernel representation for divided differences. This new approach does not only allow us to derive the same bounds, but also allows us to derive explicit formulas for several of the relevant integrals and sums, which are interesting on their own. Moreover, in contrast to Griebel & Oettershagen (2016), we measure the computational cost, not in terms of the number of indices in $\mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$, but in the actual number of points in the sparse grid $\Xi_{\boldsymbol{\omega}}(\ell, d)$.

The fourth section is devoted to sampling inequalities. We start with collecting necessary results on quasi-interpolation and Smolyak's algorithm. We then give the generic idea of how to derive sampling inequalities in rather general situations. Then the main results of this paper are derived under certain assumptions on the univariate data sets, the smoothness of the target function and the weight vector of the index set. Here, we distinguish, as is common in this context, between oversampling and nonoversampling. In particular, we will also give a general version of Theorem 2.1 and its proof, and discuss several examples.

3. Estimates on the cost

As mentioned above we understand as cost the cardinality of the anisotropic grid $\Xi^{\otimes}_{\boldsymbol{\omega}}(\ell, d)$ from (2.3). It is the goal of this section to give both lower and upper bounds on the cardinality of $\Xi^{\otimes}_{\boldsymbol{\omega}}(\ell, d)$. We will achieve this in two steps, where the first one is finding bounds for the cardinality of the index sets, in particular of $\mathscr{I}_{\boldsymbol{\omega}}(\ell, D)$ from (2.1). To a certain extend this has already been done in Griebel & Oettershagen (2016), Nobile *et al.* (2008a) and Chkifa *et al.* (2014). However, as we employ a different proof technique, we can, in many situations derive, even explicit representations instead of only bounds.

3.1 *Bounds on the cardinality of index sets*

3.1.1 *Monotone sets and basic properties.* Monotone index sets play an important role when it comes to describing anisotropic behavior. Unfortunately, the notation varies throughout the different publications. Hence, we will now collect all relevant notation and properties of monotone index sets. We will also describe specific monotone index sets, which are defined by a hyper-plane and form anisotropic simplices.

For the definitions and basic facts we mainly follow Chkifa *et al.* (2014, Section 2.1) and Nobile *et al.* (2008a, Equations 2.7, 2.10 and pp. 2425) when it comes to the specific index sets $\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$ and $\mathscr{J}_{\boldsymbol{\omega}}(\ell,d)$.

DEFINITION 3.1 An index set $\Lambda \subseteq \mathbb{N}^d$ is called *monotone* if $\boldsymbol{\nu} \in \Lambda$ and $\boldsymbol{\mu} \in \mathbb{N}^d$ with $\boldsymbol{\mu} \leqslant \boldsymbol{\nu}$ implies $\boldsymbol{\mu} \in \Lambda$.

For $1 \leqslant j \leqslant d$ we denote, as usual, the jth unit vector by $\boldsymbol{e}_j \in \mathbb{N}_0^d$, i.e., $(\boldsymbol{e}_j)_k = \delta_{jk}$. Then Λ is obviously a monotone set if and only if each $\boldsymbol{\nu} \in \Lambda$ with $\nu_j \geqslant 2$ for any $1 \leqslant j \leqslant d$ satisfies $\boldsymbol{\nu} - \boldsymbol{e}_j \in \Lambda$.

As mentioned above the set $\mathscr{I}_{\boldsymbol{\omega}}(\ell,\mathscr{D})$ is a monotone set (or admissible in the sense of Conrad & Marzouk, 2013, Definition 2.4). To see this let $\boldsymbol{\eta} \in \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$ with $\eta_d \geqslant 2$ for a $1 \leqslant j \leqslant d$. Then $\boldsymbol{\eta} - \boldsymbol{e}_j$ satisfies

$$\sum_{k=1}^{\mathrm{d}} (\boldsymbol{\eta} - \boldsymbol{e}_j - \mathbf{1})_k \omega_k \leqslant \sum_{k=1}^{d} (\eta_k - 1)\omega_k \leqslant \ell\omega_{\min}.$$

Hence, we also have $\boldsymbol{\eta} - \boldsymbol{e}_j \in \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$. In the rest of the paper we will particularly be interested in monotone index sets, which are related to anisotropic simplices. To this end let $\boldsymbol{\omega} \in \mathbb{R}_+^d$ be a weight vector and $T \in (0,\infty)$ a threshold. Following Griebel & Oettershagen (2016, Equations 2.2, 2.3 and 2.4) we define

$$\begin{aligned}
\mathscr{E}_{\boldsymbol{\omega}}(T) &:= \left\{ \boldsymbol{x} \in \mathbb{R}_{\geqslant 0}^d : \ \boldsymbol{\omega} \cdot \boldsymbol{x} = \sum_{j=1}^{d} \omega_j x_j \leqslant T \right\}, \\
\mathscr{D}_{\boldsymbol{\omega}}(T) &:= \left\{ \boldsymbol{k} \in \mathbb{N}_0^d : \ \boldsymbol{\omega} \cdot \boldsymbol{k} = \sum_{j=1}^{d} \omega_j k_j \leqslant T \right\} = \mathscr{E}_{\boldsymbol{\omega}}(T) \cap \mathbb{N}_0^d, \\
\left[\mathscr{D}_{\boldsymbol{\omega}}\right](T) &:= \bigcup_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)} \bigtimes_{j=1}^{d} [k_j, k_j+1) =: \bigcup_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)} [\boldsymbol{k}, \boldsymbol{k}+\mathbf{1}),
\end{aligned}$$

where $\mathbf{1} = (1,\ldots,1)^T \in \mathbb{R}^d$ is the constant vector having all entries as 1. We remark that these sets satisfy the following inclusions, see Griebel & Oettershagen (2016, Proof of Theorem. 2.6),

$$\mathscr{E}_{\boldsymbol{\omega}}(T) \subseteq \left[\mathscr{D}_{\boldsymbol{\omega}}\right](T) \subseteq \mathscr{E}_{\boldsymbol{\omega}}\left(T + \|\boldsymbol{\omega}\|_1\right). \tag{3.1}$$

As $[\mathscr{D}_{\boldsymbol{\omega}}](T)$ is composed of disjoint cubes of volume one the number $\#\mathscr{D}_{\boldsymbol{\omega}}(T)$ of elements in $\mathscr{D}_{\boldsymbol{\omega}}(T)$ is given by

$$\#\mathscr{D}_{\boldsymbol{\omega}}(T) = \mathrm{vol}\left(\mathscr{D}_{\boldsymbol{\omega}}(T)\right). \tag{3.2}$$

Finally, the set $\mathscr{D}_{\boldsymbol{\omega}}(T)$ is closely connected to our index set $\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$ from (2.1) as we obviously have

$$\mathscr{I}_{\boldsymbol{\omega}}(\ell,d) = \left\{ \boldsymbol{\nu} \in \mathbb{N}^d : \ (\boldsymbol{\nu} - \mathbf{1}) \cdot \boldsymbol{\omega} \leqslant \ell\omega_{\min} \right\} = \mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min}) + \{\mathbf{1}\}. \tag{3.3}$$

We will also need the complements of these sets. Here, however, it is important to note the base set within which the complements are formed. To be more precise we have

$$\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)^{\complement} = \mathbb{N}^d \setminus \mathscr{I}_{\boldsymbol{\omega}}(\ell,d), \qquad \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement} = \mathbb{N}_0^d \setminus \mathscr{D}_{\boldsymbol{\omega}}(T),$$
$$[\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement} = \mathbb{R}^d_{\geqslant 0} \setminus [\mathscr{D}_{\boldsymbol{\omega}}](T), \qquad \mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement} = \mathbb{R}^d_{\geqslant 0} \setminus \mathscr{E}_{\boldsymbol{\omega}}(T).$$

Using different base sets has, of course, certain consequences. For example, we have on the one hand

$$\begin{aligned} \mathscr{I}_{\boldsymbol{\omega}}(\ell,d) &= \left\{\boldsymbol{\lambda} \in \mathbb{N}^d : \boldsymbol{\lambda}\cdot\boldsymbol{\omega} \leqslant \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right\} \\ &\subseteq \left\{\boldsymbol{\lambda} \in \mathbb{N}_0^d : \boldsymbol{\lambda}\cdot\boldsymbol{\omega} \leqslant \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right\} \\ &= \mathscr{D}_{\boldsymbol{\omega}}\left(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right), \end{aligned}$$

but we also have on the other hand

$$\begin{aligned} \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)^{\complement} &= \left\{\boldsymbol{\lambda} \in \mathbb{N}^d : \boldsymbol{\lambda}\cdot\boldsymbol{\omega} > \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right\} \\ &\subseteq \left\{\boldsymbol{\lambda} \in \mathbb{N}_0^d : \boldsymbol{\lambda}\cdot\boldsymbol{\omega} > \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right\} \\ &= \mathscr{D}_{\boldsymbol{\omega}}\left(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right)^{\complement}. \end{aligned} \tag{3.4}$$

3.1.2 *Divided differences and integrals over simplices.* Next we generalize the calculations from Griebel & Oettershagen (2016) in order to have more general results with simpler proofs and in order to improve the readability of the manuscript. Given a certain, continuous function $f : \mathbb{R} \to \mathbb{R}$, we are interested in evaluating sums of the form

$$F(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}) := \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} f(\boldsymbol{\gamma}\cdot\boldsymbol{k}). \tag{3.5}$$

As in Griebel & Oettershagen (2016) we are mainly interested in the case $f = \exp$, but will derive a more general formula. Furthermore, we are also interested in such sums if the sum is actually taken over the complement $\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement} = \mathbb{N}_0^d \setminus \mathscr{D}_{\boldsymbol{\omega}}(T)$ of $\mathscr{D}_{\boldsymbol{\omega}}(T)$. For the proofs this will make no significant difference.

In order to evaluate sums of the form (3.5) we will employ a specific form of the Faá di Bruno formula, which gives an intrinsic relation between such sums and certain multivariate integrals over the simplices $[\mathscr{D}_{\boldsymbol{\omega}}](T)$. Then we will use the Peano kernel representation of divided differences to further reduce these integrals to univariate integrals, which then can either be computed explicitly or bounded.

To introduce the Faá di Bruno formula we need to recall the notion of a partition of a set.

Definition 3.2 Let M be a set. We call a family $\pi(M)$ of subsets of M a partition if

- $\emptyset \notin \pi(M)$,
- the elements of $\pi(M)$ are pairwise disjoint, i.e., $A, B \in \pi(M)$ implies $A \cap B = \emptyset$,
- the union of all elements covers M, i.e., $\bigcup_{A\in\pi(M)} A = M$.

We denote the family of all partitions of a (finite) set M by $\mathscr{P}_{\text{part}}(M)$.

For a partition $\pi = \pi(M)$ we denote the number of sets in π by $|\pi|$ and for $A \in \pi$ we denote the number of elements in A also by $|A|$. Moreover, if π is a partition of $M = \{1, \ldots, d\}$, then any set $A \in \pi$ represents a multi-index $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ by defining $\alpha_j = 1$ if $j \in A$ and $\alpha_j = 0$ if $j \notin A$. With this, we introduce the notation

$$D^A f := D^{\boldsymbol{\alpha}} f = \frac{\partial^{|A|} f}{\prod_{j\in A} \partial x_j}.$$

With this notation the multivariate chain rule or Faà di Bruno's formula (Hardy, 2006, Proposition 1) for $g \in C^d(\mathbb{R}^d)$ and $f \in C^d(\mathbb{R})$ can be written as

$$\frac{\partial^d}{\partial x_1 \cdots \partial x_d}(f \circ g)(\boldsymbol{x}) = \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A\in\pi} D^A g(\boldsymbol{x}). \tag{3.6}$$

This allows us to derive our first representation for specific sums of the form (3.5) that we state for sums over $\mathscr{D}_{\boldsymbol{\omega}}(T)$ and sums over $\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}$, as well.

Theorem 3.3 Let $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $T > 0$. For $g \in C^d(\mathbb{R}^d)$ and $f \in C^d(\mathbb{R})$ we have

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{d+j} f(g(\boldsymbol{k}+\boldsymbol{e})) = \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A\in\pi} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \tag{3.7}$$

and

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{d+j} f(g(\boldsymbol{k}+\boldsymbol{e})) = \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement}} \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A\in\pi} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}. \tag{3.8}$$

Proof. We only prove (3.7) as (3.8) follows in the same fashion using $[\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement} = \bigcup_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} [\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{1})$. To prove (3.7) we use (3.6), which immediately yields

$$\int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A\in\pi} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \frac{\partial^{\mathrm{d}}}{\partial x_1 \cdots \partial x_{\mathrm{d}}} (f \circ g)(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

Now, we use the fact that $[\mathscr{D}_{\boldsymbol{\omega}}](T) = \bigcup_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)}[\boldsymbol{k},\boldsymbol{k}+\boldsymbol{1})$ and that $[\boldsymbol{k},\boldsymbol{k}+\boldsymbol{1})$ are disjoint and to conclude

$$\begin{aligned}\int_{[\mathscr{D}(\boldsymbol{\omega})](T)} \frac{\partial^d}{\partial x_1\cdots\partial x_d}(f\circ g)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \int_{[\boldsymbol{k},\boldsymbol{k}+\boldsymbol{1})} \frac{\partial^d}{\partial x_1\cdots\partial x_d}(f\circ g)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \int_{k_1}^{k_1+1}\cdots\int_{k_d}^{k_d+1} \frac{\partial^d}{\partial x_1\cdots\partial x_d}(f\circ g)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\ |\boldsymbol{e}|=j}} (-1)^{d+j}(f\circ g)(\boldsymbol{k}+\boldsymbol{e}).\end{aligned}$$

Above, the last equation is easily seen by induction on d using the fundamental theorem of calculus. □

An immediate consequence of this result is the following one, which is essentially a generalization of Griebel & Oettershagen (2016, Lemma 2.1).

Corollary 3.4 Let $\boldsymbol{\omega}\in\mathbb{R}^d_+$ and $T>0$. For $f\in C^d(\mathbb{R})$ and $\boldsymbol{\gamma}\in\mathbb{R}^d$ we have

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\ |\boldsymbol{e}|=j}} (-1)^{d+j} f\left((\boldsymbol{k}+\boldsymbol{e})\cdot\boldsymbol{\gamma}\right) = \prod_{j=1}^{d}\gamma_j \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} f^{(d)}(\boldsymbol{\gamma}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof. We use $g(\boldsymbol{x}) := g_{\boldsymbol{\gamma}}(\boldsymbol{x}) = \boldsymbol{\gamma}\cdot\boldsymbol{x}$ in Theorem 3.3. For this specific g we observe that for all $\pi\in\mathscr{P}_{\text{part}}(\{1,\dots,d\})$ with $\pi\neq\{\{1\},\dots,\{d\}\} =: \pi^\star$ we obviously have at least one $A\in\pi$ with at least two elements meaning particularly $D^A g(\boldsymbol{x}) = 0$ and hence

$$\prod_{A\in\pi} D^A g_{\boldsymbol{\gamma}}(\boldsymbol{x}) = 0.$$

For the remaining partitioning $\pi^\star$ we obviously have

$$\prod_{A\in\pi^\star} D^A g_{\boldsymbol{\gamma}}(\boldsymbol{x}) = \prod_{j=1}^{d}\frac{\partial}{\partial x_j} g_{\boldsymbol{\gamma}}(\boldsymbol{x}) = \prod_{j=1}^{d}\gamma_j.$$

Inserting this into (3.7) gives the stated result. □

As mentioned above the result is a generalization of Griebel & Oettershagen (2016, Lemma 2.1). The result itself (with $\mathscr{D}_{\boldsymbol{\omega}}(T)$ replaced by $\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}$) and $c=-1$ and $\boldsymbol{\gamma}=\boldsymbol{\omega}=\boldsymbol{a}$ is as follows.

