Integration and L^2-Approximation on Tensor Products of Hilbert Spaces of Increasing Smoothness

Robin Rüßmann

Vom Fachbereich Mathematik der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation

1. Gutachter:	Prof. Dr. Klaus Ritter
	RPTU Kaiserslautern-Landau
2. Gutachter:	Prof. Dr. Gunther Leobacher
	Universität Graz, Österreich
Datum der Disputation:	17. April 2026

D-386

Contents

Introduction and Overview

Two of the most studied problems in information-based complexity are the integration problem

$$\mathrm{Int}(f) = \int_{D^d} f \, d\mu$$

and the approximation problem

$$\mathrm{App}(f) = [f]_{L^2(\mu)},$$

where $d \in \mathbb{N} \cup \{\infty\}$ and μ is given as the d-fold product measure of some probability measure μ_0 on D. In each case, $f \colon D^d \to \mathbb{R}$ is a (square-) integrable function with respect to μ. In this thesis, we mostly consider the infinite-variate case $d = \infty$. Typically, one assumes that it is known a priori that $f \in H$ for some function space H; we specifically only consider the case of $H = H(k)$ being a reproducing kernel Hilbert space (RKHS).

Neither the integration problem nor the L^2-approximation problem can usually be solved exactly. Therefore, we approximate both problems, using algorithms based on only finitely many function evaluations. Without loss of generality, we use linear algorithms A given by

$$A(f) = \sum_{i=1}^{n} f(\boldsymbol{x}_i) \cdot a_i$$

for $f \in H(k)$, where $n \in \mathbb{N}$, $\boldsymbol{x}_i \in D^d$ and $a_i \in \mathbb{R}$ for the integration problem or $a_i \in L^2(\mu_0)$ for the L^2-approximation problem, respectively. To such an algorithm we associate its cost, which is determined by the effort it takes to evaluate f at the points $\boldsymbol{x}_i$, as well as its worst-case error on the unit ball in $H(k)$.

Without precisely defining $\mathrm{cost}(A)$ and $\mathrm{error}(A, k)$ yet, we mention that we are not only interested in the performance of individual algorithms, but in the inherent complexity of Int and App on $H(k)$. We quantify this via the n-th minimal error, which is given by the smallest possible value of $\mathrm{error}(A, k)$ among all algorithms A with cost at most n, i.e.

$$e_n(k) = \inf\{\mathrm{error}(A, k) \colon \mathrm{cost}(A) \leqslant n\}.$$

In the case of d being finite, an important question is how the inherent difficulty of Int and App depends on the number of variables in D^d. For instance, one could study the case of $D = [0, 1]$ and d-variate Lebesgue-integration of functions $f \in H(k)$ that are defined on D^d. The study of $e_n(k)$ both as a function of n and d simultaneously is the subject of *tractability* studies, which is comprehensively studied in Novak and Woźniakowski (2008), Novak and Woźniakowski (2010) and Novak and Woźniakowski (2012) and is the subject of a huge number of articles.

In many cases, as d increases, both Int and App become very difficult, which is also called the curse of dimensionality. To avoid this and achieve favorable results even for a large number of variables, function space settings are studied where, loosely speaking, the influence of the j-th variable on Int or App decreases fast enough as j grows.

In the case $d = \infty$, the main focus of this thesis, we also seek to control the importance of the j-th variable as j increases, which in contrast to the finite variate case is even needed to establish the well-definedness of Int and App.

This thesis is in large parts based on the results of Gnewuch et al. (2024) and Gnewuch et al. (2026), which are the topics of Chapter 2 and 3, respectively. The results of Gnewuch et al. (2024) will be abstracted to a more general setting, of which one particular instance was considered in the paper. In Chapter 1, we describe a more general function space setting, which serves to put the results of the later chapters into a broader context.

Function Space Setting: Tensor Products of RKHSs

We describe the infinite-variate function space setting of this thesis. As univariate building blocks, let $(k_j)_{j\in\mathbb{N}}$ be a sequence of real-valued nonzero reproducing kernels k_j, which all share a domain $D \times D \neq \varnothing$. Additionally, let D be equipped with some σ-algebra, and let μ_0 be a probability measure on D. The multi- and infinite-variate spaces we consider arise as tensor products of the univariate spaces $H(k_j)$. We restrict the domain $D^{\mathbb{N}}$ appropriately to $\mathfrak{X} \subseteq D^{\mathbb{N}}$ to ensure the convergence of the infinite product

$$K(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j\in\mathbb{N}} k_j(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}$. It is known that based on functions $u_j \in H(k_j)$ with norm one for $j \in \mathbb{N}$, the domain

$$\mathfrak{X} = \{\boldsymbol{x} \in D^{\mathbb{N}} : \sum_{j \in \mathbb{N}} |k_j(x_j, x_j) - 1| < \infty \text{ and } \sum_{j \in \mathbb{N}} |u_j(x_j) - 1| < \infty\}$$

ensures the convergence as desired. The infinite-variate function spaces we consider are then always of the form $H(K)$ with K defined on the domain $\mathfrak{X} \times \mathfrak{X}$. This is motivated by the fact that in this case

$$H(K) \simeq \bigotimes_{j \in \mathbb{N}} H(k_j)^{(u_j)},$$

where the right-hand side denotes the incomplete Hilbert space tensor product according to von Neumann (1939), see also Rüßmann (2020). This enables us to utilize results for abstract tensor products for the study of the function space $H(K)$, in particular regarding embeddings into other spaces. For a brief overview of the incomplete tensor product as well as the results that are important to this thesis, we have included Appendix A.

It will be shown that $f \in H(K)$ admits an orthogonal decomposition

$$f = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} f_{\boldsymbol{\nu}},$$

where $\boldsymbol{U}$ denotes the set of all finite subsets of $\mathbb{N}$. For a fixed $\boldsymbol{\nu} \in \boldsymbol{U}$, all functions $f_{\boldsymbol{\nu}}$ obtained this way belong to a subspace of $H(K)$ that is, in essence, a finite-variate tensor product space of certain subspaces of $H(k_j)$ for finitely many j.

We put this into a broader context. In the literature on Hilbert spaces of infinite-variate functions, the spaces that are considered admit a similar decomposition. However, loosely speaking, in that setting all the underlying univariate kernels k_j are the same, and the importance of different variable groups is determined by weights $\gamma_{\boldsymbol{\nu}}$. For more details, we refer to Chapter 1.

Spaces in the second setting have been studied, for instance, in Hickernell et al. (2010), Niu et al. (2011), Kuo et al. (2010b), Wasilkowski and Woźniakowski (2011), Plaskota and Wasilkowski (2011), Wasilkowski (2012), Gnewuch (2013), Dick and

Gnewuch (2014), Wasilkowski (2013) and Kuo et al. (2017a). We remark that for the study of integration and L^2-approximation, additionally it is often assumed that the underlying univariate kernel k is anchored at some point $a \in D$, which means $k(a, a) = 0$.

We now give an overview over the thesis, focusing on the setting of each chapter along with the corresponding main results.

Infinite-Variate Square-Integrability

In Chapter 1, we study the general infinite-variate tensor product setting as described above, establishing some basic results. In particular, we study the question of when Int and App with respect to μ are actually well-defined. It is reasonable to assume that at least the univariate spaces $H(k_j)$ fulfill

$$H(k_j) \subseteq L^2(\mu_0).$$

The corresponding inclusion operator is always continuous, and we denote it by T_j. We assume

$$\sum_{j\in\mathbb{N}} |\|T_j\| - 1| < \infty,$$

which ensures a proper mode of convergence of the product $\prod_{j\in\mathbb{N}} \|T_j\|$. In the main result for this chapter, Theorem 1.4, under some additional assumptions, which ensure intuitively that $T_j u_j$ is close enough to 1 in $L^2(\mu_0)$ for large $j \in \mathbb{N}$, we establish that

$$H(K) \subseteq L^2(\mu)$$

with a continuous identical inclusion T of norm

$$\|T\| = \prod_{j\in\mathbb{N}} \|T_j\|.$$

Function Space Setting: Spaces of Increasing Smoothness

In Chapter 2, we turn to the case of the univariate building blocks k_j of the infinite-variate kernel k_j having a specific form, namely

$$k_j(x, y) = 1 + \sum_{\nu \in \mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu(x) \cdot \mathfrak{e}_\nu(y),$$

where the $\mathfrak{e}_\nu$ are square-integrable functions forming an orthonormal system in $L^2(\mu_0)$ and the $\alpha_{\nu,j}$ are suitably chosen Fourier weights. In particular, at least for the purposes of this introductory presentation, we assume that the monotonicity and summability properties

(A1) $0 < \alpha_{1,j} \leqslant \alpha_{2,j} \leqslant \ldots$ for every $j \in \mathbb{N}$,

(A2) $\sum_{\nu,j \in \mathbb{N}} \alpha_{\nu,j}^{-1} < \infty$,

(A3) $\sum_{j \in \mathbb{N}} \gamma_j < \infty$, where $\gamma_j := \sup_{\nu \in \mathbb{N}} \alpha_{\nu,1}/\alpha_{\nu,j}$,

hold. Here, we can choose $u_j = 1$ in the definition of K, and the results from Chapter 1 are applicable, so that we have

$$H(K) \subseteq L^2(\mu),$$

here even with a compact identical embedding.

The special case $\alpha_{\nu,1} = \gamma_j \cdot \alpha_{\nu,j}$ leads to the well-studied weighted tensor product case, where all univariate spaces $H(k_j)$ are the same as vector spaces, with equivalent norms. In contrast to this, in many cases, it is also sensible to consider the case where $H(k_j) \subseteq H(k_1)$ with a compact embedding, at least for large j. This corresponds to

$$\lim_{\nu \to \infty} \frac{\alpha_{\nu,1}}{\alpha_{\nu,j}} = 0$$

and we refer to spaces in this setting as *spaces of increasing smoothness.*

Spaces $H(k_j)$ of this type include Haar spaces, Walsh spaces, Korobov Spaces and certain Sobolev spaces, which have all been studied in this context in Gnewuch et al. (2019), where also the general setting has been studied under the additional assumption that

$$\mathfrak{X} = D^{\mathbb{N}}$$

for the domain $\mathfrak{X}$ of K. We also refer to the predecessor works Gnewuch et al. (2017) and Hefter and Ritter (2015).

The case

$$\mathfrak{X} \subsetneq D^{\mathbb{N}}$$

poses additional challenges, which we study in this thesis. Here, our main example, which we also study in some detail, are the Hermite spaces of finite or infinite smoothness, which were studied in this setting as function spaces in Gnewuch et al. (2022). For the study of integration and L^2-approximation on finite-variate Hermite spaces, we refer in the case of finite smoothness to Irrgeher and Leobacher (2015), Dick et al. (2018), Kazashi et al. (2023), Dũng and Nguyen (2023), and Leobacher et al. (2023). In the case of infinite smoothness, we refer to Irrgeher and Leobacher (2015), Irrgeher et al. (2015), Irrgeher et al. (2016a), and Irrgeher et al. (2016b).

Main Results: Infinite-Variate Problems in Spaces of Increasing Smoothness

On the space $H(K)$, we establish upper and lower error bounds for integration and L^2-approximation. To present them, for a sequence $\boldsymbol{\beta}$, we introduce the notation $\operatorname{decay}(\boldsymbol{\beta})$, which denotes, roughly speaking, the polynomial rate of convergence of $\boldsymbol{\beta}$ to zero; for a reproducing kernel k, we denote by $\operatorname{dec}(k)$ the decay of the sequence $(e_n(k))_{n\in\mathbb{N}}$ of n-th minimal errors. Our main result is this, see Theorem 2.9: For integration and L^2-approximation we have

$$\min\left(\operatorname{dec}(k_1), \frac{\operatorname{decay}(\boldsymbol{\gamma}) - 1}{2}\right) \leqslant \operatorname{dec}(K) \leqslant \min\left(\operatorname{dec}(k_1), \frac{\operatorname{decay}(\boldsymbol{\alpha}_1^{-1}) - 1}{2}\right),$$

where γ_j according to (A3) determines the norm of the embedding from $H(k_j - 1)$ into $H(k_1 - 1)$, and $\alpha_{1,j}^{-1}$ determines the norm of the embedding from $H(k_j)$ into $L^2(\mu_0)$. In the important case of the $\alpha_{\nu,j}$ fulfilling a polynomial or exponential growth condition, see Section 2.1.2, the upper and lower bounds match, and we get

$$\operatorname{dec}(K) = \min\left(\operatorname{dec}(k_1), \frac{\rho - 1}{2}\right),$$

where $\operatorname{decay}(\boldsymbol{\gamma}) = \rho = \operatorname{decay}(\boldsymbol{\alpha}_1^{-1})$. An intuition for this is that the error of the infinite-variate problem gets smaller if we increase the growth of the Fourier weights in the j direction, which corresponds to the spaces $H(k_j)$ increasing in smoothness faster as j grows. However, increasing this growth rate is only beneficial until $\operatorname{dec}(K) = \operatorname{dec}(k_1)$, which corresponds to the infinite-variate problem being (almost) as easy as the hardest underlying univariate problem. We therefore only need to find error bounds for univariate integration or L^2-approximation on $H(k_1)$ to establish such bounds for the corresponding infinite-variate problem.

We study one example where $\mathfrak{X} \subsetneq D^{\mathbb{N}}$ holds, namely Hermite spaces, in more detail, relying on Gnewuch et al. (2024). Here, μ_0 is the standard normal distribution on $D = \mathbb{R}$. In the case of the Fourier weights growing at a polynomial rate, which corresponds to the spaces $H(k_j)$ being of finite smoothness, we obtain

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot \min\left(2r_1, \rho - 1\right)$$

for the integration problem and

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot \min\left(r_1, \rho - 1\right)$$

for the L^2-approximation problem, where r_1 is the rate of polynomial growth of $(\alpha_{\nu,1})_{\nu \in \mathbb{N}}$. We note that this factor of 2 between the difficulty of the integration problem and that of the L^2-approximation problem was not observed in any example studied in Gnewuch et al. (2019). In the case of exponential growth of the Fourier weights, we have

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot (\rho - 1)$$

for the integration problem and the L^2-approximation problem, so the aforementioned factor of 2 disappears here.

Let us mention that the upper error bounds in this setting are obtained by MDM-algorithms (MDM stands for multivariate decomposition method), which are usually employed in spaces $H(M)$ where M is an anchored kernel. Even though K is not anchored, we establish that $H(K)$ can be embedded into such a space $H(M)$, which allows us to construct MDM-algorithms to use in $H(K)$. The kernel M plays an important role in the proofs for upper error bounds, but not in the construction of said MDM-algorithm. Therefore, an MDM-algorithm can be established on $H(K)$ without first needing to figure out information about M.

Function Space Setting: Spaces with Gaussian Kernels

We now turn to Chapter 3, which concerns integration and L^2-approximation on Hilbert spaces with Gaussian kernels, based on Gnewuch et al. (2026). In this chapter, μ_0 is again the standard normal distribution on $D = \mathbb{R}$. The univariate Gaussian kernel with shape parameter $\sigma > 0$ is given by

$$\ell_\sigma(x, y) = \exp(-\sigma(x^2 - y^2))$$

for $x, y \in \mathbb{R}$.

Based on this, for $d \in \mathbb{N} \cup \{\infty\}$ and a sequence $\boldsymbol{\sigma}$ of shape parameters $\sigma_j > 0$ fulfilling $\sum_{j=1}^d \sigma_j^2 < \infty$, we define the product kernel $L_{\boldsymbol{\sigma}}$ by

$$L_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} \ell_{\sigma_j}(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ in the case $d \in \mathbb{N}$ and for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}$ in the case $d = \infty$. Here, $\mathfrak{X}$ is the appropriate domain for the infinite-variate case we discussed before, with the unit vector $u_j \in H(\ell)$ being given by

$$u_j(x) = \exp(-\sigma_j^2 x^2)$$

for $x \in \mathbb{R}$. Interestingly, for this example the well-definedness of the infinite-variate product kernel even holds for all $\boldsymbol{x}, \boldsymbol{y} \in D^{\mathbb{N}}$. However, as we will see, the restriction to $\mathfrak{X}$ is without loss of generality for the integration and the L^2-approximation problem, while it has the advantage of giving us access to results for abstract tensor products.

In any case, we have $H(L_{\boldsymbol{\sigma}}) \subseteq L^2(\mu)$. In the finite-variate case $d \in \mathbb{N}$, the integration problem on $H(L_{\boldsymbol{\sigma}})$ has been studied before in Kuo and Woźniakowski

(2012), Kuo et al. (2017b), Karvonen and Särkkä (2019), and Karvonen et al. (2021), while the L^2-approximation problem has been studied in Fasshauer et al. (2012), see also Sloan and Woźniakowski (2018). The infinite-variate case has not been studied before.

We study the infinite-variate case for the first time, and also improve some known results in the finite-variate case.

Main Result: The Transference Principle

Recall that in Chapter 2, we studied Hermite spaces as an example. In Gnewuch et al. (2022), a strong connection was established between spaces with Gaussian kernels and Hermite spaces of infinite smoothness. For $\beta \in]0,1[$, we denote by k_β the univariate Hermite space with Fourier weights $\alpha_\nu = \beta^{-\nu}$ for $\nu \in \mathbb{N}$, so, as a reminder,

$$k_\beta(x, y) = 1 + \sum_{\nu \in \mathbb{N}} \beta^\nu \cdot h_\nu(x) \cdot h_\nu(y)$$

for $x, y \in \mathbb{R}$. For a summable sequence $\boldsymbol{\beta}$ of such base parameters $\beta_j \in]0,1[$, we denote by $K_{\boldsymbol{\beta}}$ the tensor product kernel of the univariate kernels k_{β_j}. It turns out that if the sequences $\boldsymbol{\sigma}$ and $\boldsymbol{\beta}$ are related via

$$1 - \beta_j = \frac{1}{1 + (1 + 8\sigma_j^2)^{1/2}},$$

there exists an isometric isomorphism Q from $L^2(\mu)$ onto itself which fulfills that $Q_{|H(K_{\boldsymbol{\beta}})}$ is (up to a constant) also an isometric isomorphism from $H(K_{\boldsymbol{\beta}})$ to $H(L_{\boldsymbol{\sigma}})$. Further, for $f \in H(K_{\boldsymbol{\beta}})$, one function evaluation of Qf can be calculated using only one function evaluation of f, and vice versa.

Based on this result, we establish, roughly speaking, a way to transform algorithms for integration or L^2-approximation on $H(K_{\boldsymbol{\beta}})$ into algorithms for the same problem on $H(L_{\boldsymbol{\sigma}})$ and vice versa, preserving cost and (up to a constant) error. The constant is actually given by $e_0(L_{\boldsymbol{\sigma}})$, and so we obtain

$$e_n(L_{\boldsymbol{\sigma}}) = e_0(L_{\boldsymbol{\sigma}}) \cdot e_n(K_{\boldsymbol{\beta}})$$

for every $n \in \mathbb{N}$. We mention that for the integration problem, we actually require a different relation between $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}$, given by

$$1 - \beta_j = \frac{1}{1 + 2\sigma_j^2}.$$

In both cases, $\beta_j \mapsto \sigma_j^2$ is monotonically increasing and bijective from $]0,1[$ to $]0,\infty[$, which yields, for both problems, a one-to-one-correspondence between Hermite spaces of infinite smoothness and spaces with Gaussian kernels such that the given problem is equally hard on corresponding spaces.

Main Result: Infinite-Variate Problems in Spaces with Gaussian Kernels

For the infinite-variate problems on $H(L_{\boldsymbol{\sigma}})$, we obtain the following result: Put

$$\rho := \liminf_{j \to \infty} \frac{\ln(1/\sigma_j^2)}{\ln(j)},$$

which fulfills $1 \leqslant \rho \leqslant \infty$ if the shape parameters are square-summable. We obtain

$$\operatorname{dec}(L_{\boldsymbol{\sigma}}) = \frac{1}{2}(\rho - 1),$$

see Theorem 3.22 for integration and Theorem 3.27 for L^2-approximation. This corresponds to the result for the Hermite case mentioned previously, and is of course derived from it.

Main Results: Tractability in Spaces with Gaussian or Hermite Kernels

For integration in the finite-variate case, we are able to improve some known tractability results for both Hermite spaces of infinite smoothness and spaces with Gaussian kernels. For these known results, we refer mainly to Irrgeher et al. (2015) in the Hermite case and to Kuo et al. (2017b) in the Gaussian case.

In the study of tractability, we consider sequences of function spaces $(H(M_d))_{d \in \mathbb{N}}$. In our case, for Hermite spaces, we consider a sequence $\boldsymbol{\beta}$ of base parameters, and for $d \in \mathbb{N}$ the d-variate kernel M_d is given by the product of the univariate kernels

$k_{\beta_1}, \ldots, k_{\beta_d}$. In the Gaussian case, we proceed analogously based on a sequence $\boldsymbol{\sigma}$ of shape parameters.

We also make use of the information complexity $n^{\mathrm{norm}}(\varepsilon, M_d)$ with $0 < \varepsilon < 1$ and $d \in \mathbb{N}$ for the normalized error criterion. Roughly speaking, $\varepsilon \mapsto n^{\mathrm{norm}}(\varepsilon, M_d)$ is the inverse function to $n \mapsto e_n(M_d)/e_0(M_d)$.

We introduce the following notion of tractability, see Kuo et al. (2017b). Exponential convergence (t, κ)-weak tractability, for short EC-(t, κ)-WT, holds for $(M_d)_{d \in \mathbb{N}}$ if

$$\lim_{d+\varepsilon^{-1} \to \infty} \frac{\ln(n^{\mathrm{norm}}(\varepsilon, M_d))}{d^t + \ln(\varepsilon^{-1})^\kappa} = 0.$$

In the case

$$(t, \kappa) = (1, 1),$$

we put $\gamma_j := \sigma_j$ in the Gaussian case and $\gamma_j := \beta_j$ in the Hermite case and assume that $(\gamma_j)_{j \in \mathbb{N}}$ is non-increasing to obtain the following result: EC-$(1, 1)$-WT is equivalent to $\lim_{j \to \infty} \gamma_j = 0$.

In the Hermite case, this has already been established by Irrgeher et al. (2015) combined with Irrgeher et al. (2016b). The Gaussian case was only studied previously under the additional condition that $\boldsymbol{\sigma}$ is bounded by a constant less than 1/2. The necessity of $\lim_{j \to \infty} \sigma_j = 0$ has already been established under this condition, while the only known sufficient condition was that the shape parameters converge exponentially fast to 0.

We turn our attention to the case

$$t > 1 \text{ and } \kappa > 0,$$

and here we put $\gamma_j := \sigma_j$ in the Gaussian case and $\gamma_j := 1/(1 - \beta_j)$ in the Hermite case to obtain the following result: If $(\gamma_j)_{j \in \mathbb{N}}$ is polynomially bounded, then EC-(t, κ)-WT holds for all $t > 1$ and $\kappa > 0$. See Theorems 3.18 and 3.20 for more general results that permit exponentially growing shape parameters σ_j and exponentially shrinking gaps $1 - \beta_j$ for the base parameters β_j. For Hermite kernels the case $t > 1$ has not been studied before, and for the Gaussian case it was only known that EC-(t, κ)-WT holds for all $t > 1$ and $\kappa \geqslant 1$, if $\boldsymbol{\sigma}$ is non-increasing and bounded by a constant less than 1/2, see Kuo et al. (2017b).

