Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

M. Griebel and J. Oettershagen

On tensor product approximation of analytic functions

INS Preprint No. 1512

2015

On tensor product approximation of analytic functions

Michael Griebel[a,b], Jens Oettershagen[a,*]

[a] *Institute for Numerical Simulation, University of Bonn*
[b] *Fraunhofer Institute for Algorithms and Scientific Computing*

Abstract

We prove sharp, two-sided bounds on sums of the form $\sum_{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\mathcal{D}_{\boldsymbol{a}}(T)}\exp(-\sum_{j=1}^d a_jk_j)$, where $\mathcal{D}_{\boldsymbol{a}}(T) := \{\boldsymbol{k}\in\mathbb{N}_0^d : \sum_{j=1}^d a_jk_j \le T\}$ and $\boldsymbol{a}\in\mathbb{R}_+^d$. These sums appear in the error analysis of tensor product approximation, interpolation and integration of d-variate analytic functions. Examples are tensor products of univariate Fourier-Legendre expansions [6] or interpolation and integration rules at Leja points [13, 40, 41]. Moreover, we discuss the limit $d\to\infty$, where we prove both, algebraic and sub-exponential upper bounds. As an application we consider tensor products of Hardy spaces, where we study convergence rates of a certain truncated Taylor series, as well as of interpolation and integration using Leja points.

1. Introduction

Recently, the approximation, interpolation and integration of analytic functions has drawn a lot of interest, especially in the area of uncertainty quantification [3, 4, 17, 29]. Among the most popular approaches are generalized sparse grids, which use tensor products of certain univariate approximation schemes, like orthogonal polynomial expansions [6, 16, 25], Tschebyscheff interpolation [41, 42], Gaussian and Clenshaw-Curtis quadrature [41, 42], Taylor expansions [16, 17, 55] or interpolation at Leja points [14, 13, 40, 41]. For integration problems also special quasi–Monte Carlo methods have been developed [22, 23], which are able to achieve algebraic rates of convergence $\mathcal{O}(N^{-r})$ of arbitrarily high order.[1] Moreover, there exist hybrid methods, which allow unstructured point sets to project analytic functions onto a tensor product basis by a least squares fitting approach [15, 39]. Finally, there are kernel based methods which use orthogonal projections onto a certain basis of a given reproducing kernel Hilbert space of smooth functions, see e.g. [29, 46].

In this paper, we will study approximation algorithms that employ sparse tensor products of univariate approximation schemes which on the one hand allow for exponential convergence and on the other hand are maximally nested, i.e. on each level only one additional function(al) evaluation is needed. This differs from classical approaches like, e.g., Clenshaw-Curtis quadrature [27, 43] or piecewise linear splines [10], where the number of point evaluations usually doubles from level to level. The associated sparse grid or Smolyak methods were originally tailored to function spaces with dominating, but finite mixed smoothness, e.g. H^r_{mix}. They have been thoroughly analyzed

*Corresponding author, *Email*: `oettershagen@ins.uni-bonn.de`. *Phone*: +49 228 733430.
Address: Institute for Numerical Simulation, Universität Bonn. Wegelerstr. 6, 53115 Bonn, Germany.
[1] Recently, also super-algebraic rates of convergence have been proven for analytic functions in [51].

in this setting [10, 18, 49, 57] and sharp upper and lower bounds are available. However, analytic tensor product spaces and their approximability properties are not that well understood yet, albeit there has been steady progress [5, 6, 33, 34, 37, 42, 45, 55].

To this end, we consider the general problem of approximating a bounded linear operator $I_d : \mathcal{H}^{(d)} \to \mathcal{G}$ between the d-fold tensor product of Banach spaces[2] of univariate analytic functions $\mathcal{H}^{(d)} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_d$ and a normed linear space $\mathcal{G}$. Often, I_d is also referred to as *solution operator*, see e.g. [56]. We assume that I_d has a representation as an infinite series, i.e.

$$I_d(f) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^d} \Delta_{\boldsymbol{k}}(f), \tag{1.1}$$

where $\Delta_{\boldsymbol{k}} : \mathcal{H}^{(d)} \to \mathcal{G}$ is also bounded and linear and requires the evaluation of exactly one (additional) linear functional $L_{\boldsymbol{k}} : \mathcal{H}^{(d)} \to \mathbb{R}$, i.e. $\Delta_{\boldsymbol{k}}(f) = L_{\boldsymbol{k}}(f)\varphi_{\boldsymbol{k}}$, where $\varphi_{\boldsymbol{k}} \in \mathcal{G}$.

It is natural to discretize I_d by truncating the series (1.1), i.e.

$$\mathcal{A}_T(f) := \sum_{\boldsymbol{k} \in \mathcal{F}(T)} \Delta_{\boldsymbol{k}}(f) \approx I_d(f), \tag{1.2}$$

where $\mathcal{F}(T) \subset \mathbb{N}_0^d$ is a finite index set parametrized by $T \in \mathbb{R}_{\geq 0} := \{x \in \mathbb{R} : x \geq 0\}$, which exhausts the whole $\mathbb{N}_0^d$ as $T \to \infty$ and fulfills the conditions[3]

$$\boldsymbol{k} \leq \boldsymbol{v} \wedge \boldsymbol{v} \in \mathcal{F}(T) \Rightarrow \boldsymbol{k} \in \mathcal{F}(T). \tag{1.3}$$

This means that $\mathcal{F}(T)$ has no holes and that the approximation algorithm $\mathcal{A}_T$ converges for every $f \in \mathcal{H}^{(d)}$ to I_d, as T tends to infinity. Then, the error of $\mathcal{A}_T$ can be bounded through

$$\|I_d(f) - \mathcal{A}_T(f)\|_{\mathcal{G}} \leq \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{F}(T)} \|\Delta_{\boldsymbol{k}}(f)\|_{\mathcal{G}} \leq \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{F}(T)} \|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)} \to \mathcal{G}} \|f\|_{\mathcal{H}^{(d)}}, \tag{1.4}$$

where we used the triangle-inequality and $\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)} \to \mathcal{G}} = \sup_{\|f\|_{\mathcal{H}^{(d)}} \leq 1} \|\Delta_{\boldsymbol{k}}(f)\|_{\mathcal{G}}$, which implies $\|\Delta_{\boldsymbol{k}} f\|_{\mathcal{G}} \leq \|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)} \to \mathcal{G}} \|f\|_{\mathcal{H}^{(d)}}$ by a simple scaling of the unit ball of $\mathcal{H}^{(d)}$.

Examples for the setting (1.1) are given in the following, see also Section 5.2:

Fourier-Legendre series: $I_d : \mathcal{H}^{(d)} \to L^2([-1,1]^d)$ is the embedding operator and the information is given by $L_{\boldsymbol{k}}(f) = \langle f, \varphi_{\boldsymbol{k}} \rangle_{L^2([-1,1]^d)}$, where $\varphi_{\boldsymbol{k}} = \varphi_{k_1} \ldots \varphi_{k_d}$ denotes the tensor product of univariate Legendre polynomials.

Taylor series: $I_d : \mathcal{H}^{(d)} \to \mathcal{C}_d$ is the embedding operator, where the space of continuous functions $\mathcal{C}_d := C^0([-1,1]^d)$ is endowed with the (uniform) L^∞-norm. Here, $L_{\boldsymbol{k}}(f) = \frac{1}{\boldsymbol{k}!} \frac{\partial^{|\boldsymbol{k}|_1}}{\partial \boldsymbol{x}^{\boldsymbol{k}}} f(\boldsymbol{x})_{\boldsymbol{x}=\boldsymbol{0}}$ and $\varphi_{\boldsymbol{k}} = \prod_{j=1}^d x_j^{k_j}$.

Moreover, if $\Delta_{\boldsymbol{k}}(f)$ is allowed to re-use the information $L_{\boldsymbol{v}}(f)$ for all component-wise smaller multi-indices $\boldsymbol{v} \leq \boldsymbol{k}$, i.e. $\Delta_{\boldsymbol{k}}(f) = \sum_{\boldsymbol{v} \leq \boldsymbol{k}} c_{\boldsymbol{v},\boldsymbol{k}} L_{\boldsymbol{v}}(f) \varphi_{\boldsymbol{v}}$, tensor products of hierarchical interpolation or integration rules also fit into the setting (1.1). Of course, there are then additional floating

[2] which has to be equipped with a suitable crossnorm, see e.g. [30].

[3] The notation $\boldsymbol{k} \leq \boldsymbol{v}$ is short for the component-wise relation $k_j \leq v_j$ for all $j \in \{1, \ldots, d\}$.

point and memory operations needed to evaluate $\Delta_{\boldsymbol{k}}$, but one can easily precompute [27, 44] the necessary quantities in (1.2), so that $L_{\boldsymbol{k}}$ has to be evaluated only once for each $\boldsymbol{k} \in \mathcal{F}(T)$, i.e.

$$\begin{aligned} \mathcal{A}_T(f) &= \sum_{\boldsymbol{k}\in\mathcal{F}(T)} \Delta_{\boldsymbol{k}}(f) = \sum_{\boldsymbol{k}\in\mathcal{F}(T)} \sum_{\substack{\boldsymbol{v}\in\mathcal{F}(T)\\ \boldsymbol{v}\le\boldsymbol{k}}} L_{\boldsymbol{v}}(f) c_{\boldsymbol{v},\boldsymbol{k}} \varphi_{\boldsymbol{v},\boldsymbol{k}} \\ &= \sum_{\boldsymbol{v}\in\mathcal{F}(T)} L_{\boldsymbol{v}}(f) \sum_{\substack{\boldsymbol{k}\in\mathcal{F}(T)\\ \boldsymbol{k}\ge\boldsymbol{v}}} c_{\boldsymbol{v},\boldsymbol{k}} \varphi_{\boldsymbol{v},\boldsymbol{k}} =: \sum_{\boldsymbol{v}\in\mathcal{F}(T)} L_{\boldsymbol{v}}(f) \tilde{\varphi}_{\boldsymbol{v}}. \end{aligned} \tag{1.5}$$

Examples for such a setting are as follows, see also Section 5.3 and 5.4 for more details:

Polynomial interpolation: $I_d : \mathcal{H}^{(d)} \to \mathcal{C}_d$ is the embedding operator. Here, $L_{\boldsymbol{k}}(f) = f(\boldsymbol{x}_{\boldsymbol{k}})$ are function values at points $\boldsymbol{x}_{\boldsymbol{k}} \in [-1,1]^d$ and the $\varphi_{\boldsymbol{k}}$ are elements of a hierarchical tensor product polynomial basis [13, 41].

Integration: $I_d : \mathcal{H}^{(d)} \to \mathbb{R}$ is a linear functional, e.g. $I_d(f) = \int_{[-1,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$, the information is given as point evaluations $L_{\boldsymbol{k}}(f) = f(\boldsymbol{x}_{\boldsymbol{k}})$ and $\varphi_{\boldsymbol{k}} \in \mathbb{R}$ denotes the associated product of hierarchical integration weights [27, 43].

This paper deals with the setting, where for a given vector of non-negative real numbers $\boldsymbol{a} \in \mathbb{R}^d_+$ an upper bound of the form $\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)}\to\mathcal{G}} \preceq_{d,\boldsymbol{a}} \exp(-\sum_{j=1}^d a_j k_j)$ can be shown. Here, $\preceq_{d,\boldsymbol{a}}$ is short[4] for the existence of a d- and $\boldsymbol{a}$-dependent constant $c_{d,\boldsymbol{a}} > 0$ such that

$$\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)}\to\mathcal{G}} \le c_{d,\boldsymbol{a}} \exp\left(-\sum_{j=1}^d a_j k_j\right) \qquad \text{for all } \boldsymbol{k} \in \mathbb{N}_0^d, \tag{1.6}$$

which is often the case if $\mathcal{H}^{(d)}$ is a space of analytic functions.[5] If, in addition, also a lower bound with the same $\boldsymbol{k}$-asymptotics as the upper bound (1.6) is available, one can deduce that an asymptotically optimal index set $\mathcal{F}(T)$ (with respect to the bound on the right hand side of (1.4)) is of the form

$$\mathcal{F}(T) := \mathcal{D}_{\boldsymbol{a}}(T) := \left\{ \boldsymbol{k} \in \mathbb{N}_0^d : \sum_{j=1}^d a_j k_j \le T \right\}. \tag{1.7}$$

However, such a lower bound is often hard to obtain, or not even present at all. Therefore we will here and in the following only assume (1.6) and take the index set (1.7) as given without requiring its optimality.

Here we note that in certain cases, where $\mathcal{H}^{(d)}$ and $\mathcal{G}$ are Hilbert spaces and certain orthogonality properties of $\Delta_{\boldsymbol{k}}$ are present, squaring the left hand side and the summands on the right hand side allows even for equality in (1.4). All of our following results can be directly applied to this case as well, with $2\boldsymbol{a}$ instead of $\boldsymbol{a}$.

[4] Consequently, $A \asymp_{d,\boldsymbol{a}} B$ is short for $A \preceq_{d,\boldsymbol{a}} B \wedge B \preceq_{d,\boldsymbol{a}} A$.

[5] Note that anisotropic sparse grid discretizations in mixed Sobolev spaces also may allow for such a bound [28]. Then, however, the evaluation of $\Delta_{\boldsymbol{k}}$ involves $\approx 2^{\sum_{j=1}^d k_j}$ function values, which does not fit to our cost model.

Anyway, combining (1.4) with (1.6), where the vector $\boldsymbol{a} = (a_1, \ldots, a_d)$ of positive and ordered weights is prescribed, the asymptotic convergence rate of the algorithm $\mathcal{A}_T$ with $\mathcal{F}(T)$ given in (1.7) can be bounded from above by the sum

$$\sup_{\|f\|_{\mathcal{H}^{(d)}} \leq 1} \|I_d(f) - \mathcal{A}_T(f)\|_{\mathcal{G}} \preceq_{d,\boldsymbol{a}} \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^d a_j k_j\right). \tag{1.8}$$

This motivates the study of the right hand side's precise asymptotic behavior as T tends to infinity. To this end, we will first prove in Theorem 2.6 the two-sided estimate

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^d a_j k_j\right) \asymp_{\boldsymbol{a},d} \frac{1}{(d-1)!}\Gamma(d,T), \tag{1.9}$$

where $\Gamma(d,T)$ denotes the upper incomplete Gamma function and the d- and $\boldsymbol{a}$-dependent constants are explicitly given. Next, we will relate (1.9) to the total cost of the associated approximation algorithm. Here, we only count the number of additional function(al) evaluations $L_{\boldsymbol{k}}(f)$, which corresponds to the number of elements $N = |\mathcal{D}_{\boldsymbol{a}}(T)|$ in the index set. Then we use bounds on the number of lattice points in simplices to obtain in Theorem 2.10 that

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^d a_j k_j\right) \asymp_{\boldsymbol{a},d} \exp\left(-\kappa(d)\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right) \cdot N^{\frac{d-1}{d}} \tag{1.10}$$

$$\preceq_{\boldsymbol{a},d} \exp\left(-\frac{d}{e}\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right), \tag{1.11}$$

where $\mathrm{gm}(\boldsymbol{a}) = (\prod_{j=1}^d a_j)^{1/d}$ is the geometric mean of $\boldsymbol{a}$ and $\kappa(d) := \sqrt[d]{d!} > d/e$. Note here that (1.10) improves on previous results in [6, 55].[6] We remark that (1.11) decays significantly faster than the bound for the corresponding full tensor product method, which only gives a rate of

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \{\max_j a_j k_j \leq T\}} \exp\left(-\sum_{j=1}^d a_j k_j\right) \asymp_{d,\boldsymbol{a}} \exp\left(-\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right),$$

where now $N = |\{\boldsymbol{k} \in \mathbb{N}_0^d : \max_{j=1,\ldots,d} a_j k_j \leq T\}| = \left(\prod_{j=1}^d \lfloor \frac{T}{a_j} \rfloor + 1\right)$.