Corollary 3.5 Let $\boldsymbol{\omega}\in\mathbb{R}^d_+$ and $T>0$. For $c\neq 0$ we have

$$G_c(\boldsymbol{\omega},T,d,\boldsymbol{\gamma}) := \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\gamma}\cdot\boldsymbol{k}\right) = (-c)^d \prod_{j=1}^{d}\frac{\gamma_j}{\left(1-\exp(c\gamma_j)\right)} \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.9}$$

Proof. Setting $f(x) = f_c(x) = \exp(cx)$ in Corollary 3.4 and using $f_c^{(d)}(\boldsymbol{\gamma}\cdot\boldsymbol{x}) = c^d f_c(\boldsymbol{\gamma}\cdot\boldsymbol{x})$ yield

$$\begin{aligned}
c^d \prod_{j=1}^{d} \gamma_j \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} f_c(\boldsymbol{\gamma}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \prod_{j=1}^{d} \gamma_j \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} f_c^{(d)}(\boldsymbol{\gamma}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{d+j} f_c\left((\boldsymbol{k}+\boldsymbol{e})\cdot\boldsymbol{\gamma}\right) \\
&= \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{d+j} f_c\left(\boldsymbol{k}\cdot\boldsymbol{\gamma}\right) f_c\left(\boldsymbol{e}\cdot\boldsymbol{\gamma}\right) \\
&= \left(\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} f_c\left(\boldsymbol{k}\cdot\boldsymbol{\gamma}\right)\right) \left(\sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{\mathrm{d}+j} f_c\left(\boldsymbol{e}\cdot\boldsymbol{\gamma}\right)\right) \\
&= (-1)^d G_c(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}) \left(\sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^j \prod_{i=1}^{d} f_c(\gamma_i e_i)\right).
\end{aligned}$$

Next, for $\mathscr{I} = \{1, \ldots, d\}$ and arbitrary $\boldsymbol{x} \in \mathbb{R}^d$ we have the identity

$$\prod_{i\in\mathscr{I}} (1 + x_i) = \sum_{\mathscr{J}\subseteq\mathscr{I}} \prod_{i\in\mathscr{J}} x_i = \sum_{j=0}^{d} \sum_{\substack{\mathscr{J}\subseteq\mathscr{I} \\ |\mathscr{J}|=j}} \prod_{i\in\mathscr{J}} x_i. \tag{3.10}$$

Setting $x_i = -f_c(\gamma_i e_i)$ and noting that we have for each $\mathscr{J} \subseteq \mathscr{I}$ exactly one vector $\boldsymbol{e} \in \{0, 1\}^d$ with $e_i = 1$ for $i \in \mathscr{J}$ and $e_i = 0$ for $i \notin \mathscr{J}$, we see that $f_c(\gamma_i e_i) = f_c(\gamma_i)$ if $i \in \mathscr{J}$ and $f_c(\gamma_i e_i) = 1$ if $i \notin \mathscr{J}$. Hence,

$$\prod_{i=1}^{d} \left(1 - f_c(\gamma_i e_i)\right) = \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\} \\ |\boldsymbol{e}|=j}} (-1)^j \prod_{i=1}^{d} f_c(\gamma_i e_i).$$

These altogether give

$$c^d \prod_{j=1}^{d} \gamma_j \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} f_c(\boldsymbol{\gamma}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (-1)^d G_c(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}) \prod_{j=1}^{d} \left(1 - f_c(\gamma_j)\right),$$

and hence

$$G_c(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}) = (-c)^d \prod_{j=1}^{d} \frac{\gamma_j}{\left(1 - f_c(\gamma_j)\right)} \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} f_c(\boldsymbol{\gamma} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

□

In the proof of the last corollary we have not only derived upper and lower bounds on the sum $G(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma})$, but have actually given an integral representation of the sum, which might be useful in other applications. Here, we just point out the computation of specific tail distributions in the context of risk measures, see also Hendriks & Landsman (2017). There, a representation of the tail distribution

$$P(X_1 > x_1, X_2 > x_2, \dots, X_d > x_d) = h(\boldsymbol{\lambda} \cdot \boldsymbol{x}) \tag{3.11}$$

is required, where h is a multiply monotonic function of order $d-1$. A typical example of such a function is $h(x) = (1 + \beta x) \exp(-x)$ for $\beta > 0$, see Hendriks & Landsman (2017, Example 4). The density of the tail distribution (3.11) is given as

$$p_X(\boldsymbol{x}) = (-1)^d \prod_{j=1}^{d} \lambda_j h^{(d)}(\boldsymbol{\lambda} \cdot \boldsymbol{x}), \tag{3.12}$$

see Hendriks & Landsman (2017, Equation 7). Hence, we can apply Corollary 3.4 to derive a closed form expression for the probability of the set $\left[\mathscr{D}_{\boldsymbol{\omega}}\right](T)$. To be more precise we have

$$\int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} p_X(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = (-1)^d \prod_{j=1}^{d} \lambda_j \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} h^{(d)}(\boldsymbol{\lambda} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \sum_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e} \in \{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^j h\left((\boldsymbol{k} + \boldsymbol{e}) \cdot \boldsymbol{\lambda}\right).$$

As mentioned above the previous result also holds with $\mathscr{D}_{\boldsymbol{\omega}}(T)$ replaced by $\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}$. However, later on, we need a more general result for this situation, which we will provide now.

Corollary 3.6 For $c < 0$, $\boldsymbol{\omega}, \boldsymbol{\gamma} \in \mathbb{R}^d_+$, $\boldsymbol{\rho} \in \mathbb{N}^d$ and $T > 0$ let

$$G_c^{\complement}(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}, \boldsymbol{\rho}) := \sum_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \boldsymbol{k}^{\boldsymbol{\rho}} \exp\left(c \boldsymbol{\gamma} \cdot \boldsymbol{k}\right). \tag{3.13}$$

Then, provided that $|c|\gamma_j > \log(2)\rho_j$, $1 \leqslant j \leqslant d$, we have the upper bound

$$G_c^{\complement}(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}, \boldsymbol{\rho}) \leqslant |c|^d C_2 \int_W \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c \boldsymbol{\gamma} \cdot \boldsymbol{x}) \prod_{j=1}^{d} \left(\gamma_j + \frac{\rho_j}{c x_j}\right) \mathrm{d}\boldsymbol{x},$$

and the lower bound

$$G_c^{\complement}(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}, \boldsymbol{\rho}) \geqslant |c|^d C_1 \int_W \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c \boldsymbol{\gamma} \cdot \boldsymbol{x}) \prod_{j=1}^{d} \left(\gamma_j + \frac{\rho_j}{c x_j}\right) \mathrm{d}\boldsymbol{x},$$

where the constants $C_1, C_2 > 0$ are defined as

$$C_1 := \left(\prod_{j=1}^{d}(1 - e^{c\gamma_j})\right)^{-1} \qquad \text{and} \qquad C_2 := \left(\prod_{j=1}^{d}(1 - e^{c\gamma_j + \log(2)\rho_j})\right)^{-1},$$

and where

$$W := \bigcup_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement} \cap \mathbb{N}^d} [\boldsymbol{k}, \boldsymbol{k}+\boldsymbol{1}) \subseteq [\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement}. \tag{3.14}$$

Proof. The assumptions $c < 0$ and $\boldsymbol{\gamma} \in \mathbb{R}^d_+$ ensure that all sums and integrals are well defined. We will use Theorem 3.3. However, we have to modify it accordingly. We first note that indices $\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}$ that have a zero component do not contribute to the sum $G^{\complement}_c(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}, \boldsymbol{\rho})$. Hence, we have

$$G^{\complement}_c(\boldsymbol{\omega}, T, d, \boldsymbol{\gamma}, \boldsymbol{\rho}) = \sum_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement} \cap \mathbb{N}^d} \boldsymbol{k}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma} \cdot \boldsymbol{k}).$$

For this reason we also have to alter the volume over which we integrate. Instead of using $[\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement}$ we have to use the set W defined above. Then, obviously, the proof of Theorem 3.3 yields

$$\sum_{\boldsymbol{k} \in \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement} \cap \mathbb{N}^d} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e} \in \{0,1\}^d \\ |\boldsymbol{e}|=j}} (-1)^{d+j} f(g(\boldsymbol{k}+\boldsymbol{e})) = \int_W \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A \in \pi} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}. \tag{3.15}$$

Next, it suffices to require only $g \in C^d(\mathbb{R}^d_+)$ instead of $g \in C^d(\mathbb{R}^d)$ as we only use arguments with positive components. Hence, we set again $f(x) = f_c(x) = \exp(cx)$, $x \in \mathbb{R}$, but this time $g(\boldsymbol{x}) := \boldsymbol{\gamma} \cdot \boldsymbol{x} + c^{-1}\boldsymbol{\rho} \cdot \log(\boldsymbol{x})$ in (3.15), where we use the notation $\log(\boldsymbol{x}) := (\log x_1, \ldots, \log x_d)$ for $\boldsymbol{x} \in \mathbb{R}^d_+$.

For the right-hand side of (3.15), i.e., the integral, we note, as in the proof of Corollary 3.4, that the only partition π of $\mathscr{I} = \{1, \ldots, d\}$ that has nonvanishing derivatives $D^A g(\boldsymbol{x})$ with $A \in \pi$ is given by $\pi^\star = \{\{1\}, \ldots, \{d\}\}$ and for this partition we have

$$\prod_{A \in \pi^\star} D^A g(\boldsymbol{x}) = \prod_{j=1}^{d} \frac{\partial}{\partial x_j} g(\boldsymbol{x}) = \prod_{j=1}^{d}\left(\gamma_j + \frac{\rho_j}{c x_j}\right).$$

Thus, with this g and with $f = f_c$, the right-hand side of (3.15) gives

$$\begin{aligned} &\int_W \sum_{\pi \in \mathscr{P}_{\text{part}}(\{1,\ldots,d\})} f_c^{(|\pi|)}(g(\boldsymbol{x})) \prod_{A \in \pi} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_W f_c^{(|\pi^\star|)}(g(\boldsymbol{x})) \prod_{A \in \pi^\star} D^A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\ &= \int_W c^d f_c(g(\boldsymbol{x})) \prod_{j=1}^{d}\left(\gamma_j + \frac{\rho_j}{c x_j}\right) \mathrm{d}\boldsymbol{x} = \int_W \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma} \cdot \boldsymbol{x}) \prod_{j=1}^{d}\left(c\gamma_j + \frac{\rho_j}{x_j}\right) \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{3.16}$$

using $f_c(g(\boldsymbol{x})) = f_c(\gamma \cdot \boldsymbol{x} + c^{-1}\boldsymbol{\rho} \cdot \log \boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma} \cdot \boldsymbol{x})$ in the last step.

For the left-hand side of (3.15), i.e., the sum, we use $f_c(g(\boldsymbol{k}+\boldsymbol{e})) = \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{k})\exp(c\boldsymbol{\gamma}\cdot\boldsymbol{e})(\boldsymbol{k}+\boldsymbol{e})^{\boldsymbol{\rho}}$ and derive

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}\cap\mathbb{N}^d} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\|\boldsymbol{e}|=j}} (-1)^{d+j} f_c(g(\boldsymbol{k}+\boldsymbol{e})) = \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}\cap\mathbb{N}^d} \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\|\boldsymbol{e}|=j}} (-1)^{d+j} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{k})\, e^{c\boldsymbol{\gamma}\cdot\boldsymbol{e}}(\boldsymbol{k}+\boldsymbol{e})^{\boldsymbol{\rho}}$$

$$= (-1)^d \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}\cap\mathbb{N}^d} \boldsymbol{k}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{k}) \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\|\boldsymbol{e}|=j}} (-1)^j \prod_{i=1}^{d} \left(\left(1+\frac{e_i}{k_i}\right)^{\rho_i} e^{c\gamma_i e_i}\right)$$

$$= (-1)^d \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}\cap\mathbb{N}^d} \boldsymbol{k}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{k}) \sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\|\boldsymbol{e}|=j}} \prod_{\substack{i=1\\e_i=1}}^{d} \left(-\left(1+\frac{1}{k_i}\right)^{\rho_i} e^{c\gamma_i}\right). \tag{3.17}$$

At this point we can again use (3.10), this time, with $x_i = -\left(1+\frac{e_i}{k_i}\right)^{\rho_i} e^{c\gamma_i e_i}$. This yields

$$\sum_{j=0}^{d} \sum_{\substack{\boldsymbol{e}\in\{0,1\}^d\\|\boldsymbol{e}|=j}} \prod_{\substack{i=1\\e_i=1}}^{d} (-1)\left(1+\frac{1}{k_i}\right)^{\rho_i} e^{c\gamma_i} = \prod_{j=1}^{d}\left(1-\left(1+\frac{1}{k_j}\right)^{\rho_j} e^{c\gamma_j}\right).$$

Inserting this into (3.17) and equating the result to (3.16) we end up with

$$(-1)^d \int_W \boldsymbol{x}^{\rho} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{x}) \prod_{j=1}^{d}\left(c\gamma_j + \frac{\rho_j}{x_j}\right) \mathrm{d}\boldsymbol{x} = \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}\cap\mathbb{N}^d} \boldsymbol{k}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\gamma}\cdot\boldsymbol{k}) \prod_{j=1}^{d}\left(1-\left(1+\frac{1}{k_j}\right)^{\rho_j} e^{c\gamma_j}\right).$$

Furthermore, we observe for $k_j, \rho_j \in \mathbb{N}$ that

$$1 - e^{c\gamma_j + \log(2)\rho_j} \leqslant 1 - \left(1+\frac{1}{k_j}\right)^{\rho_j} e^{c\gamma_j} \leqslant 1 - e^{c\gamma_j}, \qquad 1 \leqslant j \leqslant d,$$

and the lower bound is positive as long as $|c|\gamma_j > \log(2)\rho_j$, which gives the stated bounds with the stated constants. □

In any case we have represented sums of the form (3.5) as integrals over simplices. In many cases it is possible to further simplify such integrals and to reduce these multivariate integrals to univariate integrals by employing the Hermite–Gnocchi theorem for divided differences $[\gamma_0, \cdots, \gamma_d](f)$ of a function f, which can be found, for example, in Baxter & Brummelhuis (2011, Theorem. 4.2) and which gives for $f \in C^d(\mathbb{R})$ the identities

$$[\gamma_0, \dots, \gamma_d]\,(f) = \int_{\mathscr{E}_1(1)} f^{(d)}\left(\gamma_0 + \sum_{j=1}^{d} t_j(\gamma_j - \gamma_0)\right) \mathrm{d}\boldsymbol{t} \tag{3.18}$$

$$= \int_0^1 \int_0^{1-t_1} \cdots \int_0^{1-\sum_{j=1}^{d-1} t_j} f^{(d)}\left(\gamma_0 + \sum_{j=1}^{d} t_j(\gamma_j - \gamma_0)\right) \mathrm{d}t_d \cdots \mathrm{d}t_2\, \mathrm{d}t_1. \tag{3.19}$$