Notation

Throughout this thesis, we use the following notations.

For any Hilbert space H and $f, g \in H$, we denote the scalar product of f and g by $\langle f, g \rangle_H$ and the norm of f by $\|f\|_H$. For bounded linear operators T, their norm is denoted without subscript by $\|T\|$.

We denote finite or infinite families with bold letters and their members with the corresponding normally set letters, for instance $\boldsymbol{\tau} = (\tau_j)_{j\in J}$. We introduce the following pointwise operations for real-valued families. Let $\boldsymbol{\tau} := (\tau_j)_{j\in J}$ and $\boldsymbol{x} := (x_j)_{j\in J}$ be real-valued families and $z \in \mathbb{R}$. We put $\boldsymbol{\tau x} := (\tau_j x_j)_{j\in J}$ and $\boldsymbol{\tau}^z := (\tau_j^z)_{j\in J}$ if $\boldsymbol{\tau}$ is positive, i.e., $\tau_j > 0$ for every $j \in J$. Moreover, $\mathbf{1} - \boldsymbol{\tau} := (1 - \tau_j)_{j\in J}$. If $\boldsymbol{\tau}$ is positive and $\sum_{j\in J} |\tau_j - 1| < \infty$ then we set $\boldsymbol{\tau}_* = \prod_{j\in J} \tau_j > 0$.

We denote the asymptotic equivalence of two sequences $\boldsymbol{\tau}, \boldsymbol{\omega}$ by $\boldsymbol{\tau} \asymp \boldsymbol{\omega}$.

Chapter 1

General Setting

In this chapter, we present the framework for integration and L^2-approximation of finite- and infinite-variate functions we will study in this thesis and provide some general results. Based on this, in later chapters, we will study more specific cases. In this chapter, our main interest lies in the infinite-variate case, and many results presented here are well-known in the finite-variate case. Nevertheless, in the interest of a uniform presentation, we handle the finite- and infinite-variate case simultaneously.

1.1 Function Space Setting

The function spaces we consider throughout this thesis are always given as reproducing kernel Hilbert spaces (RKHSs) over $\mathbb{R}$. For a reproducing kernel k, the corresponding RKHS is denoted by $H(k)$. For an introduction to RKHSs, as well as basic results, we refer to Aronszajn (1950) as well as Paulsen and Raghupathi (2016). Given a real-valued reproducing kernel k on a domain $D \times D \neq \varnothing$, we denote by $H(k)$ its corresponding reproducing kernel Hilbert space.

More precisely, we consider the tensor product of d 'univariate' RKHSs, where either $d \in \mathbb{N}$ or $d = \infty$. In the case $d \in \mathbb{N}$, we put $J = \{1, \ldots, d\}$, in the case $d = \infty$ we put $J = \mathbb{N}$. In either case, let $(k_j)_{j \in J}$ be a family of reproducing kernels on a common domain $D \times D \neq \varnothing$. Additionally, we assume $k_j \neq 0$ for all $j \in J$.

In the finite-variate case $d \in \mathbb{N}$, it is straightforward to define the tensor product kernel K by

$$K(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j \in J} k_j(x_j, y_j) \tag{1.1}$$

for $\boldsymbol{x}, \boldsymbol{y} \in D^d$, which is again a reproducing kernel.

In the infinite-variate case, however, K according to (1.1) might not be well-defined on the whole domain $D^J \times D^J$. Therefore, we define an appropriate domain as follows, which is trivial in the finite-variate case: We remark that for any sequence of complex numbers $\boldsymbol{z} = (z_j)_{j \in J}$ the convergence

$$\sum_{j \in J} |z_j - 1| < \infty$$

implies the convergence of $\prod_{j \in J} z_j$, and further that $z_j = 0$ holds for at most finitely many $j \in J$ and that $\prod_{j \in J} z_j = 0$ holds if and only if $z_j = 0$ for some $j \in J$. For basic results on the convergence of infinite products of complex numbers, we refer to von Neumann (1939, Chap. 2). We mention that convergence of $\prod_{j \in J} z_j$ here is understood as convergence of the net of all finite partial products: $\prod_{j \in J} z_j$ converges to $z \in \mathbb{C}$ if for every $\varepsilon > 0$ there is a finite $I_0 \subseteq J$ such that for every finite $I \supseteq I_0$ we have $|\prod_{j \in I} z_j - z| < \varepsilon$.

Motivated by this, we define the set

$$\mathfrak{X}_0 = \{\boldsymbol{x} \in D^J \colon \sum_{j \in J} |k_j(x_j, x_j) - 1| < \infty\}.$$

For $\boldsymbol{x} \in \mathfrak{X}_0$, we have that $K(\boldsymbol{x}, \boldsymbol{x})$ according to (1.1) is well-defined, with a proper mode of convergence as described above.

However, for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}_0$, in general $K(\boldsymbol{x}, \boldsymbol{y})$ according to (1.1) is not well-defined. A minimal counterexample, of course in the case $J = \mathbb{N}$, to this is given by $D = \{-1, 1\}$ as well as $k_j(x_j, y_j) = x_j \cdot y_j$ for $j \in J$ and $x_j, y_j \in D$. In this example, we have $\mathfrak{X}_0 = D^J$, but $K(\boldsymbol{x}, -\boldsymbol{x})$ is never well-defined.

We therefore restrict our domain further in a way that is motivated by the study of abstract tensor products of Hilbert spaces. For each $j \in J$, let an $u_j \in H(k_j)$ with

$\|u_j\|_{H(k_j)} = 1$ be given. Define

$$\mathfrak{X}^{(\boldsymbol{u})} = \{\boldsymbol{x} \in \mathfrak{X}_0 \colon \sum_{j \in J} |u(x_j) - 1| < \infty\}. \tag{1.2}$$

If we assume

$$\mathfrak{X}^{(\boldsymbol{u})} \neq \varnothing, \tag{1.3}$$

$K^{(\boldsymbol{u})} = K$ as given by (1.1) is well-defined for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}^{(\boldsymbol{u})}$ and a reproducing kernel with domain $\mathfrak{X}^{(\boldsymbol{u})}$, see Theorem A.4. This allows us to consider the space $H(K)$ of real-valued functions on the domain $\mathfrak{X}^{(\boldsymbol{u})}$. Further, roughly speaking, there is a 'canonical' isometric isomorphism between $H(K)$ and the Hilbert space tensor product $\bigotimes_{j \in J} H(k_j)^{(u_j)}$, again see Theorem A.4, where this is also made more precise. Here, we just remark that this means we can make use of the theory of Hilbert space tensor products when studying $H(K)$. For a short overview regarding the tensor product of countably many Hilbert spaces according to von Neumann (1939), including the results needed for this thesis, Appendix A is included, where we also give further references.

When it is necessary to make clear the dependence of K on the k_j and u_j, we use the notation

$$\bigotimes_{j \in J} k_j = \bigotimes_{j \in J} k_j^{(u_j)} = K \tag{1.4}$$

with K being defined according to (1.1) on the domain $\mathfrak{X}^{(\boldsymbol{u})} \times \mathfrak{X}^{(\boldsymbol{u})}$.

We mention the following facts: First, the choice of $\boldsymbol{u}$ affects $\mathfrak{X}^{(\boldsymbol{u})}$ in the following way: If $\boldsymbol{v} = (v_j)_{j \in J}$ also fulfills $v_j \in H(k_j)$ with $\|v_j\|_{H(k_j)} = 1$ for all $j \in J$, we have $\mathfrak{X}^{(\boldsymbol{u})} = \mathfrak{X}^{(\boldsymbol{v})}$ if and only if

$$\sum_{j \in J} |\langle u_j, v_j \rangle_{H(k_j)} - 1| < \infty.$$

Otherwise, we have $\mathfrak{X}^{(\boldsymbol{u})} \cap \mathfrak{X}^{(\boldsymbol{v})} = \varnothing$, cf. (A.1). Furthermore, as long as $\mathfrak{X}_0 \neq \varnothing$, we can always find $\boldsymbol{u} = (u_j)_{j \in J}$ such that (1.3) is fulfilled. For any $\boldsymbol{x} \in \mathfrak{X}_0$, one can simply choose

$$u_j = \frac{k(\cdot, x_j)}{\|k(\cdot, x_j)\|_{H(k_j)}}$$

for $j \in \mathbb{N}$. In applications, however, there might be further restrictions to the choice of $\boldsymbol{u}$, see in particular Theorem 1.4.

Finally, if (1.3) is fulfilled, for any $\boldsymbol{x} \in \mathfrak{X}$, there are at most finitely many x_j with $k_j(x_j, x_j) = 0$. Since we assumed $k_j \neq 0$ for all $j \in J$, and a sequence being an element of $\mathfrak{X}^{(\boldsymbol{u})}$ does not depend on finitely many members of the sequence, there always exists $\boldsymbol{y} \in \mathfrak{X}^{(\boldsymbol{u})}$ with $K(\boldsymbol{y}, \boldsymbol{y}) \neq 0$, and therefore we have

$$H(K) \neq \{0\}. \tag{1.5}$$

Remark 1.1. For each $j \in J$, let $(e_{\eta,j})_{\eta \in N_j}$ be an orthonormal basis of $H(k_j)$. We assume $0 \in N_j$ for notational convenience and require

$$e_{0,j} = u_j$$

for every $j \in J$. Let $\boldsymbol{N}$ denote the set of all sequences $\boldsymbol{\eta} := (\eta_j)_{j \in J}$ with $\eta_j \in N_j$ for every $j \in \mathbb{N}$ and with $\{j \in \mathbb{N} \colon \eta_j \neq 0\}$ being finite. Then the functions

$$e_{\boldsymbol{\eta}} := \prod_{j \in J} e_{\eta_j, j} \tag{1.6}$$

with $\boldsymbol{\eta} \in \boldsymbol{N}$ form an orthonormal basis of $H(K)$, which follows by the corresponding statement for abstract tensor products, see Remark A.1, together with the isomorphism given by Theorem A.4, which respects the product structure.

Remark 1.2. The reproducing kernel K according to (1.1) can be written as the superposition of certain kernels of product form that only depend on finitely many variables.

More precisely, for each $j \in J$, we have that k_j^*, given by

$$k_j^*(x, y) = k_j(x, y) - u_j(x) u_j(y)$$

for $x, y \in D$, is also a reproducing kernel, the corresponding RKHS $H(k_j^*)$ being the orthogonal complement of $\operatorname{span}\{u_j\}$ in $H(k_j)$. Now, let $\boldsymbol{U}$ be the set of all finite subsets of J. For $\boldsymbol{\nu} \in \boldsymbol{U}$, define $k^*_{\boldsymbol{\nu}}$ by

$$k^*_{\boldsymbol{\nu}}(\boldsymbol{x}_{\boldsymbol{\nu}}, \boldsymbol{y}_{\boldsymbol{\nu}}) = \prod_{j \in \boldsymbol{\nu}} k^*_j(x_j, y_j)$$

for $x_{\boldsymbol{\nu}} = (x_j)_{j \in \boldsymbol{\nu}}, y_{\boldsymbol{\nu}} = (y_j)_{j \in \boldsymbol{\nu}} \in D^{\boldsymbol{\nu}}$. Finally, put

$$\mathfrak{X}^* = \{\boldsymbol{x} \in D^J \colon \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} k^*_{\boldsymbol{\nu}}(\boldsymbol{x}_{\boldsymbol{\nu}}, \boldsymbol{x}_{\boldsymbol{\nu}}) < \infty\},$$

where $\boldsymbol{x}_{\boldsymbol{\nu}} = (x_j)_{j \in \boldsymbol{\nu}}$, and define K^* by

$$K^*(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} k^*_{\boldsymbol{\nu}}(\boldsymbol{x}_{\boldsymbol{\nu}}, \boldsymbol{y}_{\boldsymbol{\nu}})$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}^*$. Since we have

$$k_j(x, y) = u_j(x) u_j(y) + k^*_j(x, y)$$

for every $j \in J$ and every $x, y \in D$, a simple calculation shows

$$\mathfrak{X}^* = \mathfrak{X} \text{ as well as } K^* = K^{(\boldsymbol{u})}.$$

For $\boldsymbol{\nu} \in \boldsymbol{U}$, we isometrically embed $H(k^*_{\boldsymbol{\nu}})$ into $H(K^{(\boldsymbol{u})})$ by mapping $f_{\boldsymbol{\nu}} \in H(k^*_{\boldsymbol{\nu}})$ to

$$f \colon \mathfrak{X} \to \mathbb{R} \colon \boldsymbol{x} \mapsto f_{\boldsymbol{\nu}}(\boldsymbol{x}_{\boldsymbol{\nu}}). \tag{1.7}$$

In this sense, we have $H(k^*_{\boldsymbol{\nu}}) \subseteq H(K^{(\boldsymbol{u})})$ as a closed subspace. Further, in this sense, for $\boldsymbol{\nu}_1 \neq \boldsymbol{\nu}_2 \in \boldsymbol{U}$, we have that $H(k^*_{\boldsymbol{\nu}_1})$ and $H(k^*_{\boldsymbol{\nu}_2})$ are orthogonal, which follows readily from Remark 1.1. For $f \in H(K^{(\boldsymbol{u})})$, we therefore have a unique decomposition

$$f = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} f_{\boldsymbol{\nu}} \tag{1.8}$$

with $f_{\boldsymbol{\nu}} \in H(k^*_{\boldsymbol{\nu}})$ for $\boldsymbol{\nu} \in \boldsymbol{U}$. Further, the $f_{\boldsymbol{\nu}}$ are pairwise orthogonal. This decomposition is most useful, since it allows us, under additional assumptions, to trace back the study of algorithmic integration and L^2-approximation on the infinite-variate space

$H(K^{(\boldsymbol{u})})$ to the study of these problems on the finite-variate spaces. See Section 2.3 for further details and references.

Remark 1.3. In the literature, a class of spaces of infinite-variate functions different from the one described previously in this section, but still admitting a function decomposition of the form (1.8) has been studied, see for example Kuo et al. (2010b), Wasilkowski and Woźniakowski (2011) and Wasilkowski (2012). We describe it here briefly.

We only consider the infinite-variate case $J = \mathbb{N}$. As a univariate building block, let $k \neq 0$ be a reproducing kernel on $D \times D$. Additionally, we assume

$$1 \notin H(k).$$

As in Remark 1.2, denote by $\boldsymbol{U}$ the set of all finite subsets of $\mathbb{N}$. For every $\boldsymbol{\nu} \in \boldsymbol{U}$, let $\gamma_{\boldsymbol{\nu}} \geqslant 0$ be some nonnegative weight. For $d \in \mathbb{N}$ we define K_d as the d-fold tensor product of k with itself, i.e.

$$K_d(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} k(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in D^d$. We use the notation $\boldsymbol{x}_{\boldsymbol{\nu}}$ as in Remark 1.2. The domain of the infinite-variate kernel we seek to construct is then given by

$$\mathfrak{X} = \{\boldsymbol{x} \in D^{\mathbb{N}} \colon \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} \gamma_{\boldsymbol{\nu}} K_{|\boldsymbol{\nu}|}(\boldsymbol{x}_{\boldsymbol{\nu}}, \boldsymbol{y}_{\boldsymbol{\nu}}) < \infty\},$$

and the infinite-variate kernel K is given by

$$K(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} \gamma_{\boldsymbol{\nu}} K_{|\boldsymbol{\nu}|}(\boldsymbol{x}_{\boldsymbol{\nu}}, \boldsymbol{y}_{\boldsymbol{\nu}}) \tag{1.9}$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}$. The kernel K is the superposition of the finite-variate kernels of tensor product form $\gamma_{\boldsymbol{\nu}} K_{|\boldsymbol{\nu}|}$. Roughly speaking, $\gamma_{\boldsymbol{\nu}}$ determines the importance of the variable group $\boldsymbol{\nu}$.

We relate this construction to our general setting. Note that K given by (1.9) is not necessarily a tensor product kernel. However, in the special case of product weights, that is, the existence of a sequence $(\gamma_j)_{j\in\mathbb{N}}$ such that

$$\gamma_{\boldsymbol{\nu}} = \prod_{j\in\boldsymbol{\nu}} \gamma_j$$

for all $\boldsymbol{\nu} \in \boldsymbol{U}$, we have, with the choice of $u_j = 1$ for all $j \in \mathbb{N}$,

$$\mathfrak{X} = \mathfrak{X}^{(\boldsymbol{u})} \text{ and } K = \bigotimes_{j\in\mathbb{N}} (1 + \gamma_j k)^{(u_j)},$$

which follows from Remark 1.2. Therefore, the setting described in this remark and our general setting both 'share' the special case of product weights.

1.2 Integration and L^2-Approximation

In the sequel, let D be equipped with some σ-algebra $\mathcal{D}$ and let μ_0 be a probability measure on $(D, \mathcal{D})$, and assume k_j to be $\mathcal{D} \bigotimes \mathcal{D}$-measurable, which implies that $H(k_j)$ consists of $\mathcal{D}$-measurable functions. Denote by $\mu = \mu^{(d)}$ the d-fold product measure of μ_0 with itself.

We first remark that the measurability of the k_j implies that $\mathfrak{X}^{(\boldsymbol{u})}$ is a measurable set in the d-fold product of $(D, \mathcal{D})$ with itself. By Kolmogorov's 0-1-law, we even get $\mu(\mathfrak{X}^{(\boldsymbol{u})}) \in \{0, 1\}$. In the sequel, we assume we are in the non-trivial case

$$\mu(\mathfrak{X}^{(\boldsymbol{u})}) = 1.$$

We assume that for each $j \in J$, all elements of $H(k_j)$ are square-integrable with respect to μ_0, and denote by T_j the inclusion mapping

$$T_j \colon H(k_j) \to L^2(\mu_0) \colon f \mapsto [f],$$

where $[f]$ is the $L^2(\mu_0)$-equivalence class of f.

Observe that T_j is always continuous, which follows easily from the closed graph theorem along with the facts that convergence of a sequence in an RKHS is equivalent to pointwise convergence and that convergence in $L^2(\mu_0)$ implies almost sure convergence of at least a subsequence.

The following result establishes sufficient conditions for $H(K)$ being a space of square-integrable functions with respect to μ. We only consider the case $J = \mathbb{N}$.

Theorem 1.4. *Assume that*

$$\sum_{j\in\mathbb{N}} |\|T_j\| - 1| < \infty. \tag{1.10}$$

Further assume that

$$\sum_{j\in\mathbb{N}} |\|T_j u_j\| - 1| < \infty \text{ and} \tag{1.11}$$

$$\sum_{j\in\mathbb{N}} \left| \int_D u_j d\mu_0 - 1 \right| < \infty. \tag{1.12}$$

Then, we have $H(K^{(\boldsymbol{u})}) \subseteq L^2(\mu)$ with a continuous identical embedding T with norm

$$\|T\| = \prod_{j\in\mathbb{N}} \|T_j\|.$$

Proof. We apply Theorem A.3, setting $H_j = H(k_j)$ as well as $G_j = L^2(\mu_0)$ and $v_j = 1$ for $j \in \mathbb{N}$. Note that (A.4) holds by (1.11) and (A.5) holds by (1.12). Thus, there exists a bounded linear operator $T^{\otimes}\colon \bigotimes_{j\in\mathbb{N}} H(k_j)^{(u_j)} \to \bigotimes_{j\in\mathbb{N}} L^2(\mu_0)^{(1)}$ fulfilling

$$T^{\otimes} \bigotimes_{j\in\mathbb{N}} f_j = \bigotimes_{j\in\mathbb{N}} [f_j]$$

for all sequences with $f_j \neq u_j$ only finitely often and

$$\|T^{\otimes}\| = \prod_{j\in\mathbb{N}} \|T_j\|.$$

Now, let $\Phi\colon \bigotimes_{j\in\mathbb{N}} H(k_j)^{(u_j)} \to H(K^{(\boldsymbol{u})})$ denote the isometric isomorphism according to Theorem A.4 and $\Psi\colon L^2(\mu) \to \bigotimes_{j\in\mathbb{N}} L^2(\mu_0)^{(1)}$ the isometric isomorphism according to Gnewuch et al. (2022, Thm. A.7). We show that

$$T = \Psi^{-1} T^{\otimes} \Phi^{-1}$$

is the embedding with the desired properties, noting that $\|T\| = \prod_{j\in\mathbb{N}} \|T_j\|$ clearly holds.

Firstly, for $n \in \mathbb{N}$, we clearly have the existence of $v_n = [\prod_{j=1}^n u_j]$ in $L^2(\mu)$. It was shown in Gnewuch et al. (2022, Thm. A.7) that $(v_n)_{n\in\mathbb{N}}$ converges to some $v \in L^2(\mu)$ and that $\bigotimes_{j\in\mathbb{N}} u_j = \Psi(v)$. For the function $\widetilde{u} = \prod_{j\in\mathbb{N}} u_j \in H(K)$, we thus get

$$v = \Psi^{-1} T^{\otimes} \Phi^{-1} \widetilde{u}.$$

On the other hand, we also have almost sure convergence of v_n to $\widetilde{u}$, so by uniqueness of the limit, we indeed have $[\widetilde{u}] = v$.

Now, let $(e_{\boldsymbol{\eta}})_{\boldsymbol{\eta}\in\boldsymbol{N}}$ be an orthonormal basis of $H(K)$ as described in Remark 1.1. In particular, $e_{\boldsymbol{\eta}}$ is given as a product of functions $e_{\eta_j,j}$, of which only finitely many of which differ from u_j. As before, we obtain

$$[e_{\boldsymbol{\eta}}] = \Psi^{-1} T^{\otimes} \Phi^{-1} (e_{\boldsymbol{\eta}}).$$

For arbitrary $f \in H(K)$, there exists a square-summable sequence $(c_{\boldsymbol{\eta}})_{\boldsymbol{\eta}\in\boldsymbol{N}}$ such that $f = \sum_{\boldsymbol{\eta}\in\boldsymbol{N}} c_{\boldsymbol{\eta}} \cdot e_{\boldsymbol{\eta}}$, which immediately yields

$$[f] = \Psi^{-1} T^{\otimes} \Phi^{-1} (f). \qquad \square$$

If the conditions of Theorem 1.4 are fulfilled, the functionals

$$\mathrm{Int}\colon H(K^{(\boldsymbol{u})}) \to \mathbb{R}\colon f \mapsto \int_{\mathfrak{X}^{(\boldsymbol{u})}} f d\mu \tag{1.13}$$

and

$$\mathrm{App}\colon H(K^{(\boldsymbol{u})}) \to L^2(\mu)\colon f \mapsto [f] \tag{1.14}$$

are well-defined, continuous linear operators. We remark that while we only consider the L^2-case in this thesis, Int is of course already well-defined if $H(K^{(\boldsymbol{u})}) \subseteq L^1(\mu)$.