We will also discuss the infinite-dimensional limit case

$$\lim_{d\to\infty} \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \setminus \mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^d a_j k_j\right), \tag{1.12}$$

assuming a certain behavior of the weights $a_j, j \in \mathbb{N}$. Here we obtain the following results for the summation problem (1.12).

[6] Note that the approach in [55] is applicable to a more general class of decay assumptions than the one we consider in (1.6). But there is a gap between the upper and lower bounds given in [55].

- If the assumption $\sum_{j=1}^{\infty} \frac{1}{e^{a_j/\beta}-1} < \infty$ is fulfilled for some $\beta > 1$, we have

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^{\infty} a_j k_j\right) \preceq_{\boldsymbol{a},\beta} N^{-(\beta-1)}, \tag{1.13}$$

  i.e. an algebraic rate $N^{-(\beta-1)}$ is attained, see also [6].

- If the stronger assumption $a_j \geq \alpha j$ holds with $\alpha > 0$, we have

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^{\infty} a_j k_j\right) \leq \frac{2}{\alpha\sqrt{\log N}} N^{1+\frac{\alpha}{4}-\frac{3}{8}\alpha\log(N)^{1/2}}, \tag{1.14}$$

  which is sub-exponential, i.e. decays faster than any algebraic rate $\mathcal{O}(N^{-r}), r \in \mathbb{R}^+$.

In order to apply this result to an infinite-dimensional approximation problem of type (1.8), one has to ensure that the constant $c_{d,\boldsymbol{a}}$ which is involved in the $\preceq$-notation does not behave pathologically as $d \to \infty$, i.e. $\lim_{d\to\infty} c_{d,\boldsymbol{a}} < \infty$. This will be discussed in more detail in Section 4. Moreover, we apply our results to the more general setting where only upper bounds of the form $\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)}\to\mathcal{G}} \preceq_{d,\boldsymbol{a},b} \prod_{j=1}^{d}(k_j+1)^b \exp\left(-\sum_{j=1}^{d} a_j k_j\right)$ with some $b \geq 0$ are given. Here, for the case of finite d, we again obtain the asymptotic upper bound

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)} \prod_{j=1}^{d}(k_j+1)^b \exp\left(-\sum_{j=1}^{d} a_j k_j\right) \preceq_{d,\boldsymbol{a},b} \exp\left(-\frac{d}{e}\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right). \tag{1.15}$$

In the infinite-dimensional setting we essentially get the rates from (1.13) and (1.14) with slight adjustments of β and α.

Possible applications of our bounds of the exponential sum (1.8) or (1.15) are given by sparse grid or tensor product approximation in certain classes of analytic functions with Legendre polynomials [6] or polynomial interpolation and quadrature at tensor products of Leja points [13, 40]. As a specific example, we will consider tensor products of Hardy spaces H_r^1, which contain functions that are analytic on discs $\mathbb{D}_r = \{z \in \mathbb{C} : |z| < r\}$ and have bounded $L^1(\partial\mathbb{D}_r)$-norm on its boundary. Here, we derive sharp, exponential bounds for the L^∞-approximation with properly truncated multivariate Taylor polynomials. Moreover, we consider interpolation and integration at sparse tensor products of Leja points.

The remainder of the article is organized as follows: In Section 2 we introduce the basic notation and derive two-sided bounds for the sum (1.8). In Section 3 we extend our analysis to the limit $d \to \infty$, assuming a certain behavior of the a_j. The fourth Section contains the generalized summation problem. Applications of our results to approximation, interpolation and integration in Hardy spaces of analytic functions are given in Section 5. We close in Section 6 with some concluding remarks.

2. Bounds from above and below for the finite-dimensional case

In the following we will abbreviate $\boldsymbol{a}^t\boldsymbol{k} := \sum_{j=1}^{d} a_j k_j$ and write (1.8) as

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} = \sum_{\boldsymbol{k}\in\mathbb{N}_0\setminus\mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^{d} a_j k_j\right), \tag{2.1}$$

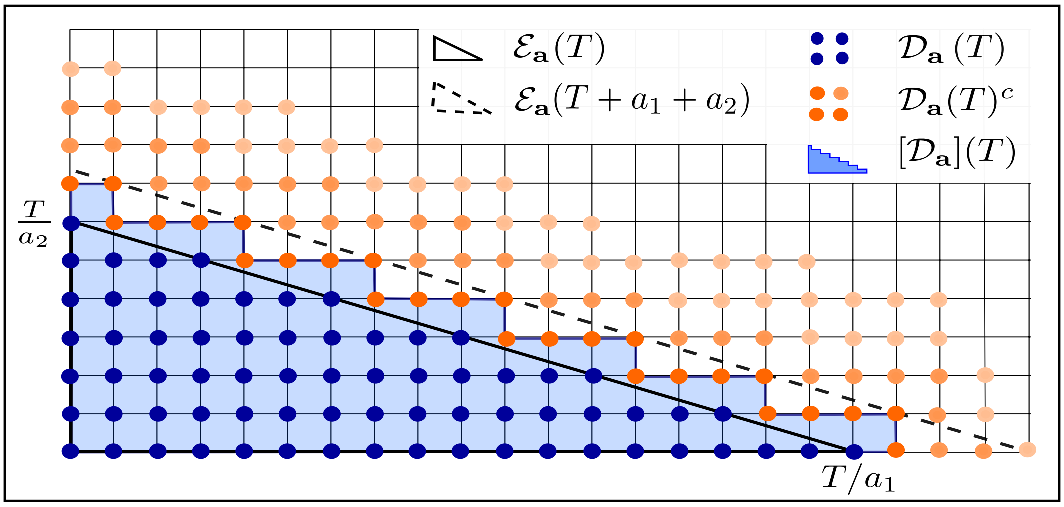

Figure 1: Illustration of the sets $\mathcal{D}_{\boldsymbol{a}}(T), \mathcal{E}_{\boldsymbol{a}}(T)$ and $[\mathcal{D}_{\boldsymbol{a}}](T)$ for $a_1 = 1$, $a_2 = 3$ and $T = 18$ in dimension $d = 2$.

where $\mathcal{D}_{\boldsymbol{a}}(T)^c \subset \mathbb{N}_0^d$ denotes the complement of

$$\mathcal{D}_{\boldsymbol{a}}(T) = \{\boldsymbol{k} \in \mathbb{N}_0^d : \sum_{j=1}^{d} a_j k_j \leq T\} \tag{2.2}$$

in $\mathbb{N}_0^d$, i.e. $\mathcal{D}_{\boldsymbol{a}}(T)^c = \mathbb{N}_0^d \setminus \mathcal{D}_{\boldsymbol{a}}(T) = \{\boldsymbol{k} \in \mathbb{N}_0^d : \sum_{j=1}^{d} a_j k_j > T\}$.

2.1. Error bound with respect to T

Our strategy will be to relate the discrete sum (2.1) to a continuous integral which we can bound from below and above. To this end, we need two auxiliary sets. The first one is the upper right part (i.e. the first hyper-octant) of the ℓ_1-ellipse with semi-axes $a_1^{-1}, \ldots, a_d^{-1}$ given by

$$\mathcal{E}_{\boldsymbol{a}}(T) := \left\{ \boldsymbol{x} \in \mathbb{R}_{\geq 0}^d : \sum_{j=1}^{d} a_j x_j \leq T \right\}, \tag{2.3}$$

where $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$. Since it holds that $\mathcal{D}_{\boldsymbol{a}}(T) = \{\lfloor \boldsymbol{x} \rfloor : \boldsymbol{x} \in \mathcal{E}_{\boldsymbol{a}}(T)\} = \mathcal{E}_{\boldsymbol{a}}(T) \cap \mathbb{N}_0^d$, one can consider $\mathcal{D}_{\boldsymbol{a}}(T)$ to be the discrete analogue of $\mathcal{E}_{\boldsymbol{a}}(T)$.

Moreover, we define

$$[\mathcal{D}_{\boldsymbol{a}}](T) := \bigcup_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} \bigotimes_{j=1}^{d} [k_j, k_j + 1) \subset \mathbb{R}_{\geq 0}^d, \tag{2.4}$$

which is the union of all blocks $\bigotimes_{j=1}^{d} [k_j, k_j + 1)$ anchored at an element $\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)$. The complements of $\mathcal{E}_{\boldsymbol{a}}(T)$ and $[\mathcal{D}_{\boldsymbol{a}}](T)$ in $\mathbb{R}_{\geq 0}^d$ are given by $\mathcal{E}_{\boldsymbol{a}}(T)^c = \{\boldsymbol{x} \in \mathbb{R}_{\geq 0}^d : \boldsymbol{a}^t \boldsymbol{x} > T\}$ and $[\mathcal{D}_{\boldsymbol{a}}](T)^c = \bigcup_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} \bigotimes_{j=1}^{d} [k_j, k_j + 1)$, respectively, see Figure 1.

Before we prove our main results we need several lemmas. First, we relate the discrete sum (2.1) to a continuous integral.

Lemma 2.1. *Let $\boldsymbol{a} \in \mathbb{R}_+^d$ and $T \in \mathbb{R}_{\geq 0}$. Then it holds that*

$$\sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}} = \left(\prod_{j=1}^d \frac{a_j}{1 - e^{-a_j}} \right) \int_{[\mathcal{D}_{\boldsymbol{a}}](T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x}. \tag{2.5}$$

Proof.

$$\begin{aligned}
\int_{[\mathcal{D}_{\boldsymbol{a}}](T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} &= \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} \int_{k_1}^{k_1+1} \cdots \int_{k_d}^{k_d+1} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} = \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} \prod_{j=1}^d \int_{k_j}^{k_j+1} e^{-a_j x_j} \, \mathrm{d}x_j \\
&= \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} \left(\prod_{j=1}^d e^{-a_j k_j} \frac{1 - e^{-a_j}}{a_j} \right) = \left(\prod_{j=1}^d \frac{1 - e^{-a_j}}{a_j} \right) \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}}. \qquad \square
\end{aligned}$$

The following lemma will be proven in Appendix A. It implies a close relationship between the incomplete Gamma function and (1.8).

Lemma 2.2. *Let $\mathcal{E}_{\boldsymbol{1}}(1)^c = \mathbb{R}_{\geq 0}^d \setminus \mathcal{E}_{\boldsymbol{1}}(1) = \{\boldsymbol{x} \in \mathbb{R}_{\geq 0}^d : \sum_{i=1}^d x_i > T\}$ denote the complement of the unit simplex in $\mathbb{R}_{\geq 0}^d$. Then it holds that*

$$\int_{\mathcal{E}_{\boldsymbol{1}}(1)^c} \exp\left(-T \sum_{j=1}^d y_j \right) \, \mathrm{d}\boldsymbol{y} = T^{-d} \frac{\Gamma(d, T)}{(d-1)!}, \tag{2.6}$$

where $\Gamma(d, T) = \int_T^\infty t^{d-1} e^{-t} \, \mathrm{d}t$ denotes the upper incomplete Gamma function.

A simple scaling of the axes leads to the following result.

Proposition 2.3. *Let $\boldsymbol{a} \in \mathbb{R}_+^d$ and $T \in \mathbb{R}_{\geq 0}$. Then it holds that*

$$\int_{\mathcal{E}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} = \left(\prod_{j=1}^d \frac{1}{a_j} \right) \frac{\Gamma(d, T)}{(d-1)!} \tag{2.7}$$

Proof. Using the change of variables $y_j = \frac{a_j}{T} x_j$ and Lemma 2.2 we obtain

$$\begin{aligned}
\int_{\mathcal{E}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} &= \left(\prod_{j=1}^d \frac{T}{a_j} \right) \int_{\mathcal{E}_{\boldsymbol{1}}(1)^c} \exp\left(-T \sum_{j=1}^d y_j \right) \, \mathrm{d}\boldsymbol{y} \\
&= \left(\prod_{j=1}^d \frac{T}{a_j} \right) T^{-d} \frac{\Gamma(d, T)}{(d-1)!}.
\end{aligned}$$

$\square$

Remark 2.4. *(Comments on the incomplete Gamma function)*

1. *If $T \in \mathbb{R}_{\geq 0}$ is fixed, $\Gamma(d, T)$ is strictly increasing in d. If on the other hand $d \in \mathbb{N}$ is fixed, $\Gamma(d, T)$ is a strictly decreasing function in T. This can easily be seen from the integral representation*

$$\Gamma(d, T) = \int_T^\infty t^{d-1} e^{-t} \, \mathrm{d}t. \tag{2.8}$$

2. *As it is pointed out in the proof of Lemma 2.2, for $d \in \mathbb{N}$ we have the equality [1]*

$$\frac{\Gamma(d,T)}{(d-1)!} = e^{-T} \sum_{k=0}^{d-1} \frac{T^k}{k!}. \tag{2.9}$$

The following bounds for $\Gamma(d,T)$ will be useful.

Lemma 2.5. *For $d \in \mathbb{N}$ and $T \in \mathbb{R}_{\geq 0}$ it holds that*

$$e^{-T} T^{d-1} \leq \Gamma(d,T).$$

Moreover, for $B > 1$ and $T \geq B(d-1)$ it holds that

$$\Gamma(d,T) \leq \frac{B}{B-1} e^{-T} T^{d-1}.$$

In particular, we have for $T \geq d$

$$e^{-T} T^{d-1} \leq \Gamma(d,T) \leq d\, e^{-T} T^{d-1}.$$

Proof. The lower bound follows from (2.9), while the upper bound is proven in [9]. □

Now we are ready to prove sharp asymptotic bounds on $\sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t \boldsymbol{k}}$.

Theorem 2.6. *For $T \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{a} \in \mathbb{R}^d_+$ there holds the upper bound*

$$\sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}} \leq \left(\prod_{j=1}^{d} \frac{e^{a_j}}{e^{a_j} - 1} \right) \cdot \frac{\Gamma(d,T)}{(d-1)!} \tag{2.10}$$

and the lower bound

$$\sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}} \geq \left(\prod_{j=1}^{d} \frac{1}{e^{a_j} - 1} \right) \frac{\Gamma(d,T)}{(d-1)!}. \tag{2.11}$$

Proof. Using Lemma 2.1, $[\mathcal{D}_{\boldsymbol{a}}](T)^c \subset \mathcal{E}_{\boldsymbol{a}}(T)^c$ and Proposition 2.3, we obtain

$$\begin{aligned}
\sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}} &= \left(\prod_{j=1}^{d} \frac{a_j}{1 - e^{-a_j}} \right) \int_{[\mathcal{D}_{\boldsymbol{a}}](T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} \leq \left(\prod_{j=1}^{d} \frac{a_j}{1 - e^{-a_j}} \right) \int_{\mathcal{E}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} \\
&= \left(\prod_{j=1}^{d} \frac{1}{1 - e^{-a_j}} \right) \frac{1}{(d-1)!} \Gamma(d,T),
\end{aligned}$$

which gives the upper bound (2.10) by noting that $(1 - e^{-a_j})^{-1} = e^{a_j}/(e^{a_j} - 1)$.
The lower bound follows similarly by using $\mathcal{E}_{\boldsymbol{a}}(T + \sum_{i=1}^{d} a_i)^c \subset [\mathcal{D}_{\boldsymbol{a}}](T)^c$.