This allows us to express specific multivariate integrals, namely those on the right-hand side of (3.19), as divided differences. For divided differences there are, however, other explicit formulas available. For example from Lesch (2017, Equation (A.5)) we have the explicit representation

$$\left[\gamma_0^{\alpha_0+1}, \dots, \gamma_K^{\alpha_K+1}\right](f) = \sum_{k=0}^{K} \frac{1}{\alpha_k!} \partial_{\gamma_k}^{\alpha_k} \left[f(\gamma_k) \prod_{\substack{j=0 \\ j \neq k}}^{K} (\gamma_k - \gamma_j)^{-(\alpha_j+1)} \right]. \tag{3.20}$$

Here the term $\gamma_j^{\alpha_j+1}$ in the divided difference means that γ_j appears $\alpha_j + 1$ times within the bracket. This also leads to an explicit representation for the function $f_c(x) = \exp(cx)$, see also Olsen *et al.* (2010, Equation (14)) for the case $c = 1$, which also follows directly from Baxter & Brummelhuis (2011, Corollary 4.5):

$$[0, \varepsilon, 2\varepsilon, \dots, d\varepsilon]\,(f_c) = \frac{1}{d!}\left(\frac{f_c(\varepsilon) - 1}{\varepsilon}\right)^d. \tag{3.21}$$

Another way of expressing divided differences uses the Peano kernel representation. Defining the compactly supported B-splines

$$\mathbb{M}\left(s \,|\, 0, \gamma_1, \dots, \gamma_d\right) := d\left[0, \gamma_1, \dots, \gamma_d\right](\cdot - s)_+^{d-1}$$

the Peano kernel representation is given by

$$\left[\gamma_0, \dots, \gamma_d\right](f) = \frac{1}{d!} \int_{\mathbb{R}} \mathbb{M}\left(t \,|\, \gamma_0, \dots, \gamma_d\right) f^{(d)}(t)\, \mathrm{d}t, \tag{3.22}$$

see, for example, de Boor (2005, Equations 47 and 48. Combining (3.22) and (3.18), where we also set $\gamma_0 = 0$, gives the identity

$$\int_{\mathscr{E}_1(1)} f^{(d)}(\boldsymbol{\gamma} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \left[0, \gamma_1, \dots, \gamma_d\right](f) = \frac{1}{d!} \int_{\mathbb{R}} \mathbb{M}\left(s \,|\, 0, \gamma_1, \dots, \gamma_d\right) f^{(d)}(s)\, \mathrm{d}s \tag{3.23}$$

for integrals of ridge functions over simplices. This, however, can be used to establish the following result.

PROPOSITION 3.7 Let $\boldsymbol{\omega} \in \mathbb{R}_+^d$, $\boldsymbol{\gamma} \in \mathbb{R}_{\geqslant 0}^d$, $c \in \mathbb{R}$ and $T > 0$. Then

$$\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\gamma} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} = T^d \left(\prod_{j=1}^d \omega_j^{-1}\right) \frac{1}{d!} \int_{\mathbb{R}} \mathbb{M}\left(s \,\middle|\, 0, \frac{\gamma_1}{\omega_1}, \ldots, \frac{\gamma_d}{\omega_d}\right) \exp(cTs)\, \mathrm{d}s. \tag{3.24}$$

Proof. We use again the notation $f_c(x) = \exp(cx)$. Then a simple variable substitution yields

$$\begin{aligned} \int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} f_c\left(\boldsymbol{\gamma} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} &= \int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp\left(\sum_{j=1}^d c\gamma_j x_j\right) \mathrm{d}\boldsymbol{x} = T^d \left(\prod_{j=1}^d \omega_j^{-1}\right) \int_{\mathcal{E}_{\mathbf{1}}(1)} \exp\left(cT \sum_{j=1}^d \frac{\gamma_j}{\omega_j} x_j\right) \mathrm{d}\boldsymbol{x} \\ &= T^d \left(\prod_{j=1}^d \omega_j^{-1}\right) \int_{\mathcal{E}_{\mathbf{1}}(1)} f_{cT}\left(\boldsymbol{\beta} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} \end{aligned}$$

with $\beta_j := \frac{\gamma_j}{\omega_j}$ for $1 \leqslant j \leqslant d$. Now we use (3.23) to derive

$$\begin{aligned} \int_{\mathscr{E}_{1}(1)} f_{cT}(\boldsymbol{\beta} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} &= (cT)^d \int_{\mathscr{E}_{1}(1)} f_{cT}^{(d)}(\boldsymbol{\beta} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \frac{(cT)^{-d}}{d!} \int_{\mathbb{R}} \mathbb{M}(s \,|\, 0, \beta_1, \ldots, \beta_d) f_{cT}^{(d)}(s)\, \mathrm{d}s \\ &= \frac{1}{d!} \int_{\mathbb{R}} \mathbb{M}(s \,|\, 0, \beta_1, \ldots, \beta_d) f_{cT}(s)\, \mathrm{d}s. \end{aligned} \tag{3.25}$$

□

A special case of Proposition 3.7 is essentially Griebel & Oettershagen (2016, Proposition 2.3), which we will rephrase as follows.

COROLLARY 3.8 Let $\boldsymbol{\omega} \in \mathbb{R}_+^d$, $c \in \mathbb{R}$ and $T > 0$. Then

$$\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\omega} \cdot \boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} = \left(\prod_{j=1}^d \omega_j^{-1}\right) \frac{1}{(d-1)!} \int_0^T s^{d-1} \exp(cs)\, \mathrm{d}s. \tag{3.26}$$

Proof. We choose $\boldsymbol{\gamma} = \boldsymbol{\omega}$ in (3.24). Using the same notation for multiple entries as in (3.20) the involved Peano kernel has the property

$$\mathbb{M}(s \,|\, 0^{d+1-j}, 1^j) = d \binom{d-1}{j-1} s^{j-1} (1-s)^{d-j} \chi_{[0,1]}(s). \tag{3.27}$$

Using this for $j = d$ in (3.24) yields

$$\begin{aligned}
\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} &= T^d \left(\prod_{j=1}^{d} \omega_j^{-1}\right) \frac{1}{d!} \int_{\mathbb{R}} \mathbb{M}\left(s \,|\, 0, 1, \ldots, 1\right) \exp(cTs)\, \mathrm{d}s \\
&= T^d \left(\prod_{j=1}^{d} \omega_j^{-1}\right) \frac{1}{(d-1)!} \int_0^1 s^{d-1} \exp(cTs)\, \mathrm{d}s \\
&= \left(\prod_{j=1}^{d} \omega_j^{-1}\right) \frac{1}{(d-1)!} \int_0^T s^{d-1} \exp(cs)\, \mathrm{d}s.
\end{aligned}$$

□

Another consequence is the following result, which was also proven in Griebel & Oettershagen (2016, Equation 2.11) in a different way.

COROLLARY 3.9 For $c < 0$ and $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $T > 0$ we have

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) \leqslant \frac{1}{(d-1)!} \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} \int_{-cT}^{\infty} t^{d-1} \exp(-t)\, \mathrm{d}t. \tag{3.28}$$

Proof. Using the substitution $t = -cs$ in the integral on the right-hand side of (3.26) yields

$$\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x} = \left(\prod_{j=1}^{d} \frac{1}{(-c)\omega_j}\right) \frac{1}{(d-1)!} \int_0^{-cT} t^{d-1} \exp(-t)\, \mathrm{d}t. \tag{3.29}$$

Next (3.9) with $\boldsymbol{\gamma} = \boldsymbol{\omega}$ gives

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathbb{N}_0^d} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) &= \lim_{T\to\infty} (-c)^d \prod_{j=1}^{d} \frac{\omega_j}{\left(1-\exp(c\omega_j)\right)} \int_{[\mathscr{D}_{\omega}](T)} \exp(c\boldsymbol{\omega}\cdot\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\
&= (-c)^d \prod_{j=1}^{d} \frac{\omega_j}{\left(1-\exp(c\omega_j)\right)} \int_{\mathbb{R}_+^d} \exp(c\boldsymbol{\omega}\cdot\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\
&= \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)}.
\end{aligned}$$

Hence, using (3.9) with $\boldsymbol{\gamma} = \boldsymbol{\omega}$ again and the above result shows

$$\begin{aligned}
0 \leqslant \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) &= \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) - \sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) \\
&= \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} - (-c)^d \prod_{j=1}^{d} \frac{\omega_j}{\left(1-\exp(c\omega_j)\right)} \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \exp(c\boldsymbol{\omega}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} \left(1 - \prod_{j=1}(-c\omega_j) \int_{[\mathscr{D}_{\boldsymbol{\omega}}](T)} \exp(c\boldsymbol{\omega}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right).
\end{aligned}$$

Since $1-\exp(c\omega_j) > 0$ for $c < 0$ and $\omega_j > 0$ we can conclude that the last term in brackets has to be non-negative. Hence, we can further enlarge it using (3.1) and (3.29), as long as the result remains non-negative. Hence, we have

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) &\leqslant \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} \left(1 - \prod_{j=1}(-c\omega_j) \int_{\mathscr{E}_{\boldsymbol{\omega}}(T)} \exp(c\boldsymbol{\omega}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right) \\
&= \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} \left(1 - \frac{1}{(d-1)!} \int_0^{-cT} t^{d-1} \exp(-t)\,\mathrm{d}t\right).
\end{aligned}$$

Next we observe

$$1 = \frac{\Gamma(d)}{\Gamma(d)} = \frac{1}{(d-1)!} \int_0^{\infty} s^{d-1} \exp(-s)\,\mathrm{d}s$$

that allows us to conclude

$$\sum_{\boldsymbol{k}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \exp\left(c\boldsymbol{\omega}\cdot\boldsymbol{k}\right) \leqslant \frac{1}{(d-1)!} \prod_{j=1}^{d} \frac{1}{\left(1-\exp(c\omega_j)\right)} \int_{-cT}^{\infty} t^{d-1} \exp(-t)\,\mathrm{d}t.$$

□

Yet another immediate consequence is the following classical result for the volume of a simplex.

COROLLARY 3.10 Let $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $T > 0$. Then

$$\operatorname{vol}\left(\mathscr{E}_{\boldsymbol{\omega}}(T)\right) = \left(\prod_{j=1}^{d} \omega_j^{-1}\right) \frac{T^d}{d!}. \tag{3.30}$$

Proof. Take the limit $c \to 0$ on both sides of (3.26) and observe

$$\lim_{c\to 0} \frac{1}{(d-1)!} \int_0^T t^{d-1} \exp(ct)\,\mathrm{d}t = \frac{1}{(d-1)!} \int_0^T s^{d-1}\,\mathrm{d}s = \frac{T^d}{d!}.$$

□

We also need integrals of the following form, where we integrate over the complement of $\mathscr{E}_{\boldsymbol{\omega}}(T)$ given by $\mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement} = \mathbb{R}^d_{\geqslant 0} \setminus \mathscr{E}_{\boldsymbol{\omega}}(T) = \{\boldsymbol{x} \in \mathbb{R}^d_{\geqslant 0} : \boldsymbol{\omega} \cdot \boldsymbol{x} > T\}$. Estimates for such quantities can also be found in Tempone & Wolfers (2018). The proof of Tempone & Wolfers (2018, Lemma A.1), however, is based on a different technique. Theorem 3.3 above is a more general result that states an identity and not only an estimate, and that is not restricted to the particular exponential-type sums. In particular, Theorem 3.3 allows us to prove Corollary 3.4 that again is an equality. Moreover, Theorem 3.3 could also be applied to the example from (3.11) with $h(\boldsymbol{\lambda} \cdot \boldsymbol{k}) = (1 + \beta \boldsymbol{\lambda} \cdot \boldsymbol{k}) \exp(\boldsymbol{\lambda} \cdot \boldsymbol{k})$, which seems not to be covered directly by Tempone & Wolfers (2018, Lemma A.1). For the sake of brevity we do not pursue this further.

LEMMA 3.11 For $c < 0$, $\boldsymbol{\rho} \in \mathbb{N}_0^d$, $\boldsymbol{\omega} \in \mathbb{R}^d_+$ and $T \geqslant 1$ we have

$$\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement}} \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\omega} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \leqslant T^{d+|\boldsymbol{\rho}|} \exp(cT) \boldsymbol{\omega}^{-(\boldsymbol{\rho}+\mathbf{1})} \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c(\|\boldsymbol{x}\|_1 - 1))\, \mathrm{d}\boldsymbol{x}.$$

Proof. Again, the change of variables of the form $x_j = Ty_j/\omega_j$, $1 \leqslant j \leqslant d$, yields

$$\begin{aligned}
\int_{\mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement}} \boldsymbol{x}^{\boldsymbol{\rho}} \exp(c\boldsymbol{\omega} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x} &= T^d \left(\prod_{j=1}^d \omega_j^{-1}\right) \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} T^{|\boldsymbol{\rho}|} \boldsymbol{\omega}^{-\boldsymbol{\rho}} \boldsymbol{y}^{\boldsymbol{\rho}} \exp\left(cT \sum_{j=1}^d y_j\right) \mathrm{d}\boldsymbol{y} \\
&= T^{d+|\boldsymbol{\rho}|} \boldsymbol{\omega}^{-(\boldsymbol{\rho}+\mathbf{1})} \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} \boldsymbol{x}^{\boldsymbol{\rho}} \exp\left(cT \sum_{j=1}^d x_j\right) \mathrm{d}\boldsymbol{x} \\
&= T^{d+|\boldsymbol{\rho}|} \exp(cT) \boldsymbol{\omega}^{-(\boldsymbol{\rho}+\mathbf{1})} \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} \boldsymbol{x}^{\boldsymbol{\rho}} \exp\left(cT\left(\sum_{j=1}^d x_j - 1\right)\right) \mathrm{d}\boldsymbol{x} \\
&\leqslant T^{d+|\boldsymbol{\rho}|} \exp(cT) \boldsymbol{\omega}^{-(\boldsymbol{\rho}+\mathbf{1})} \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} \boldsymbol{x}^{\rho} \exp\left(c\left(\sum_{j=1}^d x_j - 1\right)\right) \mathrm{d}\boldsymbol{x},
\end{aligned}$$

where the last step is valid for $c < 0$, $T \geqslant 1$ and $\sum_{j=1}^d x_j - 1 > 0$. This also guarantees the existence of the last integral. Finally, as $\boldsymbol{x} \in \mathbb{R}^d_{\geqslant 0}$, we have $\sum_{j=1}^d x_j = \|\boldsymbol{x}\|_1$, which gives the statement. □

3.2 *Bounds on the cardinality of the index set*

The considerations so far allow us to give upper and lower bounds on the number of multi-indices contained in $\mathscr{D}_{\boldsymbol{\omega}}(T)$, see also Rauhut & Schwab (2017, Equation 5.14), Griebel & Oettershagen (2016, Lemma 2.8) and Beged-dov (1972).