1.3 Computational Problems and Algorithms

For reproducing kernel Hilbert spaces of square-integrable functions, our goal is to analyze the worst case minimal errors of the integration problem and the L^2-approximation problem, that is, to analyze the approximation of the operators Int and App, respectively, within the framework of information based complexity (IBC). The class of algorithms we consider are linear algorithms based on finitely many function evaluations. We mention that for linear problems like Int and App, the restriction to linear algorithms is without loss of generality, see Novak and Woźniakowski (2008, Sec. 4.2.2) for details and references. For more precise definitions, let $E \neq \varnothing$ and let k be a reproducing kernel with domain $E \times E$. Let ρ be a probability measure on E such that the operators Int and App are well-defined with respect to ρ.

We consider algorithms A given by

$$A(f) = \sum_{i=1}^{n} f(\boldsymbol{x}_i) \cdot a_i \tag{1.15}$$

for $f \in H(k)$, where $n \in \mathbb{N}$, $\boldsymbol{x}_i \in E$ and $a_i \in \mathbb{R}$ for the integration problem or $a_i \in L^2(\rho)$ for the L^2-approximation problem, respectively. For the integration problem, these are of course also called quadrature formulas or quadrature rules. In either case, the $\boldsymbol{x}_i$ are called nodes and the a_i are called weights.

The worst-case error of A on $H(k)$ is defined as

$$\operatorname{error}(A, k) = \sup \left\{ \left| \int_E f d\rho - A(f) \right| : f \in H(k) \text{ and } \|f\|_{H(k)} \leqslant 1 \right\}$$

for the integration problem and

$$\operatorname{error}(A, k) = \sup \left\{ \|f - A(f)\|_{L^2(\rho)} : f \in H(k) \text{ and } \|f\|_{H(k)} \leqslant 1 \right\}$$

for the L^2-approximation problem, respectively.

In either case, we define the cost of an algorithm of the form (1.15) by

$$\operatorname{cost}(A) = \sum_{i=1}^{n} \operatorname{cost}(\boldsymbol{x}_i),$$

where $\mathrm{cost}\colon E \to \mathbb{N} \cup \{\infty\}$ is a mapping specifying the cost of a single function evaluation at a given node. We distinguish two cases: In the case of E being finite-variate, it is natural and indeed commonly done to consider the case $\mathrm{cost}(\boldsymbol{x}) = 1$ for each $\boldsymbol{x} \in E$. In the infinite-variate case $E \subseteq D^{\mathbb{N}}$ for some univariate domain D, we consider a less generous cost model, namely the unrestricted subspace sampling model introduced by Kuo et al. (2010a): Given a default value $a \in D$ and a non-decreasing function $\$\colon \mathbb{N} \cup \{\infty\} \to \mathbb{N} \cup \{\infty\}$, where $\$(m) = \infty$ if and only if $m = \infty$, we define

$$\mathrm{cost}(\boldsymbol{x}) = \$(|\{j \in \mathbb{N}\colon x_j \neq a\}|). \tag{1.16}$$

The set $\{j \in \mathbb{N}\colon x_j \neq a\}$ used in the definition is also called the set of active variables of $\boldsymbol{x}$. Throughout this thesis we assume that there exist $c_1, c_2 > 0$ such that

$$c_1 \cdot m \leqslant \$(m) \leqslant \exp(c_2 \cdot m) \tag{1.17}$$

for all $m \in \mathbb{N}$, cf., e.g., Kuo et al. (2010b) and Plaskota and Wasilkowski (2011). We mention that the upper bound is generous, while the lower bound is reasonable, since it takes at least m numbers to specify a point $\boldsymbol{x}$ with m components different from a.

We are not only interested in the error and cost of specific algorithms, but the inherent complexity of both computational problems. We quantify this by the n-th minimal error, which is given by

$$e_n(k) = \inf\{\mathrm{error}(A, k)\colon \mathrm{cost}(A) \leqslant n\}$$

for $n \in \mathbb{N}$. Sometimes, we instead consider the ε-complexity, given by

$$n(\varepsilon, k) = \inf\{n \in \mathbb{N}\colon e_n(k) \leqslant \varepsilon\}$$

for $\varepsilon > 0$. Roughly speaking, $e_n(k)$ and $n(\varepsilon, k)$ are inverse to one another.

To study the behavior of $e_n(k)$ as $n \to \infty$, another key quantity is the decay. First, the decay of a sequence of positive real numbers $\boldsymbol{z}$ is defined by

$$\mathrm{decay}(\boldsymbol{z}) = \sup\left\{\tau > 0\colon \sum_{j \in \mathbb{N}} z_j^{1/\tau} < \infty\right\}. \tag{1.18}$$

It is well known, see e.g. Fasshauer et al. (2012, p.250), that if $\boldsymbol{z}$ is non-decreasing, we have

$$\operatorname{decay}(\boldsymbol{z}) = \sup\{\tau > 0\colon\ \sup\{z_n \cdot n^\tau\} < \infty\},$$

so the decay is roughly the polynomial rate at which $\boldsymbol{z}$ tends to 0. In the case of n-th minimal errors, we abbreviate

$$\operatorname{dec}(k) = \operatorname{decay}((e_n(k))_{n\in\mathbb{N}}). \tag{1.19}$$

1.3.1 Algorithms in the Multivariate Case

We briefly consider the finite-variate case $d \in \mathbb{N}$. Given univariate kernels k_j for $j \in \{1,\dots,d\}$, we denote by K_d the d-variate tensor product kernel of the k_j. We introduce two well-known ways to construct algorithms for integration or L^2-approximation on $H(K_d)$, given algorithms for the corresponding univariate problems on the spaces $H(k_j)$.

Remark 1.5. Let us assume that for each $j \in \{1,\dots,d\}$, we have a number of nodes $n_j \in \mathbb{N}$, as well as an algorithm A_{j,n_j} for the univariate problem on $H(k_j)$, which is given by

$$A_{n_j,j}(f) = \sum_{i=1}^{n_j} f(\boldsymbol{x}_{i,j}) \cdot a_{i,j}$$

for $f \in H(k_j)$. We set $\boldsymbol{n} = (n_j)_{j\in\{1,\dots,d\}}$.

We introduce the full tensor product rule $A_{d,\boldsymbol{n}} = \bigotimes_{j=1}^d A_{j,n_j}$ by

$$A_{d,\boldsymbol{n}}(f) = \sum_{\substack{\boldsymbol{i}\in\mathbb{N}^d\\ \boldsymbol{i}\leqslant\boldsymbol{n}}} f(\boldsymbol{x}_{\boldsymbol{i}}) \cdot a_{\boldsymbol{i}},$$

for $f \in H(K_d)$, where $\boldsymbol{i} \leqslant \boldsymbol{n}$ is meant pointwisely and

$$\boldsymbol{x}_{\boldsymbol{i}} = (\boldsymbol{x}_{i_1,1},\dots,\boldsymbol{x}_{i_d,d}) \qquad \text{and} \qquad a_{\boldsymbol{i}} = \prod_{j=1}^d a_{i_j,j}.$$

Since the set of nodes for the d-variate algorithm is the Cartesian product of the sets of nodes for the univariate problems, this algorithm often has a prohibitively high cost. However, if the underlying univariate function spaces have high regularity, optimal or almost optimal error rates can sometimes still be achieved. Because of their relevance to Section 3.2.4, We mention Irrgeher et al. (2015), Irrgeher et al. (2016a) and Irrgeher et al. (2016b), where Hermite spaces were studied, as well as Kuo et al. (2017b) and Karvonen et al. (2021), where spaces with Gaussian kernels were studied.

Remark 1.6. Now, let us even assume that for each $j \in \{1, \dots, d\}$ and for each $n \in \mathbb{N}$, an algorithm $A_{n,j}$ for the univariate problem on $H(k_j)$ with cost n is given. For notational convenience, set $A_{0,j} = 0$. We introduce the Smolyak or sparse grid algorithm of level $L \in \mathbb{N}$ by

$$Q_L(f) = \sum_{\substack{\boldsymbol{l} \in \mathbb{N}^d \\ l_1 + \dots + l_d \leqslant L}} \bigotimes_{j=1}^{d} (A_{l_j,j} - A_{l_j-1,j})$$

for $f \in H(K_d)$.

In the case $k_1 = \dots = k_d$, if there exists $\kappa > 1$ and $c > 0$ such that

$$\operatorname{error}(A_{n,j}, k_1) \leqslant c \cdot n^{-\kappa} \tag{1.20}$$

holds for all $n \in \mathbb{N}$, Smolyak algorithms yield the following upper error bound, which can be found in Gnewuch and Wnuk (2020, Thm. 7) or can be derived from Wasilkowski and Woźniakowski (1995, Thm. 1): There exist constants $C_0, C_1 > 0$, which only depend on c and κ, such that for every $n \in \mathbb{N}_0$ and a Smolyak algorithm Q_L with cost n, we have

$$\operatorname{error}(Q_L, K_d) \leqslant C_0 C_1^{|\boldsymbol{u}|} \left(1 + \frac{\ln(n+1)}{\max(|\boldsymbol{u}| - 1, 1)}\right)^{(\kappa+1)(|\boldsymbol{u}|-1)} (n+1)^{-\kappa}. \tag{1.21}$$

Chapter 2

Spaces of Increasing Smoothness

In Chapter 1, we saw how, given a sequence $(k_j)_{j\in\mathbb{N}}$ of univariate reproducing kernels and based on some $u_j \in H(k_j)$ with norm one, their infinite tensor product $K = K^{(\boldsymbol{u})}$ is defined, as well as when $H(K)$ consists of square-integrable functions. In this chapter, we study a special class of univariate kernels k_j and corresponding infinite-variate kernels K with the goal to approximate the operators Int and App, see (1.13) and (1.14), on $H(K)$.

We introduce the class of kernels studied in this chapter, beginning by defining univariate kernels k_j. Let $\varnothing \neq D \subseteq \mathbb{R}$, let μ_0 be a Borel probability measure on D, and let μ be its d-fold product measure. Let $(\mathfrak{e}_\nu)_{\nu\in\mathbb{N}_0}$ be a sequence of square-integrable functions on D such that the corresponding L^2-equivalence classes form an orthonormal system in $L^2(\mu_0)$. Further, for a fixed $j \in \mathbb{N}$, let a sequence of Fourier weights $\boldsymbol{\alpha}_j = (\alpha_{\nu,j})_{\nu\in\mathbb{N}_0}$ be given fulfilling

$$\inf_{\nu\in\mathbb{N}_0} \alpha_{\nu,j} > 0 \text{ and } \sum_{j\in\mathbb{N}_0} \alpha_{\nu,j}^{-1}\mathfrak{e}_\nu(x)^2 < \infty$$

for all $x \in D$. It is known, see Lemma 2.1, that k_j given by

$$k_j(x,y) = \sum_{\nu\in\mathbb{N}_0} \alpha_{\nu,j}^{-1}\mathfrak{e}_\nu(x)\mathfrak{e}_\nu(y)$$

for $x, y \in D$ is a reproducing kernel on $D \times D$. We additionally assume that

$$\mathfrak{e}_0 = 1 \text{ and } \alpha_{0,j} = 1,$$

which implies

$$1 \in H(k_j) \text{ with } \|1\|_{H(k_j)} = 1.$$

This is already sufficient to define the infinite-variate kernel K. We choose

$$u_j = 1,$$

in which case the kernel $K = \bigotimes_{j \in \mathbb{N}} k_j^{(u_j)}$ is well-defined on the domain $\mathfrak{X} \times \mathfrak{X}$, where

$$\mathfrak{X} = \{\boldsymbol{x} \in D^{\mathbb{N}} \colon \sum_{\nu, j \in \mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) < \infty\}.$$

As a control on the importance of the univariate spaces $H(k_j)$ towards $H(K)$, we introduce certain summability and monotonicity assumptions, see (A1)–(A3) on page 35. We discuss these assumptions in detail later. Here, we mention two facts: Firstly, (A2) ensures that $H(K)$ is continuously identically embedded into $L^2(\mu)$, so that Int and App are well-defined. Secondly, under suitable assumptions on the Fourier weights, we have

$$H(k_j) \subseteq H(k_1)$$

with a compact identical embedding. Further, the norm of these embeddings decays to 0 sufficiently fast.

This setting of compact embeddings has been studied before under the name *increasing smoothness*. In Gnewuch et al. (2019), upper and lower bounds of dec(K), see (1.19) were obtained under the additional assumption

$$\mathfrak{X} = D^{\mathbb{N}}. \tag{2.1}$$

It was also established that multiple classical examples fall under this special setting, including Korobov spaces, Haar spaces, Walsh spaces and certain Sobolev spaces.

We study this setting without requiring (2.1). Our study is based on Gnewuch et al. (2024), where Hermite spaces with μ_0 being the standard normal distribution are studied. Hermite spaces fall under the setting discussed here and do not fulfill (2.1).

In this chapter, we generalize the results from Gnewuch et al. (2024), so that we achieve general results without needing the assumption (2.1). We also address the example of the Hermite spaces.

We proceed as follows. In Section 2.1, we define the setting in detail and give some basic embedding results. In Section 2.2, we embed $H(K)$ into a space $H(M_a^{\uparrow})$ with a reproducing kernel $M_a^{\uparrow}$ that is the tensor product of kernels anchored at some $a \in D$. We also embed a corresponding space $H(M_a^{\downarrow})$ with similar properties into $H(K)$. A particular challenge here is handling the different domains of $K, M_a^{\uparrow}$, and $M_a^{\downarrow}$ correctly. Based on these embeddings, we achieve our main results in Section 2.3, namely, for both problems, upper and lower bounds on $\mathrm{dec}(K)$. The upper and lower bounds will turn out to match for polynomially or exponentially growing Fourier weights. We also provide a sketch of how to construct MDM-Algorithms that achieve the upper error bound. Lastly, in Section 2.4, we consider the specific example of Hermite spaces.

2.1 Setting

Let $\varnothing \neq D \subseteq \mathbb{R}$ and let μ_0 be a Borel probability measure on D. In this chapter, we focus on the infinite-variate case, and we again study RKHSs whose kernels are given by tensor products of univariate kernels k_j, all defined on the same domain $D \times D$, with $j \in \mathbb{N}$. However, in this chapter we require the k_j to have the form

$$k_j(x, y) = \sum_{\nu \in \mathbb{N}_0} \alpha_{\nu,j}^{-1} \mathfrak{e}_\nu(x) \mathfrak{e}_\nu(y) \tag{2.2}$$

for $x, y \in D$, where $(\mathfrak{e}_\nu)_{\nu \in \mathbb{N}_0}$ is a family of square-integrable functions whose equivalence classes form an orthonormal system of $L^2(\mu_0)$ and $(\alpha_{\nu,j})_{\nu \in \mathbb{N}_0, j \in \mathbb{N}}$ is a family of positive Fourier weights fulfilling certain conditions which ensure, for instance, convergence in (2.2).

To put this precisely, we use a proposition due to Gnewuch et al. (2022, Prop 2.1) which we present here, fitted to the case we are interested in. As this proposition concerns the univariate case, we omit the parameter j in the notation.

Proposition 2.1. *Let $\boldsymbol{\alpha} = (\alpha_\nu)_{\nu \in \mathbb{N}_0}$ be a sequence of real numbers fulfilling*

$$\inf_{\nu \in \mathbb{N}_0} \alpha_\nu > 0.$$

Further, let $(\mathfrak{e}_\nu)_{\nu \in \mathbb{N}_0}$ be a sequence of square-integrable functions on D such that the corresponding sequence of L^2-equivalence classes is an orthonormal basis of $L^2(\mu_0)$ and

$$\sum_{\nu \in \mathbb{N}_0} \alpha_\nu^{-1} \cdot \mathfrak{e}_\nu(x)^2 < \infty \tag{2.3}$$

holds for all $x \in D$.

Then, if a sequence of real numbers $(c_\nu)_{\nu \in \mathbb{N}_0}$ fulfills

$$\sum_{\nu \in \mathbb{N}_0} \alpha_\nu^{-1} \cdot c_\nu^2 < \infty,$$

the series $\sum_{\nu \in \mathbb{N}_0} c_\nu \cdot e_\nu$ converges in $L^2(\mu_0)$ and the sequence $\sum_{\nu \in \mathbb{N}_0} c_\nu \cdot \mathfrak{e}_\nu$ converges absolutely in every point. Furthermore, k given by

$$k(x, y) = \sum_{\nu \in \mathbb{N}_0} \alpha_\nu^{-1} \mathfrak{e}_\nu(x) \mathfrak{e}_\nu(y)$$

for $x, y \in D$ defines a reproducing kernel on $D \times D$. The corresponding Hilbert space is given by

$$H(k) = \left\{ \sum_{\nu \in \mathbb{N}_0} c_\nu \cdot \mathfrak{e}_\nu \colon (c_\nu)_{\nu \in \mathbb{N}_0} \in \mathbb{R}^{\mathbb{N}_0} \textit{ with } \sum_{\nu \in \mathbb{N}_0} \alpha_\nu^{-1} \cdot c_\nu^2 < \infty \right\} \tag{2.4}$$

and

$$\langle f, g \rangle_{H(k)} = \sum_{\nu \in \mathbb{N}_0} \alpha_\nu \cdot \int_D f \cdot \mathfrak{e}_\nu d\mu_0 \cdot \int_D g \cdot \mathfrak{e}_\nu d\mu_0. \tag{2.5}$$

We remark that in Gnewuch et al. (2022), the assumption $D \subseteq \mathbb{R}$ is not required.

Throughout this chapter, based on Proposition 2.1, we consider the following setting: We assume that a family of square-integrable functions $(\mathfrak{e}_\nu)_{\nu \in \mathbb{N}_0}$ forming an orthonormal basis in $L^2(\mu_0)$ is given. Further, for each $j \in \mathbb{N}$, let $(\alpha_{\nu,j})_{\nu\in\mathbb{N}_0}$ be a sequence of Fourier weights with

$$\inf_{\nu\in\mathbb{N}_0} \alpha_{\nu,j} > 0 \tag{2.6}$$

and

$$\sum_{\nu\in\mathbb{N}_0} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu(x)^2 < \infty \tag{2.7}$$

for all $x \in D$. Further, we assume that

$$\mathfrak{e}_0 = 1 \quad \text{and} \quad \alpha_{0,j} = 1 \text{ for all } j \in \mathbb{N}. \tag{2.8}$$

Then, for each $j \in \mathbb{N}$, a reproducing kernel k_j is given by

$$k_j(x, y) = 1 + \sum_{\nu\in\mathbb{N}} \alpha_{\nu,j}^{-1} \mathfrak{e}_\nu(x)\mathfrak{e}_\nu(y)$$

for $x, y \in D$.

For the corresponding infinite-variate kernel, we set $u_j = 1$. Then, the definition of the domain $\mathfrak{X}^{(\boldsymbol{u})}$ as given by (1.2) simplifies to

$$\mathfrak{X}^{(\boldsymbol{u})} = \left\{ \boldsymbol{x} \in D^{\mathbb{N}} \colon \sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) < \infty \right\}$$

and we consider the infinite-variate kernel K given by $K(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j\in\mathbb{N}} k_j(x_j, y_j)$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}^{(\boldsymbol{u})}$. Since $\boldsymbol{u}$ is fixed throughout the chapter, we will also write

$$\mathfrak{X} = \mathfrak{X}^{(\boldsymbol{u})}.$$

In this setting, we have $\boldsymbol{x} \in \mathfrak{X}$ if and only if $(\prod_{j=1}^n k_j(x_j, x_j))_{n\in\mathbb{N}}$ converges, since $k_j(x_j, x_j) \geqslant 1$ for all $j \in \mathbb{N}$.

We occasionally employ the monotonicity and summability properties

(A1) $0 < \alpha_{1,j} \leqslant \alpha_{2,j} \leqslant \ldots$ for all $j \in \mathbb{N}$,

(A2) $\sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} < \infty$, (A3) $\sum_{j\in\mathbb{N}} \gamma_j < \infty$, where $\gamma_j := \sup_{\nu\in\mathbb{N}} \alpha_{\nu,1}/\alpha_{\nu,j}$,

which we do not assume globally in this chapter, but they or a subset of them will often be needed to establish results, e.g. concerning $\mu(\mathfrak{X})$ or embeddings.

2.1.1 Basic Embedding Results

In the following, we give some basic results and intuition about the weights $\alpha_{\nu,j}$ as well as the role of the assumptions (A1)–(A3). Firstly, for each $j \in \mathbb{N}$, it follows immediately from the definition of the scalar product on $H(k_j)$, see (2.5), that

$$(\alpha_{\nu,j}^{-1/2}\mathfrak{e}_\nu)_{\nu\in\mathbb{N}_0}$$

is an orthonormal system in $H(k_j)$, which by (2.4) is even an orthonormal basis. Intuitively speaking, assumption (A1) means that the $\mathfrak{e}_\nu$ get less important in the space $H(k_j)$ as ν increases.

We observe that $H(k_j)$ is identically embedded into $L^2(\mu_0)$. We denote the embedding operator by T_j and note that T_j is continuous with

$$\|T_j\| = \sup_{\nu\in\mathbb{N}_0} \alpha_{\nu,j}^{-1/2} \geqslant 1. \tag{2.9}$$

Under assumption (A1), this supremum is always attained at $\nu = 0$ or $\nu = 1$. Under assumption (A2), it is attained at $\nu \neq 0$ for at most finitely many $j \in \mathbb{N}$.

We briefly turn our attention to the Hilbert space adjoint operator T_j^* of T_j. For $\nu_1, \nu_2 \in \mathbb{N}_0$ we have

$$\alpha_{\nu_2,j} \cdot \langle \mathfrak{e}_{\nu_1}, \mathfrak{e}_{\nu_2} \rangle_{H(k_j)} = \langle \mathfrak{e}_{\nu_1}, \mathfrak{e}_{\nu_2} \rangle_{L^2(\mu_0)} = \langle \mathfrak{e}_{\nu_1}, T_j^* \mathfrak{e}_{\nu_2} \rangle_{H(k_j)} \tag{2.10}$$

and therefore $T_j^*\mathfrak{e}_\nu = \alpha_{\nu.j}^{-1}\mathfrak{e}_\nu$ for all $\nu \in \mathbb{N}$. It follows that

$$T_j^* T_j f = \sum_{\nu\in\mathbb{N}_0} \alpha_{\nu,j}^{-1} \cdot \langle f, \mathfrak{e}_\nu \rangle_{L^2(\mu_0)} \cdot \mathfrak{e}_\nu \tag{2.11}$$

for all $f \in H(k_j)$, see Gnewuch et al. (2019, Eqn. (14)). Therefore, the sequence of singular values of T_j is $(\alpha_{\nu,j}^{-1/2})_{\nu\in\mathbb{N}_0}$, which implies firstly that T_j is injective and secondly that T_j is compact if and only if

$$\lim_{\nu\to\infty} \alpha_{\nu,j}^{-1} = 0.$$

We remark that this is a very weak additional requirement, since we already assume (2.7) and $D \neq \varnothing$. Further, it is obviously implied by (A2).