To be precise, we have

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} &= \left(\prod_{j=1}^{d}\frac{a_j}{1-e^{-a_j}}\right)\int_{[\mathcal{D}_{\boldsymbol{a}}](T)^c} e^{-\boldsymbol{a}^t\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}\\
&\geq \left(\prod_{j=1}^{d}\frac{a_j}{1-e^{-a_j}}\right)\int_{\mathcal{E}_{\boldsymbol{a}}(T+\sum_{i=1}^d a_i)^c} e^{-\boldsymbol{a}^t\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}\\
&= \left(\prod_{j=1}^{d}\frac{1}{1-e^{-a_j}}\right)\frac{1}{(d-1)!}\Gamma(d,T+\sum_{i=1}^d a_i)\\
&= \left(\prod_{j=1}^{d}\frac{1}{1-e^{-a_j}}\right)e^{-T-\sum_{i=1}^d a_i}\sum_{k=0}^{d-1}\frac{1}{k!}(T+\sum_{i=1}^d a_i)^k\\
&\geq \left(\prod_{j=1}^{d}\frac{e^{-a_j}}{1-e^{-a_j}}\right)e^{-T}\sum_{k=0}^{d-1}\frac{1}{k!}T^k = \left(\prod_{j=1}^{d}\frac{e^{-a_j}}{1-e^{-a_j}}\right)\frac{1}{(d-1)!}\Gamma(d,T).
\end{aligned}$$

□

Moreover, we can use Lemma 2.5 to bound the incomplete Gamma function. Then we arrive at the following corollary.

Corollary 2.7. *Let $T \geq d$ and $\boldsymbol{a} \in \mathbb{R}_+^d$. Then there holds the upper bound*

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \leq d\left(\prod_{j=1}^{d}\frac{e^{a_j}}{e^{a_j}-1}\right)e^{-T}\frac{T^{d-1}}{(d-1)!}$$

and the lower bound

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \geq \left(\prod_{j=1}^{d}\frac{1}{e^{a_j}-1}\right)e^{-T}\frac{T^{d-1}}{(d-1)!}.$$

We remark that the asymptotic rate $\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \asymp e^{-T}T^{d-1}$ has been known before, see e.g. [28, 53]. However, the explicit dependence of the constants on d and $\boldsymbol{a}$ seems to be new, at least to our knowledge.

2.2. Error bound with respect to the total cost N

Next, we will relate the bound from Theorem 2.6 to the total number of elements N that are contained in $\mathcal{D}_{\boldsymbol{a}}(T)$. To this end, the following result from [7] gives an estimate for the number of lattice points in $\mathcal{D}_{\boldsymbol{a}}(T)$.

Lemma 2.8. *Let $N(T,d,\boldsymbol{a}) := |\mathcal{D}_{\boldsymbol{a}}(T)|$ denote the cardinality of $\mathcal{D}_{\boldsymbol{a}}(T)$. Then it holds that*

$$L(T,d,\boldsymbol{a}) := \frac{T^d}{d!}\prod_{j=1}^{d}a_j^{-1} \leq N(T,d,\boldsymbol{a}) \leq \frac{\left(T+\sum_{j=1}^d a_j\right)^d}{d!}\prod_{j=1}^{d}a_j^{-1} =: U(T,d,\boldsymbol{a}), \tag{2.12}$$

where the lower bound is valid only for $T \geq a_j$, $j = 1,\ldots,d$.

Remark 2.9. *This estimate is sharp in the sense of*

$$\frac{U(T,d,\boldsymbol{a})}{L(T,d,\boldsymbol{a})} \leq 1+\varepsilon \quad \text{if} \quad T \geq \frac{\sum_{j=1}^{d} a_j}{(1+\varepsilon)^{1/d}-1}.$$

Recently, in [31] the much sharper upper bound

$$N(T,d,\boldsymbol{a}) \leq \frac{T^d}{d!}\prod_{j=1}^{d} a_j^{-1} \prod_{j=1}^{d}\left(1+\frac{j a_j}{T}\right) \tag{2.13}$$

has been obtained. However, for a complexity result we later on need to express the cost bound in terms of the parameter T, see (2.16)*. This works for the bounds from Lemma 2.8 but, at least to our knowledge, is not possible for* (2.13) *with general $\boldsymbol{a} \in \mathbb{R}_+^d$.*

In the following we will write

$$\mathrm{gm}(\boldsymbol{a}) := \sqrt[d]{\prod_{j=1}^{d} a_j} \tag{2.14}$$

for the geometric mean of the elements of the vector $\boldsymbol{a} \in \mathbb{R}_+^d$ and

$$\kappa(d) := \sqrt[d]{d!}. \tag{2.15}$$

Then, rearranging the inequalities from Lemma 2.8 yields

$$\begin{aligned} T &\leq \sqrt[d]{N(T,d,\boldsymbol{a})}\,\kappa(d)\mathrm{gm}(\boldsymbol{a}) \\ T &\geq \sqrt[d]{N(T,d,\boldsymbol{a})}\,\kappa(d)\mathrm{gm}(\boldsymbol{a}) - \sum_{j=1}^{d} a_j. \end{aligned} \tag{2.16}$$

Now we are ready to bound $\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}}$ with respect to the total number of lattice points $N = N(T,d,\boldsymbol{a})$ that are contained in $\mathcal{D}_{\boldsymbol{a}}(T)$.

Theorem 2.10. *Let $N = |\mathcal{D}_{\boldsymbol{a}}(T)|$. Then we have the following, asymptotically optimal estimate for* (1.8) *with respect to N:*

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \asymp_{d,\boldsymbol{a}} \exp\left(-\kappa(d)\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right)\cdot N^{\frac{d-1}{d}}. \tag{2.17}$$

In particular, we have for $T \geq d$ or $N \geq \left(\frac{d+\sum_{j=1}^{d} a_j}{\kappa(d)\cdot\mathrm{gm}(\boldsymbol{a})}\right)^d$, the estimate

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \leq de\,\mathrm{gm}(\boldsymbol{a})^{d-1}\left(\prod_{j=1}^{d}\frac{e^{a_j}}{1-e^{-a_j}}\right)\exp\left(-\kappa(d)\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right)\cdot N^{\frac{d-1}{d}} \tag{2.18}$$

and, for $T \geq a_d$ or $N \geq \left(\frac{a_d}{\kappa(d)\mathrm{gm}(\boldsymbol{a})}\right)^d$, the estimate

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \geq \mathrm{gm}(\boldsymbol{a})^{d-1}\left(\prod_{j=1}^{d}\frac{1}{e^{a_j}-1}\right)\exp\left(-\kappa(d)\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right)N^{\frac{d-1}{d}}. \tag{2.19}$$

Proof. For the upper bound we use (2.10), (2.16), Lemma 2.5 and finally $\frac{\kappa(d)^{d-1}}{(d-1)!} \le e$, i.e.

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} &\le \left(\prod_{j=1}^d \frac{1}{1-e^{-a_j}}\right)\cdot\frac{\Gamma(d,T)}{(d-1)!}\\
&\le \left(\prod_{j=1}^d \frac{1}{1-e^{-a_j}}\right)\cdot\frac{1}{(d-1)!}\Gamma\left(d,\sqrt[d]{N}\,\kappa(d)\mathrm{gm}(\boldsymbol{a})-\sum_{j=1}^d a_j\right)\\
&\le d\left(\prod_{j=1}^d \frac{1}{1-e^{-a_j}}\right)\exp\left(-\sqrt[d]{N}\,\kappa(d)\cdot\mathrm{gm}(\boldsymbol{a})+\sum_{j=1}^d a_j\right)\\
&\quad\cdot\frac{1}{(d-1)!}\left(\sqrt[d]{N}\,\kappa(d)\mathrm{gm}(\boldsymbol{a})-\sum_{j=1}^d a_j\right)^{d-1}\\
&\le d\left(\prod_{j=1}^d \frac{e^{a_j}}{1-e^{-a_j}}\right)\exp\left(-\kappa(d)\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right)\frac{1}{(d-1)!}\left(\sqrt[d]{N}\,\kappa(d)\cdot\mathrm{gm}(\boldsymbol{a})\right)^{d-1}\\
&\le d\,e\left(\prod_{j=1}^d \frac{e^{a_j}}{1-e^{-a_j}}\right)\exp\left(-\kappa(d)\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right)\cdot N^{\frac{d-1}{d}}\cdot\mathrm{gm}(\boldsymbol{a})^{d-1}.
\end{aligned}$$

Now we prove the lower bound. Here, due to (2.11), (2.16) and Lemma 2.5, we obtain

$$\begin{aligned}
\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} &\ge \left(\prod_{j=1}^d \frac{e^{-a_j}}{1-e^{-a_j}}\right)\frac{1}{(d-1)!}\Gamma(d,T)\\
&\ge \left(\prod_{j=1}^d \frac{e^{-a_j}}{1-e^{-a_j}}\right)\frac{1}{(d-1)!}\Gamma\left(d,\sqrt[d]{N}\,\kappa(d)\mathrm{gm}(\boldsymbol{a})\right)\\
&\ge \left(\prod_{j=1}^d \frac{e^{-a_j}}{1-e^{-a_j}}\right)\exp\left(-\kappa(d)\sqrt[d]{N}\,\mathrm{gm}(\boldsymbol{a})\right)\frac{1}{(d-1)!}\left(\sqrt[d]{N}\,\kappa(d)\cdot\mathrm{gm}(\boldsymbol{a})\right)^{d-1}\\
&\ge \left(\prod_{j=1}^d \frac{1}{e^{a_j}-1}\right)\exp\left(-\kappa(d)\sqrt[d]{N}\,\mathrm{gm}(\boldsymbol{a})\right)N^{\frac{d-1}{d}}\mathrm{gm}(\boldsymbol{a})^{d-1}.
\end{aligned}$$

□

If we use Stirling's approximation to bound $\kappa(d) > \frac{d}{e}$, we obtain the following asymptotic upper bound, which improves a result in [6].

Corollary 2.11.

$$\sum_{\boldsymbol{k}\in\mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t\boldsymbol{k}} \preceq_{d,\boldsymbol{a}} \exp\left(-\frac{d}{e}\,\mathrm{gm}(\boldsymbol{a})\,\sqrt[d]{N}\right). \tag{2.20}$$

3. Upper bounds for the infinite-dimensional case

So far, we have dealt with the finite-dimensional setting $d < \infty$. We now turn to the infinite-dimensional case. To this end, let $0 < a_1 \leq a_2 \leq \dots$ be an infinite, ordered sequence of positive weights, i.e. $(a_j)_{j=1}^{\infty} =: \boldsymbol{a} \in \mathbb{R}_+^{\infty}$. First, one should ask the question whether an infinite-dimensional problem makes sense at all, as one has to ensure that already for $\mathcal{D}_{\boldsymbol{a}}(T) = \emptyset$, i.e. $T < 0$, the sum

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^{\infty}} e^{-\boldsymbol{a}^t \boldsymbol{k}} := \sum_{\boldsymbol{k} \in \mathbb{N}_0^{\infty}} \exp\left(-\sum_{j=1}^{\infty} a_j k_j \right) < \infty, \tag{3.1}$$

which, in the setting of tensor product approximation corresponds to the error of the zero-algorithm. Here, the sum runs over $\mathbb{N}_0^{\infty}$, which denotes the countable set of all non-negative integer sequences with only finitely many non-zero elements.

Using $\log(x) \leq x - 1$, we note that

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^{\infty}} \exp\left(-\sum_{j=1}^{\infty} a_j k_j \right) = \prod_{j=1}^{\infty} \left(\sum_{k=0}^{\infty} e^{-a_j k} \right) = \prod_{j=1}^{\infty} \frac{e^{a_j}}{e^{a_j} - 1} \leq \exp\left(\sum_{j=1}^{\infty} \frac{1}{e^{a_j} - 1} \right).$$

Therefore, the condition

$$\sum_{j=1}^{\infty} \frac{1}{e^{a_j} - 1} < \infty \tag{3.2}$$

implies (3.1) and hence the well-posedness of the problem.

3.1. Logarithmic growth

First, we consider a rather weak assumption on $(a_j)_{j=1}^{\infty}$, i.e. the finiteness of (3.2). Here, the following result from [26, Thm 2.5], which is an improvement of Stechkin's Lemma, will be useful.

Lemma 3.1. *Let $0 \leq p \leq q$ and $(c_j)_{j \in \mathbb{N}}$ be a countable set of positive numbers, whose m largest elements are $\{c_1, \dots, c_m\}$. Then*

$$\left(\sum_{j > m} (c_j)^q \right)^{1/q} \leq \left(\left(\frac{p}{q} \right)^{\frac{p}{q}} \left(1 - \frac{p}{q} \right)^{1 - p/q} \right)^{\frac{1}{p}} m^{-\frac{1}{p} + \frac{1}{q}} \left(\sum_{j=1}^{\infty} (c_j)^p \right)^{1/p}.$$

Now we show an algebraic upper bound, which is a slight improvement over [6].

Theorem 3.2. *Let the infinite, ordered sequence $\boldsymbol{a} = (a_j)_{j=1}^{\infty}$, $a_j > 0$ and $T \in \mathbb{R}_{\geq 0}$ be given. If there exists a real number $\beta > 1$ such that*

$$M(\boldsymbol{a}, \beta) := \sum_{j=1}^{\infty} \frac{1}{e^{a_j/\beta} - 1} < \infty \tag{3.3}$$

is true, it holds that

$$\sum_{\boldsymbol{k} \in \mathbb{N}_0^{\infty} \setminus \mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^{\infty} a_j k_j \right) \leq \frac{1}{\beta} \exp(\beta\, M(\boldsymbol{a}, \beta))\ N^{-(\beta - 1)}, \tag{3.4}$$

where as usual $N = |\mathcal{D}_{\boldsymbol{a}}(T)| < \infty$ denotes the cardinality of $\mathcal{D}_{\boldsymbol{a}}(T)$.

Proof. We note that $\mathcal{D}_{\boldsymbol{a}}(T)$ contains the N largest elements of $(e^{-\boldsymbol{a}^t\boldsymbol{k}})_{\boldsymbol{k}\in\mathbb{N}_0^\infty}$. Therefore, we can apply Lemma 3.1 with $q=1$ and $p=\frac{1}{\beta}$ to obtain

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\sum_{j=1}^\infty a_jk_j} \le \frac{1}{\beta}\left(\frac{\beta-1}{\beta}\right)^{\beta-1} N^{-(\beta-1)}\left(\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty} e^{-\frac{1}{\beta}\sum_{j=1}^\infty a_jk_j}\right)^\beta .$$

Then, the claim follows from $((\beta-1)/\beta)^{\beta-1}\in(1,1/e)$ for $\beta\in(1,\infty)$ and

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty} e^{-\frac{1}{\beta}\sum_{j=1}^\infty a_jk_j} = \prod_{j=1}^\infty \frac{e^{a_j/\beta}}{e^{a_j/\beta}-1} \le \exp\left(\sum_{j=1}^\infty \frac{1}{e^{a_j/\beta}-1}\right).$$

□

Remark 3.3. *The bound in Theorem 3.2 is especially relevant if one assumes logarithmic growth of $\boldsymbol{a}$, i.e. $a_j\ge\alpha\log(j+1)$ with $\alpha>\beta>1$.*

3.2. Linear growth

If stronger assumptions on $(a_j)_{j=1}^\infty$ are made, e.g. $a_j\ge\alpha j$, the Stechkin estimate only allows for algebraic convergence, which, however, can be arbitrarily high. Here, it is possible to refine the approach to obtain a more precise result, i.e. a sub-exponential rate of convergence. To this end, similar to [6], the basic strategy will be to choose β depending on N, i.e. β_N, such that the estimate $\exp(\beta_N\, M(\boldsymbol{a},\beta_N))\; N^{-(\beta_N-1)}$ from Theorem 3.2 gets approximately minimized for a given N.