COROLLARY 3.12 Let $\boldsymbol{\omega} \in \mathbb{R}^d_+$ and $T > 0$. Then

$$\left(\prod_{j=1}^d \omega_j^{-1}\right) \frac{T^d}{d!} \leqslant \#\mathscr{D}_{\boldsymbol{\omega}}(T) \leqslant \left(\prod_{j=1}^d \omega_j^{-1}\right) \frac{(T + \|\boldsymbol{\omega}\|_1)^d}{d!}. \tag{3.31}$$

Proof. The inclusions given in (3.1) and the relation from (3.2) immediately yield

$$\operatorname{vol}\left(\mathscr{E}_{\boldsymbol{\omega}}(T)\right) \leqslant \#\mathscr{D}_{\boldsymbol{\omega}}(T) = \operatorname{vol}\left(\left[\mathscr{D}_{\boldsymbol{\omega}}\right](T)\right) \leqslant \operatorname{vol}\left(\mathscr{E}_{\boldsymbol{\omega}}(T + \|\boldsymbol{\omega}\|_1)\right),$$

which gives together with (3.30) the stated inequalities. □

Finally, we are now in the position to bound the number of indices in $\mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$. The following result can also, albeit with a different proof, be found in Nobile *et al.* (2008a, Remark 3.7). It is a direct consequence of (3.31) and the relation (3.3) between $\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})$ and $\mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$.

PROPOSITION 3.13 For $\ell \in \mathbb{N}$ and for $\boldsymbol{\omega} \in \mathbb{R}^d_+$ the number of indices in $\mathcal{I}_{\boldsymbol{\omega}}(\ell, d)$ can be bounded by

$$\frac{(\ell\omega_{\min})^d}{d!}\prod_{j=1}^{d}\omega_j^{-1} \leqslant \#\mathcal{I}_{\boldsymbol{\omega}}(\ell, d) = \#\mathcal{D}_{\boldsymbol{\omega}}(\ell\omega_{\min}) \leqslant \frac{\left(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right)^d}{d!}\prod_{j=1}^{d}\omega_j^{-1}. \tag{3.32}$$

From this we can also directly see that

$$\left(\ell\omega_{\min}\right)^d\left(1 + d\frac{\|\boldsymbol{\omega}\|_1}{\ell\omega_{\min}}\right) \leqslant \left(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right)^d = \left(\ell\omega_{\min}\right)^d\left(1 + \frac{\|\boldsymbol{\omega}\|_1}{\ell\omega_{\min}}\right)^d$$

and hence for $\ell \geqslant \frac{\|\boldsymbol{\omega}\|_1}{\omega_{\min}}$ that

$$\left(\ell\omega_{\min}\right)^d \leqslant \left(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1\right)^d = \left(\ell\omega_{\min}\right)^d 2^d,$$

which has also been observed in Nobile *et al.* (2008a, Remark 3.7).

3.3 *Estimates on the number of grid points*

After estimating the number of indices in $\mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ we proceed now with the main task of bounding the number of points in our sparse grid (2.3), i.e., in

$$\varXi^{\otimes}_{\boldsymbol{\omega}}(\ell, d) := \bigcup_{\boldsymbol{\lambda}\in\mathscr{J}_{\boldsymbol{\omega}}(\ell,d)} \varXi^{\otimes}_{\boldsymbol{\lambda}} = \bigcup_{\boldsymbol{\lambda}\in\mathscr{J}_{\boldsymbol{\omega}}(\ell,d)} \varXi^{(1)}_{\lambda_1}\times\cdots\times\varXi^{(d)}_{\lambda_d}.$$

From this representation and the definition of $\mathscr{J}_{\boldsymbol{\omega}}(\ell, d)$ we immediately have (see also Haji-Ali *et al.*, 2018, Section 5.1) that

$$\#\varXi^{\otimes}_{\boldsymbol{\omega}}(\ell, d) = \sum_{\boldsymbol{\lambda}\in\mathscr{J}_{\boldsymbol{\omega}}(\ell,d)} \#\varXi^{\otimes}_{\boldsymbol{\lambda}} \leqslant \sum_{\boldsymbol{\lambda}\in\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)} \#\varXi^{\otimes}_{\boldsymbol{\lambda}} \leqslant \sum_{\boldsymbol{\lambda}\in\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)}\prod_{j=1}^{d} N^{(j)}(\lambda_j), \tag{3.33}$$

if $N^{(j)}(\lambda_j)$ denotes the number of points in $\varXi^{(j)}_{\lambda_j}$. To employ this bound further we need to make some assumptions on the asymptotic behavior of the $N^{(j)}(\lambda_j)$.

3.3.1 *Bounds in the situation of linear oversampling.*

Here we mainly follow Haji-Ali *et al.* (2018). Throughout this subsection we will assume that $\boldsymbol{\omega} \in \mathbb{N}^d$ and that the number of points $N^{(j)}(\lambda_j)$ of $\varXi^{j}_{\lambda_j}$ is of the form

$$N^{(j)}(\lambda_j) = \lambda_j + b_j, \tag{3.34}$$

with a given number $b_j \in \mathbb{N}_0$, i.e., we consider linear oversampling. A more general choice would be $N^{(j)}(\lambda_j) = a_j\lambda_j + b_j$ with $a_j \in \mathbb{N}$. However, as we have

$$\prod_{j=1}^{d} N^{(j)}(\lambda_j) = \prod_{j=1}^{d} \left(a_j\lambda_j + b_j\right) = \left(\prod_{j=1}^{d} a_j\right) \prod_{j=1}^{d} \left(\lambda_j + \frac{b_j}{a_j}\right),$$

we can simply choose $a_j = 1$ for all $1 \leqslant j \leqslant d$ without restriction. Then we have

$$\prod_{j=1}^{d} N^{(j)}(\lambda_j) = \prod_{j=1}^{d} \left(\lambda_j + b_j\right) = \prod_{j=1}^{d} \exp\left(\log\left(\lambda_j + b_j\right)\right) \leqslant \prod_{j=1}^{d} \exp\left(\left\lceil \log\left(\lambda_j + b_j\right)\right\rceil\right).$$

Hence, if we define $\nu_j := \left\lceil \log\left(\lambda_j + b_j\right)\right\rceil$, $1 \leqslant j \leqslant d$, then this implies

$$\prod_{j=1}^{d} N^{(j)}(\lambda_j) \leqslant \prod_{j=1}^{d} \exp\left(\nu_j\right) = \exp(\boldsymbol{\nu} \cdot \mathbf{1}).$$

Moreover, $\boldsymbol{\lambda} \in \mathscr{D}_{\boldsymbol{\omega}}(T)$and $\boldsymbol{\omega} \in \mathbb{N}^d$ leads to

$$\begin{aligned}\frac{\boldsymbol{\nu} \cdot \mathbf{1}}{\|\boldsymbol{\omega}\|_1} &= \frac{\sum_{j=1}^{d} \nu_j}{\sum_{j=1}^{d} \omega_j} \leqslant \frac{\sum_{j=1}^{d} \left(\log\left(\lambda_j + b_j\right) + 1\right)}{\sum_{j=1}^{d} \omega_j} \leqslant \frac{\sum_{j=1}^{d} \omega_j \log\left(\lambda_j + b_j\right)}{\sum_{j=1}^{d} \omega_j} + 1 \\ &\leqslant \log\left(\frac{\sum_{j=1}^{d} \omega_j \left(\lambda_j + b_j\right)}{\sum_{j=1}^{d} \omega_j}\right) + 1 \leqslant \log\left(\frac{T + \sum_{j=1}^{d} \omega_j b_j}{\sum_{j=1}^{d} \omega_j}\right) + 1,\end{aligned}$$

due to Jensen's inequality for the concave function $x \mapsto \log x$. This means in particular that $\boldsymbol{\lambda} \in \mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})$ implies $\boldsymbol{\nu} \in \mathscr{D}_{\mathbf{1}}(S)$ with

$$S := \|\boldsymbol{\omega}\|_1 \left(\log\left(\frac{\ell\omega_{\min} + \boldsymbol{\omega} \cdot \boldsymbol{b}}{\|\boldsymbol{\omega}\|_1}\right) + 1\right). \tag{3.35}$$

This gives us the bound

$$\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d) \leqslant \sum_{\boldsymbol{\lambda} \in \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)} \prod_{j=1}^{d} N^{(j)}(\lambda_j) \leqslant \sum_{\boldsymbol{\nu} \in \mathscr{D}_{\mathbf{1}}(S)} \exp(\boldsymbol{\nu} \cdot \mathbf{1}).$$

If we now apply (3.9) by setting $\boldsymbol{\omega}$ there to $\mathbf{1}$ and with $c = 1$, $\boldsymbol{\gamma} = \mathbf{1}$ and $T = S$ then we find

$$\begin{aligned}\sum_{\boldsymbol{\nu}\in\mathscr{D}_{\mathbf{1}}(S)} \exp(\boldsymbol{\nu}\cdot\mathbf{1}) &= (e-1)^{-d}\int_{[\mathscr{D}_{\mathbf{1}}](S)} \exp(\mathbf{1}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leqslant (e-1)^{-d}\int_{\mathscr{E}_{\mathbf{1}}(S+\|\boldsymbol{\omega}\|_1)} \exp(\mathbf{1}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \frac{1}{(e-1)^d(d-1)!}\int_0^{S+\|\boldsymbol{\omega}\|_1} s^{d-1}\exp(s)\,\mathrm{d}s,\end{aligned}$$

using also (3.1) and Corollary 3.8. The integral on the right-hand side can further be bounded using the following lemma.

LEMMA 3.14 Let $d \in \mathbb{N}$. For $T \geqslant 0$ we have

$$\int_0^T \exp(s)s^{d-1}\mathrm{d}s \leqslant \exp(T)T^{d-1},$$

and for $T \geqslant d$ we have

$$\int_T^\infty s^{d-1}\exp(-s)\,\mathrm{d}s \leqslant d\exp(-T)T^{d-1}. \tag{3.36}$$

Proof. The mean value theorem guarantees the existence of a $\xi \in [0, T]$ such that

$$\int_0^T \exp(s)s^{d-1}\mathrm{d}s = \xi^{d-1}\int_0^T \exp(s)\mathrm{d}s = \xi^{d-1}[\exp(T)-1] \leqslant T^{d-1}\exp(T),$$

which gives the bound for the first integral. The second bound was given in Griebel & Oettershagen (2016, Lemma 2.5). □

The first result of this lemma hence yields

$$\begin{aligned}\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d) &\leqslant \sum_{\boldsymbol{\nu}\in\mathscr{D}_{\mathbf{1}}(S)} \exp(\boldsymbol{\nu}\cdot\mathbf{1}) \leqslant \frac{1}{(e-1)^d(d-1)!}\int_0^{S+\|\boldsymbol{\omega}\|_1} s^{d-1}\exp(s)\,\mathrm{d}s\\ &\leqslant \frac{1}{(e-1)^d(d-1)!}\exp(S+\|\boldsymbol{\omega}\|_1)(S+\|\boldsymbol{\omega}\|_1)^{d-1}.\end{aligned}$$

Recalling the definition of S from (3.35) we have proven the following result.

LEMMA 3.15 For $\boldsymbol{\omega} \in \mathbb{N}^d$ and under the Assumption (3.34) on the number of points in the grids $\Xi_{\lambda_j}^{(j)}$ the number of points in the anisotropic tensor grid $\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)$ can be bounded by

$$\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d) \leqslant \frac{\exp(2\|\boldsymbol{\omega}\|_1)\|\boldsymbol{\omega}\|_1^{d-1}}{(e-1)^d(d-1)!}\left[\log\left(\frac{\ell\omega_{\min}+\boldsymbol{\omega}\cdot\boldsymbol{b}}{\|\boldsymbol{\omega}\|_1}+2\right)\right]^{d-1}\left(\frac{\ell\omega_{\min}+\boldsymbol{\omega}\cdot\boldsymbol{b}}{\|\boldsymbol{\omega}\|_1}\right)^{\|\boldsymbol{\omega}\|_1}.$$

Another direct consequence of the above considerations is the following one, which simply follows by setting $b_j = 0$, $1 \leqslant j \leqslant d$.

COROLLARY 3.16 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{N}^d$ and $\ell \in \mathbb{N}$. Then

$$\begin{aligned}\sum_{\boldsymbol{\lambda}\in\Lambda} \boldsymbol{\lambda}^{\mathbf{1}} = \sum_{\boldsymbol{\lambda}\in\Lambda}\prod_{j=1}^{d}\lambda_j \;&\leqslant\; \frac{e^{2\|\boldsymbol{\omega}\|_1}\|\boldsymbol{\omega}\|_1^{d-1}}{(e-1)^d(d-1)!}\left[\log\left(\frac{\ell\omega_{\min}}{\|\boldsymbol{\omega}\|_1}+2\right)^{d-1}\left(\frac{\ell\omega_{\min}}{\|\boldsymbol{\omega}\|_1}\right)\right]^{\|\boldsymbol{\omega}\|_1}\\ &=\; C(\boldsymbol{\omega}, d)\ell^{\|\boldsymbol{\omega}\|_1}\left[\log(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)\right]^{d-1}.\end{aligned}$$

3.3.2 *Bounds for exponential grids.* We now want to bound the number of points in $\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d)$ if the number of points in the univariate grids $\Xi_{\lambda_j}^{(j)}$ grow exponentially. Hence, we now make the following assumption.

ASSUMPTION 3.17 There exist universal constants $C_1, C_2 > 0$ and for each $1 \leqslant j \leqslant d$ there is a constant $\eta_j > 0$ such that the cardinality $N^{(j)}(k)$ of $\Xi_k^{(j)}$ satisfies

$$C_1 2^{\eta_j k} \leqslant N^{(j)}(k) \leqslant C_2 2^{\eta_j k} \tag{3.37}$$

for all $1 \leqslant j \leqslant d$ and $k \in \mathbb{N}$.

PROPOSITION 3.18 Let $\boldsymbol{\omega}, \boldsymbol{\eta} \in \mathbb{R}_+^d$ and let $\alpha_j := \omega_j/\eta_j$, $1 \leqslant j \leqslant d$, and let $\alpha_{\min} = \min \alpha_j$. Then, under Assumption 3.17, the number of points in $\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d)$ can be bounded from above by

$$\begin{aligned}\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d) \;&\leqslant\; \frac{C_2^d \log(2)^{d-1}\alpha_{\min}}{(d-1)!}\prod_{j=1}^{d}\frac{\alpha_j^{-1}}{2^{\omega_j/\alpha_j}-1}2^{2\|\boldsymbol{\omega}\|_1/\alpha_{\min}}(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)^{d-1}2^{\ell\omega_{\min}/\alpha_{\min}}\\ &=\; C(\boldsymbol{\omega}, \boldsymbol{\alpha})\ell^{d-1}2^{\ell\omega_{\min}/\alpha_{\min}}.\end{aligned}$$

Moreover, if $1 \leqslant j^\star \leqslant d$ is an index for which $\omega_{j^\star} = \min \omega_j = \omega_{\min}$ holds, then the number of points can be bounded from below by

$$\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d) \geqslant C_1^d\, 2^{\|\boldsymbol{\eta}\|_1}\, 2^{\ell\omega_{\min}/\alpha_{j^\star}}.$$

Proof. We start with the upper bound. Using (3.33), (3.3) and Assumption 3.17, we can conclude that

$$
\begin{aligned}
\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d) &\leqslant \sum_{\boldsymbol{\lambda}\in\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)} \prod_{j=1}^{d} N^{(j)}(\lambda_j) \leqslant C_2^d \sum_{\boldsymbol{\lambda}\in\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)} \prod_{j=1}^{d} 2^{\eta_j\lambda_j} \\
&\leqslant C_2^d \sum_{\boldsymbol{\lambda}\in\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)} 2^{\boldsymbol{\eta}\cdot\boldsymbol{\lambda}} = C_2^d \sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})+\{\mathbf{1}\}} 2^{\boldsymbol{\eta}\cdot\boldsymbol{\lambda}} \\
&= C_2^d\, 2^{\|\boldsymbol{\eta}\|_1} \sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\boldsymbol{\omega}_{\min})} 2^{\boldsymbol{\eta}\cdot\boldsymbol{\lambda}} \\
&= C_2^d\, 2^{\|\boldsymbol{\eta}\|_1} \log(2)^d \prod_{j=1}^{d} \frac{\eta_j}{2^{\eta_j}-1} \int_{[\mathscr{D}_{\boldsymbol{\omega}}](\ell\omega_{\min})} 2^{\boldsymbol{\eta}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x} \\
&\leqslant C_2^d\, 2^{\|\boldsymbol{\eta}\|_1} \log(2)^d \prod_{j=1}^{d} \frac{\eta_j}{2^{\eta_j}-1} \int_{\mathscr{E}_{\boldsymbol{\omega}}(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)} 2^{\boldsymbol{\eta}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x},
\end{aligned}
\tag{3.38}
$$

where we have also used (3.9) with $c = \log(2)$ and $\boldsymbol{\gamma}$ replaced by $\boldsymbol{\eta}$ and $[\mathscr{D}_{\boldsymbol{\omega}}](T) \subseteq \mathscr{E}_{\boldsymbol{\omega}}(T + \|\boldsymbol{\omega}\|_1)$ from (3.1).