In Section 2.2, we are also interested in identical embeddings $S_{j.i}$ from $H(k_j)$ into $H(k_i)$, where $i < j \in \mathbb{N}$. Analogous to our discussion on T_j, we get the following results. Firstly, $S_{j,i}$ is well-defined as well as continuous if and only if

$$\sup_{\nu\in\mathbb{N}} \frac{\alpha_{\nu,i}}{\alpha_{\nu,j}} < \infty,$$

in which case

$$\|S_{j,i}\|^2 = \sup_{\nu\in\mathbb{N}_0} \frac{\alpha_{\nu,i}}{\alpha_{\nu,j}} \geqslant 1. \tag{2.12}$$

This norm makes the meaning of assumption (A3) clear: Note that $H(k_j)$ is the orthogonal sum of the space $H(1)$ of constant functions and the space

$$H(k_j - 1) = \left\{ f \in H(k_j) \colon \int_D f d\mu_0 = 0 \right\}, \tag{2.13}$$

so $\gamma_j^{1/2}$ as given in (A3) is the norm of the identical embedding of $H(k_j - 1)$ into the space $H(k_1 - 1)$. Therefore, γ_j as given in (A3) being finite assures that $H(k_j - 1)$ is always embedded into $H(k_1 - 1)$, the norm of the embedding being given by $\gamma_j^{1/2}$, and (A3) as a whole ensures that the norms of the embeddings from $H(k_j - 1)$ to $H(k_1 - 1)$ converge to 0 sufficiently fast.

Secondly $S_{j,i}$, if it is well-defined, is always injective and thirdly is compact if and only if

$$\lim_{\nu\to\infty} \frac{\alpha_{\nu,i}}{\alpha_{\nu,j}} = 0.$$

This setting of compact identical embeddings is also known as *increasing smoothness.*

We conclude this section by studying the infinite-variate embedding from $H(K)$ into $L^2(\mu)$. This was shown in the case of $\mathfrak{e}_\nu$ being the Hermite polynomial of degree ν in Gnewuch et al. (2022, Lem. 3.8, Prop. 3.10 and Lem. 3.12). Here, we present a different proof, in part based on the results from Section 1.1.

Theorem 2.2. *Assume (A2) is fulfilled. Then $H(K)$ is identically embedded into $L^2(\mu)$, the norm of T given by*

$$\|T\| = \prod_{j\in\mathbb{N}} \sup_{\nu\in\mathbb{N}_0} \alpha_{\nu,j}^{-1/2}.$$

Furthermore, T is injective and compact, and we have

$$\mu(\mathfrak{X}) = 1.$$

Proof. To show $\mu(\{\sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) < \infty\}) = 1$, it suffices to show

$$\int_{D^\mathbb{N}} \sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) d\mu < \infty.$$

This is indeed true, since, by monotone convergence,

$$\int_{D^\mathbb{N}} \sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) d\mu = \sum_{\nu,j\in\mathbb{N}} \int_D \alpha_{\nu,j}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) d\mu_0 = \sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} < \infty.$$

We apply Theorem 1.4. Note that, as previously discussed, the embedding T_j from $H(k_j)$ into $L^2(\mu_0)$ always exists, is continuous as well as injective, and its operator norm is given by (2.9). Further, (A2) ensures that $\|T_j\| \neq 1$ only finitely often, which in turn implies (1.10). Note that (1.11) and (1.12) are trivilally fulfilled here with $u_j = 1$ for $j \in \mathbb{N}$. Therefore, by Theorem 1.4, T is indeed continuous with the norm given as claimed.

For ease of notation, we do not distinguish between square-integrable functions and their equivalence classes in the following. To show compactness, we make use of the adjoint operator T^* and the eigenvalues of T^*T again, cf. (2.10) and (2.11). Let $\boldsymbol{U}$ be the set of all finite subsets of $\mathbb{N}$. For $\boldsymbol{\nu} \in \boldsymbol{U}$, we define $\mathfrak{e}_{\boldsymbol{\nu}}$ and $\beta_{\boldsymbol{\nu}}$ by

$$\mathfrak{e}_{\boldsymbol{\nu}}(\boldsymbol{x}) = \prod_{j\in\boldsymbol{\nu}} \mathfrak{e}_{\boldsymbol{\nu}_j}(x_j) \text{ and } \beta_{\boldsymbol{\nu}} = \prod_{j\in\boldsymbol{\nu}} \alpha_{\boldsymbol{\nu}_j,j}^{-1/2}$$

for all $\boldsymbol{x} \in \mathfrak{X}$. Then it follows from Remark A.1 that $(\mathfrak{e}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\boldsymbol{U}}$ is an orthonormal basis of $L^2(\mu)$ and from Remark 1.1 that $(\beta_{\boldsymbol{\nu}} \cdot \mathfrak{e}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\boldsymbol{U}}$ is an orthonormal basis of $H(K)$. Analogously to (2.10), we get

$$T^* \mathfrak{e}_{\boldsymbol{\nu}} = \beta_{\boldsymbol{\nu}}^{-1} \mathfrak{e}_{\boldsymbol{\nu}}$$

for all $\boldsymbol{\nu} \in \boldsymbol{U}$, which leads to the decomposition

$$T^*Tf = \sum_{\boldsymbol{\nu}\in\boldsymbol{U}} \beta_{\boldsymbol{\nu}}^{-1} \cdot \langle f, \mathfrak{e}_{\boldsymbol{\nu}} \rangle_{L^2(\mu)} \cdot \mathfrak{e}_{\boldsymbol{\nu}}.$$

Since

$$\beta_{\boldsymbol{\nu}} \leqslant (\sup_{j\in\boldsymbol{\nu}} \alpha_{\nu_j,j}^{-1})^{|\boldsymbol{\nu}|}$$

holds for all $\boldsymbol{\nu} \in \boldsymbol{U}$, the summability condition (A2) implies the convergence of $(\beta_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\boldsymbol{U}}$ to 0 and therefore compactness of T. Since $\beta_{\boldsymbol{\nu}} \neq 0$ for all $\boldsymbol{\nu} \in \boldsymbol{U}$, we get injectivity of T. □

Remark 2.3. We briefly consider only two kernels say k_1 and k_2, and assume that $H(k_2)$ is compactly identically embedded into $H(k_1)$, or equivalently

$$\lim_{\nu\to\infty} \frac{\alpha_{\nu,1}}{\alpha_{\nu,2}} = 0.$$

We mention briefly a result on interpolation between these spaces due to Gnewuch et al. (2022, Rem. 2.3). Let $0 < \theta < 1$ and define the Fourier weights

$$\alpha_\nu = \alpha_{\nu,2}^{\theta} \cdot \alpha_{\nu,1}^{1-\theta}. \tag{2.14}$$

For the kernel $k^{(\theta)}$ given by

$$k^{(\theta)}(x, y) = 1 + \sum_{\nu\in\mathbb{N}} \alpha_\nu^{-1} \mathfrak{e}_\nu(x)\mathfrak{e}_\nu(y)$$

for $x, y \in D$, we obtain the following result. The space $H(k^{(\theta)})$ is the same as the interpolation space $K_{\theta,2}(H(k_2), H(k_1)) = J_{\theta,2}(H(k_2), H(k_1))$ obtained with quadratic interpolation via the K- or the J-method. We do not explain these methods further here, only mentioning the following consequence which will be of some relevance to us in Section 2.4.3: For $f \in H(k_2)$, we have

$$\|f\|_{H(k^{(\theta)})} \leqslant \|f\|_{H(k_2)}^{1-\theta} \cdot \|f\|_{H(k_1)}^{\theta},$$

see Chandler-Wilde et al. (2015, Lem. 2.5). For details on the interpolation of Hilbert spaces, we refer to Chandler-Wilde et al. (2015).

2.1.2 Weights of Polynomial and Exponential Growth

In this section, we study the two types of Fourier weights (PG) and (EG) which are defined as follows.

(PG) Let $r_j > 1$ for $j \in \mathbb{N}$ such that

$$\sum_{j \in \mathbb{N}} 2^{-r_j} < \infty \tag{2.15}$$

and

$$r_1 = \inf_{j \in \mathbb{N}} r_j. \tag{2.16}$$

The Fourier weights with a *polynomial growth* are given by

$$\alpha_{\nu,j} := (\nu + 1)^{r_j}$$

for $\nu, j \in \mathbb{N}$.

(EG) Let $r_j, b_j > 0$ for $j \in \mathbb{N}$ such that (2.15) and (2.16) are satisfied and

$$b_1 = \inf_{j \in \mathbb{N}} b_j. \tag{2.17}$$

The Fourier weights with a *(sub-)exponential growth* are given by

$$\alpha_{\nu,j} := 2^{r_j \cdot \nu^{b_j}}$$

for $\nu, j \in \mathbb{N}$.

Note that in the case (EG), the weights can be parametized differently by varying the base of exponentiation instead of the r_j in the exponent: Let $b_j, \beta_j > 0$ be given for $j \in \mathbb{N}$ such that (2.17) is satisfied and

$$\sum_{j \in \mathbb{N}} \beta_j < \infty.$$

Define Fourier weights

$$\alpha_{\nu,j} = \beta_j^{-\nu^{b_j}}$$

for $\nu, j \in \mathbb{N}$. Then,

$$\alpha_{\nu,j} = 2^{-\nu^{b_j} \log_2(\beta_j)}$$

and $r_j = \log_2(\beta_j)$ satisfies (2.15), and therefore, we are in the case (EG) (after potentially reordering the kernels to ensure $r_1 = \inf_{j \in \mathbb{N}} r_j$). This different presentation of the case (EG) will be most useful in Chapter 3, where we establish a close relation between Hermite spaces in the case (EG) and RKHSs with Gaussian kernels, while the original presentation is more useful in this chapter and also sometimes allows us to consider (PG) and (EG) simultaneously.

Lemma 2.4. *In both cases,* (PG) *and* (EG)*, we have* (A1)*,* (A2)*, and* (A3) *with*

$$\gamma_j = 2^{r_1 - r_j}$$

for every $j \in \mathbb{N}$.

Proof. Obviously, (A1) is satisfied.

Furthermore, we have $\sum_{\nu \in \mathbb{N}} \alpha_{\nu,j}^{-1} < \infty$ for every $j \in \mathbb{N}$, which is also obvious. To show (A2), it therefore suffices to show that there exists an integer $j_0 \in \mathbb{N}$ such that

$$\sum_{\nu \in \mathbb{N},\ j \geqslant j_0} \alpha_{\nu,j}^{-1} < \infty.$$

This was shown in Gnewuch et al. (2022, Lem. 3.17), utilizing in both cases the inequality

$$\sum_{n \geqslant n_0} n^{-\tau} \leqslant n_0^{-\tau} + \int_{n_0}^{\infty} n^{-\tau} d\tau = n_0^{-\tau} \left(\frac{n_0}{\tau - 1} \right) \tag{2.18}$$

for $n_0 \in \mathbb{N}_0$ and $\tau > 1$, which yields the claim for (PG) immediately and for (EG) since there exists a $c > 0$ such that

$$\alpha_{\nu,j}^{-1} \leqslant j^{-c\nu^{b_1}}$$

for all $\nu, j \in \mathbb{N}$.

Regarding (A3), the supremum in the definition of (A3) is attained for $\nu = 1$ in both cases. In the case (PG) we obtain

$$\gamma_j = \sup_{\nu \in \mathbb{N}} (\nu + 1)^{r_1 - r_j} = 2^{r_1 - r_j}$$

for every $j \in \mathbb{N}$ from (2.16), so that (A3) follows from (2.15) in this case. In the sequel we consider the case (EG), where we have

$$\gamma_j = \sup_{\nu \in \mathbb{N}} 2^{r_1 \cdot \nu^{b_1} - r_j \cdot \nu^{b_j}}.$$

Using (2.16) and (2.17) we obtain

$$2^{r_1 - r_j} \leqslant \gamma_j \leqslant \sup_{\nu \in \mathbb{N}} 2^{(r_1 - r_j) \cdot \nu^{b_1}} = 2^{r_1 - r_j}$$

for every $j \in \mathbb{N}$, and (A3) follows from (2.15) also in the case (EG). □

Remark 2.5. Recall from Section 2.1.1 that for $i, j \in \mathbb{N}$ we have $H(k_j) \subsetneq H(k_i)$ with a compact identical embedding if and only if

$$\lim_{\nu \to \infty} \frac{\alpha_{\nu,i}}{\alpha_{\nu,j}} = 0. \tag{2.19}$$

In the case (PG) we have (2.19) if and only if $r_i < r_j$; observe that the latter holds for a fixed $i \in \mathbb{N}$ if j is sufficiently large. In the case (EG) we have (2.19) if and only if $b_i < b_j$ or $b_i = b_j$ and $r_i < r_j$; observe that the latter holds for $i = 1$ if j is sufficiently large.

2.2 Anchored Spaces and Embeddings

Fix some $a \in D$. A reproducing kernel m with domain $D \times D$ is called *anchored* at a if $m(a,a) = 0$. It is straightforward to see that this is equivalent to $f(a) = 0$ for all $f \in H(m)$.

A convenient setting to study infinite-variate integration and L^2-approximation is given if we have a sequence $(m_j)_{j\in\mathbb{N}}$ of reproducing kernels on $D \times D$, where

$$m_j = \tilde{\gamma}_j \cdot m$$

for suitable positive numbers $\tilde{\gamma}_j$ and a kernel m that is anchored at $a \in D$. In this case, in Section 2.3, we will study so-called *multidimensional decomposition methods (MDMs)* on $H(M)$ where the kernel

$$M = \bigotimes_{j\in\mathbb{N}} (1 + m_j)^{(1)}$$

is defined on an appropriate domain as laid out in Chapter 1. In this case, the orthogonal decomposition

$$f = \sum_{\boldsymbol{\nu}\in\boldsymbol{U}} f_{\boldsymbol{\nu}}$$

of $f \in H(M)$ described in Remark 1.2 is an *anchored decomposition*, which, in particular, allows the evaluation of $f_{\boldsymbol{\nu}}$ at a point $\boldsymbol{x}_{\boldsymbol{\nu}}$ using an acceptably small number of function evaluations of f at points $\boldsymbol{y}$ with at most $|\boldsymbol{\nu}|$ components not equal to a. We make this more precise: Assume that the constant sequence $\boldsymbol{a} = (a, a, \ldots)$ is in the domain $\mathfrak{X}$ of M and therefore, so is every sequence which differs from $\boldsymbol{a}$ in at most finitely many points. Since m is anchored at a, we have

$$f_{\boldsymbol{\nu}}(\boldsymbol{x}) = \sum_{\boldsymbol{\lambda}\subseteq\boldsymbol{\nu}} (-1)^{|\boldsymbol{\nu}\setminus\boldsymbol{\lambda}|} \cdot f(\boldsymbol{x}_{\boldsymbol{\lambda}}, \boldsymbol{a}_{\boldsymbol{\lambda}^c}) \tag{2.20}$$

for every $\boldsymbol{x} \in \mathfrak{X}$, where $f(\boldsymbol{x}_{\boldsymbol{\lambda}}, \boldsymbol{a}_{\boldsymbol{\lambda}^c})$ denotes the value of f at the point $\boldsymbol{y}$ given by $y_j := x_j$ if $j \in \boldsymbol{\lambda}$ and $y_j = a$ otherwise, see Kuo et al. (2010a, Exmp. 2.3). In particular, $f_\varnothing$ is constant and we have

$$f_\varnothing = f(\boldsymbol{a}). \tag{2.21}$$

We conclude that a finite number of function values of f suffice to obtain a function value of $f_{\boldsymbol{\nu}}$; the corresponding number $2^{|\boldsymbol{\nu}|}$ is acceptable as long as $|\boldsymbol{\nu}|$ is sufficiently small. The anchored component $f_{\boldsymbol{\nu}}$ of f has the property that

$$f_{\boldsymbol{\nu}}(\boldsymbol{x}_{\boldsymbol{\lambda}}) = 0 \text{ if } x_j = a \text{ for some } j \in \boldsymbol{\lambda}. \tag{2.22}$$

See Kuo et al. (2010a, Thm. 2.1) for a general result on the decomposition of multivariate functions.

In this setting, the decomposition $f = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} f_{\boldsymbol{\nu}}$ is called the *anchored function decomposition*, which is also known as *cut HDMR*, where HDMR stands for 'high-dimensional model representation', cf. Rabitz and Alış (1999) and Kuo et al. (2010a). It will turn out that in this setting of the anchored decomposition, MDMs are almost optimal algorithms for integration as well as L^2-approximation, see Section 2.3.2 combined with Theorem 2.9.

However, we aim to find such algorithms for spaces in the setting as given by Section 2.1, where $k_j - 1$ is not necessarily anchored. Indeed, it follows immediately from the definition of k_j that $k_j(a, a) = 0$ if and only if $\mathfrak{e}_\nu(a) = 0$ for all $\nu \in \mathbb{N}$. In the case of Hermite spaces, which we will study in Section 2.4, $\mathfrak{e}_1$ and $\mathfrak{e}_2$ do not share a root.

More generally, if $\mathfrak{e}_\nu(a) = 0$ is true for all $\nu \in \mathbb{N}$, by (2.13) we have for all $f \in H(K)$ and $\varnothing \neq \boldsymbol{\nu} \in \boldsymbol{U}$ that $f_{\boldsymbol{\nu}}$ has integral 0, which implies

$$\int_{\mathfrak{X}} f d\mu = f_\varnothing$$

and, by (2.21),

$$\int_{\mathfrak{X}} f d\mu = f(\boldsymbol{a}).$$

In other words, the integration problem is trivial in this case. This result is well known, see, e.g., Hickernell et al. (2010, Rem. 3).

Throughout this section and Section 2.3.1, we may therefore assume that for some $\nu_0 \in \mathbb{N}$, we have $\mathfrak{e}_{\nu_0}(a) \neq 0$. For ease of presentation, we actually only consider the case

$$\mathfrak{e}_1(a) \neq 0, \tag{2.23}$$

which is essentially without loss of generality. Formally, we may replace $\mathfrak{e}_1$ by $\mathfrak{e}_{\nu_0}$ and the corresponding Fourier weights $\alpha_{1,j}$ by $\alpha_{\nu_0,j}$ in our results. We comment on the (non-existent) impact of (2.23) on the L^2-approximation problem in Remark 2.11.

To achieve upper error bounds, in this section, we study the embedding of $H(K)$ into the product of suitably chosen anchored spaces. To achieve lower bounds, we then study the embedding of such a product into $H(K)$.

The following theorem, which is based on Gnewuch et al. (2024, Thm. 3.2), describes how $H(K)$ can be embedded (not identically, but via a restriction of the domain) into a suitable space as described above. Despite the more general setting at hand here, the proof is very similar to that of Gnewuch et al. (2024, Thm. 3.2). We set

$$c^{\uparrow}(a) = 1 + k_1(a, a) \tag{2.24}$$

for $a \in D$.

Theorem 2.6. *Assume that (A1)–(A3) hold. For every $a \in D$ there exists a reproducing kernel*

$$m_a^{\uparrow} \colon D \times D \to \mathbb{R}$$

with the following properties:

(i) We have

$$H(k_1) = H(1 + m_a^{\uparrow})$$

as vector spaces, and

$$m_a^{\uparrow}(a, a) = 0.$$

The operator norms of the identical embeddings $T_j \colon H(k_j) \hookrightarrow H(1 + \gamma_j m_a^{\uparrow})$ and $V \colon H(m_a^{\uparrow}) \hookrightarrow H(k_1)$ satisfy

$$\|T_j\| \leqslant \left(1 + \gamma_j c^{\uparrow}(a)\right)^{1/2}$$

for every $j \in \mathbb{N}$ as well as

$$\|V\| \leqslant c^{\uparrow}(a).$$

(ii) The maximal domain of the reproducing kernel

$$M_a^{\uparrow}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j \in \mathbb{N}} (1 + \gamma_j \cdot m_a^{\uparrow}(x_j, y_j)) \tag{2.25}$$

is equal to

$$\mathfrak{X}^{\uparrow} = \left\{ \boldsymbol{x} \in D^{\mathbb{N}} \colon \sum_{\nu, j \in \mathbb{N}} \gamma_j \alpha_{\nu,1}^{-1} \cdot \mathfrak{e}_\nu^2(x_j) < \infty \right\}. \tag{2.26}$$

Further, we have

$$\mu(\mathfrak{X}^{\uparrow}) = 1 \tag{2.27}$$

and

$$\{\boldsymbol{x} \in D^{\mathbb{N}} \colon \boldsymbol{x} \text{ constant except for finitely many terms}\} \subseteq \mathfrak{X}^{\uparrow} \subseteq \mathfrak{X}. \tag{2.28}$$

(iii) We have

$$\{f|_{\mathfrak{X}^{\uparrow}} \colon f \in H(K)\} \subseteq H(M_a^{\uparrow}) \subseteq L^2(\mu),$$

and the operator norm of the restriction $T \colon H(K) \to H(M_a^{\uparrow})$, $f \mapsto f|_{\mathfrak{X}^{\uparrow}}$ satisfies

$$\|T\| \leqslant C^{\uparrow}(a) := \prod_{j \in \mathbb{N}} (1 + \gamma_j \cdot c^{\uparrow}(a))^{1/2}. \tag{2.29}$$

Proof. Recall from (2.13) that $H(k_1)$ is the orthogonal sum of the spaces $H(1)$ and $H(k^*)$, where $k^* := k_1 - 1$. Moreover, we have $\gamma_1 = 1$. We put

$$\xi(f) := \int_D f d\mu_0. \tag{2.30}$$

In a first embedding step we consider the reproducing kernels

$$k_j^{\uparrow} := 1 + \gamma_j k^*$$

for $j \in \mathbb{N}$. We have

$$H(k_j^{\uparrow}) = H(k_1)$$

as vector spaces, and

$$\|f\|_{H(k_j^{\uparrow})}^2 = |\xi(f)|^2 + \gamma_j^{-1} \cdot \|f\|^2$$

for $f \in H(k_1)$, where

$$\|f\|^2 := \sum_{\nu \in \mathbb{N}} \alpha_{\nu,1} \cdot \langle f, \mathfrak{e}_\nu \rangle_{L^2(\mu_0)}^2,$$

cf. Gnewuch et al. (2022, Prop. 2.1). Since $\frac{\alpha_{\nu,1}}{\gamma_j} \leqslant \alpha_{\nu,j}$ holds, we have $H(k_j) \subseteq H(k_j^{\uparrow})$ with an identical embedding

$$T_j^{(1)} \colon H(k_j) \hookrightarrow H(k_j^{\uparrow})$$

of norm one.