Theorem 3.4. *Let $(a_j)_{j=1}^\infty$ fulfill $a_j\ge\alpha j$ for some $\alpha>0$ and all $j\in\mathbb{N}$. Then it holds for $N>\exp(4/\alpha^2)$ that*

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} \exp\left(-\sum_{j=1}^\infty a_jk_j\right) \le \frac{2}{\alpha\sqrt{\log(N)}} N^{1+\frac{\alpha}{4}-\frac{3}{8}\alpha\log(N)^{1/2}}, \tag{3.5}$$

which decays faster than any algebraic rate N^{-r}, $r\in\mathbb{R}_+$.

Proof. First, we note that

$$M(\alpha,\beta)=\sum_{j=1}^\infty\frac{1}{e^{a_j/\beta}-1}\le\max_{j\in\mathbb{N}}\left\{1+\frac{1}{e^{a_j/\beta}-1}\right\}\cdot\sum_{j=1}^\infty e^{-a_j/\beta}.$$

Next, we use $a_j\ge\alpha j$ and the monotonicity of $e^{-\frac{\alpha}{\beta}j}$ to obtain

$$\sum_{j=1}^\infty e^{-a_j/\beta}\le\sum_{j=1}^\infty e^{-\frac{\alpha}{\beta}j}\le\int_0^\infty e^{-\frac{\alpha}{\beta}x}\,\mathrm{d}x=\frac{\beta}{\alpha},$$

which implies that $M(\boldsymbol{a},\beta)<\infty$ for every $\beta>1$. Moreover, using the Taylor expansion of exp we obtain $\exp(\alpha/\beta)-1\ge\alpha/\beta$ and hence

$$\max_{j\in\mathbb{N}}\left\{1+\frac{1}{e^{\alpha j/\beta}-1}\right\}=1+\frac{1}{e^{\alpha/\beta}-1}\le1+\frac{\beta}{\alpha}.$$

Therefore, the bound

$$\exp(\beta\, M(\boldsymbol{a},\beta)) \le \exp\left(\beta\left(1+\frac{1}{e^{\alpha/\beta}-1}\right)\sum_{j=1}^{\infty} e^{-\frac{\alpha}{\beta}j}\right) \le \exp\left(\frac{\beta^2}{\alpha}\left(1+\frac{\beta}{\alpha}\right)\right)$$

holds for arbitrary $\beta > 1$ and the β-dependent estimate (3.4) becomes

$$\frac{1}{\beta}\exp(\beta\, M(\boldsymbol{a},\beta))N^{-(\beta-1)} \le \frac{1}{\beta}\exp\left(\frac{\beta^2}{\alpha}+\frac{\beta^3}{\alpha^2}-(\beta-1)\log(N)\right). \tag{3.6}$$

Now, for a given $N > \exp(4/\alpha^2)$, (3.6) is approximately minimized (with respect to β) by

$$\beta_N = \frac{\alpha}{2}\sqrt{\log(N)} > 1, \tag{3.7}$$

which we insert into (3.6) to obtain

$$\frac{1}{\beta_N}\exp(\beta_N\, M(\boldsymbol{a},\beta_N))N^{-(\beta_N-1)} \le \frac{2}{\alpha\sqrt{\log(N)}}\exp\left(\log(N)\left(1+\frac{\alpha}{4}\right)-\frac{3}{8}\alpha\log(N)^{3/2}\right).$$

□

This shows that even for infinite dimensional approximation problems sub-exponential rates of convergence are possible. Similar results in a somewhat different setting can also be found in [54]. Of course, there is still faster growth behavior of $(a_j)_{j=1}^{\infty}$ imaginable. But note that it has been proven in [37] that, no matter how fast a_j grows in j, it is not possible to achieve exponential convergence. Thus, a sub-exponential rate of convergence is the best one can hope for. However, the precise sub-exponential asymptotics for a specific, super-logarithmic growing sequence $(a_j)_{j=1}^{\infty}$ is an open problem and will be analyzed in more detail in the future.

4. Extension to a more general summation problem

When applying our results on (2.1) to real approximation problems, see Section 5.3, one sometimes only has bounds of the form

$$\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)}\to\mathcal{G}} \le \exp\left(-\sum_{j=1}^{d} a_j k_j\right)\prod_{j=1}^{d}\rho_j(k_j) =: e^{-\boldsymbol{a}^t\boldsymbol{k}}\boldsymbol{\rho}(\boldsymbol{k}), \tag{4.1}$$

where $a_j > 0$ and $\rho_j : \mathbb{N}_0 \to \mathbb{R}_+$. It is the goal of this section to establish bounds on (4.1) that fit into the setting analyzed in Section 2. This means that there is some $\hat{\boldsymbol{a}} \le \boldsymbol{a}$ such that

$$\|\Delta_{\boldsymbol{k}}\|_{\mathcal{H}^{(d)}\to\mathcal{G}} \le e^{-\boldsymbol{a}^t\boldsymbol{k}}\prod_{j=1}^{d}\rho_j(k_j) \le c_{d,\boldsymbol{a}}\, e^{-\hat{\boldsymbol{a}}^t\boldsymbol{k}} \text{ and } \lim_{d\to\infty} c_{d,\boldsymbol{a}} < \infty. \tag{4.2}$$

To this end, we note that (4.2) is fulfilled for $\rho_j(k_j) \le \tilde{c}_j \in \mathbb{R}_+$ if $\tilde{C} := \prod_{j\in\mathbb{N}}\tilde{c}_j < \infty$. Then $c_{d,\boldsymbol{a}} = \tilde{C}$ and $\hat{\boldsymbol{a}} = \boldsymbol{a}$. We generalize this to

$$\rho_j(k_j) \le \tilde{c}_j \cdot \begin{cases} 1 & \text{if } k_j = 0 \\ \bar{c}_j\,(k_j+1)^b & \text{if } k_j \ge 1 \end{cases}, \quad \text{with } \tilde{C} := \prod_{j=1}^{\infty}\tilde{c}_j < \infty, \tag{4.3}$$

$0 \le b < \infty$, $\tilde{c}_j \ge 1$ and $1 \le \bar{c}_j \le \bar{C} := \max_j \bar{c}_j < \infty$ for all $j \in \mathbb{N}$.[7]

Here, we note that for every given $\delta > 0$, it is possible to find $\hat{c}_{j,\delta} > 0$ such that

$$\bar{c}_j(k_j+1)^b \exp(-a_j k_j) \le \hat{c}_{j,\delta} \exp(-\hat{a}_j\, k_j) \quad \text{for all } k_j \in \mathbb{N}_0 \tag{4.4}$$

with $\hat{a}_j = (1-\delta)a_j$. Therefore, in the finite-dimensional case $d < \infty$ we have with $c_{d,\boldsymbol{a}} = \prod_{j=1}^d \tilde{c}_j \hat{c}_{j,\delta}$ the upper bound

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t\boldsymbol{k}}\,\boldsymbol{\rho}(\boldsymbol{k}) \le c_{d,\boldsymbol{a}} \sum_{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\mathcal{D}_{\boldsymbol{a}}(T)} \exp(-\sum_{j=1}^d \hat{a}_j k_j) \tag{4.5}$$

$$\asymp_{d,\boldsymbol{a},\delta} \ \exp\left(-\kappa(d)(1-\delta)\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right) N^{\frac{d-1}{d}}, \tag{4.6}$$

which, for $\delta \in (0, 1-\frac{d}{e}\kappa(d)^{-1})$, yields the following corollary.

Corollary 4.1. *If* (4.3) *and* (4.4) *hold for sufficiently small* δ *and all* $j = 1,\dots,d$ *we have the asymptotic upper bound*

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t\boldsymbol{k}}\,\rho(\boldsymbol{k}) \preceq_{d,\boldsymbol{a},b} \ \exp\left(-\frac{d}{e}\mathrm{gm}(\boldsymbol{a})\sqrt[d]{N}\right).$$

As d tends to infinity, usually $\prod_{j=1}^d \hat{c}_{j,\delta}$ and hence $c_{d,\boldsymbol{a}}$ grows exponentially. Therefore we need to refine the approach. To this end, we use the fact that for $k_j = 0$ we have a sharper bound. Here, the idea is to choose a $\delta > 0$, such that, with $\hat{a}_j := (1-\delta)a_j < a_j$, it holds that

$$\bar{C}\,(k_j+1)^b \exp(-a_j\,k_j) \le 1\cdot\exp(-\hat{a}_j\,k_j) \text{ for all } k_j \ge 1. \tag{4.7}$$

This is true for all $k_j \ge 1$ if

$$\delta a_j \ge \log(\bar{C}) + b\log(2), \tag{4.8}$$

which follows from re-arranging (4.7) and noting that $\log(k+1)/k$ is monotonously decreasing for $k \ge 1$. Since in the infinite-dimensional setting the ordered sequence of weights $\boldsymbol{a} = (a_j)_{j\in\mathbb{N}}$ has to tend to infinity we obtain that for every $\delta > 0$, there exists a number $d(\delta) \in \mathbb{N}$, such that for all $j > d(\delta)$ the inequality (4.8) and hence (4.7) can be fulfilled.

Therefore, we can now bound for $\boldsymbol{k} \in \mathbb{N}_0^\infty$

$$\begin{aligned}\prod_{j=1}^\infty e^{-a_jk_j}\rho_j(k_j) &\le \left(\prod_{j=1}^\infty \tilde{c}_j\right)\left(\prod_{\substack{j=1\\k_j\ge1}}^\infty \bar{c}_j(k_j+1)^b e^{-a_jk_j}\right)\\ &\le \left(\prod_{j=1}^\infty \tilde{c}_j\right)\left(\prod_{\substack{j=1\\k_j\ge1}}^{d(\delta)} \hat{c}_{j,\delta}e^{-a_j(1-\delta)k_j}\right)\left(\prod_{\substack{j=d(\delta)+1\\k_j\ge1}}^\infty e^{-a_j(1-\delta)k_j}\right)\\ &\le \tilde{C}\left(\prod_{j=1}^{d(\delta)}\hat{c}_{j,\delta}\right)\cdot\prod_{j=1}^\infty e^{-a_j(1-\delta)k_j}.\end{aligned}$$

[7] This seems technical at this point, but we will see in Section 5 that (4.3) is a realistic setting for certain choices of $\|\Delta_k\|_{H^1_r\to\mathcal{G}}$.

Hence, we obtain that, with $c_{\infty,\boldsymbol{a}} = \tilde{C}\left(\prod_{j=1}^{d(\delta)} \hat{c}_{j,\delta}\right) < \infty$, it holds that

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t\boldsymbol{k}}\rho(\boldsymbol{k}) \leq c_{\infty,\boldsymbol{a}} \sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-(1-\delta)\,\boldsymbol{a}^t\boldsymbol{k}},$$

which leads to the following corollary.

Corollary 4.2.
Let $\rho = \prod_j \rho_j$ fulfill (4.3).

(i) Let $\beta > 1$ such that for a given $\delta > 0$ it holds $M((1-\delta)\boldsymbol{a},\beta)) < \infty$. Then

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t\boldsymbol{k}}\,\rho(\boldsymbol{k}) \leq \tilde{C}\left(\prod_{j=1}^{d(\delta)} \hat{c}_{j,\delta}\right) e^{\beta\, M((1-\delta)\boldsymbol{a},\beta)}\, N^{-(\beta-1)} \preceq_{\boldsymbol{a},\delta,\rho} N^{-(\beta-1)}.$$

(ii) If $a_j \geq \alpha j$ for $\alpha > 0$ it holds for every positive $\hat{\alpha} < \alpha$ that

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^\infty\setminus\mathcal{D}_{\boldsymbol{a}}(T)} e^{-\boldsymbol{a}^t\boldsymbol{k}}\,\rho(\boldsymbol{k}) \preceq_{\hat{\alpha}} (\log N)^{-\frac{1}{2}} N^{1+\frac{\hat{\alpha}}{4}-\frac{3}{8}\hat{\alpha}\log(N)^{1/2}}.$$

Remark 4.3. *The term $\bar{c}_j(k_j+1)^b$ in* (4.7) *can be generalized to arbitrary polynomials of degree b in the variable (k_j+1). This follows easily, because every such polynomial can be bounded by $\bar{c}_j(k_j+1)^b$ with sufficiently large $\bar{c}_j$.*

Our approach can also be extended to expressions of the form $(2k_j+1)$ or $\sqrt{2k_j+1}$, which occur in bounds for the Legendre-coefficients of functions that are analytic in Bernstein ellipses or certain discs [6, 55]. Moreover, it can be further generalized by allowing the exponents b to depend on the different coordinate directions, provided that certain conditions on the sequence of exponents $(b_j)_{j\in\mathbb{N}}$ are fulfilled.

5. Applications

As an application we will consider approximation, interpolation and integration in tensor products of Hardy spaces of analytic functions on $[-1,1]$. To this end, we will first discuss univariate Hardy spaces and properties of their (infinite) tensor products. Then we prove a Taylor theorem for the approximation in multivariate Hardy spaces with sharp, two-sided error bounds using our previous results from Sections 2 and 3. Furthermore, we deal with polynomial interpolation at so-called Leja points, where the results from Section 4 will be useful. Finally, we study integration at Leja points.

5.1. Tensor products of Hardy spaces

Following [24], we define the scale of univariate Hardy spaces H_r^p with $r \in (1,\infty)$ and $p \in [1,\infty]$ as the set of functions that are holomorphic on the open disc $\mathbb{D}_r = \{z \in \mathbb{C} : |z| < r\}$ and have bounded L^p-norms on all circles of radius $t < r$, i.e. $(\frac{1}{2\pi}\int_0^{2\pi} \left|f\left(te^{\mathrm{i}\varphi}\right)\right|^p\, \mathrm{d}\varphi)^{1/p} < \infty$ for all $0 \leq t < r$.

Then, functions $f \in H_r^p$ can be extended point-wise to the boundary of $\mathbb{D}_r$ almost everywhere and their norm[8] is given by

$$\|f\|_{H_r^p} = \left(\frac{1}{2\pi r} \int_{\partial \mathbb{D}_r} |f(z)|^p d|z| \right)^{\frac{1}{p}} = \left(\frac{1}{2\pi} \int_0^{2\pi} |f\left(re^{\mathrm{i}\varphi}\right)|^p d\varphi \right)^{\frac{1}{p}}, \tag{5.1}$$

i.e. by the L^p-norm on the circle bounding $\mathbb{D}_r$ or equivalently on the torus $\mathbb{T} = [0, 2\pi)$. Therefore, analogously to the scale of $L^p(\mathbb{T})$-spaces, H_r^1 are Banach spaces and we have the inclusion $H_r^q \subset H_r^p$ for $1 \le p < q \le \infty$, see e.g. [24]. Moreover, for $1 < t < r$ it holds $H_r^1 \subset H_t^p$.

Because of $r > 1$, it holds that $[-1, 1] \subset \mathbb{D}_r$ and we can study approximation, interpolation and integration problems on this interval for functions $f \in H_r^1$. Note that, instead of fixing the interval and varying the radius r of the disc, it is also possible (and common in the literature) to fix $r = 1$ and study approximation problems on intervals $[-\tau, \tau]$ for $0 < \tau < 1$. This is, up to normalization, equivalent to our setting with $\tau = 1/r$.