So far we have not yet employed the connections between $\boldsymbol{\omega}$ and $\boldsymbol{\eta}$ given by $\eta_j = \omega_j/\alpha_j$. This yields

$$
\boldsymbol{\eta}\cdot\boldsymbol{x} = \sum_{j=1}^{d} \eta_j x_j = \sum_{j=1}^{d} \frac{1}{\alpha_j}\omega_j x_j \leqslant \frac{1}{\alpha_{\min}}\boldsymbol{\omega}\cdot\boldsymbol{x},
$$

which leads to

$$
\#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d) \leqslant C_2^d\, 2^{\|\boldsymbol{\omega}\|_1/\alpha_{\min}} \log(2)^d \prod_{j=1}^{d} \frac{\omega_j\alpha_j^{-1}}{2^{\omega_j/\alpha_j}-1} \int_{\mathscr{E}_{\boldsymbol{\omega}}(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)} 2^{\frac{1}{\alpha_{\min}}\boldsymbol{\omega}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}.
$$

Next we can use Corollary 3.8 with $T = \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1$ and $c = \log(2)/\alpha_{\min}$ that gives

$$
\int_{\mathscr{E}_{\boldsymbol{\omega}}(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)} 2^{\frac{1}{\alpha_{\min}}\boldsymbol{\omega}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(d-1)!}\prod_{j=1}^{d}\omega_j^{-1} \int_0^{\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1} s^{d-1}\exp(s\log(2)/\alpha_{\min})\,\mathrm{d}s.
$$

For the latter integral we can use the upper bound of Lemma 3.14 in the form

$$
\int_0^T s^{d-1}\exp(cs)\,\mathrm{d}s \leqslant \frac{T^{d-1}}{c}\exp(cT)
$$

with $T = \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1$ and $c = \log(2)/\alpha_{\min}$ to derive

$$\int_0^{\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1} s^{d-1}\exp(s\log(2)/\alpha_{\min})\,\mathrm{d}s \leqslant (\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1)^{d-1}\frac{\alpha_{\min}}{\log(2)}2^{(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1)/\alpha_{\min}}.$$

Inserting this back into the previous estimates gives the stated upper bound on $\#\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)$.

For the lower bound on $\#\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)$ we note that

$$\boldsymbol{\nu}^\star = \mathbf{1} + \ell\boldsymbol{e}_{j^\star} = (1,\ldots,1,\ell+1,1,\ldots,1)^T \in \mathscr{J}_{\boldsymbol{\omega}}(\ell,d) = \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)\setminus\mathscr{I}\left(\ell - \frac{\|\boldsymbol{\omega}\|_1}{\omega_{\min}}, d\right),$$

since we have on the one hand $(\boldsymbol{\nu}^\star - \mathbf{1})\cdot\boldsymbol{\omega} = \ell\omega_{\min} \leqslant \ell\omega_{\min}$ implying $\boldsymbol{\nu}^\star \in \mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$ and on the other hand $(\boldsymbol{\nu}^\star - \mathbf{1})\cdot\boldsymbol{\omega} = \ell\omega_{\min} > \ell\boldsymbol{\omega}_{\min} - \|\boldsymbol{\omega}\|_1$ implying $\boldsymbol{\nu}^\star \notin \mathscr{I}(\ell - \|\boldsymbol{\omega}\|_1/\omega_{\min}, d)$. This gives the obvious bound

$$\#\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d) = \#\bigcup_{\boldsymbol{\lambda}\in\mathscr{J}_{\omega}(\ell,d)}\varXi_{\boldsymbol{\lambda}}^{\otimes} \geqslant \#\varXi_{\boldsymbol{\nu}^\star}^{\otimes} = \prod_{j=1}^{d} N^{(j)}(\nu_j^\star) \geqslant C_1^d\, 2^{\|\boldsymbol{\eta}\|_1}\, 2^{\ell\eta_{j^\star}},$$

where we have used Assumption 3.17. The lower bound then follows from $\eta_j = \omega_j/\alpha_j$ and the choice of $j^\star$. □

This shows in particular $\log_2(\#\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)) \geqslant \log_2(C_1^d 2^{\|\boldsymbol{\eta}\|_1}) + \ell\omega_{\min}/\alpha_{j^\star}$ or

$$\ell \leqslant C(d,\boldsymbol{\omega},\boldsymbol{\alpha})\log\#\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d),$$

taking also the connection between $\boldsymbol{\eta}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$ into account.

4. Sampling inequalities

After deriving bounds on the number of points in the anisotropic sparse grid $\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)$ we will now proceed to state and prove sampling inequalities on $\varXi_{\boldsymbol{\omega}}^{\otimes}(\ell,d)$. These sampling inequalities are proven by employing a Smolyak construction and by comparison to best approximation processes. Hence, we will start this section by discussing univariate operators and then proceed to the Smolyak construction in a rather general way before we will state and prove our sampling inequalities.

4.1 *Univariate quasi-interpolation operators*

In the following we will consider univariate or multivariate spaces of low dimensions and associated quasi-interpolation operators.

As outlined in Section 2, we let $\Omega^{(j)} \subseteq \mathbb{R}^{n_j}$ be a bounded domain with a sufficiently smooth boundary. For two normed spaces $\mathscr{H}^{\beta_j}(\Omega^{(j)}) \subseteq \mathscr{H}^{\alpha_j}(\Omega^{(j)})$, with $\mathscr{H}^{\beta_j}(\Omega^{(j)}) \subseteq C(\Omega^{(j)})$, we consider the embedding operators

$$\iota^{(j)} : \mathscr{H}^{\beta_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)}).$$

Usually, $H^{\beta_j}(\Omega^{(j)})$ will be a Sobolev space of order β_j, but for the time being, it can also be a more general space. Furthermore, for each j we consider a nested sequence of linear subspaces

$$\emptyset = \Pi_0^{(j)} \subseteq \Pi_1^{(j)} \subseteq \cdots \subseteq \Pi_j^{(j)} \subseteq \Pi_{j+1}^{(j)} \subseteq \cdots \subseteq \mathscr{H}^{\beta_j}(\Omega^{(j)})$$

and denote the dimension of $\Pi_k^{(j)}$ by $M^{(j)}(k) := \dim \Pi_k^{(j)}$, yielding a nondecreasing sequence $M^{(j)} : \mathbb{N}_0 \to \mathbb{N}_0$. With these finite-dimensional spaces come best approximation operators $\mathscr{A}_k^{(j)} : \mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \Pi_k^{(j)} \subseteq \mathscr{H}^{\alpha_j}(\Omega^{(j)})$ defined by

$$f \mapsto \arg \min_{p \in \Pi_k^{(j)}} \|f - p\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)})}.$$

For simplicity, we will denote the restriction of $\mathscr{A}_k^{(j)}$ to the subspace $\mathscr{H}^{\beta_j}(\Omega^{(j)})$ again simply by $\mathscr{A}_k^{(j)}$. We will assume that these restrictions provide good approximations to the embedding operator and that we have a quantitative error bound of the form

$$\left\| \iota^{(j)} - \mathscr{A}_k^{(j)} \right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})} \leqslant \rho(\boldsymbol{\alpha}; k), \qquad k \in \mathbb{N}, \tag{4.1}$$

with $\lim_{k\to\infty} \rho(\boldsymbol{\alpha}; k) = 0$. This directly yields a density result of the form

$$\lim_{k\to\infty} \left\| \iota^{(j)} - \mathscr{A}_k^{(j)} \right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})} = 0, \tag{4.2}$$

i.e., we have that $\{\Pi_k^{(j)}\}_k$ is dense in $\mathscr{H}^{\beta_j}(\Omega^{(j)})$ with respect to the $\mathscr{H}^{\alpha_j}(\Omega^{(j)})$-norm. Furthermore, if we additionally set $\mathscr{A}_0^{(j)} \equiv 0$ then we can rewrite the approximation operators using the difference operators $\Delta_k(\mathscr{A}^{(j)}) := \mathscr{A}_k^{(j)} - \mathscr{A}_{k-1}^{(j)} : \mathscr{H}^{\beta_j}(\Omega^{(j)}) \to H^{\alpha_j}(\Omega^{(j)})$ as

$$\mathscr{A}_J^{(j)} = \sum_{k=1}^{J} \left(\mathscr{A}_k^{(j)} - \mathscr{A}_{k-1}^{(j)} \right) = \sum_{k=1}^{J} \Delta_k(\mathscr{A}^{(j)})$$

and observe that, with this notation, (4.2) yields

$$\iota^{(j)} = \lim_{J\to\infty} \mathscr{A}_J^{(j)} = \sum_{k=1}^{\infty} \Delta_k(\mathscr{A}^{(j)}). \tag{4.3}$$

Next, for each j and each $k \in \mathbb{N}$, we fix a sequence $N^{(j)} : \mathbb{N}_0 \to \mathbb{N}_0$ and consider a family of discrete nested sets of points $\Xi_k^{(j)} = \left\{ \xi_{1;k}^{(j)}, \ldots, \xi_{N^{(j)}(k);k}^{(j)} \right\} \subseteq \Omega^{(j)}$ for $k \in \mathbb{N}$ with $\Xi_{k-1}^{(j)} \subseteq \Xi_k^{(j)}$ and $\Xi_0^{(j)} = \emptyset$. Based on these point sets we consider *quasi-interpolation operators* (see also Nobile *et al.*, 2008a, Equation 2.5)

$$\mathscr{Q}_k^{(j)} : C(\Omega^{(j)}) \to C(\Omega^{(j)}), \qquad f \mapsto \mathscr{Q}_k^{(j)}(f) = \sum_{n=1}^{N^{(j)}(k)} f(\xi_{n;k}^{(j)}) \phi_{n;k}^{(j)},$$

where $\phi_{n;k}^{(j)} \in C(\Omega^{(j)})$ for $1 \leqslant n \leqslant N^{(j)}(k)$ and $k \in \mathbb{N}$ are given functions. Furthermore, we define $\mathscr{Q}_0^{(j)} \equiv 0$. Again, we will interpret $Q_k^{(j)}$ as operators $Q_k^{(j)} : \mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})$. The corresponding difference operators are again defined for $k \in \mathbb{N}$ as $\Delta_k(\mathscr{Q}^{(j)}) := \mathscr{Q}_k^{(j)} - \mathscr{Q}_{k-1}^{(j)}$. The same telescoping sum argument as above shows that we can recover the quasi-interpolation operators from the difference operators via

$$\mathscr{Q}_J^{(j)} = \left(\mathscr{Q}_J^{(j)} - \mathscr{Q}_{J-1}^{(j)}\right) + \left(\mathscr{Q}_{J-1}^{(j)} - \mathscr{Q}_{J-2}^{(j)}\right) + \cdots + \left(\mathscr{Q}_1^{(j)} - \mathscr{Q}_0^{(j)}\right) = \sum_{k=1}^{J} \Delta_k(\mathscr{Q}^{(j)}). \tag{4.4}$$

We assume that these quasi-interpolation operators $\mathscr{Q}_k^{(j)}$ are also exact on the finite-dimensional subspaces $\Pi_k^{(j)} \subseteq \mathscr{H}^{\beta_j}(\Omega^{(j)}) \subseteq C(\Omega^{(j)})$, i.e., that they satisfy

$$\mathscr{Q}_k^{(j)}(p)(x) = \sum_{n=1}^{N^{(j)}(k)} p(\xi_{n;k}^{(j)})\phi_{n;k}^{(j)}(x), \qquad x \in \Omega^{(j)}, \quad p \in \Pi_k^{(j)}. \tag{4.5}$$

For a function $f \in \mathscr{H}^{\beta_j}(\Omega^{(j)})$ we have $\mathscr{A}_k^{(j)}(\iota^{(j)}f) \in \Pi_k^{(j)}$ showing $\mathscr{Q}_k^{(j)}(\mathscr{A}_k^{(j)}(\iota^{(j)}f)) = \mathscr{A}_k^{(j)}(\iota^{(j)}f)$ so that this directly gives rise to an error estimate of the form

$$\begin{aligned}
\left\|(\mathrm{Id} - \mathscr{Q}_k^{(j)})(\iota^{(j)}f)\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)})} &\leqslant \left\|(\mathrm{Id} - \mathscr{Q}_k^{(j)})\left[\iota^{(j)}f - \mathscr{A}_k^{(j)}(\iota^{(j)}f) + \mathscr{A}_k^{(j)}(\iota^{(j)}f)\right]\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)})} \\
&= \left\|(\mathrm{Id} - \mathcal{Q}_k^{(j)})\left[\iota^{(j)}f - \mathcal{A}_k^{(j)}(\iota^{(j)}f)\right]\right\|_{\mathcal{H}^{\alpha_j}(\Omega^{(j)})} \\
&\leqslant \left(1 + \left\|\mathscr{Q}_k^{(j)}\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})}\right)\left\|\iota^{(j)}f - \mathscr{A}_k^{(j)}(\iota^{(j)}f)\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)})} \\
&\leqslant \left(1 + \left\|\mathscr{Q}_k^{(j)}\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})}\right)\rho(\boldsymbol{\alpha}; k)\,\|u\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)})},
\end{aligned}$$

which is of particular interest as long as the *Lebesgue constant* $\left\|\mathscr{Q}_k^{(j)}\right\|_{\mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})}$ remains bounded or at least controllable as a function of k.