Next, we describe a second embedding step, which depends on $a \in D$. This step is essentially due to Gnewuch et al. (2019, Lem. 3.13), but with slightly different assumptions on the weights. Our goal is to show that there exists a reproducing kernel $\widehat{m}_a$ that satisfies $H(k_1) = H(1 + \widehat{m}_a)$ and $\widehat{m}_a(a, a) = 0$ as well as $H(1) \cap H(\widehat{m}_a) = \{0\}$ and

$$\|f\|_{H(1+\gamma\widehat{m}_a)}^2 = |f(a)|^2 + \gamma^{-1} \cdot \|f\|^2$$

for every $f \in H(k_1)$ and every $\gamma > 0$. On $H(k_1)$, we introduce the seminorm $\|\cdot\|_a$ given by

$$\|f\|_a = |f(a)|$$

for $f \in H(k_1)$. Then, the following two conditions that were introduced in Gnewuch et al. (2017) are fulfilled:

(B1) $k^* \neq 0$ is a reproducing kernel on some set $D \times D$ and $H(k^*) \cap H(1) = \{0\}$ and

(B2) $\|1\|_a = 1$ and there exists a $c > 0$ such that $\|f\|_a \leqslant \|f\|_{(k^*)}$ for all $f \in H(k^*)$.

For (B2), one can choose c to be the operator norm of the evaluation functional at the point a in the nontrivial case that this is not the zero functional. It was shown in Gnewuch et al. (2017, Rem. 2.5) via the equivalence of certain norms that (B1) and

(B2) imply, in particular, that on $H(k_1)$ equipped with the norm given by

$$\|f\|_{\tilde{a}}^2 = |f(a)|^2 + \|f\|$$

for $f \in H(k_1)$, all function evaluations are continuous. Further, it is clear that this norm is induced by a scalar product. Therefore, $H(k_1)$ equipped with this norm is a RKHS, and we denote the corresponding kernel by m. Finally, let $\widehat{m}_a$ be the reproducing kernel of the closed subspace of $H(m)$ containing all f with $f(a) = 0$. Then $\widehat{m}_a$ has all required properties.

Now, let $f \in H(k_1)$. Since

$$|f(a) - \xi(f)| \leqslant k_1(a,a)^{1/2} \cdot \|f - \xi(f)\|_{H(k_1)} = k_1(a,a)^{1/2} \cdot \|f\|,$$

we obtain

$$\begin{aligned}
\|f\|_{H(1+\gamma\widehat{m}_a)}^2 &\leqslant (|\xi(f)| + |f(a) - \xi(f)|)^2 + \gamma^{-1} \cdot \|f\|^2 \\
&\leqslant \left(|\xi(f)| + k_1(a,a)^{1/2} \cdot \|f\|\right)^2 + \gamma^{-1} \cdot \|f\|^2 \\
&\leqslant (1+\gamma) \cdot |\xi(f)|^2 + \left(1 + \gamma^{-1}\right) \cdot k_1(a,a) \cdot \|f\|^2 + \gamma^{-1} \cdot \|f\|^2 \\
&\leqslant (1+\gamma) \cdot \left(|\xi(f)|^2 + \gamma^{-1} \cdot k_1(a,a) \cdot \|f\|^2 + \gamma^{-1} \cdot \|f\|^2\right) \\
&= (1+\gamma) \cdot \left(|\xi(f)|^2 + \gamma^{-1} c^{\uparrow}(a) \cdot \|f\|^2\right).
\end{aligned}$$

For $f \in H(\widehat{m}_a)$ we also obtain

$$\|f\|_{H(k_1)}^2 = |f(a) - \xi(f)|^2 + \|f\|^2 \leqslant c^{\uparrow}(a) \cdot \|f\|^2.$$

Put $m_a^{\uparrow} := c^{\uparrow}(a) \cdot \widehat{m}_a$. Clearly, $H(k_1) = H(1 + m_a^{\uparrow})$ and $m_a^{\uparrow}(a,a) = 0$. Furthermore,

$$\|f\|_{H(1+\gamma \cdot m_a^{\uparrow})} \leqslant (1 + \gamma c^{\uparrow}(a))^{1/2} \cdot \|f\|_{H(1+\gamma k*)}$$

for every $f \in H(k_1)$ and every $\gamma > 0$, cf. Gnewuch et al. (2017, Thm. 2.1) and its proof. Moreover,

$$\|f\|_{H(k_1)} \leqslant c^{\uparrow}(a) \cdot \|f\|_{H(m_a^{\uparrow})}$$

for every $f \in H(m_a^{\uparrow})$. Consequently,

$$\|T_j\| \leqslant \|T_j^{(1)}\| \cdot \|T_j^{(2)}\| \leqslant (1 + \gamma_j c^{\uparrow}(a))^{1/2}$$

for the identical embedding

$$T_j^{(2)} \colon H(k_j^{\uparrow}) \hookrightarrow H(1 + \gamma_j \cdot m_a^{\uparrow})$$

as well as $\|V\| \leqslant c^{\uparrow}(a)$. This completes the proof of (i).

Due to the equivalence of $\|\cdot\|_{H(k_1)}$ and $\|\cdot\|_{H(1+m_a^{\uparrow})}$, there exists a constant $c > 0$ such that

$$m_a^{\uparrow}(x,x) \leqslant c \cdot (1 + k^*(x,x))$$

and

$$k^*(x,x) \leqslant c \cdot (1 + m_a^{\uparrow}(x,x))$$

for all $x \in D$. Let $\boldsymbol{x} \in D^{\mathbb{N}}$. Employing (A3) we conclude that

$$\sum_{j \in \mathbb{N}} \gamma_j m_a^{\uparrow}(x_j, x_j) < \infty \qquad \Leftrightarrow \qquad \boldsymbol{x} \in \mathfrak{X}^{\uparrow}.$$

Therefore $M_a^{\uparrow}$ as defined in (2.25) is a reproducing kernel with maximal domain $\mathfrak{X}^{\uparrow}$.

Analogous to Theorem 2.2, $\mu(\mathfrak{X}^{\uparrow}) = 1$ follows from the properties (A2) and (A3). Further, $\mathfrak{X}^{\uparrow} \subseteq \mathfrak{X}$ is obvious by definition of γ_j and $\{\boldsymbol{x} \in D^{\mathbb{N}} \colon \boldsymbol{x} \text{ constant}\} \subseteq \mathfrak{X}^{\uparrow}$ follows from (2.3) along with (A3).

We claim that

$$\{f|_{\mathfrak{X}^{\uparrow}} \colon f \in H(K)\} \subseteq H(M_a^{\uparrow}).$$

Indeed, Theorem A.5, together with (A3) and applied with $u_j := 1$ for every $j \in \mathbb{N}$, $\mathfrak{X}^{(\boldsymbol{u})} := \mathfrak{X}$, and $\mathfrak{Y}^{(\boldsymbol{u})} := \mathfrak{X}^{\uparrow}$, yield the claim as well as (2.29).

Since

$$\int_D m_a^{\uparrow}(x,x) d\mu_0(x) \leqslant c \cdot \int_D k_1(x,x) d\mu_0(x) = c \cdot \left(1 + \sum_{\nu \in \mathbb{N}} \alpha_{\nu,1}^{-1}\right) < \infty,$$

which follows from (A2), we obtain

$$\int_{\mathfrak{X}^\uparrow} M_a^\uparrow(\boldsymbol{x},\boldsymbol{x})d\mu(\boldsymbol{x}) = \prod_{j\in\mathbb{N}} \left(1+\gamma_j c^\uparrow(a)\int_D m_a^\uparrow(x,x)d\mu_0(x)\right) < \infty$$

using (A3). Consequently, $H(M_a^\uparrow) \subseteq L^2(\mu)$. □

We remark that in Gnewuch et al. (2024, Thm. 3.2), in the case of Hermite spaces it was even shown that $\mathfrak{X}^\uparrow$ contains all bounded sequences. This seems not to be possible in the setting at hand, without any knowledge about the growth of $\mathfrak{e}_\nu(x)$ as x or ν tend to infinity. However, it is also not needed in the forthcoming analysis.

Remark 2.7. We describe the decomposition of $M_a^\uparrow$ as given in Remark 1.2. Let $\boldsymbol{U}$ denote the set of finite subsets of $\mathbb{N}$. For $\boldsymbol{\nu} \in \boldsymbol{U}$ let

$$\gamma_{\boldsymbol{\nu}} := \prod_{j\in\boldsymbol{u}} \gamma_j$$

and

$$m_{a,\boldsymbol{\nu}}^\uparrow(\boldsymbol{x},\boldsymbol{y}) := \prod_{j\in\boldsymbol{\nu}} m_a^\uparrow(x_j,y_j), \qquad \boldsymbol{x},\boldsymbol{y}\in\mathfrak{X}^\uparrow.$$

The reproducing kernel $M_a^\uparrow$ according to Theorem 3.24 is of the form

$$M_a^\uparrow = \sum_{\boldsymbol{\nu}\in\boldsymbol{U}} \gamma_{\boldsymbol{u}}\cdot m_{a,\boldsymbol{\nu}}^\uparrow.$$

This also fits into the situation from Remark 1.3 of weighted tensor product kernels with weights $\gamma_{\boldsymbol{u}}$ of product form. We also have that $m_a^\uparrow(a,a) = 0$, i.e., $m_a^\uparrow$ is anchored at a.

We show a corresponding result for an embedding into $H(K)$ rather than from it. Theorem 2.6 allows us to study upper bounds for integration and L^2-approximation on $H(K)$ by instead considering $H(M_a^\uparrow)$, Theorem 2.8 will allow us to study lower bounds of the error of integration and L^2-approximation on $H(K)$. In comparison to Gnewuch et al. (2024, Thm. 3.7), some changes are needed in the second embedding step to suit our more general setting in Theorem 2.8 and its proof.

We introduce the set

$$\mathfrak{X}^{\downarrow} := \Big\{ \boldsymbol{x} \in D^{\mathbb{N}} \colon \sum_{j \in \mathbb{N}} \alpha_{1,j}^{-1} \cdot \mathfrak{e}_1(x_j)^2 < \infty \Big\},$$

which will turn out to be the maximal domain for $M_a^{\downarrow}$. Clearly we have $\mathfrak{X} \subseteq \mathfrak{X}^{\downarrow} \subseteq D^{\mathbb{N}}$, so that $\mu(\mathfrak{X}^{\downarrow}) = 1$ follows from $\mu(\mathfrak{X}) = 1$.

For $a \in D$ we put

$$c^{\downarrow}(a) := (1 + \alpha_{1,1} + a^2)^{-1}.$$

Theorem 2.8. *Assume that* (A1)–(A3) hold. For every $a \in D$ there exists a reproducing kernel

$$m_a^{\downarrow} : D \times D \to \mathbb{R}$$

with the following properties:

(i) We have

$$H(1 + m_a^{\downarrow}) = \operatorname{span}\{\mathfrak{e}_0, \mathfrak{e}_1\}$$

as vector spaces and

$$m_a^{\downarrow}(a, a) = 0.$$

(ii) The maximal domain of the reproducing kernel

$$M_a^{\downarrow}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j \in \mathbb{N}} (1 + \alpha_{1,j}^{-1} \cdot m_a^{\downarrow}(x_j, y_j))$$

is equal to $\mathfrak{X}^{\downarrow}$.

(iii) We have

$$\{ f|_{\mathfrak{X}} \colon f \in H(M_a^{\downarrow}) \} \subseteq H(K)$$

and the operator norm of the restriction $T \colon H(M_a^{\downarrow}) \to H(K)$, $f \mapsto f|_{\mathfrak{X}}$ satisfies

$$\|T\| \leqslant \prod_{j \in \mathbb{N}} (1 + \alpha_{1,j}^{-1})^{1/2}.$$

Proof. We begin by describing two embedding steps. Let $\xi(f)$ be given by (2.30). In the first embedding step we consider for $j \in \mathbb{N}$ the reproducing kernels $k_j^{\downarrow}$ given by

$$k_j^{\downarrow}(x,y) := 1 + \alpha_{1,j}^{-1} \cdot \mathfrak{e}_1(x) \cdot \mathfrak{e}_1(y)$$

for $x, y \in D$. Clearly,

$$H(k_j^{\downarrow}) = \operatorname{span}\{\mathfrak{e}_0, \mathfrak{e}_1\}$$

as vector spaces, and

$$\|f\|_{H(k_j^{\downarrow})}^2 = |\xi(f)|^2 + \alpha_{1,j} \cdot \|f\|^2$$

for every $f \in H(k_1)$, where

$$\|f\| := |\langle f, \mathfrak{e}_1 \rangle_{L^2(\mu_0)}|.$$

Hence $H(k_j^{\downarrow}) \subseteq H(k_j)$ with an identical embedding

$$T_j^{(3)} \colon H(k_j^{\downarrow}) \hookrightarrow H(k_j)$$

of norm one.

For the second embedding step, we introduce the kernel $\breve{m}_a$, given by

$$\breve{m}_a(x,y) := (\mathfrak{e}_1(x) - \mathfrak{e}_1(a)) \cdot (\mathfrak{e}_1(y) - \mathfrak{e}_1(a))$$

for $x, y \in D$. Obviously, $H(1 + \breve{m}_a) = \operatorname{span}\{\mathfrak{e}_0, \mathfrak{e}_1\}$ and $\breve{m}_a(a,a) = 0$ as well as $H(1) \cap H(\breve{m}_a) = \{0\}$. Furthermore, for every $f \in \operatorname{span}\{\mathfrak{e}_0, \mathfrak{e}_1\}$ there is some $y \in D$ such that $f = f(a) + \breve{m}_a(\cdot, y)$, which yields

$$\|f\|_{H(1+\gamma\breve{m}_a)}^2 = f(a)^2 + \gamma^{-1} \cdot (\mathfrak{e}_1(y) - \mathfrak{e}_1(a))^2 = f(a)^2 + \gamma^{-1} \cdot \langle f, \mathfrak{e}_1 \rangle_{L^2(\mu_0)}^2.$$

Let $f \in \operatorname{span}\{h_0, h_1\}$. Since

$$|\xi(f) - f(a)| \leqslant k_1^{\downarrow}(a,a)^{1/2} \cdot \|f - \xi(f)\|_{H(k_1^{\downarrow})} = (\alpha_{1,1} + a^2)^{1/2} \cdot \|f\|,$$

we have

$$\begin{aligned}
\|f\|^2_{H(k_j^\downarrow)} &\leqslant (|f(a)| + |\xi(f) - f(a)|)^2 + \alpha_{1,j} \cdot \|f\|^2 \\
&\leqslant \left(|f(a)| + (\alpha_{1,1} + a^2)^{1/2} \cdot \|f\|\right)^2 + \alpha_{1,j} \cdot \|f\|^2 \\
&\leqslant \left(1 + \alpha_{1,j}^{-1}\right) \cdot |f(a)|^2 + (1 + \alpha_{1,j}) \cdot (\alpha_{1,1} + a^2) \cdot \|f\|^2 + \alpha_{1,j} \cdot \|f\|^2 \\
&\leqslant \left(1 + \alpha_{1,j}^{-1}\right) \cdot \left(|f(a)|^2 + c^\downarrow(a)^{-1} \cdot \alpha_{1,j} \cdot \|f\|^2\right) \\
&= \left(1 + \alpha_{1,j}^{-1}\right) \cdot \|f\|^2_{H(1 + c^\downarrow(a) \cdot \alpha_{1,j}^{-1} \cdot \check{m}_a)}.
\end{aligned}$$

Put $m_a^\downarrow := c^\downarrow(a) \cdot \check{m}_a$. We get

$$\|T_j^{(4)}\| \leqslant \left(1 + \alpha_{1,j}^{-1}\right)^{1/2}$$

for the norm of the identical embedding

$$T_j^{(4)} \colon H(1 + \alpha_{1,j}^{-1} \cdot m_a^\downarrow) \hookrightarrow H(k_j^\downarrow).$$

Due to the equivalence of the norms $\|\cdot\|_{H(k_1^\downarrow)}$ and $\|\cdot\|_{H(1+m_a^\downarrow)}$, there exists a constant $c > 0$ such that

$$m_a^\downarrow(x,x) \leqslant c \cdot (1 + \alpha_{1,1}^{-1} \cdot \mathfrak{e}_1^2(x))$$

and

$$\alpha_{1,1}^{-1} \cdot \mathfrak{e}_1^2(x) \leqslant c \cdot (1 + m_a^\downarrow(x,x))$$

for all $x \in D$. This in combination with (A2) gives us

$$\sum_{j \in \mathbb{N}} \alpha_{1,j}^{-1} \cdot m_a^\downarrow(x_j, x_j) < \infty \qquad \Leftrightarrow \qquad \boldsymbol{x} \in \mathfrak{X}^\downarrow.$$

This shows that $M_a^\downarrow$ as defined in (ii) has maximal domain $\mathfrak{X}^\downarrow$.

Furthermore,

$$T_j^{(3)} \circ T_j^{(4)} \colon H(1 + \alpha_{1,j}^{-1} \cdot m_a^\downarrow) \hookrightarrow H(k_j)$$

is continuous with norm bounded by $(1 + \alpha_{1,j}^{-1})^{1/2}$. Due to (A2) and Theorem A.5

applied with $u_j := 1$, $\mathfrak{X}^{(\boldsymbol{u})} := \mathfrak{X}^{\downarrow}$, $\mathfrak{Y}^{(\boldsymbol{u})} := \mathfrak{X}$, $K^{(\boldsymbol{u})} := M_a^{\downarrow}$, and $L^{(\boldsymbol{u})} := K$ yields

$$\{f|_{\mathfrak{X}} \colon f \in H(M_a^{\downarrow})\} \subseteq H(K)$$

with the norm of the restriction operator T bounded by $\prod_{j\in\mathbb{N}}(1+\alpha_{1,j}^{-1})^{1/2}$. □

In analogy to Remark 2.7, we have a decomposition of $M_a^{\downarrow}$ given by

$$M_a^{\downarrow} = \sum_{\boldsymbol{\nu}\in\boldsymbol{U}} \gamma_{\boldsymbol{\nu}} \cdot m_{a,\boldsymbol{\nu}}^{\downarrow}, \tag{2.31}$$

where

$$\gamma_{\boldsymbol{\nu}} := \prod_{j\in\boldsymbol{\nu}} \alpha_{1,j}^{-1}$$

and

$$m_{a,\boldsymbol{\nu}}^{\downarrow}(\boldsymbol{x},\boldsymbol{y}) := \prod_{j\in\boldsymbol{\nu}} m_a^{\downarrow}(x_j,y_j),$$

for $\boldsymbol{x},\boldsymbol{y} \in \mathfrak{X}^{\downarrow}$.

2.3 Integration and L^2-Approximation in Infinitely Many Variables

We study integration and L^2-approximation with respect to the infinite-variate product measure μ of the univariate measure μ_0 for functions from the space $H(K)$, which depend on infinitely many variables. More precisely, since the functions in $H(K)$ have maximal domain $\mathfrak{X}$, we consider the restriction of μ onto $\mathfrak{X}$ instead of μ itself. However, we assume that the corresponding Fourier weights satisfy (A1)–(A3), which implies $\mu(\mathfrak{X}) = 1$.

2.3.1 Upper and Lower Error Bounds

In the sequel we fix a constant sequence $\boldsymbol{a} = (a, a, \dots) \in \mathfrak{X}^{\uparrow}$. Note that by Theorem 2.6(ii), all sequences that are constant except finitely often are contained in $\mathfrak{X}^{\uparrow}$.

Based on $\boldsymbol{a}$, we employ the unrestriced subspace sampling model as described in Section 1.3, with the cost of a function evaluation being given by (1.16) and under the condition (1.17). Therefore, any algorithm with finite cost only evaluates functions at points $\boldsymbol{x}$ that differ from $\boldsymbol{a}$ only finitely often. Recall also the definition of the decay, see (1.18), which quantifies, roughly speaking, the polynomial rate of convergence of a sequence to 0.

In the following main result, cf. Gnewuch et al. (2024, Thm. 4.9), we present an upper bound and a lower bound for the decay $\mathrm{dec}(K)$ of the n-th minimal errors $e_n(K)$. We use $\boldsymbol{\alpha}_1^{-1}$ and $\boldsymbol{\gamma}$ to denote the sequences $(\alpha_{1,j}^{-1})_{j\in\mathbb{N}}$ and $(\gamma_j)_{j\in\mathbb{N}}$, which satisfy $\mathrm{decay}(\boldsymbol{\alpha}_1^{-1}) \geqslant 1$ and $\mathrm{decay}(\boldsymbol{\gamma}) \geqslant 1$, see (A2) and (A3). Recall that lower bounds on $\mathrm{dec}(K)$ correspond to upper error bounds and vice versa.

Theorem 2.9. *For integration and L^2-approximation we have*

$$\min\left(\mathrm{dec}(k_1), \frac{\mathrm{decay}(\boldsymbol{\gamma}) - 1}{2}\right) \leqslant \mathrm{dec}(K) \leqslant \min\left(\mathrm{dec}(k_1), \frac{\mathrm{decay}(\boldsymbol{\alpha}_1^{-1}) - 1}{2}\right).$$

Proof. Due to Theorem 2.6, we have $H(k_1) = H(1 + m_a^{\uparrow})$. Hence the closed graph theorem ensures that the norms of $H(k_1)$ and $H(1+m_a^{\uparrow})$ are equivalent, which implies that for some $c_1, c_2 > 0$ we have

$$c_1 \cdot e_n(k_1) \leqslant e_n(1 + m_a^{\uparrow}) \leqslant c_2 \cdot e_n(k_1),$$

which in turn implies

$$\mathrm{dec}(k_1) = \mathrm{dec}(1 + m_a^{\uparrow}).$$

Furthermore, due to Theorem 2.6, the restriction map $f \mapsto f|_{\mathfrak{X}^{\uparrow}}$ is a continuous linear map from $H(K)$ into $H(M_a^{\uparrow})$. Due to (2.27), we have $f = f|_{\mathfrak{X}^{\uparrow}}$ in $L^2(\mu)$ for all $f \in H(K)$ and thereby

$$e_n(K) \leqslant C \cdot e_n(M_a^{\uparrow})$$

for all $n \in \mathbb{N}$, with C denoting the norm of the restriction map. Therefore the lower bound on $\mathrm{dec}(K)$ follows from

$$\mathrm{dec}(M_a^{\uparrow}) \geqslant \min\left(\mathrm{dec}(1 + m_a^{\uparrow}), \frac{\mathrm{decay}(\boldsymbol{\gamma}) - 1}{2}\right),$$

a result that was established for superpositions of weighted tensor products of an anchored kernel through the use of MDM-Algorithms, see Plaskota and Wasilkowski (2011, Thm. 2) for integration and Wasilkowski (2012, Cor. 9) for L^2-approximation and cf. Remark 2.7. See Section 2.3.2 for further details.

Due to Theorem 2.8 and the closed graph theorem, we have that the restriction $f \mapsto f|_{\mathfrak{X}}$ is a continuous linear map from $H(M_a^{\downarrow})$ into $H(K)$. Since $\mu(\mathfrak{X}) = \mu(\mathfrak{X}^{\downarrow}) = 1$, we obtain $f = f|_{\mathfrak{X}}$ in $L^2(\mu)$ for all $f \in H(M_a^{\downarrow})$. Consequently,

$$e_n(K) \geqslant c \cdot e_n(M_a^{\downarrow})$$

for all $n \in \mathbb{N}$, with c denoting the norm of the restriction map. Noting that

$$e_n(K) \geqslant e_n(k_1)$$

for all $n \in \mathbb{N}$, it suffices to prove the upper bound

$$\operatorname{dec}(M_a^{\downarrow}) \leqslant \frac{\operatorname{decay}(\boldsymbol{\alpha}_1^{-1}) - 1}{2} \tag{2.32}$$

to establish the desired upper bound for $\operatorname{dec}(K)$.