In the context of approximation theory, functions of this class have been studied for quite some time, e.g. the case $r > 1$ and $p = 2$ in [38, 48], $r \ge 1$ and $p = \infty$ in [35] or $r = 1$ and $p \in (1, \infty]$ in [2, 50]. Moreover, in [21] a closely related space was considered. In this paper we will concentrate on the case $p = 1$, but due to the inclusion $H_r^p \subset H_r^1$ for $p > 1$ all upper bounds hold for higher p as well.[9]

The next lemma will be useful to give upper bounds on the norm of linear functionals in H_r^1. Here, the main tool is Cauchy's integral formula

$$f(x) = \frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_t} \frac{f(z)}{z - x} \, \mathrm{d}z, \qquad \text{where } x \in \mathbb{D}_t \text{ and } f \in H_r^1, \tag{5.2}$$

which holds for all $t < r$ and due to [24, Thm 3.6] also for $t = r$.

Lemma 5.1. *Let $L : H_r^1 \to \mathbb{C}$ be a bounded linear functional. Then it holds that*

$$|L(f)| \le r \max_{\theta \in [0, 2\pi)} \left| L\left(\frac{1}{re^{\mathrm{i}\theta} - \cdot} \right) \right| \|f\|_{H_r^1}. \tag{5.3}$$

Proof. Using Cauchy's integral formula we obtain

$$\begin{aligned}
|L(f)| &= \left| L\left(\frac{1}{2\pi \mathrm{i}} \int_{\partial \mathbb{D}_r} \frac{f(z)}{z - \cdot} \, \mathrm{d}z \right) \right| \le \frac{1}{2\pi} \int_{\partial \mathbb{D}_r} |f(z)| \left| L\left(\frac{1}{z - \cdot} \right) \right| \, \mathrm{d}|z| \\
&\le r \sup_{z \in \partial \mathbb{D}_r} \left| L\left(\frac{1}{z - \cdot} \right) \right| \cdot \frac{1}{2\pi r} \int_{\partial \mathbb{D}_r} |f(z)| \, \mathrm{d}|z| \\
&= r \sup_{\theta \in [0, 2\pi)} \left| L\left(\frac{1}{re^{\mathrm{i}\theta} - \cdot} \right) \right| \|f\|_{H_r^1}.
\end{aligned}$$

□

[8] In the case $p = \infty$ we have the obvious modification $\|f\|_{H_r^\infty} = \operatorname{ess\,sup}_{z \in \partial \mathbb{D}_r} |f(z)|$.

[9] For the case $p > 1$ it might nevertheless be possible to obtain improved p-dependent estimates.

This lemma easily yields upper bounds for $f \in H^1_r$ in the uniform norm, i.e.

$$\|f\|_{L^\infty} := \max_{t\in[-1,1]} |f(t)| \le r \max_{t\in[-1,1]} \sup_{\theta\in[0,2\pi)} \left|\frac{1}{re^{\mathrm{i}\theta}-t}\right| \cdot \|f\|_{H^1_r} = \frac{r}{r-1}\cdot \|f\|_{H^1_r}. \tag{5.4}$$

Moreover, integration with respect to the normalized uniform measure $\frac{1}{2}\,\mathrm{d}x$ on $[-1,1]$ can also be bounded from above by the same quantity

$$\left|\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x\right| \le \frac{1}{2} 2 \max_{t\in[-1,1]} |f(t)| \le \frac{r}{r-1}\cdot \|f\|_{H^1_r}. \tag{5.5}$$

This leads us to the next proposition, which establishes two-sided bounds for the norms of the linear operators $I_1^{\mathrm{emb}} : H^1_r \to \mathcal{C}_1 := C^0([-1,1])$ and $I_1^{\mathrm{int}} : H^1_r \to \mathbb{R}$, given by

$$I_1^{\mathrm{emb}}(f) = f \quad \text{and} \quad I_1^{\mathrm{int}}(f) = \frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x \tag{5.6}$$

with operator norms

$$\|I_1^{\mathrm{emb}}\|_{H^1_r\to\mathcal{C}_1} = \sup_{\|f\|_{H^1_r}\le 1} \|f\|_{L^\infty([-1,1])} \quad \text{and} \quad \|I_1^{\mathrm{int}}\|_{H^1_r\to\mathbb{R}} = \sup_{\|f\|_{H^1_r}\le 1} \left|\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x\right|. \tag{5.7}$$

Proposition 5.2. *It holds that*

$$1 \le \|I_1^{\mathrm{int}}\|_{H^1_r\to\mathbb{R}} \le \|I_1^{\mathrm{emb}}\|_{H^1_r\to\mathcal{C}_1} \le \frac{r}{r-1}.$$

Proof. It follows from (5.7) that the upper bound has already been shown in (5.4) and (5.5). The lower bounds follow from $f(z) \equiv 1$ belonging to the unit ball of H^1_r. □

Next, we consider algebraic tensor products of univariate Hardy spaces, whose completion with respect to the norm

$$\|f\|_{H^1_{\boldsymbol{r}}} := \frac{1}{(2\pi)^d}\int_{[0,2\pi)^d} \left|f\left(r_1 e^{\mathrm{i}\varphi_1},\dots,r_d e^{\mathrm{i}\varphi_d}\right)\right|\,\mathrm{d}\boldsymbol{\varphi}, \tag{5.8}$$

will be denoted by

$$H^1_{\boldsymbol{r}} := H^1_{r_1}\otimes\ldots\otimes H^1_{r_d}, \qquad \text{where } \boldsymbol{r}\in\mathbb{R}^d_{>1}. \tag{5.9}$$

Obviously, the norm (5.8), which is the L^1-norm on $\bigotimes_{j=1}^d \partial\mathbb{D}_{r_j}$, is a reasonable crossnorm [30]. We note that, by Hartogs' theorem [36], every d-variate function that is holomorphic in each variable separately is also jointly holomorphic. This implies that every function $f\in H^1_{\boldsymbol{r}}$ is also in $\mathrm{Hol}(\mathbb{D}_{\boldsymbol{r}})$, i.e. in the set of functions holomorphic in the polydisc $\mathbb{D}_{\boldsymbol{r}} := \bigotimes_{j=1}^d \mathbb{D}_{r_j}$. Moreover, algebraic polynomials are dense in both, $H^1_{\boldsymbol{r}}$ and the set of functions which are holomorphic in $\mathbb{D}_{\boldsymbol{r}}$ and have finite $H^1_{\boldsymbol{t}}$-norm with $\boldsymbol{t} = s\,\boldsymbol{r}$ for all scaling factors $0\le s<1$, see [47]. Therefore we can conclude that the tensor product space $H^1_{\boldsymbol{r}}$ from (5.9) indeed coincides with

$$\left\{ f\in\mathrm{Hol}(\mathbb{D}_{\boldsymbol{r}}) \text{ and } \sup_{0\le s<1}\int_{[0,2\pi)} \left|f\left(s\,r_1 e^{\mathrm{i}\varphi_1},\dots,s\,r_d e^{\mathrm{i}\varphi_d}\right)\right|\,\mathrm{d}\boldsymbol{\varphi} < \infty \right\}. \tag{5.10}$$

Now we consider the tensor product operators $I_d^{\mathrm{emb}} := \bigotimes_{j=1}^d I_1^{\mathrm{emb}} : H_{\boldsymbol{r}}^1 \to \mathcal{C}_d := C^0([-1,1]^d)$ and $I_d^{\mathrm{int}} := \bigotimes_{j=1}^d I_1^{\mathrm{int}} : H_{\boldsymbol{r}}^1 \to \mathbb{R}$. Their norms are bounded from above and below by

$$1 \leq \|I_d^{\mathrm{int}}\|_{H_{\boldsymbol{r}}^1 \to \mathbb{R}} \leq \|I_d^{\mathrm{emb}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} \leq \prod_{j=1}^d \frac{r_j}{r_j - 1},$$

which follows from the fact that the constant function $f(\boldsymbol{z}) = 1$ belongs to the unit ball of $H_{\boldsymbol{r}}^1$ and from the following proposition, which is proven in Appendix B.

Proposition 5.3. *Let $P = P_1 \otimes \ldots \otimes P_d : H_{\boldsymbol{r}}^1 \to \mathcal{G}$ be a tensor product of linear and bounded operators $P_j : H_{r_j}^1 \to \mathcal{G}_j$, where $\mathcal{G} \in \{\mathbb{R}, \mathcal{C}_d\}$ and $\mathcal{G}_j \in \{\mathbb{R}, \mathcal{C}_1\}$, i.e. either a linear functional or a bounded linear map into the space of continuous functions. Then it holds that*

$$\|P\|_{H_{\boldsymbol{r}}^1 \to \mathcal{G}} \leq \prod_{j=1}^d \|P_j\|_{H_{r_j}^1 \to \mathcal{G}_j}. \tag{5.11}$$

The next proposition gives a condition for the well-posedness of both approximation problems in the limiting case $d \to \infty$.

Proposition 5.4. *Let $r_j \in (1, \infty)$ for all $j \in \mathbb{N}$. Then, if*

$$\sum_{j=1}^\infty \frac{1}{r_j - 1} < \infty, \tag{5.12}$$

the limit operators $\lim_{d\to\infty} \bigotimes_{j=1}^d I_1^{\mathrm{emb}}$ and $\lim_{d\to\infty} \bigotimes_{j=1}^d I_1^{\mathrm{int}}$ have finite norm, i.e.

$$1 \leq \lim_{d\to\infty} \|I_d^{\mathrm{int}}\|_{H_{\boldsymbol{r}}^1 \to \mathbb{R}} \leq \lim_{d\to\infty} \|I_d^{\mathrm{emb}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} < \infty.$$

Proof. The claim follows from

$$\prod_{j=1}^\infty \|I_1^{\mathrm{int}}\|_{H_{r_j}^1 \to \mathbb{R}} \leq \prod_{j=1}^\infty \|I_1^{\mathrm{emb}}\|_{H_{r_j}^1 \to \mathcal{C}_1} \leq \prod_{j=1}^\infty \frac{r_j}{r_j - 1} = \prod_{j=1}^\infty \left(1 + \frac{1}{r_j - 1}\right) \leq \exp\left(\sum_{j=1}^\infty \frac{1}{r_j - 1}\right),$$

which is finite if (5.12) holds. □

Thus we have established that both integration with respect to the uniform probability measure $\bigotimes_{j=1}^d \frac{1}{2}\,\mathrm{d}x_j$ and approximation in the L^∞-norm of the cube $[-1,1]^d$ are well-defined problems, even as d tends to infinity, as long as the sequence of radii $(r_j)_{j\in\mathbb{N}}$ satisfies the condition (5.12).

5.2. Approximation in multivariate Hardy spaces using Taylor series

For $\boldsymbol{r} \in \mathbb{R}_{>1}^d$ and $\boldsymbol{x} \in [-1,1]^d$, we can express $f : H_{\boldsymbol{r}}^1 \to \mathbb{R}$ by the multivariate Taylor series

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} f^{(\boldsymbol{k})}(\boldsymbol{0}) \frac{1}{\boldsymbol{k}!} \boldsymbol{x}^{\boldsymbol{k}} := \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} f^{(\boldsymbol{k})}(\boldsymbol{0}) \prod_{j=1}^d \frac{1}{k_j!} x_j^{k_j},$$

where $f^{(\boldsymbol{k})}(\mathbf{0}) = D_{\boldsymbol{k}} f(\mathbf{0}) = \frac{\partial^{|\boldsymbol{k}|_1}}{\partial x_1^{k_1} \ldots \partial x_d^{k_d}} f(\boldsymbol{x})_{\boldsymbol{x}=\mathbf{0}}$. Hence, using the notation (1.1) of the introduction we can decompose the embedding operator $I_d^{\mathrm{emb}} : H_{\boldsymbol{r}}^1 \to \mathcal{C}_d := C^0([-1,1]^d)$ into

$$I_d^{\mathrm{emb}}(f) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^d} \Delta_{\boldsymbol{k}}(f), \qquad \text{with } \Delta_{\boldsymbol{k}}(f) = \frac{f^{(\boldsymbol{k})}(\mathbf{0})}{\boldsymbol{k}!}\, \boldsymbol{x}^{\boldsymbol{k}}.$$

Now, Lemma 5.1 yields the upper bound

$$\begin{aligned}
\|\Delta_{\boldsymbol{k}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} &= \sup_{\|f\|_{H_{\boldsymbol{r}}^1} \le 1} \left\| \frac{f^{(\boldsymbol{k})}(\mathbf{0})}{\boldsymbol{k}!}\, \boldsymbol{x}^{\boldsymbol{k}} \right\|_{L^\infty} = \sup_{\|f\|_{H_{\boldsymbol{r}}^1} \le 1} \frac{|f^{(\boldsymbol{k})}(\mathbf{0})|}{\boldsymbol{k}!} \\
&\le \prod_{j=1}^d \left(\frac{r_j}{k_j!} \max_{\theta \in [0,2\pi)} \left| \frac{d^{k_j}}{dt^{k_j}} \frac{1}{re^{\mathrm{i}\theta} - t} \right|_{t=0} \right) = \prod_{j=1}^d r_j^{-k_j}.
\end{aligned}$$

Moreover, using that $\boldsymbol{z} \mapsto \prod_{j=1}^d r_j^{-k_j} z_j^{k_j}$ belongs to the unit ball of $H_{\boldsymbol{r}}^1$, we obtain the lower bound

$$\|\Delta_{\boldsymbol{k}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} = \sup_{\|f\|_{H_{\boldsymbol{r}}^1} \le 1} \frac{|f^{(\boldsymbol{k})}(\mathbf{0})|}{\boldsymbol{k}!} \ge \prod_{j=1}^d r_j^{-k_j}.$$

Altogether, we thus have obtained

$$\|\Delta_{\boldsymbol{k}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} = \prod_{j=1}^d r_j^{-k_j} = \exp\left(-\sum_{j=1}^d \log(r_j) k_j \right). \tag{5.13}$$

Therefore, with $\boldsymbol{a} = \log(\boldsymbol{r})$, which is meant component-wise, we can bound the error of the approximation algorithm

$$\mathcal{A}_{T,\boldsymbol{a}}(f) := \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} \frac{f^{(\boldsymbol{k})}(\mathbf{0})}{\boldsymbol{k}!}\, \boldsymbol{x}^{\boldsymbol{k}}$$

by

$$\|I_d^{\mathrm{emb}}(f) - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty([-1,1]^d)} \le \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)^c} e^{-\boldsymbol{a}^t \boldsymbol{k}} \|f\|_{H_{\boldsymbol{r}}^1}. \tag{5.14}$$

For the univariate setting $d = 1$ this gives

$$\|f - \mathcal{A}_{T,a_1}(f)\|_{L^\infty} \le \sum_{k_1 = T+1}^{\infty} r_1^{-k_1} \|f\|_{H_r^1} = \frac{r_1^{-T}}{r_1 - 1} \|f\|_{H_r^1}. \tag{5.15}$$

For $d \ge 2$ we can invoke the Theorems 2.10, 3.2 and 3.4 to obtain the following result.

Corollary 5.5. *Let $T \in \mathbb{R}_{\ge 0}$ and $\boldsymbol{r} \in \mathbb{R}_{>1}^d$ with $1 < r_1 \le r_2 \le \ldots \le r_d$. The algorithm $\mathcal{A}_{T,\boldsymbol{a}}(f) = \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} \Delta_{\boldsymbol{k}}(f)$, where $\boldsymbol{a} = \log(\boldsymbol{r})$ and $\Delta_{\boldsymbol{k}} = \frac{f^{(\boldsymbol{k})}(\mathbf{0})}{\boldsymbol{k}!} \boldsymbol{x}^{\boldsymbol{k}}$, approximates $f \in H_{\boldsymbol{r}}^1$ in the $L^\infty([-1,1]^d)$-norm with the following asymptotic rates of convergence. Here, $N = |\mathcal{D}_{\boldsymbol{a}}(T)|$ denotes the number of derivative evaluations.*

(i) For $d < \infty$ and $T \geq d$ we have

$$\sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \leq C_{d,\boldsymbol{a}} \exp\left(-\kappa(d)\mathrm{gm}(\boldsymbol{a}) \sqrt[d]{N}\right) N^{\frac{d-1}{d}},$$

where $C_{d,\boldsymbol{a}} = de\, \mathrm{gm}(\boldsymbol{a})^{d-1} \left(\prod_{j=1}^d \frac{e^{a_j}}{1-e^{-a_j}}\right)$.