4.2 *Tensor products and abstract sampling inequalities*

After looking at essentially the univariate case we will now turn our attention to the multivariate case by employing tensor products and the Smolyak construction. Based upon this we will give a first abstract sampling inequality. Using the notation provided in the last subsection we start by introducing tensor product spaces, see also Aubin (2000) for more details,

$$\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes}) := \bigotimes_{j=1}^{d} \mathscr{H}^{\beta_j}(\Omega^{(j)}) \subseteq \mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes}) := \bigotimes_{j=1}^{d} \mathscr{H}^{\alpha_j}(\Omega^{(j)}).$$

Based on a finite monotone set $\Lambda \subseteq \mathbb{N}^d$, we can define Smolyak's formula for a general sequence of operator $\mathscr{B} = \{\mathscr{B}^{(j)}_{\lambda_j}\}$, where the indices satisfy $1 \leqslant j \leqslant d$ and $\boldsymbol{\lambda} \in \Lambda$ and where $\mathscr{B}^{(j)}_{\lambda_j} : \mathscr{H}^{\alpha_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})$, as

$$\mathbb{S}^{\otimes}_{\Lambda}(\mathscr{B}) \equiv \sum_{\boldsymbol{\lambda}\in\Lambda} \Delta_{\boldsymbol{\lambda}}(\mathscr{B}^{\otimes}) := \sum_{\boldsymbol{\lambda}\in\Lambda} \bigotimes_{j=1}^{d} \Delta_{\lambda_j}(\mathscr{B}^{(j)}) : \mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes}). \tag{4.6}$$

As in the previous section we will denote the restriction of $\mathbb{S}^{\otimes}_{\Lambda}(\mathscr{B})$ to $\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$ again simply by $\mathbb{S}^{\otimes}_{\Lambda}(\mathscr{B})$.

Applying this to the family of operators $\mathscr{A} = \{\mathscr{A}^{(j)}_k\}$ from the last subsection and using (4.3) we can directly derive the identity

$$\iota^{\otimes} = \sum_{\boldsymbol{\lambda}\in\mathbb{N}^d} \Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})$$

for the embedding operator $\iota^{\otimes} : \mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes})$.

Furthermore, we have by (4.5) that the quasi-interpolation operators $\mathscr{Q}^{(j)}_k$ are exact on $\Pi^{(j)}_k$. As the the spaces $\{\Pi^{(j)}_k\}_k$ are nested we can derive from Conrad & Marzouk (2013, Theorem 3.2) that the space $\Pi^{\otimes}_{\Lambda}$ on which Smolyak's construction is exact, i.e., on which we have

$$\mathbb{S}^{\otimes}_{\Lambda}(\mathscr{Q})(p) = \sum_{\boldsymbol{\lambda}\in\Lambda} \Delta_{\boldsymbol{\lambda}}(\mathscr{Q}^{\otimes})(p) = \sum_{\boldsymbol{\lambda}\in\Lambda} \bigotimes_{j=1}^{d} \Delta_{\lambda_j}(\mathscr{Q}^{(j)})(p) = \sum_{\boldsymbol{\lambda}\in\Lambda} \bigotimes_{j=1}^{d} \left(\mathscr{Q}^{(j)}_{\lambda_j} - \mathscr{Q}^{(j)}_{\lambda_j-1}\right)(p) = p$$

for all $p \in \Pi^{\otimes}_{\Lambda}$, satisfies

$$\bigcup_{\boldsymbol{\lambda}\in\Lambda} \bigotimes_{j=1}^{d} \Pi^{(j)}_{\lambda_j} \subseteq \Pi^{\otimes}_{\Lambda}. \tag{4.7}$$

From the definition we have

$$\iota^{\otimes} - \mathbb{S}^{\otimes}_{\Lambda}(\mathscr{A}) = \sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{(\otimes)}) : \mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes}). \tag{4.8}$$

In order to derive explicit bounds we can reduce this to the component case via

$$\begin{aligned}
\left\|\iota^{\otimes} - \mathbb{S}^{\otimes}_{\Lambda}(\mathscr{A})\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})\to\mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes})} &= \left\|\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{(\otimes)})\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})\to\mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes})} \\
&\leqslant \sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{(\otimes)})\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})\to\mathscr{H}^{\boldsymbol{\alpha}}(\Omega^{\otimes})} \\
&\leqslant \sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \prod_{j=1}^{d} \left\|\Delta_{\lambda_j}(\mathscr{A}^{(j)})\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)})\to\mathscr{H}^{\alpha_j}(\Omega^{j})}.
\end{aligned}$$

Compare also Chkifa *et al.* (2014, Remark 2.3) for a result for polynomial interpolation and Schillings & Schwab (2013, Theorem 5.2) for numerical quadrature. For $f \in \mathscr{H}^{\boldsymbol\beta}(\Omega^{\otimes})$ this gives, as in the univariate case, directly rise to an error estimate of the form

$$\begin{aligned}
&\left\|(\mathrm{Id} - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q}))(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \\
&\quad\leqslant \left\|(\mathrm{Id} - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q}))\left[\iota^{\otimes} f - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{A})(\iota^{\otimes} f) + \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{A})(\iota^{\otimes} f)\right]\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \\
&\quad= \left\|(\mathrm{Id} - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q}))\left[\iota^{\otimes} f - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{A})(\iota^{\otimes} f)\right]\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \\
&\quad\leqslant \left(1 + \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}\right) \left\|\iota^{\otimes} f - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{A})(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \\
&\quad\leqslant \left(1 + \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}\right) \sum_{\boldsymbol\lambda \in \Lambda^{\complement}} \prod_{j=1}^{d} \left\|\Delta_{\lambda_j}(\mathscr{A}^{(j)})\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})} \|f\|_{\mathscr{H}^{\boldsymbol\beta}(\Omega^{\otimes})}.
\end{aligned}$$

This immediately allows us to conclude a general abstract sampling inequality.

Proposition 4.1 Under the assumptions made throughout this section we have for $f \in \mathscr{H}^{\boldsymbol\beta}(\Omega^{\otimes})$ the abstract sampling inequality

$$\begin{aligned}
&\left\|\iota^{\otimes} f\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \\
&\quad\leqslant \left(1 + \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}\right) \sum_{\boldsymbol\lambda \in \Lambda^{\complement}} \prod_{j=1}^{d} \left\|\Delta_{\lambda_j}(\mathscr{A}^{(j)})\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to \mathscr{H}^{\alpha_j}(\Omega^{(j)})} \|f\|_{\mathscr{H}^{\boldsymbol\beta}(\Omega^{\otimes})} \\
&\qquad + \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}.
\end{aligned}$$

Proof. This simply follows from the considerations above and

$$\left\|\iota^{\otimes} f\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \leqslant \left\|\iota^{\otimes} f - \mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} + \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}.$$

□

Note that the bound in Proposition 4.1 can further be bounded by employing

$$\left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})(\iota^{\otimes} f)\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \leqslant \left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes}) \to \mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})} \left\|\iota^{\otimes} f\right\|_{\mathscr{H}^{\boldsymbol\alpha}(\Omega^{\otimes})}.$$

However, as we will see in the next subsection, there is a better estimate in specific situations.

4.3 *Specific sampling inequalities*

We will now apply the results of the last subsection to derive specific, more concrete sampling inequalities. To this end we will make some additional assumptions and specifications:

- The function spaces $\mathscr{H}^{\alpha_j}(\Omega^{(j)})$ are given by $\mathscr{H}^{\alpha_j}(\Omega^{(j)}) = C_b(\Omega^{(j)})$ for $1 \leqslant j \leqslant d$. The function spaces $\mathscr{H}^{\beta_j}(\Omega^{(j)})$ can either be $\mathscr{H}^{\beta_j}(\Omega^{(j)}) = C_b^{\beta_j}(\Omega^{(j)})$ or $\mathscr{H}^{\beta_j}(\Omega^{(j)}) = H^{\beta_j}(\Omega^{(j)})$. Though, in principle, other tensor product spaces are also possible. In the case of Sobolev spaces the embedding

$\mathscr{H}^{\beta_j}(\Omega^{(j)}) \subseteq L_\infty(\Omega^{(j)})$ will be satisfied by the Sobolev embedding theorem, which we, from now on, will implicitly assume to hold.

- We will assume that the underlying monotone set Λ is of the form $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $\ell \in \mathbb{N}$.
- We will assume that the univariate grid $\{\Xi_k^{(j)}\}_k$ are nested and their numbers of points $N^{(j)}(k)$ satisfy Assumption 3.17, i.e., we have $C_1 2^{\eta_j k} \leqslant N^{(j)}(k) \leqslant C_2 2^{\eta_j k}$. The latter is only required when expressing the bounds in terms of the number of points rather than in the threshold ℓ.

When employing Proposition 4.1 we need to find bounds on the three expressions

$$\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}, \qquad \|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})(\iota^\otimes f)\|_{L_\infty(\Omega^\otimes)}, \tag{4.9}$$

$$\sum_{\boldsymbol{\lambda}\in\Lambda^\complement} \|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^\otimes)\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \leqslant \sum_{\boldsymbol{\lambda}\in\Lambda^\complement}\prod_{j=1}^d \left\|\Delta_{\lambda_j}(\mathscr{A}^{(j)})\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)})\to L_\infty(\Omega^{(j)})}. \tag{4.10}$$

The second term in (4.9) can obviously be reduced to the first term. However, by the following observation a more sophisticated bound is possible. As each operator $\mathscr{Q}_{\lambda_j}^{(j)}$ requires only the values at the points $\Xi_{\lambda_j}^{(j)}$ of the function it is applied to the operator $\mathbb{S}_\Lambda^\otimes(\mathscr{Q})$, when applied to $\iota^\otimes f$ with $f \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)$, requires only the values of f at the grid $\Xi_{\boldsymbol{\omega}}^\otimes(\ell, d)$, see also Nobile *et al.* (2008a, p. 2419). In particular, if two functions f and $\widetilde{f}$ have the same values at $\Xi_{\boldsymbol{\omega}}^\otimes(\ell, d)$ they satisfy $\mathbb{S}_\Lambda^\otimes(\mathscr{Q})(\iota^\otimes f) = \mathbb{S}_\Lambda^\otimes(\mathscr{Q})(\iota^\otimes \widetilde{f})$. Hence, for a given $f \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)$ we may choose $\widetilde{f} \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)$ with $f|\Xi_{\boldsymbol{\omega}}^\otimes(\ell, d) = \widetilde{f}|\Xi_{\boldsymbol{\omega}}^\otimes(\ell, d)$ and $\|\iota^\otimes \widetilde{f}\|_{L_\infty(\Omega^\otimes)} = \|f\|_{\ell_\infty(\Xi_{\boldsymbol{\omega}}^\otimes(\ell,d))}$. This shows

$$\begin{aligned}
\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})(\iota^\otimes f)\|_{L_\infty(\Omega^\otimes)} &= \|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})(\iota^\otimes \widetilde{f})\|_{L_\infty(\Omega^\otimes)} \leqslant \|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}\|\iota^\otimes \widetilde{f}\|_{L_\infty(\Omega^\otimes)} \\
&= \|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}\|f\|_{\ell_\infty(\Xi_{\boldsymbol{\omega}}^\otimes(\ell,d))},
\end{aligned}$$

so that we indeed only have to determine an upper bound on $\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}$ and on the term in (4.10). The general bound in Proposition 4.1 hence becomes

$$\begin{aligned}
\left\|\iota^\otimes f\right\|_{L_\infty(\Omega^\otimes)} \leqslant{}& \left(1 + \left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}\right) \sum_{\boldsymbol{\lambda}\in\Lambda^\complement} \left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^\otimes)\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \|f\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)} \\
&+ \left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \|f\|_{\ell_\infty(\Xi_{\boldsymbol{\omega}}^\otimes(\ell,d))}.
\end{aligned} \tag{4.11}$$

To achieve this we need to make some more assumptions on the families of operators $\mathscr{Q}$ and $\mathscr{A}$. Hence, from now on, we will make the following assumptions on the quasi-interpolation operators $\mathcal{Q}_k^{(j)}$.

ASSUMPTION 4.2 There exist a universal constant $C_L > 1$ and a constant vector $\boldsymbol{\rho} \in \mathbb{N}_0^d$ such that the operators $\mathcal{Q}_k^{(j)}$ are bounded by

$$\left\|\mathcal{Q}_k^{(j)}\right\|_{L_\infty(\Omega^{(j)})\to L_\infty(\Omega^{(j)})} \leqslant C_L k^{\rho_j} \tag{4.12}$$

for all $k \in \mathbb{N}_0$ and all $1 \leqslant j \leqslant d$.

Next we assume that the best approximation operators $\mathscr{A}_k^{(j)}$ satisfy the following type of Jackson inequality.

ASSUMPTION 4.3 There exist a constant $C_A > 0$ and constant vectors $\boldsymbol{\gamma} \in \mathbb{R}_+^d$ and $\boldsymbol{\nu} \in \mathbb{N}_0^d$ such that

$$\left\| \iota^{(j)} - \mathscr{A}_k^{(j)} \right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to L_\infty(\Omega^{(j)})} \leqslant C_A 2^{-\gamma_j k} k^{\nu_j} \tag{4.13}$$

for all $k \in \mathbb{N}_0$ and all $1 \leqslant j \leqslant d$.

Obviously, the constant γ_j will depend on the smoothness β_j of the Sobolev spaces $\mathscr{H}^{\beta_j}(\Omega^{(j)})$, but for the time being we suppress this dependency. However, it is important to see the relation to the number of points $N^{(j)}(k)$ if Assumption 3.17 is satisfied. In this situation we have $2^{-k} \leqslant C_2^{1/\eta_j}[N^{(j)}(k)]^{-1/\eta_j}$ and $k \leqslant [\log_2 N^{(j)}(k) - \log_2 C_1]/\eta_j$, hence

$$\begin{aligned} \left\| \iota^{(j)} - \mathscr{A}_k^{(j)} \right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to L_\infty(\Omega^{(j)})} &\leqslant C_A 2^{-\gamma_j k} k^{\nu_j} \leqslant C_A C_2^{\gamma_j/\eta_j} [N^{(j)}(k)]^{-\gamma_j/\eta_j} \left(\frac{\log_2(N^{(j)}(k))}{\eta_j} \right)^{\nu_j}, \\ &\leqslant C[N^{(j)}(k)]^{-\gamma_j/\eta_j} \log^{\nu_j}(N^{(j)}(k)) \end{aligned} \tag{4.14}$$

with a constant C depending on $\boldsymbol{\eta}, \boldsymbol{\gamma}, \boldsymbol{\nu}$, so that the quotient $\alpha_j := \gamma_j/\eta_j$ should reflect the smoothness β_j of the space.

Assumption 4.3 particularly means that we also have

$$\left\| \iota^{(j)} - \mathscr{A}_k^{(j)} \right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)}) \to L_\infty(\Omega^{(j)})} \to 0, \qquad k \to \infty.$$

Assumptions 3.17 and 4.2 allow us to bound the terms in (4.9). We will do this only in the two specific situations $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$ as they describe the most relevant cases. The first case is often a result of *oversampling*, i.e., when we employ more points in our grids than necessary for reproducing the finite-dimensional approximation spaces. The second case corresponds often to *nonoversampling* where the number of grid points and the dimension of the univariate subspaces coincide. Nonetheless, the techniques provided in Section 3 would also allow us to deal with more general situations.

LEMMA 4.4 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $\ell \in \mathbb{N}$. Assume that Assumption 4.2 is satisfied with $\boldsymbol{\rho} \in \mathbb{N}_0^d$. Then $\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\|_{L_\infty(\Omega^\otimes) \to L_\infty(\Omega^\otimes)}$ can be bounded as follows.