In the case of L^2-approximation, this is an established result, ultimately based on the study of algorithms that do not only allow finitely many function evaluations of f, but even the evaluation of finitely many linear bounded functionals at f. We refer to Wasilkowski (2012, Cor. 9).

In the case of integration, we follow an idea due to Kuo et al. (2010b, Sec. 3.3): Let $N \in \mathbb{N}$, and let Q be any algorithm of cost at most N, given by

$$Q(f) = \sum_{i=1}^{k} f(\boldsymbol{x}_i) \cdot b_i$$

for $f \in H(M_a^{\downarrow})$ with $b_i \in \mathbb{R}$ and $\boldsymbol{x}_i$ being different from $\boldsymbol{a}$ only finitely often. By rearranging the terms according to which variable groups of $\boldsymbol{x}_i$ differ from $\boldsymbol{a}$ and renaming the weights b_i and nodes $\boldsymbol{x}_i$, we get the representation

$$Q(f) = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} \sum_{i=1}^{n_{\boldsymbol{\nu}}} f(\boldsymbol{x}_i^{(\boldsymbol{\nu})}) \cdot b_i^{(\boldsymbol{\nu})},$$

where $\boldsymbol{U}$ is the set of all finite subsets of $\mathbb{N}$ and $\boldsymbol{x}_i^{(\boldsymbol{\nu})}$ differs from $\boldsymbol{a}$ exactly at the entries given by $\boldsymbol{\nu}$ for all $\boldsymbol{\nu} \in \boldsymbol{U}$ and $i = 1 \ldots, n_{\boldsymbol{\nu}}$. In the sequel, we use both representations. We now construct a function from $H(M_a^{\downarrow})$ to 'fool' Q.

First, define

$$G(N) = \sup \left\{ \sum_{i=1}^{m} \ell_i \colon m, \ell_i \in \mathbb{N} \text{ and } \sum_{i=1}^{m} \$(\ell_i) \leqslant N \right\}.$$

Because of the lower bound from (1.17), we have

$$G(N) \leqslant c_1 \cdot N.$$

Denote by

$$\boldsymbol{U}_Q = \{\boldsymbol{\nu} \in \boldsymbol{U} \colon n_{\boldsymbol{\nu}} \neq 0\}$$

the set of all variable groups for which there is some node $\boldsymbol{x}_i$ that differs from $\boldsymbol{a}$ exactly at the entries given by $\boldsymbol{\nu}$ and by

$$J_Q = \bigcup_{\boldsymbol{\nu} \in \boldsymbol{U}_Q} \boldsymbol{\nu}$$

the set of all variables used by Q. Clearly, we have

$$|J_Q| \leqslant G(N) \leqslant c_1 \cdot N.$$

We define the univariate function g by

$$g(x) = \mathfrak{e}_1(x) - \mathfrak{e}_1(a) \text{ for } x \in D,$$

which is clearly an element of $H(m_a^{\downarrow})$, and the infinite-variate function f by

$$f(\boldsymbol{x}) = \left(\sum_{j \notin J_Q} \alpha_{1,j}^{-1} \right)^{-1/2} \sum_{j \notin J_Q} \alpha_{1,j}^{-1} g(x_j) \text{ for } \boldsymbol{x} \in \mathfrak{X}^{\downarrow},$$

which is an element of $H(M_a^{\downarrow})$, which follows from the decomposition (2.31). We also get $Q(f) = 0$ by definition of J_Q and (2.22). Further, we have $\int g d\mu_0 = -e_1(a)$ and $\|g\|^2_{1+\alpha_{1,j}^{-1} m_a^{\downarrow}} = \alpha_{1,j}^{-1}$, which implies $\|f\|_{M_a^{\downarrow}} = 1$ and

$$\left| \int f d\mu \right| = \mathfrak{e}_1(a) \cdot \left(\sum_{j \in J_Q} \alpha_{1,j}^{-1} \right)^{1/2},$$

and thereby the desired bound. □

We give a short explanation of the result of Theorem 2.9. Since $H(k_j)$ is continuously embedded into $H(k_1)$ for every $j \in \mathbb{N}$, both the univariate integration problem and L^2-approximation problem are, up to a constant, at least as hard on $H(k_1)$ as on $H(k_j)$. Therefore, $\mathrm{dec}(k_1)$ quantifies how hard our hardest univariate problem is. With the help of our result, to achieve upper or lower error bounds for the infinite-variate integration or L^2-approximation problem on $H(K)$, one needs only to study the corresponding univariate problem on $H(k_1)$.

The sequences $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}_1^{-1}$ both quantify, loosely speaking, how fast the importance of the spaces $H(k_j)$ towards $H(K)$ diminishes, since the first sequence determines the norms of the embeddings from $H(k_j - 1)$ into $H(k_1 - 1)$, while the latter determines the norms of the embeddings from $H(k_j)$ into $L^2(\mu_0)$. Our conclusion is that the difficulty of the integration problem and the L^2-approximation problem can be controlled just in terms of the difficulty of the hardest underlying univariate problem, as well as growth properties of the weights. If the weights grow fast enough, the infinite-variate problem is no harder than the hardest underlying univariate problem.

A natural question is whether or not the upper and lower bounds in Theorem 2.9 match. Obviously, we always have $1 \leqslant \mathrm{decay}(\boldsymbol{\gamma}) \leqslant \mathrm{decay}(\boldsymbol{\alpha}_1^{-1})$. We will establish that in the important cases (PG) and (EG) we even have $\mathrm{decay}(\boldsymbol{\gamma}) = \mathrm{decay}(\boldsymbol{\alpha}_1^{-1})$, so that our upper and lower bounds match, see Remark 2.10.

For completeness' sake, we present an abstract example of Fourier weights where $\boldsymbol{\alpha}_1^{-1}$ decays infinitely fast, while $\boldsymbol{\gamma}$ decays very slowly. To this end, let $\varepsilon > 0$. For $\nu, j \in \mathbb{N}$, define

$$\alpha_{\nu,j} = \begin{cases} 2^j \cdot \nu^{1+\varepsilon}, & \text{if } \nu < j \\ 2^\nu j^{1+\varepsilon}, & \text{otherwise.} \end{cases}$$

Informally speaking, for a fixed j, the sequence $(\alpha_{\nu,j}^{-1})_{\nu\in\mathbb{N}}$ falls only slightly faster than linear until $\nu = j$, then falls exponentially fast. For each $j \in \mathbb{N}$, the supremum in the definition of γ_j is attained at $\nu = j$, so that

$$\gamma_j = j^{-(1+\varepsilon)}.$$

Therefore, we have

$$\text{decay}(\boldsymbol{\alpha}_1^{-1}) = \infty \quad \text{and} \quad \text{decay}(\gamma) = 1 + \varepsilon$$

in this example. It is readily checked that (A1)–(A3) hold. We are not aware of an example in the literature where these Fourier weights are a reasonable choice.

Remark 2.10. We return to the cases (PG) and (EG) from Section 2.1.2. In this case, the upper and lower bounds do match as seen in the following.

We have

$$\gamma_j = 2^{r_1 - r_j} = 2^{r_1} \cdot \alpha_{1,j}^{-1}$$

for every $j \in \mathbb{N}$ in both cases, (PG) and (EG), see Lemma 2.4, and therefore

$$\text{decay}(\boldsymbol{\gamma}) = \rho = \text{decay}(\boldsymbol{\alpha}_1^{-1}). \tag{2.33}$$

Together with Theorem 2.9 this implies

$$\text{dec}(K) = \min\left(\text{dec}(k_1), \frac{\rho - 1}{2}\right) \tag{2.34}$$

for integration and for L^2-approximation.

Therefore, the decay of the minimal errors on $H(K)$ is fully determined by the decay of the minimal errors on $H(k_1)$, which corresponds to the hardest underlying univariate problem, as well as ρ, which quantifies the growth of smoothness of $H(k_j)$ as j increases. If ρ is big enough, the infinite-variate problem on $H(K)$ is no more difficult that the univariate problem on $H(k_1)$.

We conclude this section with the following remark.

Remark 2.11. We discuss the impact of the assumption (2.23) on our upper and lower bounds in the case of L^2-approximation. The lower bounds on $\operatorname{dec}(K)$, which correspond to upper error bounds, are clearly not affected by this.

Regarding upper bounds on $\operatorname{dec}(K)$, we observe that Theorem 2.8 is not impacted by this assumption; the continuous embedding $H(M_a^{\downarrow}) \subseteq H(K)$ holds regardless. While the bound (2.32) does not hold for the integration problem without assuming (2.23), it does still hold for the L^2-approximation problem, as Wasilkowski (2012, Cor. 9) still holds. Therefore, the bounds from Theorem 2.9 hold for the L^2-approximation problem regardless of whether (2.23) is fulfilled, while for the integration problem, we actually require (2.23), but the problem is trivial otherwise.

2.3.2 Multivariate Decomposition Methods

Let us now describe the algorithms that yield the lower bounds on the polynomial decay rate $\operatorname{dec}(K)$ of the n-th minimal errors for integration and L^2-approximation on $H(K)$. These kinds of algorithms were first called *changing dimension algorithms*, see Kuo et al. (2010b), and are now known as *multivariate decomposition methods (MDMs)*.

MDMs operating on $H(M_a^{\uparrow})$

At first, we consider the reproducing kernels $m_a^{\uparrow}$ and $M_a^{\uparrow}$ with a fixed $a \in D$, as introduced in Theorem 2.6. Recall

$$H(1) \cap H(m_a^{\uparrow}) = \{0\} \quad \text{and} \quad M_a^{\uparrow} = \bigotimes_{j \in \mathbb{N}} (1 + \gamma_j m_a^{\uparrow}),$$

which yields the anchored function decomposition from Remark 2.7. MDMs based on the anchored function decomposition as described in Section 2.2 were first considered in Kuo et al. (2010b); see, e.g., Wasilkowski and Woźniakowski (2011), Plaskota and Wasilkowski (2011), Wasilkowski (2012), Gnewuch (2013), Dick and Gnewuch (2014), Wasilkowski (2013), Kuo et al. (2017a) for subsequent work.

For $f \in H(M_a^\uparrow)$ and $\boldsymbol{\nu} \in \boldsymbol{U}$, denote by $f_{\boldsymbol{\nu}}$ the orthogonal decomposition of f into $H(m_{a,\boldsymbol{\nu}}^\uparrow)$. Observe that $f_{\boldsymbol{\nu}}$ only depends on the variables with indices in $\boldsymbol{\nu}$, and

$$\|f\|_{H(M_a^\uparrow)}^2 = \sum_{\boldsymbol{\nu} \in \boldsymbol{U}} \gamma_{\boldsymbol{\nu}}^{-1} \cdot \|f_{\boldsymbol{\nu}}\|_{H(m_{a,\boldsymbol{\nu}}^\uparrow)}^2. \tag{2.35}$$

In principle, we may study integration and L^2-approximation on each of the spaces $H(m_{a,\boldsymbol{\nu}}^\uparrow)$ with $\boldsymbol{\nu} \in \boldsymbol{U} \backslash \{\varnothing\}$. Formally the underlying measure is the infinite product μ, but since the elements of $H(m_{a,\boldsymbol{\nu}}^\uparrow)$ only depend on the variables with indices in $\boldsymbol{\nu}$, we are actually dealing with integration and L^2-approximation with respect to the $|\boldsymbol{\nu}|$-dimensional standard normal distribution. Cf. (1.7) for an isometric embedding from a finite-variate tensor product space into an infinite-variate one in a more general setting.

To construct an MDM, we have to choose a finite set $\mathcal{A}$ of non-empty elements of $\boldsymbol{U}$, i.e., of finite sets of variables, and for each $\boldsymbol{\nu} \in \mathcal{A}$ an algorithm $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}$ for integration or L^2-approximation of functions from $H(m_{a,\boldsymbol{\nu}}^\uparrow)$, which uses $n_{\boldsymbol{\nu}} \in \mathbb{N}$ function values of each input function $f_{\boldsymbol{\nu}} \in H(m_{a,\boldsymbol{\nu}}^\uparrow)$. We assume that the algorithms $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}$ are of the form (1.15). For L^2-approximation we assume additionally that also the approximating functions $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}(f_{\boldsymbol{\nu}})$ only depend on the variables with indices in $\boldsymbol{\nu}$. The corresponding MDM on $H(M_a^\uparrow)$ is given by

$$A(f) := f(\boldsymbol{a}) + \sum_{\boldsymbol{\nu} \in \mathcal{A}} A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}(f_{\boldsymbol{\nu}}). \tag{2.36}$$

For notational convenience we put $n_{\boldsymbol{\nu}} := 0$ for $\boldsymbol{\nu} \in \boldsymbol{U} \backslash (\mathcal{A} \cup \{\varnothing\})$ and $A_{\boldsymbol{\nu},0} := 0$. For every choice of $\mathcal{A}$ and of algorithms $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}$ for $\boldsymbol{\nu} \in \mathcal{A}$ we obtain

$$\mathrm{error}^2(A, M_a^\uparrow) \leqslant \sum_{\boldsymbol{\nu} \in \boldsymbol{U} \backslash \{\varnothing\}} \gamma_{\boldsymbol{\nu}} \cdot \mathrm{error}^2(A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}, m_{a,\boldsymbol{\nu}}^\uparrow) \tag{2.37}$$

directly from (2.35) and (2.37). This was first shown in Plaskota and Wasilkowski (2011, p. 513) for integration; the same argument applies for L^2-approximation. We add that

$$\mathrm{error}(A_{\boldsymbol{\nu},0}, m^{\uparrow}_{a,\boldsymbol{\nu}}) = \big(\mathrm{error}(0, m^{\uparrow}_{a})\big)^{|\boldsymbol{\nu}|}$$

for the error of the zero algorithm, i.e., for the operator norm of the integration functional or the L^2-embedding operator. Since $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}$ is applied to the anchored components $f_{\boldsymbol{\nu}}$ of $f \in H(M^{\uparrow}_a)$ in (2.36), we obtain from (2.20) that evaluating $f_{\boldsymbol{\nu}}$ at a point $\boldsymbol{x}_{\boldsymbol{\nu}}$ needs at most $2^{|\boldsymbol{\nu}|}$ function evaluations of f, each at nodes $\boldsymbol{x}$ with at most $|\boldsymbol{\nu}|$ entries not equal to a, which yields the bound

$$\mathrm{cost}(A) \leqslant \$(0) + \sum_{\boldsymbol{\nu}\in\mathcal{A}} n_{\boldsymbol{\nu}} \cdot 2^{|\boldsymbol{\nu}|} \cdot \$(|\boldsymbol{\nu}|) \tag{2.38}$$

on the cost of A, see, e.g., Plaskota and Wasilkowski (2011, p. 512).

We add that in order to obtain good finite-variate algorithms $A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}$ and to further gain control on $\mathrm{error}^2(A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}, m^{\uparrow}_{a,\boldsymbol{\nu}})$ in (2.37), we may use Smolyak's construction, only requiring good algorithms on the space $H(m^{\uparrow}_a)$. See Remark 1.6 for a brief explanation as well as two references.

For suitable MDMs we thus have an explicit upper bound for the worst-case error on the unit ball in $H(M^{\uparrow}_a)$ and an explicit upper bound for the cost. See Plaskota and Wasilkowski (2011) and Wasilkowski (2013) for a detailed analysis in a general setting and for the almost optimal choice of $\mathcal{A}$ and of the algorithms $A_{\boldsymbol{u},n_{\boldsymbol{u}}}$ for $\boldsymbol{u} \in \mathcal{A}$.

MDMs operating on $H(K)$

The study of MDMs operating on $H(M^{\uparrow}_a)$, along with the embedding of $H(K)$ into $H(M^{\uparrow}_a)$, see Theorem 2.6, already gives us the upper error bound obtained in Theorem 2.9. However, we wish to construct MDMs operating on $H(K)$ working only with the kernels k_1 and K, relegating the kernels $m^{\uparrow}_a$ and $M^{\uparrow}_a$ to only being relevant in the proofs and not having to construct them explicitly.

For a detailed explanation, more knowledge about the $\mathfrak{e}_\nu$ is needed, for instance, some (asymptotic) bounds for $\mathfrak{e}_\nu(x)$ as ν or x vary. Therefore, we sketch a general idea of how to proceed here. For a more detailed analysis in the case of Hermite spaces with (PG) or (EG), see Gnewuch et al. (2024, Sec.4.3.3).

We assume that $\mathrm{dec}(k_1) > 0$, which yields the existence of a sequence algorithms

A_n for integration or L^2-approximation, respectively, on the space $H(k_1)$, using at most n function evaluations and with the error bound

$$\mathrm{error}(A_n, k_1) \leqslant c \cdot n^{-\kappa}$$

for some constants $c > 0$ and $\kappa > 1$. Although we do not know $m_a^{\uparrow}$ explicitly, we still know that $H(m_a^{\uparrow})$ is continuously embedded in $H(k_1)$ and that the corresponding embedding constant is at most $c^{\uparrow}(a)$, see Theorem 2.6(i). Hence we have

$$\mathrm{error}(A_n, m_a^{\uparrow}) \leqslant c^{\uparrow}(a) \cdot \mathrm{error}(A_n, k_1),$$

and the upper bound on $\mathrm{error}(A_n, m_a^{\uparrow})$ may be used to obtain upper bounds for the errors of Smolyak algorithms $A_{\boldsymbol{u},n_{\boldsymbol{u}}}$ operating on the spaces $H(m_{a,\boldsymbol{u}}^{\uparrow})$ as seen in (1.21). Furthermore,

$$\mathrm{error}(0, m_a^{\uparrow}) \leqslant c^{\uparrow}(a)$$

for integration and

$$\mathrm{error}(0, m_a^{\uparrow}) \leqslant c^{\uparrow}(a) \cdot \max(1, \alpha_{1,1}^{-1/2})$$

for L^2-approximation. Finally, we employ Theorem 2.6.(iii) to return to the space $H(K)$: Since the embedding of $H(K)$ into $H(M_a^{\uparrow})$ by restriction is continuous with operator norm bounded by $C^{\uparrow}(a)$, we obtain

$$\mathrm{error}^2(A, K) \leqslant \left(C^{\uparrow}(a)\right)^2 \cdot \sum_{\boldsymbol{\nu} \in \boldsymbol{U} \setminus \{\varnothing\}} \gamma_{\boldsymbol{\nu}} \cdot \mathrm{error}^2(A_{\boldsymbol{\nu},n_{\boldsymbol{\nu}}}, m_{a,\boldsymbol{\nu}}^{\uparrow}) \tag{2.39}$$

from (2.37). Compared to MDMs on $H(M_a^{\uparrow})$, in the present setting it is desirable to gain some control on the constants $c^{\uparrow}(a)$ and $C^{\uparrow}(a)$ in the error bounds. In particular, we want both constants to be as small as possible through a good choice of a.

We may therefore consider the following optimization problem: For a given error tolerance $\varepsilon > 0$, determine $\mathcal{A}$ and $n_{\boldsymbol{\nu}}$ for $\boldsymbol{\nu} \in \mathcal{A}$ such that the upper bound for $\mathrm{error}(A, K)$ is at most ε and the upper bound (2.38) for $\mathrm{cost}(A)$ is as small as possible. This problem has been studied in, e.g., Plaskota and Wasilkowski (2011) and Wasilkowski (2013). Again, we also refer to Gnewuch et al. (2024, Sec. 4.3.3).

For the algorithm A given by the solution of this optimization problem, the error bound (2.39), further bounded by the error bounds for Smolyak algorithms given by (1.21), and together with the bound on the cost given by 2.38, yield the upper error bound in Theorem 2.9.

2.4 Hermite Spaces

In this section, we study one specific class of examples of spaces with increasing smoothness, namely Hermite spaces on the domain $D = \mathbb{R}$. We also consider

$$\mu_0 = N(0,1)$$

to be the standard normal distribution throughout this section. In particular, for $\nu \in \mathbb{N}_0$, the underlying basis function

$$\mathfrak{e}_\nu = h_\nu$$

is given by the Hermite polynomial of degree ν, while for the Fourier weights, we will mostly consider the cases (PG) and (EG). Interestingly, in this setting (PG) and (EG) coincide with classical notions of finite or infinite smoothness, respectively, of functions $f \in H(k_j)$.

This setting is of special interest for two reasons. Firstly, we will see that

$$\mathfrak{X} \subsetneq \mathbb{R}^{\mathbb{N}},$$

which is in contrast to other examples that were studied in Gnewuch et al. (2019). Secondly, the study of integration and L^2-approximation on Hermite spaces, particularly in the case (EG), will be instrumental to the study of these problems on spaces with Gaussian kernels in Chapter 3. In particular, we will establish a way to transform algorithms for a given problem on a Hermite space into an algorithm for the same problem on a corresponding space with a Gaussian kernel, preserving cost and, up to a constant, error.

By the general results we have already established, particularly in Section 2.3.1, we basically only need to find $\operatorname{dec}(k_1)$ or achieve good bounds on it to study $\operatorname{dec}(K)$.

Therefore the majority of this section, up to and including Section 2.4.3, is only concerned with the study of the univariate case, and for this we drop the index j in our notation. In the univariate case, we mostly survey known results.

At the end of the section, we turn to the infinite-variate case again, and will be able to exactly determine $\operatorname{dec}(K)$ in the cases (PG) and (EG).

Finite-variate Hermite spaces with different kinds of Fourier weights have been studied, e.g., in the following papers on integration and approximation problems. Fourier weights with a polynomial growth are considered in Irrgeher and Leobacher (2015), Dick et al. (2018), Kazashi et al. (2023), Dũng and Nguyen (2023), and Leobacher et al. (2023). We will survey some of the results in the univariate case in Section 2.4.3.

Fourier weights with a (sub-)exponential growth have been studied in Irrgeher and Leobacher (2015), Irrgeher et al. (2015), Irrgeher et al. (2016a), and Irrgeher et al. (2016b). We will revisit this case in Section 3.2.3 as well as Section 3.2.4, since it is closely related to the setting of RKHSs with Gaussian kernels, which we study in Chapter 3.

2.4.1 The Univariate Case

For $\nu \in \mathbb{N}_0$, we denote the Hermite polynomial of degree ν by h_ν. We give some of the basic definitions and properties found in Gnewuch et al. (2022, Sec. 3.1).