(ii) For $d = \infty$ we have for every $\beta > 1$ that ensures $\sum_{j=1}^\infty \frac{1}{e^{a_j/\beta}-1} < \infty$

$$\lim_{d\to\infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \leq C_{\boldsymbol{a}}\, N^{-(\beta-1)},$$

where the constant is $C_{\boldsymbol{a}} = \exp(\beta\, M(\boldsymbol{a}, \beta))$, see (3.3)*.*

(iii) For $d = \infty$ and $r_j \geq \alpha j$ with $\alpha > 0$ we have

$$\lim_{d\to\infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \leq \frac{2}{\alpha\sqrt{\log N}} N^{1+\frac{\alpha}{4}-\frac{3}{8}\alpha \log(N)^{1/2}}.$$

5.3. Interpolation in multivariate Hardy spaces at Leja points

In classical interpolation schemes, like Tschebyscheff interpolation, the number of points usually has to be doubled from level to level in order to obtain a nested sequence of interpolation operators. In the multivariate setting, this requires $\approx 2^{|\boldsymbol{k}|}$ additional function(al) values for the evaluation of $\Delta_{\boldsymbol{k}}$. However, we aim here at an approach, where only a single point is added to the set of interpolation nodes, which, in the multivariate setting, leads to just 1 additional function(al) value for $\Delta_{\boldsymbol{k}}$ – independent of $\boldsymbol{k} \in \mathbb{N}_0^d$. To this end, so-called *Leja sequences* are a natural choice, which will be employed in the following.

5.3.1. General discussion

In the preceding section we showed that the proper truncation of a multivariate Taylor series gives an excellent rate of convergence for approximating functions in multivariate Hardy spaces. However, in many practical applications, derivative-information is not available or expensive to obtain. Therefore, we now study the use of point-evaluations at pairwise-distinct points $\boldsymbol{x}_{\boldsymbol{k}} \in [-1,1]^d, \boldsymbol{k} \in \mathbb{N}_0^d$ as information for the approximation of $I_d^{\mathrm{emb}} : H^1_{\boldsymbol{r}} \to \mathcal{C}_d := C^0([-1,1]^d)$. As before, the error shall be measured in L^∞. To this end, we choose a sequence of pairwise distinct univariate points $(\xi_i)_{i=0}^\infty$ and use points from its d-fold tensor product, i.e. $\boldsymbol{x}_{\boldsymbol{k}} = (\xi_{k_1}, \ldots, \xi_{k_d})$ for $\boldsymbol{k} \in \mathbb{N}_0^d$.

In order to construct a suitable basis, we will follow [5, 14, 13] and use the univariate polynomial interpolation operators $\mathcal{I}_n$, given by

$$\mathcal{I}_n f(x) = \sum_{i=0}^n f(\xi_i)\ell_{i,n}(x), \qquad \text{where } \ell_{i,n} = \prod_{\substack{l=0\\ l\neq i}}^n \frac{x-\xi_l}{\xi_i - \xi_l}$$

denotes the Lagrange polynomial associated to the points $\xi_0, \ldots, \xi_n$. Then we define the hierarchical interpolation operator $\Delta_k : H^1_r \to \mathcal{C}_1 = C^0([-1,1])$ as $\Delta_0 f(x) \equiv f(\xi_0)\cdot 1$ and for $k \in \mathbb{N}$

$$\Delta_k f(x) = \mathcal{I}_k f(x) - \mathcal{I}_{k-1} f(x) = \sum_{i=0}^{k-1} f(\xi_i)(\ell_{i,k}(x) - \ell_{i,k-1}(x)) + f(\xi_k)\ell_{k,k}(x).$$

Hence, we obtain the hierarchical decomposition $\mathcal{I}_n f(x) = \sum_{k=0}^{n} \Delta_k f(x)$.

In order to bound $\|\Delta_k\|_{H_r^1 \to \mathcal{C}_1}$, we can use our previous result on the approximability by Taylor polynomials as an upper bound for the best approximation with polynomials $p \in \mathbb{P}_k$ of degree k in combination with the Lebesgue constant $\Lambda_k := \sup_{\|f\|_{L^\infty} \le 1} \|\mathcal{I}_k(f)\|_{L^\infty} = \| \sum_{i=0}^{k} |\ell_{i,k}(\cdot)| \|_{L^\infty([-1,1])}$ of the sequence $(\xi_i)_{i=0}^k$, i.e. for $k \geq 1$ and $f \in H_r^1$

$$\begin{aligned}\|\Delta_k(f)\|_{L^\infty} \le& \|(I_1^{\mathrm{emb}} - \mathcal{I}_k)(f)\|_{L^\infty} + \|(I_1^{\mathrm{emb}} - \mathcal{I}_{k-1})(f)\|_{L^\infty} \\ \le& (1+\Lambda_k) \inf_{p \in \mathbb{P}_k} \|f-p\|_{L^\infty} + (1+\Lambda_{k-1}) \inf_{p \in \mathbb{P}_{k-1}} \|f-p\|_{L^\infty} \\ \le& 2(1+\max\{\Lambda_k, \Lambda_{k-1}\}) \inf_{p \in \mathbb{P}_{k-1}} \|f-p\|_{L^\infty} \\ \le& \frac{2r}{r-1}(1+\max\{\Lambda_k, \Lambda_{k-1}\}) r^{-k} \|f\|_{H_r^1},\end{aligned}$$

where the last inequality follows from

$$\inf_{p \in \mathbb{P}_{k-1}} \|f-p\|_{L^\infty} \le \left\| \sum_{m=k}^{\infty} \frac{f^{(m)}(0)}{m!} x^m \right\|_{L^\infty} \le \sum_{m=k}^{\infty} r^{-m} \|f\|_{H_r^1} = \frac{r}{r-1} r^{-k} \|f\|_{H_r^1}.$$

Hence, we have established that for $k \geq 1$ it holds that

$$\|\Delta_k\|_{H_r^1 \to \mathcal{C}_1} = \sup_{\|f\|_{H_r^1} \le 1} \|\Delta_k(f)\|_{L^\infty} \le \frac{2r}{r-1}(1+\max\{\Lambda_k, \Lambda_{k-1}\}) r^{-k}.$$

For the case $k=0$ we have $\|\Delta_0\|_{H_r^1 \to \mathcal{C}_1} = \|\mathcal{I}_0\|_{H_r^1 \to \mathcal{C}_1} = \sup_{\|f\|_{H_r^1} \le 1} f(\xi_0) \cdot 1 \le \frac{r}{r-\xi_0}$, which gives $\|\Delta_0\|_{H_r^1 \to \mathcal{C}_1} = \frac{r}{r-1}$ for $\xi_0 = 1$ and $\|\Delta_0\|_{H_r^1 \to \mathcal{C}_1} = 1$ for $\xi_0 = 0$. If an estimate of the form $\Lambda_k \le c(k+1)^b$ would be available,[10] we could bound

$$\|\Delta_k\|_{H_r^1 \to \mathcal{C}_1} \le \frac{r}{r-1} 2(1+c(k+1)^b) r^{-k} \le \frac{r}{r-1} 2(1+c)(k+1)^b r^{-k}. \tag{5.16}$$

Consequently for the multivariate hierarchical operators $\Delta_{\boldsymbol{k}} : H_{\boldsymbol{r}}^1 \to \mathcal{C}_d = C^0([-1,1]^d)$, $\boldsymbol{k} \in \mathbb{N}_0^d$, which are simply tensor products of the univariate ones, we have

$$\Delta_{\boldsymbol{k}}(f) = \bigotimes_{j=1}^{d} \Delta_{k_j}(f) \quad \text{and} \quad \|\Delta_{\boldsymbol{k}}\|_{H_{\boldsymbol{r}}^1 \to \mathcal{C}_d} \le \prod_{j=1}^{d} r_j^{-k_j} \frac{r_j}{r_j - 1} \begin{cases} 1 & k_j = 0, \\ 2(1+c)(k_j+1)^b & k_j \geq 1. \end{cases} \tag{5.17}$$

This fits into the setting of Corollaries 4.1 and 4.2 with $\tilde{c}_j = \frac{r_j}{r_j - 1}$ and $\bar{c}_j = 2(1+c)$. To this end, we will discuss so-called Leja sequences, which allow for $b = 2$ and $c = 8\sqrt{2}$.

5.3.2. Leja points

For a given, compact set $\mathcal{S} \subset \mathbb{C}$ and starting point $\xi_0 \in \mathcal{S}$, the associated Leja sequence $(\xi_i)_{i=0}^\infty \subset \mathcal{S}$ stems from a certain recursive optimization process. It is defined by

$$\xi_{m+1} = \arg\max_{z \in \mathcal{S}} \left| \prod_{i=0}^{m} (z - \xi_i) \right|. \tag{5.18}$$

[10] Sometimes, even direct estimates of $\|\Delta_k\|_{H_r^1 \to \mathcal{C}_1}$, which circumvent the usage of the triangle inequality, can be obtained, see [14].

We will consider two types of Leja sequences, where we will keep the notation from [11, 12, 14]. First, we consider the classical Leja sequence (L) on $\mathcal{S} = [-1, 1]$. We set $\xi_0 = 0$ and iteratively compute the point sets (5.18). The second approach comes from [11], see also [12, 14] and uses Leja points on the complex unit circle, i.e. $\mathcal{S} = \mathbb{D}_1$. Then, those points are projected onto the real axis, with elimination of possible doubles. The resulting sequence will be called projected $\Re$-Leja sequence.

For the $\Re$-Leja sequence there is a closed-form solution of the optimization problem (5.18) for $\mathcal{S} = \mathbb{D}_1$ available. Then, the elements of the $\Re$-Leja sequence are given by $\xi_k = \cos(\pi \sum_{l=0}^{s} \kappa_l 2^{-l})$, where κ_l are the binary digits of $k = \sum_{l=0}^{s} \kappa_l 2^l$, see [11, 14]. Consequently, bounds for the Lebesgue constant Λ_n of the projected $\Re$-Leja sequence $\xi_0, \ldots, \xi_n$ could be derived. In [11] it was bounded by $\mathcal{O}(\log(n+1)n^3)$. This was improved in [12] to $\Lambda_n \leq 5(n+1)^2 \log(n+1)$ and later in [14] to $\Lambda_n \leq 8\sqrt{2}(n+1)^2$.

Hence, following (5.17), we can invoke Corollaries 4.1 and 4.2 with $\tilde{c}_j = \frac{r_j}{r_j - 1}$ and $\bar{c}_j = 2(1+8\sqrt{2})$ and have the following bounds for multivariate interpolation at tensor products of the projected $\Re$-Leja sequences.

Corollary 5.6. *Let $T \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{r} \in \mathbb{R}^d_{>1}$ with $1 < r_1 \leq r_2 \leq \ldots$ and $d \in \mathbb{N} \cup \{\infty\}$. The algorithm $\mathcal{A}_{T,\boldsymbol{a}}(f) = \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} \Delta_{\boldsymbol{k}}(f)$, where $\boldsymbol{a} = \log(\boldsymbol{r})$ and $\Delta_{\boldsymbol{k}}$ is given in (5.17) with $(\xi_i)_{i=0}^{\infty}$ being the $\Re$-Leja sequence, approximates $f \in H^1_{\boldsymbol{r}}$ in the L^∞-norm with the following asymptotic rates of convergence.*

(i) For $d < \infty$ and $\delta \in (0, 1 - \frac{d}{e}\kappa(d)^{-1})$ we have

$$\sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \preceq_{d,\boldsymbol{a},\delta} N^{\frac{d-1}{d}} \exp\left(-\kappa(d)(1-\delta)\mathrm{gm}(\boldsymbol{a}) \sqrt[d]{N}\right)$$

$$\preceq_{d,\boldsymbol{a}} \quad \exp\left(-\frac{d}{e}\mathrm{gm}(\boldsymbol{a}) \sqrt[d]{N}\right).$$

(ii) For $d = \infty$ we have for every $\beta > 1$ that ensures $\sum_{j=1}^{\infty} \frac{1}{e^{(1-\delta)a_j/\beta} - 1} < \infty$

$$\lim_{d \to \infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \preceq_{\boldsymbol{a}} N^{-(\beta-1)}.$$

(iii) For $d = \infty$ and $r_j \geq \alpha j$ we have for all $0 < \hat{\alpha} < \alpha$

$$\lim_{d \to \infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \|f - \mathcal{A}_{T,\boldsymbol{a}}(f)\|_{L^\infty} \preceq_{\hat{\alpha}} (\log N)^{-\frac{1}{2}} N^{1 + \frac{\hat{\alpha}}{4} - \frac{3}{8}\hat{\alpha} \log(N)^{1/2}}.$$

Unfortunately, there is no closed-form solution of the optimization problem (5.18) with $\mathcal{S} = [-1, 1]$, i.e. for the (L)-Leja sequence. Therefore, the points of the (L)-sequence can only be obtained by numerical computations, which makes tight bounds for their Lebesgue constant difficult to obtain. Here, we are only aware of [52], where a sub-exponential asymptotic behavior of Λ_n was shown for the (L)-sequence. This means that $\lim_{n\to\infty}(\Lambda_n)^{1/n} \to 1$, which is of course weaker than the algebraic bound for the $\Re$-sequence. However, in [12] numerical evidence was given that the (L)-sequence in fact allows for a substantially smaller Lebesgue constant, i.e. $\mathcal{O}(n)$, than the $\Re$-sequence. If this could be proven, we would have analogue results to Corollary 5.6 for (L)-Leja points, albeit with much better values of δ and $\hat{\boldsymbol{a}}$. Thus, using the (L)-Leja points would practically lead to improved convergence behavior.

5.4. Integration in Hardy spaces using Leja points

Next, we study numerical integration with respect to the normalized uniform measure of functions $f \in H^1_{\boldsymbol{r}}$, i.e. approximating the linear functional[11] $I_d^{\text{int}}(f) = \frac{1}{2^d}\int_{[-1,1]^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. Of course, since we already obtained a polynomial approximation/interpolation $\mathcal{A}_{T,\boldsymbol{a}}(f)$ to f, we could simply compute $I_d^{\text{int}}(\mathcal{A}_{T,\boldsymbol{a}}(f))$ and would be done. But for the sake of completeness, we will here go all the way from univariate quadrature rules to a multivariate integration method.

First, we briefly discuss the univariate case. Here, we prove that quadrature rules which are based on a polynomial degree of exactness achieve exponential convergence in H^1_r if $r > 1$.