- If $\boldsymbol{\rho} = \mathbf{0}$ then

$$\left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes) \to L_\infty(\Omega^\otimes)} \leqslant (2C_L)^d \frac{(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1)^d}{d!} \prod_{j=1}^d \omega_j^{-1}. \tag{4.15}$$

- If $\boldsymbol{\rho} = \mathbf{1}$ and $\boldsymbol{\omega} \in \mathbb{N}^d$ then there is a constant $C = C(\boldsymbol{\omega}, d)$ such that

$$\left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes) \to L_\infty(\Omega^\otimes)} \leqslant C(\boldsymbol{\omega}, d) \ell^{\|\boldsymbol{\omega}\|_1} \left[\log(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1)\right]^{d-1}. \tag{4.16}$$

Proof. In both cases we can start with estimating

$$\begin{aligned}
\left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} &= \left\|\sum_{\boldsymbol{\lambda}\in\Lambda}\Delta_{\boldsymbol{\lambda}}(\mathscr{Q}^\otimes)\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \\
&= \left\|\sum_{\boldsymbol{\lambda}\in\Lambda}\prod_{j=1}^d\left(\mathscr{Q}_{\lambda_j}^{(j)}-\mathscr{Q}_{\lambda_j-1}^{(j)}\right)\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \\
&\leqslant \sum_{\boldsymbol{\lambda}\in\Lambda}\prod_{j=1}^d\left\|\mathscr{Q}_{\lambda_j}^{(j)}-Q_{\lambda_j-1}^{(j)}\right\|_{L_\infty(\Omega^{(j)})\to L_\infty(\Omega^{(j)})} \\
&\leqslant \sum_{\boldsymbol{\lambda}\in\Lambda}\prod_{j=1}^d\left(C_L\left[\lambda_j^{\rho_j}+(\lambda_j-1)^{\rho_j}\right]\right) \\
&\leqslant (2C_L)^d\sum_{\boldsymbol{\lambda}\in\Lambda}\prod_{j=1}^d\lambda_j^{\rho_j}=(2C_L)^d\sum_{\boldsymbol{\lambda}\in\Lambda}\boldsymbol{\lambda}^{\boldsymbol{\rho}}.
\end{aligned}$$

In the first case, i.e., if $\boldsymbol{\rho}=\mathbf{0}$, we can conclude

$$\left\|\mathbb{S}_\Lambda^\otimes(\mathscr{Q})\right\|_{L_\infty(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \leqslant (2C_L)^d\#\Lambda\leqslant(2C_L)^d\frac{\left(\ell\omega_{\min}+\|\boldsymbol{\omega}\|_1\right)^d}{d!}\prod_{j=1}^d\omega_j^{-1},$$

where we have used (3.32). The case $\boldsymbol{\rho}=\mathbf{1}$ follows directly from Corollary 3.16. □

Next, we need to bound the term in (4.10), i.e., the sum over the norms of $\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^\otimes)$. This can be done as follows.

Lemma 4.5 Let $\Lambda=\mathscr{I}_{\boldsymbol{\omega}}(\ell,d)$ with $\boldsymbol{\omega}\in\mathbb{R}_+^d$ and $\ell\in\mathbb{N}$. If Assumption 4.3 is satisfied with $\boldsymbol{\gamma}\in\mathbb{R}_+^d$ and $\boldsymbol{\nu}\in\mathbb{N}_0^d$ then

$$\sum_{\boldsymbol{\lambda}\in\Lambda^\complement}\left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^\otimes)\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)\to L_\infty(\Omega^\otimes)}\leqslant C_A^d\prod_{j=1}^d\left[1+2^{-\gamma_j}\right]\sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})^\complement}\exp(-\log(2)\boldsymbol{\gamma}\cdot\boldsymbol{\lambda})(\boldsymbol{\lambda}+\mathbf{1})^{\boldsymbol{\nu}}. \tag{4.17}$$

Proof. We first bound the summands directly, using Assumption 4.3 as follows:

$$\begin{aligned}
&\left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^\otimes)\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^\otimes)\to L_\infty(\Omega^\otimes)} \\
&\leqslant \prod_{j=1}^d\left\|\Delta_{\lambda_j}(\mathscr{A}^{(j)})\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)})\to L_\infty(\Omega^0)} \\
&= \prod_{j=1}^d\left\|\mathscr{A}_{\lambda_j}^{(j)}-\mathscr{A}_{\lambda_j-1}^{(j)}\right\|_{\mathscr{H}^{\beta_j}(\Omega^{(j)})\to L_\infty(\Omega^{(j)})}
\end{aligned}$$

$$\leqslant \prod_{j=1}^{d}\left[C_A 2^{-\gamma_j\lambda_j}\lambda_j^{\nu_j} + C_A 2^{-\gamma_j(\lambda_j-1)}(\lambda_j-1)^{\nu_j}\right]$$

$$\leqslant \prod_{j=1}^{d}\left[C_A(1+2^{-\gamma_j})2^{-\gamma_j(\lambda_j-1)}\lambda_j^{\nu_j}\right]$$

$$= C_A^d\prod_{j=1}^{d}\left[1+2^{-\gamma_j}\right]\prod_{j=1}^{d}\left[2^{-\gamma_j(\lambda_j-1)}\lambda_j^{\nu_j}\right] = C_A^d\prod_{j=1}^{d}\left[1+2^{-\gamma_j}\right]2^{-\boldsymbol{\gamma}\cdot(\boldsymbol{\lambda}-\mathbf{1})}\boldsymbol{\lambda}^{\boldsymbol{\nu}}.$$

Hence, we have

$$\begin{aligned}\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}}\left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})\right\|_{\mathscr{H}^{\beta}(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})} &= \sum_{\boldsymbol{\lambda}\in\left(\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})+\{\mathbf{1}\}\right)^{\complement}}\left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})\right\|_{\mathscr{H}^{\beta}(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})}\\ &\leqslant C_A^d\prod_{j=1}^{d}(1+2^{-\gamma_j})\sum_{\boldsymbol{\lambda}\in\left(\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})+\{\mathbf{1}\}\right)^{\complement}}2^{-\boldsymbol{\gamma}\cdot(\boldsymbol{\lambda}-\mathbf{1})}\boldsymbol{\lambda}^{\boldsymbol{\nu}} \qquad (4.18)\\ &= C_A^d\prod_{j=1}^{d}\left[1+2^{-\gamma_j}\right]\sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})^{\complement}}2^{-\boldsymbol{\gamma}\cdot\boldsymbol{\lambda}}(\boldsymbol{\lambda}+\mathbf{1})^{\boldsymbol{\nu}}\\ &= C_A^d\prod_{j=1}^{d}\left[1+2^{-\gamma_j}\right]\sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})^{\complement}}\exp(-\log(2)\boldsymbol{\gamma}\cdot\boldsymbol{\lambda})(\boldsymbol{\lambda}+\mathbf{1})^{\boldsymbol{\nu}}.\end{aligned}$$

□

Note that we have $\gamma_j > 0$ and hence $1 + 2^{-\gamma_j} \leqslant 2$ for all $1 \leqslant j \leqslant d$ so that we could bound the remaining product by 2^d.

Next we need to bound the final sum in the statement of the last lemma. We will do this only in the special case of a weight vector $\boldsymbol{\omega}$ chosen as $\boldsymbol{\omega} = \boldsymbol{\gamma}$, where $\boldsymbol{\gamma}$ is the vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_d)^T$ from Assumption 4.3. Moreover, we will discuss only the cases where $\boldsymbol{\nu}$ from Assumption 4.3 satisfies either $\boldsymbol{\nu} = \mathbf{0}$ or $\boldsymbol{\nu} \in \mathbb{N}^d$. We start with the former.

Proposition 4.6 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{R}_+^d$ and $\ell \in \mathbb{N}$ satisfying $\ell \geq d/(\omega_{\min}\log(2))$. Let Assumption 4.3 with $\boldsymbol{\gamma} = \boldsymbol{\omega}$ and $\boldsymbol{\nu} = \mathbf{0}$ be satisfied. Then

$$\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}}\left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})\right\|_{\mathscr{H}^{\beta}(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})} \leqslant \frac{C_A^d d\log(2)^{d-1}}{(d-1)!}\left(\frac{1+2^{-\omega_{\min}}}{1-2^{-\omega_{\min}}}\right)^d\omega_{\min}^{d-1}\ell^{d-1}2^{-\ell\omega_{\min}}.$$

Proof. We use Corollary 3.9, i.e., (3.28) with $T = \ell\omega_{\min}$ and $c = -\log(2)$ and then the second statement of Lemma 3.14, i.e., (3.36) with $T = \log(2)\ell\omega_{\min} \geqslant d$. This yields

$$\sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(\ell\omega_{\min})^{\complement}} \exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{\lambda}) \leqslant \frac{1}{(d-1)!}\prod_{j=1}^{d}\frac{1}{\left(1-2^{-\omega_j}\right)}\int_{\log(2)\ell\omega_{\min}}^{\infty} s^{d-1}\exp(-s)\,\mathrm{d}s$$
$$\leqslant \frac{1}{(d-1)!}\prod_{j=1}^{d}\frac{1}{(1-2^{-\omega_j})}d2^{-\ell\omega_{\min}}(\log(2)\ell\omega_{\min})^{d-1}.$$

Inserting this into (4.17) and using the fact that

$$\frac{1+2^{-\omega_j}}{1-2^{\omega_j}} \leqslant \frac{1+2^{-\omega_{\min}}}{1-2^{-\omega_{\min}}}, \qquad 1 \leqslant j \leqslant d,$$

gives the stated result after rearranging the terms. □

Next we deal with the case $\boldsymbol{\nu} \in \mathbb{N}^d$. Here we have first of all the following result, which is, due to its more general nature, not as sharp as the one in Proposition 4.6.

PROPOSITION 4.7 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\ell \in \mathbb{N}$ and $\boldsymbol{\omega} \in \mathbb{R}^d_+$. Let Assumption 4.3 with $\boldsymbol{\nu} \in \mathbb{N}^d$ and $\boldsymbol{\gamma} = \boldsymbol{\omega}$ be satisfied. If $\omega_j > \nu_j$ for $1 \leqslant j \leqslant d$ then there is a constant $C = C(\boldsymbol{\omega}, \boldsymbol{\nu}) > 0$ such that

$$\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})\right\|_{\mathscr{H}^{\beta}(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})} \leqslant C\ell^{d+|\boldsymbol{\nu}|}2^{-\ell\omega_{\min}}.$$

Proof. We begin by directly bounding the sum in (4.18) using Corollary 3.6 with $c = -\log(2)$. Employing also $\boldsymbol{\omega} = \boldsymbol{\gamma}$ and, by (3.4), $\Lambda^{\complement} = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)^{\complement} \subseteq \mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}$ with $T = \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1$, we can conclude that

$$\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} 2^{-\boldsymbol{\omega}\cdot(\boldsymbol{\lambda}-\mathbf{1})}\boldsymbol{\lambda}^{\boldsymbol{\nu}} \leqslant 2^{\|\boldsymbol{\omega}\|_1}\sum_{\boldsymbol{\lambda}\in\mathscr{D}_{\boldsymbol{\omega}}(T)^{\complement}} \exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{\lambda})\boldsymbol{\lambda}^{\boldsymbol{\nu}}$$
$$\leqslant \log(2)^d C_2 2^{\|\boldsymbol{\omega}\|_1}\int_W \boldsymbol{x}^{\boldsymbol{\nu}}\exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{x})\prod_{j=1}^{d}\left(\omega_j + \frac{\nu_j}{(-\log(2))x_j}\right)\mathrm{d}\boldsymbol{x}$$
$$\leqslant \log(2)^d C_2 2^{\|\boldsymbol{\omega}\|_1}\int_W \boldsymbol{x}^{\boldsymbol{\nu}}\exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{x})\prod_{j=1}^{d}\left(\omega_j + \frac{\nu_j}{\log(2)x_j}\right)\mathrm{d}\boldsymbol{x}.$$

Here the set W from (3.14) contains particularly only cubes $[\boldsymbol{k}, \boldsymbol{k}+\mathbf{1})$ with $\boldsymbol{k} \in \mathbb{N}^d$, which means that any $\boldsymbol{x} \in W$ satisfies $\boldsymbol{x} \geqslant \mathbf{1}$. This and our assumption $\boldsymbol{\omega} \geqslant \boldsymbol{\nu}$ allow us first to further bound the product under the integral by

$$\prod_{j=1}^{d}\left(\omega_j + \frac{\nu_j}{\log(2)x_j}\right) \leqslant \prod_{j=1}^{d}\left(\omega_j + \frac{\nu_j}{\log(2)}\right) \leqslant \boldsymbol{\omega}^{\mathbf{1}} \prod_{j=1}^{d}\left(1 + \frac{1}{\log(2)}\right) \leqslant \boldsymbol{\omega}^{\mathbf{1}} \frac{(1+\log(2))^d}{\log(2)^d}.$$

This leads then to the bound

$$\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} 2^{-\boldsymbol{\omega}\cdot(\boldsymbol{\lambda}-\mathbf{1})}\boldsymbol{\lambda}^{\boldsymbol{\nu}} \leqslant C_2(1+\log(2))^d 2^{\|\boldsymbol{\omega}\|_1}\boldsymbol{\omega}^{\mathbf{1}} \int_W \boldsymbol{x}^{\boldsymbol{\nu}} \exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

As we have non-negative integrands we can make use of $W \subseteq [\mathscr{D}_{\boldsymbol{\omega}}](T)^{\complement} \subseteq \mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement}$ and then use Lemma 3.11 to derive, with $T = \ell\omega_{\min} + \|\boldsymbol{\omega}\|_1$,

$$\begin{aligned}\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} 2^{-\boldsymbol{\omega}\cdot(\boldsymbol{\lambda}-\mathbf{1})}\boldsymbol{\lambda}^{\boldsymbol{\nu}} &\leqslant C_2(1+\log(2))^d 2^{\|\boldsymbol{\omega}\|_1}\boldsymbol{\omega}^{\mathbf{1}} \int_{\mathscr{E}_{\boldsymbol{\omega}}(T)^{\complement}} \exp(-\log(2)\boldsymbol{\omega}\cdot\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\nu}}\mathrm{d}\boldsymbol{x} \\ &\leqslant C_2(1+\log(2))^d 2^{\|\boldsymbol{\omega}\|_1}\boldsymbol{\omega}^{\mathbf{1}} T^{d+|\boldsymbol{\nu}|}2^{-T}\omega^{-(\boldsymbol{\nu}+\mathbf{1})} \int_{\mathscr{E}_{\mathbf{1}}(1)^{\complement}} \boldsymbol{x}^{\boldsymbol{\nu}} 2^{-\|\boldsymbol{x}\|_1+1}\mathrm{d}\boldsymbol{x} \\ &= C(\boldsymbol{\nu},\boldsymbol{\omega})\ell^{d+|\boldsymbol{\nu}|}2^{-\ell\omega_{\min}}.\end{aligned}$$

□

At this point we can give a first summary of our achievements. As mentioned above we only consider the two cases $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$. Again, we start with the former.

THEOREM 4.8 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{R}^d_+$ and $\ell \in \mathbb{N}$. Let Assumptions 4.2 and 4.3 with $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\gamma} = \boldsymbol{\omega}$ be satisfied. Then there is a constant $C = C(\boldsymbol{\omega}, d)$, such that for all $f \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$ and all $\ell \geqslant d/(\omega_{\min}\log(2))$, we have

$$\left\|\iota^{\otimes} f\right\|_{L_\infty(\Omega^{\otimes})} \leqslant C C_L^d C_A^d \ell^{2d-1} 2^{-\ell\omega_{\min}} \|f\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})} + C C_L^d \ell^d \|f\|_{\ell_\infty(\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell,d))}, \tag{4.19}$$

where C_A and C_L are the constants from (4.13) and (4.12), respectively.