The Hermite polynomials arise by orthonormalizing the monomials in $L^2(\mu_0)$. An explicit representation is given by

$$h_\nu(x) = \sqrt{\nu!} \sum_{k=0}^{\lfloor \nu/2 \rfloor} (-1)^k \frac{x^{\nu-2k}}{2^k k! (\nu-2k)!}$$

for $\nu \in \mathbb{N}_0$ and $x \in \mathbb{R}$, see Szegö (1975, Eqn.(5.5.4)). In particular, we have

$$h_0(x) = 1, \quad h_1(x) = x \quad \text{and} \quad h_2(x) = \frac{1}{\sqrt{2}}(x^2 - 1) \tag{2.40}$$

for $x \in \mathbb{R}$. In particular, the first part of (2.8) is fulfilled. In Gnewuch et al. (2022, Lem. 3.1), several asymptotic properties for the scale of Hermite polynomials were given. Here, we only mention Cramér's inequality

$$\sup_{\nu \in \mathbb{N}_0} |h_\nu(x)| \leqslant \exp(x^2/4) \tag{2.41}$$

for all $x \in \mathbb{R}$, which will be of use later.

Let $(\alpha_\nu)_{\nu \in \mathbb{N}_0}$ be a sequence of Fourier weights fulfilling $\inf_{\nu \in \mathbb{N}_0} \alpha_\nu > 0$ and

$$\sum_{\nu \in \mathbb{N}_0} \alpha_\nu^{-1} h_\nu(x)^2 < \infty \tag{2.42}$$

for all $x \in \mathbb{R}$. Further, we assume

$$\alpha_0 = 1.$$

Then, by Proposition 2.1, we have that k, given by

$$k(x,y) = 1 + \sum_{\nu \in \mathbb{N}} \alpha_\nu^{-1} h_\nu(x) h_\nu(y) \tag{2.43}$$

for $x, y \in \mathbb{R}$, is a reproducing kernel. We call $H(k)$ a univariate Hermite space.

A sufficient condition for (2.42) is given by

$$\sum_{\nu \in \mathbb{N}} \alpha_\nu^{-1} \nu^{-1/2} < \infty, \tag{2.44}$$

which is also necessary under the additional assumption of $\boldsymbol{\alpha}$ being non-decreasing, see Gnewuch et al. (2022, Lem. 3.3) and the comment following it.

In the sequel, we consider a non-decreasing sequence $\boldsymbol{\alpha} := (\alpha_\nu)_{\nu \in \mathbb{N}}$ of Fourier weights satisfying (2.44) and the corresponding univariate Hermite kernel.

Hermite spaces behave monotonically on the main diagonal in the following sense, see Gnewuch et al. (2024, Lemma B.1).

Lemma 2.12. *For $x, y \in \mathbb{R}$ with $|x| \leqslant |y|$ we have*

$$k(x,x) \leqslant k(y,y).$$

Proof. We observe that $k(x,x) = k(-x,-x)$ holds for all $x \in \mathbb{R}$, since each of the Hermite polynomials is either an even or an odd function. Thus, it suffices to show that $x \mapsto k(x,x)$ is non-decreasing on $\mathbb{R}^+$. To this end, we use two recurrence relations of Hermite polynomials, namely

$$h'_\nu(x) = \nu^{1/2} \cdot h_{\nu-1}(x) \quad \text{and} \quad h_\nu(x) = \nu^{-1/2} \cdot (x h_{\nu-1}(x) - h'_{\nu-1}(x))$$

for $\nu \in \mathbb{N}$ and $x \in \mathbb{R}$, see Szegö (1975, Eqn. (5.5.10)). This implies

$$(h_\nu^2)'(x) = 2x h_{\nu-1}^2(x) - (h_{\nu-1}^2)'(x)$$

and thus, inductively,

$$(h_\nu^2)'(x) = 2x \cdot \sum_{\kappa=1}^{\nu} (-1)^{\nu-\kappa} \cdot h_{\kappa-1}^2(x). \tag{2.45}$$

For $n \in \mathbb{N}$ and $x \in \mathbb{R}_+$ we consider the partial sum

$$f_n(x) := \sum_{\nu=0}^{n} \alpha_\nu^{-1} \cdot h_\nu^2(x).$$

By (2.45) we have

$$\begin{aligned} f'_n(x) &= 2x \cdot \sum_{\nu=1}^{n} \alpha_\nu^{-1} \sum_{\kappa=1}^{\nu} (-1)^{\nu-\kappa} h_{\kappa-1}^2(x) \\ &= 2x \cdot \sum_{\kappa=1}^{n} h_{\kappa-1}^2(x) \sum_{\nu=0}^{n-\kappa} (-1)^\nu \alpha_{\nu+\kappa}^{-1}. \end{aligned}$$

Since $\boldsymbol{\alpha}$ is non-decreasing, the inner sum on the right-hand side is always nonnegative, and therefore f_n is non-decreasing on $\mathbb{R}^+$. Finally, $f_n(x)$ converges to $k(x,x)$ for every $x \in \mathbb{R}^+$. □

Let us briefly recall the general setting. When discussing error bounds for MDM-algorithms, constants $c^\uparrow(a)$ and $C^\uparrow(a)$ showed up, see, for instance (2.39), which were given by (2.24) and (2.29). Lemma 2.12 reveals that, in the case of the infinite-variate product of univariate Hermite spaces, these bounds are minimal for $a = 0$. Further, in the case $a = 0$ a bound for these constants based on the values $h_{2\nu}(0)$ for $\nu \in \mathbb{N}$

can be established, see Gnewuch et al. (2024, Rem. 3.3).

2.4.2 Smoothness of Functions from Univariate Hermite Spaces

We discuss some smoothness results for functions from the space $H(k)$.

Recall the cases (PG) and (EG) from the infinite-variate setting. Here, we look at a univariate counterpart, so let $r > 1$ and

$$\alpha_\nu = (\nu + 1)^r$$

for polynomial growth or $r, b > 0$ and

$$\alpha_\nu = 2^{r \cdot \nu^b}$$

for exponential growth.

We will see that polynomial growth of the Fourier weights corresponds to finite smoothness for functions from $H(k)$, with the smoothness increasing as r increases. Similarly, exponential growth of the Fourier weights corresponds to infinite smoothness, which in a sense still increases as b increases.

In the case of polynomial growth, we follow Leobacher et al. (2023, Sec. 2), where actually, even a multivariate result is obtained. We only give the univariate case here. We consider the special case $r \in \mathbb{N}$; additionally, let a weight $0 < \gamma \leqslant 1$ be given. We define the Fourier weights $\widetilde{\alpha}_\nu$ via their inverses

$$\widetilde{\alpha}_\nu^{-1} = \begin{cases} \gamma/\nu!, \text{ if } 1 \leqslant \nu \leqslant e \\ \gamma \cdot ((\nu - r)!/\nu!), \text{ if } \nu > r, \end{cases}$$

with $\widetilde{\alpha}_0 = 1$, and let $\widetilde{k}$ be the corresponding Hermite kernel. Firstly, we have

$$\gamma/\nu^r \leqslant \widetilde{\alpha}_\nu^{-1} \leqslant \gamma \cdot (r/\nu)^r$$

for all $\nu \in \mathbb{N}$, see Leobacher et al. (2023, Lem. 2), so that the sequences $\boldsymbol{\alpha}$ and $\widetilde{\boldsymbol{\alpha}}$ are equivalent. This implies $H(k) = H(\widetilde{k})$ as vector spaces, with equivalent norms.

Secondly, the space $H(\widetilde{k})$ is the Sobolev space of all absolutely continuous functions that are at least r times weakly differentiable and whose weak derivatives up to order r are square-integrable with respect to μ_0; here γ acts as a weight for the norm. Clearly, the second result is then also true for $H(k)$, up to equivalence of norms.

In the case of exponential growth, the elements of $H(k)$ are real analytic if $b \geqslant 1/2$, and they belong to the Gevrey class of index $(2b)^{-1}$ if $0 < b < 1/2$. See Gnewuch et al. (2022, Exmp. 3.6, Lem. 3.7) for further details and references. For $s \geqslant 1$, the Gevrey class with index s is defined as the space of infinitely often differentiable functions f fulfilling the following property: For every compact $K \subseteq \mathbb{R}$, there exists a $C > 0$ such that

$$\sup_{x \in K} |f^{(n)}(x)| \leqslant C^{n+1} \cdot n!^s$$

for all $n \in \mathbb{N}_0$. Observe that this inequality holding in the case $s = 1$ is equivalent to f being analytic.

2.4.3 Integration and L^2-Approximation of Functions of a Single Variable

We now study univariate integration and L^2-approximation on the space $H(k)$. Regarding the sequence of Fourier weights $\boldsymbol{\alpha}$, we only consider the case

$$\operatorname{decay}(\boldsymbol{\alpha}^{-1}) \geqslant 1. \tag{2.46}$$

Of course, for k to be well-defined it would be sufficient to assume $\operatorname{decay}(\boldsymbol{\alpha}) \geqslant 1/2$, see (2.44). However, for our ultimate goal of studying infinite-variate problems in the cases (PG) and (EG), it suffices to consider the univariate case assuming (2.46). Note that to our knowledge, in the case $\operatorname{decay}(\boldsymbol{\alpha}^{-1}) < 1$, no upper error bounds have been established so far.

As a reminder, for two sequences of Fourier weights $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ with $\alpha_{\nu,1} < \alpha_{\nu,2}$ for $\nu \in \mathbb{N}$, we have

$$H(k_{\boldsymbol{\alpha}_1}) \subsetneq H(k_{\boldsymbol{\alpha}_2}) \tag{2.47}$$

with a continuous identical embedding of norm one, see (2.12).

For integration, we proceed as follows: First, we study the particular case $k^{[r]} := k$ with

$$\alpha_\nu := (\nu + 1)^r, \tag{2.48}$$

with $r \geqslant 1$. In the even more special case $r \in \mathbb{N}$, we survey results from Dick et al. (2018) and Dũng and Nguyen (2023) to achieve matching upper and lower bounds on $e_n(k^{[r]})$ based on smoothness properties on the functions in $H(k^{[r]})$. We generalize these results to the case of arbitrary $r > 1$ via the interpolation of Hilbert spaces, cf. Remark 2.3. Then, results for $H(k)$ under the assumption (2.46) follow from the results for $H(k^{[r]})$ with the help of (2.47).

For L^2-approximation, we utilize a result from Dolbeault et al. (2023) which essentially says that algorithms based on general linear information are not more powerful than algorithms based on function evaluation.

Integration

For Fourier weights according to (2.48) with $r \in \mathbb{N}$ even the asymptotic behavior of the n-th minimal errors for integration is known: There exist $c_1, c_2 > 0$ such that

$$c_1 \cdot n^{-r} \leqslant e_n(k^{[r]}) \leqslant c_2 \cdot n^{-r} \tag{2.49}$$

holds for all $n \in \mathbb{N}$, see Dick et al. (2018) and Dũng and Nguyen (2023). We remark that both papers actually considered the d-variate case of the d-fold tensor product of the kernel $k^{[r]}$ and the d-dimensional standard normal distribution for some $d \in \mathbb{N}$, which results in the additional factor $(\ln(n))^{(d-1)/2}$ in (2.49). The corresponding lower bound is due to Dick et al. (2018, Thm. 1), and the matching upper bound is due to Dũng and Nguyen (2023, Thm. 2.3); upper bounds involving a further logarithmic factor have already been established in Dick et al. (2018, Cor. 1) for $d \in \mathbb{N}$ and in Kazashi et al. (2023, Thm. 4.5) for $d = 1$.

We briefly sketch how asymptotically optimal quadrature formulas are constructed in Dũng and Nguyen (2023). Consider the interval $I := [-1/2, 1/2]$ and the Lebesgue measure λ on I. Denote by W^r the Sobolev space of absolutely continuous functions on I that are r times weakly differentiable, and whose weak derivatives up to order r

are square-integrable with respect to λ. We equip this space with its usual norm. Our goal is to use asymptotically optimal algorithms on W^r to construct asymptotically optimal algorithms on $H(k^{[r]})$.

To this end, denote by φ the density function of the standard normal distribution, given by

$$\varphi(x) := (2\pi)^{-1/2} \cdot \exp(-x^2/2)$$

for $x \in \mathbb{R}$. Further, we define the following integer shifts: For $f\colon \mathbb{R} \to \mathbb{R}$ and $\ell \in \mathbb{Z}$ we use $f(\cdot + \ell)$ to denote the function $x \mapsto f(x + \ell)$ on the domain I.

The idea for integration on $H(k^{[r]})$ now is to shift I to $I + \ell$ for only finitely many ℓ, and consider integration on the union of these shifted intervals, instead of on $\mathbb{R}$. More precisely, for $f \in H(k^{[r]})$, we apply asymptotically optimal algorithms for integration on W^r to integrands of the form $(f\varphi)(\cdot + \ell)$.

See Dũng and Nguyen (2023, Eqn. 2.11) for the following key result.

Lemma 2.13. *For every $r \in \mathbb{N}$ and every $\delta \in]0, 1/4[$ there exists $c > 0$ such that*

$$\|(f\varphi)(\cdot + \ell)\|_{W^r} \leqslant c \cdot \exp(-\delta\ell^2) \cdot \|f\|_{H(k^{[r]})}$$

for every $f \in H(k^{[r]})$ and every $\ell \in \mathbb{Z}$.

It is well known that the n-th minimal errors for integration on W^r with respect to λ are of the order n^{-r}. In the sequel, for every $m \in \mathbb{N}$, let A'_m be a quadrature formula based on m function evaluations fulfilling

$$\left| \int_I f d\lambda - A'_m(f) \right| \leqslant c \cdot m^{-r} \cdot \|f\|_{W^r}$$

for some $c > 0$ independent of m and for every $f \in W^r$. We mention that here one can for instance use quadrature formulas based on polynomial interpolation of degree $r - 1$ and equidistant nodes in I.

For every $L \in \mathbb{N}$ and every sequence $\boldsymbol{m} := (m_\ell)_{|\ell|<L}$ in $\mathbb{N}$ we obtain a quadrature formula

$$A_{L,\boldsymbol{m}}(f) := \sum_{|\ell|<L} A'_{m_\ell}((f \cdot \varphi)(\cdot + \ell))$$

on $H(k^{[r]})$.

See Dũng and Nguyen (2023, p. 8) for the following fact, which immediately follows from Lemma 2.13 and the continuity of $f \mapsto \int_I f d\lambda$ on W^r.

Lemma 2.14. *For every $r \in \mathbb{N}$ and every $\delta \in \,]0, 1/4[$ there exists $c > 0$ such that*

$$e(A_{L,\boldsymbol{m}}, k^{[r]}) \leqslant c \cdot \left(\sum_{|\ell|<L} m_\ell^{-r} \cdot \exp(-\delta \ell^2) + \sum_{|\ell| \geqslant L} \exp(-\delta \ell^2) \right)$$

for all $L \in \mathbb{N}$ and $\boldsymbol{m}$ as before.

In the sequel, we fix some $\delta := \,]0, 1/4[$. For $n \geqslant 2$ we choose

$$L_n := \left\lceil \left(\frac{r}{\delta} \cdot \ln(n) \right)^{1/2} \right\rceil$$

and

$$m_{\ell,n} := \left\lceil n \cdot \exp\left(-\frac{\delta}{2r} \cdot \ell^2 \right) \right\rceil,$$

and we use $A_n := A_{L_n, \boldsymbol{m}_n}$ to denote the corresponding quadrature formula on $H(k^{[r]})$.

See Dũng and Nguyen (2023, Thm. 2.1) for the following result, which yields the upper bound in (2.49) and is derived from Lemma 2.14 in a straightforward way.

Theorem 2.15. *For the integration problem on $H(k^{[r]})$ the following holds. For every $r \in \mathbb{N}$ there exists $c > 0$ such that the worst-case error and the number of nodes of A_n satisfy*

$$e(A_n, k^{[r]}) \leqslant c \cdot n^{-r}$$

and

$$\sum_{|\ell|<L_n} m_{\ell,n} \leqslant c \cdot n,$$

respectively, for every $n \in \mathbb{N}$.

Remark 2.16. Let $\overline{r} \in \mathbb{N}$. Since L_n and $m_{\ell,n}$ are non-decreasing functions of r, we may easily construct a sequence of quadrature formulas that is asymptotically optimal simultaneously for all $r \in \{1, \ldots, \overline{r}\}$.

Now, we turn to the case of $r > 1$ without requiring $r \in \mathbb{N}$. Using interpolation of Hilbert spaces and Theorem 2.15, we obtain the following result.

Theorem 2.17. *For the integration problem on $H(k)$ the following holds.*

(*i*) *If* $\mathrm{decay}(\boldsymbol{\alpha}^{-1}) > 1$ *then*

$$\mathrm{dec}(k) \geqslant \mathrm{decay}(\boldsymbol{\alpha}^{-1}).$$

(*ii*) *For Fourier weights of the form* (2.48) *we have*

$$\mathrm{dec}(k^{[r]}) = \mathrm{decay}(\boldsymbol{\alpha}^{-1}) = r$$

if $r \geqslant 1$ *and*

$$\mathrm{dec}(k^{[r]}) \leqslant \mathrm{decay}(\boldsymbol{\alpha}^{-1}) = r$$

if $1/2 < r < 1$.

Proof. First, we consider Fourier weights according to (2.48), where we obtain

$$\mathrm{dec}(k^{[r]}) = r$$

for $r \in \mathbb{N}$ immediately from (2.49).

Now, let $r \geqslant 1$, and let $\underline{r}, \overline{r} \in \mathbb{N}$ be given such that $\underline{r} \leqslant r \leqslant \overline{r}$. Then there exists a uniquely determined $\theta \in]0, 1[$ such that

$$r = (1 - \theta)\underline{r} + \theta\overline{r},$$

cf. (2.14). Then, as described in Remark 2.3, we obtain

$$(\nu + 1)^r = ((\nu + 1)^{\underline{r}})^{1-\theta} \cdot \left((\nu + 1)^{\overline{r}}\right)^{\theta},$$

for the Fourier weights of the spaces $H(k^{[r]})$, $H(k^{[\underline{r}]})$, and $H(k^{[\overline{r}]})$, as well as

$$\|f\|_{k^{[r]}} \leqslant (\|f\|_{k^{[\underline{r}]}})^{1-\theta} \cdot (\|f\|_{k^{[\overline{r}]}})^{\theta}$$

for $f \in H(k^{[\overline{r}]})$, which directly implies

$$\mathrm{error}(A, k^{[r]}) \leqslant \mathrm{error}(A, k^{[\underline{r}]})^{1-\theta} \cdot \mathrm{error}(A, k^{[\overline{r}]})^{\theta}$$

for every quadrature formula A.

Using Remark 2.16, we obtain

$$e_n(k^{[r]}) \leqslant e_n(k^{[\underline{r}]})^{1-\theta} \cdot e_n(k^{[\overline{r}]})^{\theta}.$$

This further implies

$$\mathrm{dec}(k^{[r]}) \geqslant (1-\theta)\,\mathrm{dec}(k^{[\underline{r}]}) + \theta\,\mathrm{dec}(k^{[\overline{r}]}) = r. \tag{2.50}$$

On the other hand, using (2.50) in the equivalent form

$$\mathrm{dec}(k^{[\underline{r}]}) \leqslant \frac{1}{1-\theta} \cdot \big(\mathrm{dec}(k^{[r]}) - \theta\,\mathrm{dec}(k^{[\overline{r}]})\big)$$

with $r, \overline{r} \in \mathbb{N}$ and $\underline{r} \geqslant 1$ shows that

$$\mathrm{dec}(k^{[\underline{r}]}) \leqslant \underline{r}.$$

Finally, we consider any non-decreasing sequence $\boldsymbol{\alpha}$ of Fourier weights fulfilling (2.46) and the corresponding kernel k. For every $1 \leqslant r < \mathrm{decay}(\boldsymbol{\alpha}^{-1})$ the space $H(k)$ is continuously embedded into $H(k^{[r]})$, and therefore

$$\mathrm{dec}(k) \geqslant \mathrm{dec}(k^{[r]}) = r$$

which implies $\mathrm{dec}(k) \geqslant \mathrm{decay}(\boldsymbol{\alpha}^{-1})$. □

L^2-Approximation

For Fourier weights of the form (2.48) the asymptotic behavior of the n-th minimal errors for L^2-approximation is even known for all $r > 1$, where we have

$$e_n(k^{[r]}) \asymp n^{-r/2}, \tag{2.51}$$

see Dũng and Nguyen (2023, Thm. 3.5).

As for integration, the multivariate case with the d-fold tensor product of the kernel $k^{[r]}$ and the d-dimensional standard normal distribution is studied for L^2-approximation, too, in Dũng and Nguyen (2023), which results in the additional factor $(\ln(n))^{(d-1)\cdot r/2}$ in (2.51).

In the further analysis of the approximation problem we use general results on upper bounds for the minimal worst-case error for L^2-approximation on separable RKHSs using function values due to Dolbeault et al. (2023). The results are also employed in Dũng and Nguyen (2023).

Before stating the corresponding theorem, we provide an auxiliary lemma concerning the decay of $\boldsymbol{\beta} := (\beta_n)_{n\in\mathbb{N}}$ given by

$$\beta_n := \left(\frac{1}{n}\sum_{\nu\geqslant n}\alpha_\nu^{-1}\right)^{1/2}$$

again under the assumption

$$\sum_{\nu\in\mathbb{N}}\alpha_\nu^{-1} < \infty. \tag{2.52}$$

Lemma 2.18. *If* (2.52) *is satisfied then*

$$\operatorname{decay}(\boldsymbol{\beta}) = \operatorname{decay}(\boldsymbol{\alpha}^{-1/2}).$$

Proof. Due to (2.52), we have $\operatorname{decay}(\boldsymbol{\alpha}^{-1}) \geqslant 1$.

The weak discrete Stechkin inequality asserts that for $q > 1$ there exist constants $c, C > 0$ depending only on q such that for all non-increasing sequences $(\omega_\nu)_{\nu\in\mathbb{N}}$ of positive real numbers we have

$$c^{-1}\sup_{n\geqslant 1} n\left(\frac{1}{n}\sum_{\nu\geqslant n}\omega_\nu^q\right)^{1/q} \leqslant \sup_{n\geqslant 1} n\omega_n \leqslant C\sup_{n\geqslant 1} n\left(\frac{1}{n}\sum_{\nu\geqslant n}\omega_\nu^q\right)^{1/q}.$$

For this and related inequalities including optimal constants, see Jahn and Ullrich (2021). Applying this two-sided estimate with $q := 2\tau$ and $\omega_\nu := \alpha_\nu^{-1/(2\tau)}$ shows that $\operatorname{decay}(\boldsymbol{\alpha}^{-1/2}) > \tau$ if and only if $\operatorname{decay}(\boldsymbol{\beta}) > \tau$, provided that $\tau > 1/2$. Hence $\operatorname{decay}(\boldsymbol{\beta}) = \operatorname{decay}(\boldsymbol{\alpha}^{-1/2})$ follows in the case $\operatorname{decay}(\boldsymbol{\alpha}^{-1}) > 1$.