Theorem 5.7. *If the n-point quadrature formula $Q_n(f) = \sum_{i=0}^{n-1} w_{i,n} f(\xi_i)$ integrates polynomials of degree $(\mu - 1) \in \mathbb{N}$ exactly, its quadrature error in the univariate Hardy space H^1_r can be bounded by*

$$\left|\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x - Q_n(f)\right| \leq \left(1 + \sum_{i=0}^{n-1} |w_{i,n}|\right) \frac{r}{r-1}\, r^{-\mu} \cdot \|f\|_{H^1_r}.$$

Proof. Taylor's theorem yields for an arbitrary $f \in H^1_r$

$$f(z) = \sum_{j=0}^{\mu-1} \frac{f^{(j)}(0)}{j!} z^j + S_{\mu-1}(f)(z),$$

where, according to (5.15), the remainder $S_{\mu-1}(f) = \sum_{j=\mu}^{\infty} \frac{f^{(j)}(0)}{j!} z^j$ is bounded by

$$\|S_{\mu-1}(f)\|_{L^\infty} \leq \frac{r}{r-1} r^{-\mu} \|f\|_{H^1_r}.$$

Therefore we have for the normalized uniform measure $\frac{1}{2}\,\mathrm{d}x$ on $[-1,1]$

$$\begin{aligned}
|I_1(f) - Q_n(f)| &= \left|\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x - \sum_{i=0}^{n-1} w_{i,n} f(\xi_i)\right| \\
&= \left|\frac{1}{2}\int_{-1}^{1} S_{\mu-1}(f)(x)\,\mathrm{d}x - \sum_{i=0}^{n-1} w_{i,n} S_{\mu-1}(f)(\xi_i)\right| \\
&\leq \frac{1}{2}\int_{-1}^{1} |S_{\mu-1}(f)(x)|\,\mathrm{d}x + \sum_{i=0}^{n-1} |w_{i,n}||S_{\mu-1}(f)(\xi_i)| \\
&\leq \left(1 + \sum_{i=0}^{n-1} |w_{i,n}|\right) \frac{r}{r-1} r^{-\mu} \cdot \|f\|_{H^1_r}.
\end{aligned}$$

□

As one can see, the sum over the absolute values of the quadrature weights $\nu(n) := \sum_{i=0}^{n-1} |w_{i,n}|$ plays now an important role in the error estimate. In the following we will discuss the usage of the (L)-Leja and the projected $\Re$-Leja sequence as quadrature points $(\xi_i)_{i=0}^{n-1}$. The associated weights

[11] Note here that the 2^{-d} factor is necessary to obtain a normalized measure. For integration on $[0,1]^d$ or $[-1/2, 1/2]^d$ it could be omitted. However, the Hardy spaces would have to be transformed then as well.

$w_{0,n}, \ldots, w_{n-1,n}$ are determined such that they achieve a degree of polynomial exactness of at least $(n-1)$, i.e.

$$\frac{1}{2}\int_{-1}^{1} x^k \,\mathrm{d}x = \sum_{i=0}^{n-1} w_{i,n}\xi_i^k \quad \text{for all } k = 0, 1, \ldots, n-1, \tag{5.19}$$

which is equivalent to the exact integration of the corresponding Lagrange polynomials, i.e. $w_{i,n} = \frac{1}{2}\int_{-1}^{1} \ell_{i,n-1}(x)\,\mathrm{d}x$.

In contrast to classical Newton-Cotes formulae for equidistant points, the quadrature weights associated to Leja points (5.19) are more stable, which means that $\nu(n) := \sum_{i=0}^{n-1} |w_{i,n}|$ does not grow exponentially. This directly enters the error estimate in Theorem 5.7.

In order to bound $\nu(n)$, one can use estimates on the Lebesgue constant Λ_{n-1} for L^∞ polynomial interpolation at the sequence of nested quadrature points. To this end, we have the following relation between $\nu(n)$ and Λ_n.

$$\nu(n+1) = \sum_{i=0}^{n} \left|\frac{1}{2}\int_{-1}^{1} \ell_{i,n}(x)\,\mathrm{d}x\right| \le \frac{1}{2}\int_{-1}^{1}\sum_{i=0}^{n} |\ell_{i,n}(x)|\,\mathrm{d}x \le \max_{x\in[-1,1]}\sum_{i=0}^{n} |\ell_{i,n}(x)| = \Lambda_n. \tag{5.20}$$

Recalling the bounds for Λ_n of the $\Re$-Leja points from the previous subsection then yields the following corollary of Theorem 5.7 with $\mu = n$.

Corollary 5.8. *Let $(\xi_i)_{i=1}^\infty$ be the $\Re$-Leja sequence with associated quadrature weights $w_{i,n}$, $i = 0, \ldots, n-1$ given by* (5.19)*. Then, the quadrature error in H_r^1 can be bounded by*

$$\left|\frac{1}{2}\int_{-1}^{1} f(x)\,\mathrm{d}x - \sum_{i=0}^{n-1} w_{i,n} f(\xi_i)\right| \le \left(1 + 8\sqrt{2}n^2\right)\frac{r}{r-1}\, r^{-n}\cdot \|f\|_{H_r^1}. \tag{5.21}$$

Unfortunately, such a theoretical bound of $\nu(n)$ does not yet exist for the (L)-sequence. However, it is easy to compute the weights for finite n with a computer. In Figure 2, similar to [40], we give the values of $\nu(n)$ for $n = 1, \ldots, 1000$. They show that for all practically relevant numbers of points, it holds that $\nu(n) \in [1, 1.2]$. If this would be true for all $n \in \mathbb{N}$, i.e. $\nu(n) \le c_L$, we could directly apply our results from Section 2 to tensor products of (L)-Leja points, since by Theorem 5.7 we have for $n \in \{1, \ldots, 1000\}$

$$\left|\int_{-1}^{1} f(x)\,\mathrm{d}x - \sum_{i=0}^{n-1} w_{i,n} f(\xi_i)\right| \le (1 + c_L)\frac{r}{r-1}\, r^{-n}\cdot \|f\|_{H_r^1}.$$

5.4.1. Integration in tensor products of Hardy spaces

Now we define the hierarchical quadrature rules $\Delta_0(f) = Q_1(f) = f(\xi_0)$ and

$$\Delta_k(f) := Q_{k+1}(f) - Q_k(f) = \sum_{i=0}^{k-1} f(\xi_i)(w_{i,k} - w_{i,k-1}) + w_{k,k} f(\xi_k),\, k \ge 1, \tag{5.22}$$

which, by the way, could also be obtained by integrating the hierarchical interpolation operator from the last section.

Now, let $\xi_0 \in [-1, 1]$ be the first point of any Leja-sequence, which always has weight $w_{0,1} = 1$. Using Lemma 5.1 we obtain for $\Delta_0(f) = Q_1(f) = 1 f(\xi_0)$ that $\|\Delta_0\|_{H_r^1\to\mathbb{R}} \in [1, \frac{r}{r-1}]$.

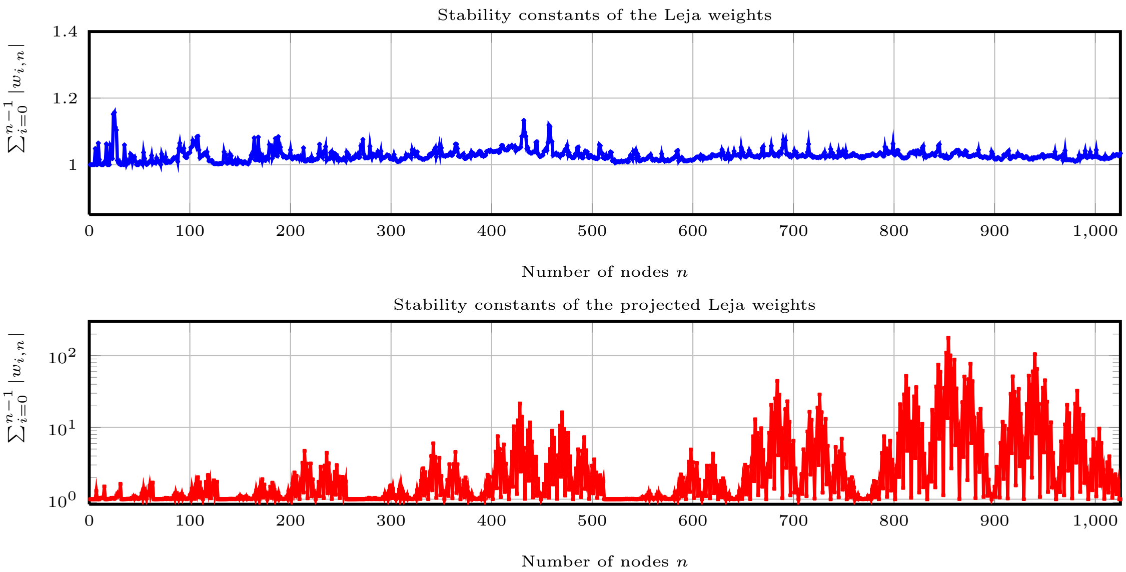

Figure 2: Stability constants $\nu(n) = \sum_{i=0}^{n-1} |w_{i,n}|$ of the weights of the (L)-Leja sequence on $[-1,1]$ (top) and of the projected $\Re$-Leja sequence (bottom).

Moreover, for $k \geq 1$ we can use Corollary 5.8 to estimate

$$\begin{aligned}|\Delta_k(f)| \leq& |(I_1 - Q_k)(f)| + |(I_1 - Q_{k+1})(f)| \leq 2\left(1 + 8\sqrt{2}(k+1)^2\right) \frac{r}{r-1}\ r^{-k}\ \|f\|_{H^1_r} \\ \leq& \left(2 + 16\sqrt{2}\right)(k+1)^2 \frac{r}{r-1}\ r^{-k}\ \|f\|_{H^1_r}\end{aligned}$$

and therefore the norm of $\Delta_{\boldsymbol{k}} = \bigotimes_{j=1}^d \Delta_{k_j}$ by

$$\|\Delta_{\boldsymbol{k}}\|_{H^1_{\boldsymbol{r}} \to \mathbb{R}} \leq \prod_{j=1}^d r_j^{-k_j} \frac{r_j}{r_j - 1} \begin{cases} 1 & k_j = 0, \\ (2 + 16\sqrt{2})\,(k_j+1)^2 & k_j \geq 1. \end{cases} \tag{5.23}$$

Now, the Corollaries 4.1 and 4.2 can be applied with $\boldsymbol{a} = \log(\boldsymbol{r})$ to yield the final error estimates.

Corollary 5.9. *Let $T \in \mathbb{R}_{\geq 0}$ and $\boldsymbol{r} \in \mathbb{R}^d_{>1}$ with $1 < r_1 \leq r_2 \leq \ldots$ and $d \in \mathbb{N} \cup \{\infty\}$. The algorithm $\mathcal{A}_{T,\boldsymbol{a}}(f) = \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{a}}(T)} \Delta_{\boldsymbol{k}}(f)$, where $\boldsymbol{a} = \log(\boldsymbol{r})$ and $\Delta_{\boldsymbol{k}}$ as the d-fold tensor product of (5.22) with $(\xi_i)_{i=0}^\infty$ being the $\Re$-Leja sequence, approximates the integral $I_d^{\mathrm{int}}(f)$ of $f \in H^1_{\boldsymbol{r}}$ with the following asymptotic rates of convergence.*

(i) For $d < \infty$ and $\delta \in (0, 1 - \frac{d}{e}\kappa(d)^{-1})$ we have

$$\begin{aligned}\sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \left| I_d^{\mathrm{int}}(f) - \mathcal{A}_{T,\boldsymbol{a}}(f) \right| &\preceq_{d,\boldsymbol{a},\delta} \exp\left(-\kappa(d)(1-\delta)\mathrm{gm}(\boldsymbol{a}) \sqrt[d]{N}\right)\ N^{\frac{d-1}{d}} \\ &\preceq_{d,\boldsymbol{a}} \quad \exp\left(-\frac{d}{e}\mathrm{gm}(\boldsymbol{a}) \sqrt[d]{N}\right).\end{aligned}$$

(ii) For $d = \infty$ we have for every $\beta > 1$ that ensures $\sum_{j=1}^\infty \frac{1}{e^{(1-\delta)a_j/\beta} - 1} < \infty$

$$\lim_{d \to \infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}} \leq 1} \left| I_d^{\mathrm{int}}(f) - \mathcal{A}_{T,\boldsymbol{a}}(f) \right| \preceq_{\boldsymbol{a}} N^{-(\beta - 1)}.$$

(iii) For $d = \infty$ and $r_j \geq \alpha j$ we have for all $0 < \hat{\alpha} < \alpha$

$$\lim_{d\to\infty} \sup_{\|f\|_{H^1_{\boldsymbol{r}}}\leq 1} \left| I_d^{\text{int}}(f) - \mathcal{A}_{T,\boldsymbol{a}}(f) \right| \preceq_{\hat{\alpha}} (\log N)^{-\frac{1}{2}} N^{1+\frac{\hat{\alpha}}{4}-\frac{3}{8}\hat{\alpha}\log(N)^{1/2}}.$$

Moreover, due to the numerical results given in Figure 2 for the (L)-Leja sequence we believe that the optimal worst-case integration error in $H^1_{\boldsymbol{r}}$ can be bounded from above by the same asymptotic convergence rate that is achieved by the Taylor series for L^∞-approximation, i.e.

Conjecture. For $d < \infty$ the optimal worst-case integration error in $H^1_{\boldsymbol{r}}$ can be bounded by

$$\inf_{\substack{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\in[-1,1]^d \\ w_1,\ldots,w_N\in\mathbb{R}}} \sup_{\|f\|_{H^1_{\boldsymbol{r}}}\leq 1} \left| I_d^{\text{int}}(f) - \sum_{i=1}^N w_i f(\boldsymbol{x}_i) \right| \preceq_{d,\boldsymbol{a}} \exp\left(-\kappa(d)\text{gm}(\boldsymbol{a})\sqrt[d]{N}\right) \; N^{\frac{d-1}{d}}.$$

6. Concluding remarks

In the present paper we derived matching upper and lower asymptotic bounds for sums of the form (1.8) in finite dimensions, with explicitly given constants. In the infinite-dimensional setting we gave both, algebraic and sub-exponential upper bounds. Applications are given by tensor product approximation of multivariate analytic functions. As a case study we investigated multivariate approximation by Taylor polynomials as well as interpolation and integration at certain tensor products of Leja points. Here, we only considered functions that are analytic in polydiscs, which has applications in certain model problems from uncertainty quantification [16, 29]. However, our results are also directly applicable to functions with more general domains of analyticity, e.g. polyellipses [55]. Here, one needs results on best approximation by univariate polynomials to functions bounded and analytic in certain ellipses. This can be achieved by estimates of Legendre coefficients [19] or bounds on Tschebyscheff interpolation [8]. Moreover, we gave numerical evidence that quadrature at (L)-Leja points on $[-1, 1]$ allows for exponential convergence without any additional algebraic terms. Note finally the numerical experiments in [40], which suggest that it is possible to construct Leja points for integration on $\mathbb{R}$ with respect to the Gaussian measure that exhibit similar favorable stability properties and thus similar sub-exponential convergence rates. This is important for stochastic and parametric diffusion problems with log-normal distributed diffusion coefficients, see e.g. [32].

Acknowledgments This research was supported by the *Collaborative Research Center 1060* of the DFG. The authors would like to thank Tino Ullrich, André Uschmajew and the anonymous referees for several helpful remarks on this manuscript.

A. Proof of Lemma 2.2

Let $\mathcal{E}_{\boldsymbol{1}}(1) := \{\boldsymbol{x} \in \mathbb{R}^d_{\geq 0} : \sum_{j=1}^d x_j \leq 1\}$ denote the d-dimensional unit simplex. We want to prove for $T \in \mathbb{R}_{\geq 0}$ and $d \in \mathbb{N}$

$$\int_{\mathbb{R}^d_{\geq 0}\setminus\mathcal{E}_{\boldsymbol{1}}(1)} \exp\left(-T\sum_{j=1}^d y_j\right) \,\mathrm{d}\boldsymbol{y} = T^{-d}\frac{\Gamma(d,T)}{(d-1)!}. \tag{A.1}$$

First, we need two auxiliary results.