Proof. We know from Proposition 4.6 that

$$\sum_{\boldsymbol{\lambda}\in\Lambda^{\complement}} \left\|\Delta_{\boldsymbol{\lambda}}(\mathscr{A}^{\otimes})\right\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})} \leqslant C_A^d C(\omega_{\min}, d)\ell^{d-1}2^{-\ell\omega_{\min}}$$

and, using $\boldsymbol{\omega} = \boldsymbol{\gamma}$, from Lemma 4.4 that

$$\left\|\mathbb{S}_{\Lambda}^{\otimes}(\mathscr{Q})\right\|_{L_\infty(\Omega^{\otimes})\to L_\infty(\Omega^{\otimes})} \leqslant C_L^d C(\boldsymbol{\omega}, d)\ell^d.$$

The statement then easily follows from (4.11). □

Before dealing with the case $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$ we want to express the above result (4.19) using rather the number of points $N := \#\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d)$ in our sparse grid $\Xi_{\boldsymbol{\omega}}^{\otimes}(\ell, d)$ than the parameter ℓ. To this end we assume now that Assumption 3.17 holds and we recall that we expect a connection between the vectors $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ of the form $\alpha_j := \gamma_j/\eta_j \geqslant 1$ for $1 \leqslant j \leqslant d$, as this quotient will reflect the smoothness of β_j of our Sobolev spaces, see (4.14). As we also have $\boldsymbol{\gamma} = \boldsymbol{\omega}$ we will make the assumption that $\eta_j = \omega_j/\alpha_j$ for $1 \leqslant j \leqslant d$. Hence, the bounds from Proposition 3.18 can be used. We have in particular

$$2^{-\ell\omega_{\min}} \leqslant C\ell^{(d-1)\alpha_{\min}} N^{-\alpha_{\min}} \qquad \text{and} \qquad \ell \leqslant C \log N. \tag{4.20}$$

With this we have the following final result in the situation of $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$.

Theorem 4.9 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\boldsymbol{\omega} \in \mathbb{R}^d_+$ and $\ell \in \mathbb{N}$. Let Assumptions 3.17, 4.2 and 4.3 with $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\gamma} = \boldsymbol{\omega}$ be satisfied. Let $\alpha_j := \omega_j/\eta_j$ for $1 \leqslant j \leqslant d$ and let $\alpha_{\min} := \min \alpha_j$. Then there is a constant $C = C(\boldsymbol{\omega}, d, \boldsymbol{\alpha})$, such that for all $f \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$, $\boldsymbol{\beta} \in \mathbb{N}^d$ and all $\ell \geqslant d/(\omega_{\min}\log(2))$, we have

$$\left\| \iota^{\otimes} f \right\|_{L_\infty(\Omega^{\otimes})} \leqslant C(\log N)^{2d-1+(d-1)\alpha_{\min}} N^{-\alpha_{\min}} \|f\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})} + C(\log N)^d \|f\|_{\ell_\infty(\Xi^{\otimes}_{\boldsymbol{\omega}}(\ell,d))}.$$

In the same fashion we can now establish a sampling inequality in the case of $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$. Here we have the following result.

Theorem 4.10 Let $\Lambda = \mathscr{I}_{\boldsymbol{\omega}}(\ell, d)$ with $\ell \in \mathbb{N}$ and $\boldsymbol{\omega} \in \mathbb{N}^d$, satisfying $\omega_j \geqslant 2$ for $1 \leqslant j \leqslant d$. Let Assumptions 3.17, 4.2 and 4.3 with $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$ and $\boldsymbol{\gamma} = \boldsymbol{\omega}$ be satisfied. Let $\alpha_j = \omega_j/\eta_j$, $1 \leqslant j \leqslant d$. Then, there is a constant $C = C(\boldsymbol{\omega}, d) > 0$ such that for all $f \in \mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$ we have

$$\begin{aligned}
\|\iota^{\otimes} f\|_{L_\infty(\Omega^{\otimes})} &\leqslant C\ell^{\|\boldsymbol{\omega}\|_1+2d} \log(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1)^{d-1} 2^{-\ell\omega_{\min}} \|f\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})} \\
&\quad + C\ell^{\|\boldsymbol{\omega}\|_1} \log(\ell\omega_{\min} + \|\boldsymbol{\omega}\|_1)^{d-1} \|f\|_{\ell_\infty(\Xi^{\otimes}_{\boldsymbol{\omega}}(\ell,d))} \\
&\leqslant C(\log N)^{\|\boldsymbol{\omega}\|_1+2d+\alpha_{\min}(d-1)} (\log\log N)^{d-1} N^{-\alpha_{\min}} \|f\|_{\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})} \\
&\quad + C(\log N)^{\|\boldsymbol{\omega}\|_1} (\log\log N)^{d-1} \|f\|_{\ell_\infty(\Xi^{\otimes}_{\boldsymbol{\omega}}(\ell,d))}.
\end{aligned}$$

Proof. The first bound follows again from the general bound (4.11) using, this time, Proposition 4.7 and the second part of Lemma 4.4. The second bound then follows again using (4.20), the relation between N and ℓ. □

4.4 *Examples*

Our first example concerns uniform grids. We consider $\Omega^{(j)} = I = [-1, 1]$ for $1 \leqslant j \leqslant d$ and choose $\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$ to be $\mathcal{H}^{\boldsymbol{\beta}}(\Omega^{\otimes}) = C_b^{\boldsymbol{\beta}}(\Omega^{\otimes})$ with $\boldsymbol{\beta} \in \mathbb{N}^d$. From Wendland (2005, Proposition 11.7) we know that a univariate grid size h, satisfying $h \leqslant Cm^{-2}$ allows us to define a stable quasi-interpolation operator $\mathcal{Q}_k^{(j)}$ reproducing polynomials with degree less than or equal to m with a Lebesgue constant bounded by 2. For m of the form $m = m_k = 2^{k-1}$ the univariate points are chosen to be the uniform gridded points with grid size $h_k := 2^{-2(k-1)+\lceil \log_2(C) \rceil}$ to satisfy $h_k \leqslant Cm_k^{-2}$. To be more precise they are defined by

$$\Xi_k^{(j)} \equiv \Xi_k = \left\{ \xi_{n;k}^{(j)} = -1 + 2(n-1)h_k, 1 \leqslant n \leqslant N^{(j)}(k) = \frac{1}{h_k} + 1 \right\}.$$

As the number of points is hence given by $N_k^{(j)} = 2^{2(k-1)-\lceil\log_2 C\rceil+1}$ we can, in Assumption 3.17, choose $C_1 = 2^{-2-\lceil\log_2(C)\rceil}$ and $C_2 = 2^{-1-\lceil\log_2(C)\rceil}$ and $\eta_j = 2$. By construction we have $\Pi_k^{(j)} = \pi_{2^{k-1}}$ as the univariate polynomials of degree at most 2^{k-1} for $k \in \mathbb{N}$. Furthermore, we have also by construction that $\|\mathscr{Q}_k^{(j)}\|_{L_\infty(I)\to L_\infty(I)} \leqslant 2$, i.e., that in Assumption 4.2 we have $\boldsymbol{\rho} = \mathbf{0}$. Finally, we have for the best approximation operator $\mathscr{A}_k^{(j)} : C_b^{\beta_j}(I) \to C_b(I)$ the estimate

$$\|\iota^{(j)} - \mathscr{A}_k^{(j)}\|_{W_\infty^{\beta_j}(I)\to L_\infty(I)} \leqslant C2^{-\beta_j k},$$

which implies for Assumption 4.3 we have $\boldsymbol{\gamma} = \boldsymbol{\beta}$ and $\boldsymbol{\nu} = \mathbf{0}$.

We end this section with an example, which is based on material provided in Rieger & Wendland (2017) and which also explains why we are particularly interested in the cases $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\rho} = \boldsymbol{\nu} = \mathbf{1}$. Again, we let $\Omega^{(j)} = I = [-1, 1]$ for $1 \leqslant j \leqslant d$ and choose $\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes})$ to be $\mathscr{H}^{\boldsymbol{\beta}}(\Omega^{\otimes}) = C_b^{\boldsymbol{\beta}}(\Omega^{\otimes})$ with $\boldsymbol{\beta} \in \mathbb{N}^d$. The univariate points are chosen to be the Chebyshev points defined by

$$\Xi_k^{(j)} \equiv \Xi_k = \left\{\xi_{n;k}^{(j)} = -\cos\left(\frac{\pi(n-1)}{m_{k+p}-1}\right), 1 \leqslant n \leqslant m_{k+p}\right\},$$

where $p \in \mathbb{N}_0$ is, if $p > 0$, an *oversampling offset* and $m_{k+p} = 2^{k+p-1} + 1 = N^{(j)}(k)$ the number of points. These sets are obviously nested. In the sense of Assumption 3.17 we have $C_1 = C_1(p) = 2^{p-1}$, $C_2 = C_2(p) = 2^p$ and, most importantly, $\eta_j = 1$, $1 \leqslant j \leqslant d$.

We choose the finite-dimensional spaces $\Pi_k^{(j)} = \pi_{m_k-1}$ as the univariate polynomials of degree at most $m_k - 1 = 2^{k-1}$, $k \in \mathbb{N}$. Then, we know from Rieger & Wendland (2017) that we have operators $\mathscr{Q}_k^{(j)}$, satisfying

$$\|\mathscr{Q}_k^{(j)}\|_{L_\infty(I)\to L_\infty(I)} \leqslant \begin{cases} 1 + \frac{5\pi}{2^{p+1}} =: C_L(p) & \text{if } p \geqslant 1, \\ \frac{2}{\pi}\log(m_k - 1) + 1 \leqslant Ck & \text{if } p = 0. \end{cases}$$

This means that we have, in the sense of Assumption 4.2, for $1 \leqslant j \leqslant d$,

$$\rho_j = \begin{cases} 0 & \text{if } p \geqslant 1, \\ 1 & \text{if } p = 0. \end{cases}$$

Finally, also from Rieger & Wendland (2017), we have a best approximation operator $\mathscr{A}_k^{(j)} : C_b^{\beta_j}(I) \to C_b(I)$ which satisfies

$$\|\iota^{(j)} - \mathscr{A}_k^{(j)}\|_{W_\infty^{\beta_j}(I)\to L_\infty(I)} \leqslant \begin{cases} C2^{-\beta_j k} & \text{if } p \geqslant 1, \\ C2^{-\beta_j k}k & \text{if } p = 0. \end{cases}$$

In the sense of Assumption 4.3 we have for $1 \leqslant j \leqslant d$ in both cases $\gamma_j = \beta_j$ and then $\nu_j = 0$ if $p \geqslant 1$, and $\nu_j = 1$ if $p = 0$.

Hence, we can apply Theorems 4.9 and 4.10 with $\boldsymbol{\gamma} = \boldsymbol{\omega} = \boldsymbol{\beta} = \boldsymbol{\alpha}$ and $\boldsymbol{\eta} = \mathbf{1}$ for all $p \in \mathbb{N}_0$, and the just defined $\boldsymbol{\rho}$ and $\boldsymbol{\nu}$.

Funding

Christian Rieger would like to thank the Deutsche Forschungsgemeinschaft (DFG) for partial support through the CRC 1060, The Mathematics of Emergent Effects.

References

Arcangéli, R., De Silanes, M. C. L. & Torrens, J. J. (2007) An extension of a bound for functions in Sobolev spaces, with applications to (m, s)-spline interpolation and smoothing. *Numer. Math.*, **107**, 181–211.

Aubin, J. (2000) *Applied Functional Analysis*, 2nd edn. New York: Wiley.

Baxter, B. & Brummelhuis, R. (2011) Functionals of exponential brownian motion and divided differences. *J. Comput. Appl. Math.*, **236**, 424–433. International Workshop on Multivariate Approximation and Interpolation with Applications (MAIA 2010).

Beged-dov, A. G. (1972) Lower and upper bounds for the number of lattice points in a simplex. *SIAM J. Appl. Math.*, **22**, 106–108.

Cheney, E. W. & Light, W. P. (1985) *Approximation Theory in Tensor Product Spaces*. Berlin: Springer.

Chkifa, A., Cohen, A. & Schwab, C. (2014) High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs. *Found. Comput. Math.*, **14**, 601–633.

Conrad, P. R. & Marzouk, Y. M. (2013) Adaptive Smolyak pseudospectral approximations. *SIAM J. Sci. Comput.*, **35**, A2643–A2670.

De Boor, C. (2005) Divided differences. *Surveys in Approximation Theory 46*, **1**, 46–69.

Griebel, M., Oeltz, D. & Vassilevski, P. (2006) Space-time approximation with sparse grids. *SIAM J. Sci. Comput.*, **28**, 701–727.

Griebel, M. & Oeltz, D. (2007) A sparse grid space-time discretization scheme for parabolic problems. *Computing*, **81**, 1–34.

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

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

Hardy, M. (2006) Combinatorics of partial derivatives. *Electr. J. Comb.*, **13**, 1–13.

Hendriks, H. & Landsman, Z. (2017) A generalization of multivariate Pareto distributions: tail risk measures, divided differences and asymptotics. *Scand. Actuar. J.*, **2017**, 785–803.

Lesch, M. (2017) Divided differences in noncommutative geometry: rearrangement lemma, functional calculus and expansional formula. *J. Noncommut. Geom.*, **11**, 193–223.

Madych, W. R. (2006) An estimate for multivariate interpolation. *J. Approx. Theory*, **142**, 116–128.

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

Nobile, F., Tempone, R. & Webster, C. G. (2008b) A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, **46**, 2309–2345.

Olsen, P., Goel, V., Micchelli, C. & Hershey, J. (2010) Modeling posterior probabilities using the linear exponential family. *11th Annual Conference of the International Speech Communication Association 2010 (INTERSPEECH 2010)*. International *Speech Communication* Association (ISCA), pp. 2994–2997.

Rauhut, H. & Schwab, C. (2017) Compressive sensing Petrov–Galerkin approximation of high-dimensional parametric operator equations. *Math. Comp.*, **86**, 661–700.

Rieger, C. & Wendland, H. (2017) Sampling inequalities for sparse grids. *Numer. Math.*, **136**, 439–466.

Schillings, C. & Schwab, C. (2013) Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. *Inverse Problems*, **29**, 065011.

Schreiber, A. (2000) Die Methode von Smolyak bei der Multivariaten Interpolation. *Ph.D. Thesis*, Germany: Universität Göttingen.

Steinwart, I. & Christmann, A. (2008) *Support Vector Machines*. New York: Springer.

Tempone, R. & Wolfers, S. (2018) Smolyak's algorithm: a powerful black box for the acceleration of scientific computations. *Sparse Grids and Applications–Miami 2016* (J. Garcke, D. Pflüger, C. G. Webster & G. Zhang eds). Lecture Notes in Computational Science and Engineering, vol. 123. Cham: Springer.

Wendland, H. (2005) *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge, UK: Cambridge University Press.

Wendland, H. & Rieger, C. (2005) Approximate interpolation with applications to selecting smoothing parameters. *Numer. Math.*, **101**, 643–662.