Lemma A.1. *(Hermite–Genocci formula)*

Let $f : \mathbb{R} \to \mathbb{R}$ be at least d-times continuously differentiable. Then the divided differences of f associated to $x_0, \dots, x_d \in \mathbb{R}$ can be written as

$$f[x_0, x_1, \dots, x_d] = \int_{\mathcal{E}_{\mathbf{1}}(1)} f^{(d)} \left(x_0 + \sum_{j=1}^{d} y_j (x_j - x_0) \right) \, \mathrm{d}\boldsymbol{y}. \tag{A.2}$$

The proof of Lemma A.1 can be found in e.g. [20].

Lemma A.2. *It holds that*

$$\int_{\mathcal{E}_{\mathbf{1}}(1)} \exp \left(-T \sum_{j=1}^{d} y_j \right) \, \mathrm{d}\boldsymbol{y} = T^{-d} \left(1 - e^{-T} \left(\sum_{k=0}^{d-1} \frac{T^k}{k!} \right) \right). \tag{A.3}$$

Proof. We apply Lemma A.1 to the function $f(t) = \exp(t)$ where the points are given by $x_0 = 0$ and $x_1 = x_2 = \ldots = x_d = -T$. It remains to show that

$$\exp[0, \underbrace{-T, \dots, -T}_{d \text{ times}}] = T^{-d} \left(1 - e^{-T} \left(\sum_{k=0}^{d-1} \frac{T^k}{k!} \right) \right), \tag{A.4}$$

which we will accomplish by induction.

To this end, let us assume that (A.4) holds. Then, with $x_{d+1} = -T$ and using the identity $f[x_1, \dots, x_{d+1}] = f^{(d)}(-T)/d!$, we get

$$\begin{aligned} \exp[x_0, \dots, x_d, x_{d+1}] &= \frac{\exp[x_1, \dots, x_{d+1}] - \exp[x_0, \dots, x_d]}{-T} \\ &= \frac{\frac{1}{d!} e^{-T} - T^{-d} \left(1 - e^{-T} \left(\sum_{k=0}^{d-1} \frac{T^k}{k!} \right) \right)}{-T} \\ &= T^{-(d+1)} \left(1 - e^{-T} \left(\sum_{k=0}^{d} \frac{T^k}{k!} \right) \right). \end{aligned}$$

□

Now we are in the position to show (A.1). To this end, we note that it holds [1]

$$e^{-T} \sum_{k=0}^{d-1} \frac{T^k}{k!} = \frac{\Gamma(d, T)}{(d-1)!} = \frac{1}{(d-1)!} \int_T^{\infty} t^{d-1} e^{-t} \, \mathrm{d}t, \tag{A.5}$$

where $\Gamma(d, T)$ denotes the upper incomplete Gamma function.

Hence, using

$$\begin{aligned} \int_{\mathbb{R}^d_{\geq 0} \setminus \mathcal{E}_{\mathbf{1}}(1)} \exp \left(-T \sum_{j=1}^{d} y_j \right) \, \mathrm{d}\boldsymbol{y} &= \left(\int_{\mathbb{R}_{\geq 0}} \exp(-Ty) \, \mathrm{d}y \right)^d - \int_{\mathcal{E}_{\mathbf{1}}(1)} \exp \left(-T \sum_{j=1}^{d} y_j \right) \, \mathrm{d}\boldsymbol{y} \\ &= T^{-d} - \int_{\mathcal{E}_{\mathbf{1}}(1)} \exp \left(-T \sum_{j=1}^{d} y_j \right) \, \mathrm{d}\boldsymbol{y} \end{aligned}$$

we arrive at the desired equality.

□

B. Proof of Proposition 5.3

We prove the claim for $d = 2$, i.e. for $P = P_1 \otimes P_2$. The higher-dimensional setting follows by recursion. Moreover, it is enough to consider functions $h(x, y) = \sum_{i=1}^{n} f_i(x) g_i(y)$, where $n \in \mathbb{N}$, $f_1, \ldots, f_n \in H^1_{r_1}$ and $g_1, \ldots, g_n \in H^1_{r_2}$ are arbitrary.

Here, we prove the claim for bounded linear operators $P_j : H^1_{r_j} \to \mathcal{C}_1, j \in \{1, 2\}$. To this end, we abbreviate $\|P_j\| = \|P_j\|_{H^1_{r_j} \to \mathcal{C}_1}$ and $\|P_1 \otimes P_2\| = \|P_1 \otimes P_2\|_{H^1_{\mathbf{r}} \to \mathcal{C}_2}$. Furthermore, we use the notation $P_j^{(x)} h(x, y)$ and $P_j^{(y)} h(x, y)$ to denote the application of P_j to the variable x or y, respectively. Then, it holds for all fixed $y \in [-1, 1]$ that

$$P_2^{(y)} h(\cdot, y) = \sum_{i=1}^{n} f_i(\cdot) P_2^{(y)} g_i(y) \ \in H^1_{r_1}$$

and, by Fatou's Theorem [24],

$$h(r_1 e^{\mathrm{i}\varphi_1}, \cdot) = \sum_{i=1}^{n} f_i(r_1 e^{\mathrm{i}\varphi_1}) g_i(\cdot) \in H^1_{r_2} \quad \text{for almost all } \varphi_1 \in [0, 2\pi).$$

Therefore we obtain

$$\begin{aligned}
\|(P_1 \otimes P_2) h\| &= \sup_{(x,y) \in [-1,1]^2} |(P_1^{(x)} \otimes P_2^{(y)}) h(x, y)| = \sup_{y \in [-1,1]} \sup_{x \in [-1,1]} \left| P_1^{(x)} \left(P_2^{(y)} h(x, y) \right) \right| \\
&\leq \sup_{y \in [-1,1]} \|P_1\| \left\| P_2^{(y)} h(\cdot, y) \right\|_{H^1_{r_1}} \\
&= \|P_1\| \sup_{y \in [-1,1]} \left(\frac{1}{2\pi} \int_{[0,2\pi)} \left| P_2^{(y)} h(r_1 e^{\mathrm{i}\varphi_1}, y) \right| \, \mathrm{d}\varphi_1 \right) \\
&\leq \|P_1\| \left(\frac{1}{2\pi} \int_{[0,2\pi)} \left| \|P_2\| \|h(r_1 e^{\mathrm{i}\varphi_1}, \cdot)\|_{H^1_{r_2}} \right| \, \mathrm{d}\varphi_1 \right) \\
&= \|P_1\| \, \|P_2\| \left(\frac{1}{2\pi} \int_{[0,2\pi)} \left| \frac{1}{2\pi} \int_{[0,2\pi)} \left| h(r_1 e^{\mathrm{i}\varphi_1}, r_2 e^{\mathrm{i}\varphi_2}) \right| \, \mathrm{d}\varphi_2 \right| \, \mathrm{d}\varphi_1 \right) \\
&= \|P_1\| \, \|P_2\| \, \|h\|_{H^1_{\mathbf{r}}}.
\end{aligned}$$

The proof for the case $P = P_1 \otimes P_2$, where $P_j : H^1_{r_j} \to \mathbb{R}$ are bounded linear functionals, works exactly in the same way by exploiting the recursive nature of the L^1-norm. □

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied Mathematics Series. Dover Publications, 1964.

[2] J.-E. Andersson and B. Bojanov. A note on the optimal quadrature in H^p. *Quadrature*, 308:301–308, 1984.

[3] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 52(2):317–355, 2007.

[4] I. Babuška, R. Tempone, and G. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.

[5] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12(4):273–288, 2000.

[6] J. Beck, F. Nobile, L. Tamellini, and R. Tempone. Convergence of quasi-optimal stochastic Galerkin methods for a class of PDEs with random coefficients. *Computers and Mathematics with Applications*, 67(4):732 – 751, 2014.

[7] A. Beged-Dov. Lower and upper bounds for the number of lattice points in a simplex. *SIAM J. on Applied Mathematics*, 22(1):106–108, 1972.

[8] S. Bernstein. Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné. *Mémoires publiés par la Classe des Sci. Acad. de Belgique*, 2:1–103, 1912.

[9] J. Borwein and O. Chan. Inequalities for the incomplete Gamma function. *Math. Inequaities. and Applications*, 12(1):115–121, 2009.

[10] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:1–123, 2004.

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

[12] A. Chkifa. On the Lebesgue constant of Leja sequences for the complex unit disk and of their real projection. *J. of Approximation Theory*, 166:176–200, 2013.

[13] A. Chkifa, A. Cohen, and C. Schwab. High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs. *Foundations of Computational Mathematics*, 14(4):601–633, 2013.

[14] A. Cohen and A. Chkifa. *Sampling Theory, a Renaissance: Compressive Sensing and Other Developments*, chapter On the Stability of Polynomial Interpolation Using Hierarchical Sampling, pages 437–458. Springer International Publishing, Cham, 2015.

[15] A. Cohen, M. Davenport, and D. Leviatan. On the Stability and Accuracy of Least Squares Approximations. *Foundations of Computational Mathematics*, 13(5):819–834, 2013.

[16] A. Cohen, R. DeVore, and C. Schwab. Convergence rates of best n-term Galerkin approximations for a class of elliptic sPDEs. *Foundations of Computational Mathematics*, 10(6):615–646, 2010.

[17] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Analysis and Applications*, 09(01):11–47, 2011.

[18] D. Dũng and T. Ullrich. Lower bounds for the integration error for multivariate functions with mixed smoothness and optimal Fibonacci cubature for functions on the square. *to appear in Math. Nachrichten*, 2014. arXiv 1311.1563.

[19] P. Davis. *Interpolation and Approximation.* Dover Puplications, New York, 1975.

[20] P. Deuflhard and A. Hohmann. *Numerische Mathematik: Eine algorithmisch orientierte Einführung*, volume 1. De Gruyter, Berlin, 2008.

[21] J. Dick. A Taylor space for multivariate integration. *Monte Carlo Methods Appl.*, 12:99–112, 2006.

[22] J. Dick. Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.*, 46:1519–1553, 2008.

[23] J. Dick, F. Kuo, Q. Le Gia, D. Nuyens, and C. Schwab. Higher order QMC Petrov–Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. on Numerical Analysis*, 52(6):2676–2702, 2014.

[24] P. Duren. *Theory of H_p spaces.* Dover books on mathematics. Courier Corporation, 2000.

[25] D. Dng and M. Griebel. Hyperbolic cross approximation in infinite dimensions. *Journal of Complexity*, 33:55–88, 2016. Also available as INS Preprint No. 1501.

[26] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing.* Applied and Numerical Harmonic Analysis. Birkhäuser, 2013.

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

[28] M. Griebel and H. Harbrecht. A note on the construction of L-fold sparse tensor product spaces. *Constructive Approximation*, 38(2):235–251, 2013. Also available as INS Preprint No. 1205.

[29] M. Griebel and C. Rieger. Reproducing kernel Hilbert spaces for parametric partial differential equations. 2015. submitted. Also available as INS preprint no. 1511.

[30] W. Hackbusch. *Tensor Spaces and Numerical Tensor Calculus.* Springer Series in Computational Mathematics. Springer-Verlag Berlin Heidelberg, 2012.

[31] A.-L. Haji-Ali, H. Harbrecht, M. Peters, and M. Siebenmorgen. Novel results for the anisotropic sparse quadrature and their impact on random diffusion problems. *Preprint 2015-27 Mathematisches Institut Universität Basel*, 2015.

[32] H. Harbrecht, M. Peters, and M. Siebenmorgen. Multilevel accelerated quadrature for PDEs with log-normal distributed random coefficient. *Preprint 2013-18, Mathematisches Institut, Universität Basel*, 2013.

[33] A. Hinrichs, E. Novak, and M. Ullrich. On weak tractability of the Clenshaw–Curtis Smolyak algorithm. *Journal of Approximation Theory*, 183:31 – 44, 2014.

[34] C. Irrgeher, P. Kritzer, G. Leobacher, and F. Pillichshammer. Integration in Hermite spaces of analytic functions. *J. Complexity*, 31:380–404, 2015.

[35] M. Kowalski, A. Werschulz, and H. Woniakowski. Is Gauss quadrature optimal for analytic functions? *Numerische Mathematik*, 47(1):89–98, 1985.

[36] G. Krantz. *Function Theory of Several Complex Variables.* AMS Chelsea Publishing, Providence, Rhode Island,, 1992.

[37] P. Kritzer, F. Pillichshammer, and H. Wozniakowski. Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces. *Math. Comp.*, 83(287):1189–1206, 2014.

[38] F. Larkin. Optimal approximation in Hilbert spaces with reproducing kernel functions. *Math. Comp*, 24(112):911–921, 1970.

[39] G. Migliorati, F. Nobile, E. von Schwerin, and R. Tempone. Analysis of discrete L^2 projection on polynomial spaces with random evaluations. *Foundations of Computational Mathematics*, 14(3):419–456, 2014.

[40] A. Narayan and J. Jakeman. Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation. *SIAM J. on Scientific Computing*, 36(6):A2952–A2983, 2014.

[41] F. Nobile, L. Tamellini, and R. Tempone. Convergence of quasi-optimal sparse grid approximation of Hilbert-valued functions: application to random elliptic PDEs. MATHICSE Technical Report Nr. 12.2014, 2014.

[42] F. Nobile, R. Tempone, and C. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. on Numerical Analysis*, 46(5):2411–2442, 2008.

[43] E. Novak and K. Ritter. High dimensional integration of smooth functions over cubes. *Numerische Mathematik*, 75(1):79–97, 1996.

[44] S. Paskov. Average case complexity of multivariate integration for smooth functions. *J. of Complexity*, 9:291–312, 1993.

[45] K. Petras. On the Smolyak cubature error for analytic functions. *Advances in Computational Mathematics*, 12(1):71–93, 2000.

[46] C. Rieger and B. Zwicknagl. Sampling inequalities for infinitely smooth functions, with applications to interpolation and machine learning. *Advances in Computational Mathematics*, 32(1):103–129, 2010.

[47] W. Rudin. *Function theory in polydiscs.* W.A. Benjamin, Inc., New York, 1969.

[48] R. Schaback. Optimal nodes for interpolation in Hardy spaces. *Mathematische Zeitschrift*, 178:169–178, 1982.

[49] S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963.

[50] F. Stenger. Optimal convergence of minimum norm approximations in H^p. *Numerische Mathematik*, 362:345–362, 1978.

[51] K. Suzuki. Accelerating convergence and tractability of multivariate integration for infinitely differentiable functions. arXiv:1505.02003, 2015.

[52] R. Taylor and V. Totik. Lebesgue constants for Leja points. *IMA J. of Numerical Analysis*, 30(2):462–486, 2010.

[53] V. Temlyakov. Approximation of functions with bounded mixed derivative. *Trudy Mat. Inst. Steklov.*, 178:1–112, 1986. English transl. in Proc. Steklov Inst. Math. 178 (1986).

[54] R. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. of Numerical Analysis*, 27(2):232–261, 2007.

[55] H. Tran, C. Webster, and G. Zhang. Analysis of quasi-optimal polynomial approximations for parameterized PDEs with deterministic and stochastic coefficients. 2015. arXiv:1508.01821.

[56] J. Traub, H. Woźniakowski, and G. Wasilkowski. *Information-Based Complexity.* Academic Press., 1988.

[57] G. Wasilkowski and H. Woźniakowski. Weighted tensor product algorithms for linear multivariate problems. *J. Complexity*, 15:402–447, 1999.