In the borderline case $\operatorname{decay}(\boldsymbol{\alpha}^{-1}) = 1$ we use

$$\beta_n^2 \leqslant \frac{1}{n} \sum_{\nu \in \mathbb{N}} \alpha_\nu^{-1}$$

and (2.52) to conclude that $\operatorname{decay}(\boldsymbol{\beta}) \geqslant 1/2 = \operatorname{decay}(\boldsymbol{\alpha}^{-1/2})$. As observed above, $\operatorname{decay}(\boldsymbol{\beta}) > 1/2$ would imply $\operatorname{decay}(\boldsymbol{\alpha}^{-1/2}) > 1/2$. Hence we have $\operatorname{decay}(\boldsymbol{\beta}) = 1/2 = \operatorname{decay}(\boldsymbol{\alpha}^{-1/2})$. □

Theorem 2.19. *For L^2-approximation on $H(k)$ the following holds. If* (2.52) *is satisfied then*

$$\operatorname{dec}(k) = \operatorname{decay}(\boldsymbol{\alpha}^{-1/2}).$$

Proof. The sequence of singular values of the embedding of $H(k)$ into $L^2(\mu_0)$ is $\boldsymbol{\alpha}^{-1/2}$. Since the singular values, as the minimal worst-case errors of L^2-approximation using general linear information, are lower bounds for the minimal worst-case errors of L^2-approximation using function values, more precisely, since

$$e_n(k) \geqslant \alpha_n^{-1/2}$$

for all $n \in \mathbb{N}$, the inequality $\operatorname{dec}(k) \leqslant \operatorname{decay}(\boldsymbol{\alpha}^{-1/2})$ follows.

For the reverse inequality, we use Dolbeault et al. (2023, Thm. 1), which shows that there exists a universal constant $c \in \mathbb{N}$ such that

$$e_{cn}(k) \leqslant \beta_n$$

for every separable RKHS $H(k)$ with square-summable singular values α_ν of its identical embedding into any L^2-space. In the present case the square-summability is guaranteed by (2.52). Since the n-th minimal errors $e_n(k)$ form a non-increasing sequence, the decay $\operatorname{dec}(k)$ of $(e_n(k))_{n \in \mathbb{N}_0}$ is the same as the decay of the subsequence $(e_{cn}(k))_{n \in \mathbb{N}_0}$. We conclude that $\operatorname{dec}(k) \geqslant \operatorname{decay}(\boldsymbol{\beta})$, and it remains to observe Lemma 2.18. □

Remark 2.20. The approach yielding the error bound from Dolbeault et al. (2023, Thm. 1) is based on a least square estimator using independent random sample points drawn with respect to a suitable density. Thus, the proof is non-constructive, as it only ensures the existence of a good deterministic algorithm using function values. Of course, it would be desirable to have explicit constructions for sample points achieving this univariate error decay.

2.4.4 Functions of Infinitely Many Variables

We now return to the infinite-variate case. Let μ be the infinite product of the univariate standard normal distribution μ_0 with itself. For $\nu, j \in \mathbb{N}$, let Fourier weights $\alpha_{\nu,j}$ be given fulfilling (A1)–(A3). In particular, the univariate Hermite kernels k_j given by

$$k_j(x,y) := 1 + \sum_{\nu\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot h_\nu(x) \cdot h_\nu(y)$$

for $x, y \in \mathbb{R}$, are well-defined, and so is the infinite-variate kernel K given by

$$K(\boldsymbol{x},\boldsymbol{y}) := \prod_{j\in\mathbb{N}} k_j(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}$, on the maximal domain

$$\mathfrak{X} := \left\{\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}} \colon \sum_{\nu,j\in\mathbb{N}} \alpha_{\nu,j}^{-1} \cdot h_\nu^2(x_j) < \infty\right\}.$$

We also call $H(K)$ an (infinite-variate) Hermite space.

Of course, we have $H(K) \subseteq L^2(\mu)$ with a compact identical embedding and $\mu(\mathfrak{X}) = 1$, which follows from Theorem 2.2 but in this specific example was first established in Gnewuch et al. (2022, Lem. 3.8, Prop. 3.10, and Lem. 3.12). In Gnewuch et al. (2022, Prop. 3.10), it was also established that $\mathfrak{X} \subsetneq \mathbb{R}^{\mathbb{N}}$, which follows immediately from the fact that that at least one of the Hermite polynomials is unbounded (of course, this is true for all except h_0); further it was established that

$$\ell_\infty(\mathbb{N}) \subsetneq \mathfrak{X} \tag{2.53}$$

using Cramér's inequality (2.41).

Remark 2.21. In Gnewuch et al. (2022, Prop. 3.19), it was established that in the case (EG) with $b_1 \geqslant 1$, we have

$$\mathfrak{X} = \Big\{ \boldsymbol{x} \in \mathbb{R}^{\mathbb{N}} \colon \sum_{j \in \mathbb{N}} 2^{-r_j} \cdot x_j^2 < \infty \Big\}. \tag{2.54}$$

In the case (PG), no such simple characterization of $\mathfrak{X}$ is known. However, we are able to show the inclusion

$$\underline{\mathfrak{X}} := \{ \boldsymbol{x} \in \mathbb{R}^{\mathbb{N}} \colon \text{there exists } c > 0 \text{ with } |x_j| \leqslant c \cdot \ln(j)^{1/2} \text{ for all } j \in \mathbb{N} \} \subseteq \mathfrak{X}.$$

This is of interest, since we actually have $\mu(\underline{\mathfrak{X}}) = 1$, and so we can characterize almost all elements of $\mathfrak{X}$. To show $\mu(\underline{\mathfrak{X}}) = 1$, one can apply the Borel-Cantelli lemma to the sets

$$A_n = [c \cdot \ln(n)^{1/2}, \infty[$$

for $n \geqslant 3$ and some $c > \sqrt{2}$, bearing in mind that for any $x \geqslant 0$, the probability $\mu_0([x, \infty[)$ is bounded by $(1/x) \cdot \exp(-x^2/2)$.

To actually show $\underline{\mathfrak{X}} \subseteq \mathfrak{X}$, we proceed as follows. We remind ourselves of the definition of the case (PG), in particular

$$\alpha_{\nu,j} = (\nu + 1)^{r_j}$$

with $\inf_{j \in \mathbb{N}} r_j = r_1 > 1$. Let $\boldsymbol{x} \in \underline{\mathfrak{X}}$, and let $c > 0$ such that $|x_j| \leqslant c \cdot \ln(j)^{1/2}$.

First, fix $\nu \in \mathbb{N}$. Since h_ν is a polynomial of degree ν, there exists some $B > 0$ and some $j_0 \in \mathbb{N}$, such that for all $j > j_0$ we have

$$\alpha_{\nu,j}^{-1} \cdot h_\nu(x_j)^2 \leqslant B \cdot \alpha_{\nu,j}^{-1} \cdot x_j^{2\nu} \leqslant B \cdot (\nu + 1)^{-r_1} \cdot \ln(j)^\nu \cdot c^{2\nu},$$

which implies the convergence

$$\sum_{j \in \mathbb{N}} \alpha_{\nu,j}^{-1} \cdot h_\nu(x_j)^2 < \infty. \tag{2.55}$$

Now, for fixed $\nu_0, j \in \mathbb{N}$, utilizing (2.41) and (2.18), we have

$$\sum_{\nu \geqslant \nu_0} \alpha_{\nu,j}^{-1} \cdot h_\nu(x_j)^2 \leqslant \exp(x_j^2/4)^2 \cdot \nu_0^{-r_1}(1 + \nu_0/(r_1 - 1)) = j^{c^2/2 - r_1 \cdot \nu_0}(1 + \nu_0/(r_1 - 1)),$$

and the exponent of j is smaller than -1 for ν_0 big enough. If this is indeed fulfilled, we obtain

$$\sum_{j \in \mathbb{N}} \sum_{\nu \geqslant \nu_0} \alpha_{\nu,j}^{-1} \cdot h_\nu(x_j)^2 < \infty,$$

which, together with (2.55) for every fixed ν, implies $\boldsymbol{x} \in \mathfrak{X}$.

2.4.5 Integration and L^2-Approximation of Functions of Infinitely Many Variables

Closing out this chapter, we give results for integration and L^2-approximation on an infinite-variate Hermite space $H(K)$ in the cases (PG) and (EG). Theorem 2.9, and in the cases (PG) and (EG) more specifically (2.33), combined with the results from Section 2.4.3 for the univariate case, yields the following main result, cf. Gnewuch et al. (2024, Cor.4.10). As a reminder, we put

$$\rho := \liminf_{j \to \infty} \frac{r_j \cdot \ln(2)}{\ln(j)} \geqslant 1,$$

which quantifies the asymptotic behavior of $\alpha_{1,j}$ as j tends to ∞.

Corollary 2.22. *In the case (PG) we have*

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot \min(2r_1, \rho - 1)$$

for the integration problem and

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot \min(r_1, \rho - 1)$$

for the L^2-approximation problem. In the case (EG) we have

$$\operatorname{dec}(K) = \tfrac{1}{2} \cdot (\rho - 1)$$

for the integration problem and the L^2-approximation problem.

Proof. As already seen in (2.34), we have

$$\operatorname{dec}(K) = \min\left(\operatorname{dec}(k_1), \frac{\rho - 1}{2}\right) \tag{2.56}$$

for integration and L^2-approximation in both cases (PG) and (EG).

The decay of $(\alpha_{\nu,1}^{-1})_{\nu\in\mathbb{N}}$ is equal to r_1 in the case (PG) and equal to ∞ in the case (EG). Using Theorems 2.17 and 2.19 we obtain $\operatorname{dec}(k_1) = r_1$ for integration and $\operatorname{dec}(k_1) = r_1/2$ for L^2-approximation in the case (PG), while $\operatorname{dec}(k_1) = \infty$ for both problems in the case (EG). □

Remark 2.23. Corollary 2.22 reveals that r_1, which is the minimal smoothness among all Hermite spaces $H(k_j)$ of univariate functions, and ρ, which concerns the growth of the smoothness as $j \to \infty$, determine the decay of the minimal errors on Hermite spaces $H(K)$ with Fourier weights of a polynomial or (sub-)exponential growth. On the one hand, if ρ is sufficiently large, then these minimal errors decay as fast as the minimal errors in the univariate case on the space $H(k_1)$; for (PG) this means $\rho \geqslant 2r_1 + 1$ or $\rho \geqslant r_1 + 1$, while we need $\rho = \infty$ in the case (EG). On the other hand, we have $\operatorname{dec}(K) = 0$ if $\rho = 1$.

Chapter 3

Spaces with Gaussian Kernels

In this chapter, we study integration and L^2-approximation on RKHSs with Gaussian kernels. The underlying probability measure is the d-fold product measure of the standard normal distribution μ_0, where $d \in \mathbb{N} \cup \{\infty\}$. The univariate Gaussian kernel with shape parameter $\sigma > 0$ is given by

$$\ell_\sigma(x, y) := \exp\left(-\sigma^2 \cdot (x-y)^2\right)$$

for $x, y \in \mathbb{R}$. Based on a finite or infinite sequence $\boldsymbol{\sigma}$ of positive shape parameters, the d-variate Gaussian kernel with those parameters is given as

$$L_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{d} \ell_{\sigma_j}(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ in the finite-variate case and $\boldsymbol{x}, \boldsymbol{y}$ being in an appropriately chosen domain $\mathfrak{X} \subsetneq \mathbb{R}^{\mathbb{N}}$ in the infinite-variate case.

This chapter is based on Gnewuch et al. (2026). We study the infinite-variate case $d = \infty$, establishing matching upper and lower bounds on the decay rate $\operatorname{dec}(L_{\boldsymbol{\sigma}})$ of the n-th minimal errors of integration or L^2-approximation on $L_{\boldsymbol{\sigma}}$.

In the finite-variate case $d \in \mathbb{N}$, the integration problem on $H(L_{\boldsymbol{\sigma}})$ has been studied before in Kuo and Woźniakowski (2012), Kuo et al. (2017b), Karvonen and Särkkä (2019), and Karvonen et al. (2021), while the L^2-approximation problem has been studied in Fasshauer et al. (2012), see also Sloan and Woźniakowski (2018).

In the case of integration, we are able to improve on some of the known results, particularly regarding tractability.

Our contributions are based on the fact that there is a close relation between spaces with Gaussian kernels and Hermite spaces with Fourier weights of exponential growth, the latter of which we introduced in Section 2.4. In particular, if the Fourier weights of the Hermite kernel and the shape parameters of the Gaussian kernel are suitably related, there exists an isometric isomorphism Q on $L^2(\mu)$ that can be restricted to an isometric isomorphism between the two RKHSs. The isomorphism Q was established in Gnewuch et al. (2022).

Making use of this isomorphism, we establish a way to construct algorithms on the space with a Gaussian kernel from algorithms on the Hermite space and vice versa, preserving error (up to a constant) and cost.

This has the consequence that any error bound that is already established for one of the two spaces, can also be applied to the other. In this context, we will discuss some of the results of the articles mentioned above, as well as corresponding results for Hermite spaces with infinite smoothness, which have been studied in Irrgeher and Leobacher (2015), Irrgeher et al. (2015), Irrgeher et al. (2016a), and Irrgeher et al. (2016b).

We proceed as follows. In Section 3.1, we introduce spaces with Gaussian kernels and study some basic properties, including their domains. We also introduce the aforementioned isometric isomorphism here. Section 3.2 is concerned with the integration problem. We establish at first a way to transfer algorithms between spaces with Gaussian kernels and Hermite spaces. Based on this, we first study the finite-variate case and then the infinite-variate case. Section 3.3 is concerned with the L^2-approximation problem. Here, too we first establish a way to transfer algorithms, which works different than in the case of integration. Based on this, we study infinite-variate L^2-approximation, since the finite-variate case is largely solved.

This chapter follows closely Gnewuch et al. (2026), with the presentation fitted to the rest of the thesis and a few details added; we will cite the main results more explicitly. We also put the discussion of $L_{\boldsymbol{\sigma}}$ in the infinite-variate case more clearly into context of Chapter 1.

Throughout this chapter, we often consider the finite-variate and the infinite-variate case simultaneously. As notational convention, let $d \in \mathbb{N} \cup \{\infty\}$, and let $J = \{1, \ldots, d\}$ in the case $d \in \mathbb{N}$ and $J = \mathbb{N}$ in the case $d = \infty$.

3.1 The Function Space Setting

We study spaces with Gaussian kernels and their relation to Hermite spaces of infinite smoothness. In this section, we mainly present known results from Steinwart et al. (2006), see also Steinwart and Christmann (2008) and Minh (2010), and from Gnewuch et al. (2022), most importantly a particular one-to-one correspondence between both types of RKHSs.

3.1.1 Univariate Gaussian Kernels

The univariate Gaussian kernel ℓ_σ with shape parameter $\sigma > 0$ is defined by

$$\ell_\sigma(x, y) := \exp\left(-\sigma^2 \cdot (x - y)^2\right)$$

for $x, y \in \mathbb{R}$. Despite a different use in stochastic analysis, the Hilbert space $H(\ell_\sigma)$ will be called a Gaussian space throughout this chapter.

For the analysis of Gaussian spaces we refer to Steinwart et al. (2006), see also Steinwart and Christmann (2008, Sec. 4.4). In particular, each function $f \in H(\ell_\sigma)$ is the real part of an entire function g restricted to the real line, where g belongs to the complex reproducing kernel Hilbert space with kernel ℓ_σ extended to $\mathbb{C}$ in the obvious way. Since

$$\begin{aligned} |f(x) - f(y)|^2 &= \langle f, \ell_\sigma(\cdot, x) - \ell_\sigma(\cdot, y)\rangle^2_{H(\ell_\sigma)} \\ &\leqslant \|f\|^2_{H(\ell_\sigma)} \cdot \|\ell_\sigma(\cdot, x) - \ell_\sigma(\cdot, y)\|^2_{H(\ell_\sigma)} \\ &= 2\,\|f\|^2_{H(\ell_\sigma)} \cdot (1 - \ell_\sigma(x, y)) \end{aligned}$$

for all $x, y \in \mathbb{R}$, the functions from $H(\ell_\sigma)$ are bounded. No Gaussian space contains any non-zero polynomial, and obviously ℓ_σ is translation-invariant, i.e., $\ell_\sigma(x, y) = \ell_\sigma(x - y, 0)$. Moreover, we have the following monotonicity property of Gaussian spaces:

$$\sigma_1 < \sigma_2 \Rightarrow H(\ell_{\sigma_1}) \subsetneq H(\ell_{\sigma_2}) \tag{3.1}$$

with a non-compact continuous identical embedding of norm $\sqrt{\sigma_2/\sigma_1}$.

3.1.2 Tensor Products of Gaussian Kernels

We study the tensor product kernel of univariate Gaussian kernels, in particular we establish the appropriate domain in the infinite-variate case. This will turn out to be a weighted ℓ^2-space, which motivates us to establish the following result.

For any sequence $\boldsymbol{\omega} := (\omega_j)_{j\in J}$ with $\omega_j \geqslant 0$ for every $j \in J$ we use $\ell^2(\boldsymbol{\omega})$ to denote the corresponding weighted ℓ^2-space, i.e.,

$$\ell^2(\boldsymbol{\omega}) := \left\{ (x_j)_{j\in J} \in \mathbb{R}^J : \sum_{j\in J} \omega_j \cdot x_j^2 < \infty \right\}. \tag{3.2}$$

In addition to the weighted ℓ^2-space $\ell^2(\boldsymbol{\omega})$, we also consider the space ℓ^∞ of all bounded sequences in $\mathbb{R}$. In the following lemma, we just consider the case $d = \infty$.

Lemma 3.1. *For every positive sequence $\boldsymbol{\omega}$ with $\sum_{j\in\mathbb{N}} \omega_j < \infty$ we have*

$$\ell^\infty \subsetneq \ell^2(\boldsymbol{\omega}) \subsetneq \mathbb{R}^{\mathbb{N}} \qquad \textit{and} \qquad \mu(\ell^2(\boldsymbol{\omega})) = 1.$$

Proof. The first statement obviously holds true, and the second statement follows from

$$\int_{\mathbb{R}^{\mathbb{N}}} \sum_{j\in\mathbb{N}} \omega_j \cdot x_j^2 d\mu(\boldsymbol{x}) = \sum_{j\in\mathbb{N}} \omega_j \cdot \int_{\mathbb{R}} x_j^2 d\mu_0(x_j) = \sum_{j\in\mathbb{N}} \omega_j < \infty,$$

which implies that $\sum_{j\in\mathbb{N}} \omega_j \cdot x_j^2 < \infty$ holds for μ-almost every $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$. □

Let $\boldsymbol{\sigma} > 0$ be a sequence of shape parameters. Throughout this chapter, we will mostly consider the case

$$\sum_{j\in J} \sigma_j^2 < \infty \tag{3.3}$$

of square-summable shape parameters.

In contrast to other settings we considered in this thesis, here it is ensured that

$$\widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j \in J} \ell_{\sigma_j}(x_j, y_j) = \exp(-\sum_{j \in J} \sigma_j^2 \cdot (x_j - y_j)^2) \tag{3.4}$$

is well-defined for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^J$, with the convention that $\exp(-\infty) := 0$. Nevertheless, it makes sense to restrict the domain to a subset $\mathfrak{X} \subseteq \mathbb{R}^J$ following the general tensor product setting, as described in Chapter 1. We discuss this in the following, and it will turn out $\mathfrak{X} = \ell^2(\boldsymbol{\sigma})$ is the correct choice. Of course, for $d < \infty$ the assumption (3.3) is trivially satisfied and $\mathbb{R}^d = \ell^2(\boldsymbol{\sigma}^2)$.

We establish some basic results regarding $\widetilde{L}_{\boldsymbol{\sigma}}$ and $\ell^2(\boldsymbol{\sigma})$.

Lemma 3.2. *(i) For all sequences $\boldsymbol{\sigma}$ fulfilling* (3.3)*, we have $\mu(\ell^2(\boldsymbol{\sigma}^2)) = 1$.*

(ii) For all sequences $\boldsymbol{\sigma}$, we have $H(\widetilde{L}_{\boldsymbol{\sigma}}) \subseteq L^2(\mu)$ with a compact identical embedding.

(iii) For all sequences $\boldsymbol{\sigma}$, we have $\widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{y}, \boldsymbol{z}) = 0$ and $h(\boldsymbol{z}) = 0$ for $\boldsymbol{y} \in \ell^2(\boldsymbol{\sigma}^2)$ and $\boldsymbol{z} \in \mathbb{R}^J \backslash \ell^2(\boldsymbol{\sigma}^2)$, where $h \in H(\widetilde{L}_{\boldsymbol{\sigma}})$ denotes the representer of $f \mapsto \int_{\mathbb{R}^J} f d\mu$ on $H(\widetilde{L}_{\boldsymbol{\sigma}})$.

Proof. Lemma 3.1 yields (i) in the non-trivial case $d = \infty$. Since $\widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{x}) = 1$ for every $\boldsymbol{x} \in \mathbb{R}^J$, the kernel $\widetilde{L}_{\boldsymbol{\sigma}}$ has finite trace, i.e.,

$$\int_{\mathbb{R}^J} \widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{x}) d\mu(\boldsymbol{x}) < \infty.$$

The latter implies (ii), see, e.g., Steinwart and Scovel (2012, Lemma 2.3).

We establish (iii) in the non-trivial case $d = \infty$. Let $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$, $\boldsymbol{y} \in \ell^2(\boldsymbol{\sigma}^2)$, and $\boldsymbol{z} \in \mathbb{R}^{\mathbb{N}} \backslash \ell^2(\boldsymbol{\sigma}^2)$. Obviously, $\widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{y}, \boldsymbol{z}) = 0$. The representer h of integration on $H(\widetilde{L}_{\boldsymbol{\sigma}})$ satisfies

$$h(\boldsymbol{x}) = \int_{\ell^2(\boldsymbol{\sigma}^2)} \widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}) d\mu(\boldsymbol{y}),$$

and therefore $h(\boldsymbol{z}) = 0$. □

A natural question in the case $d = \infty$ is then whether the restriction of the domain from $\mathbb{R}^J$ to $\ell^2(\boldsymbol{\sigma})$ in any way impacts the study of the integration or the

L^2-approximation problem. Lemma 3.2.(iii) reveals that this is not the case if (3.3) holds.

We describe how the choice of $\mathfrak{X} = \ell^2(\boldsymbol{\sigma})$ fits into the tensor product setting from Chapter 1.

Remark 3.3. Recall that the appropriate domain for an infinite tensor product kernel relies on unit vectors $u_j \in H(\ell_{\sigma_j})$ that are, loosely speaking, close to 1. Since $1 \notin H(\ell_{\sigma_j})$ for any choice of shape parameter, we choose u_j given by

$$u_j(x) = \ell_{\sigma_j}(x, 0)$$

for all $x \in \mathbb{R}$, cf. also Gnewuch et al. (2022, Sec. 4.2). Clearly, we have

$$\|u_j\|^2_{H(\ell_{\sigma_j})} = \ell_{\sigma_j}(0, 0) = 1.$$

Two elementary calculations show

$$\int_{\mathbb{R}} u_j d\mu_0 = \left(\frac{1}{2\sigma_j^2 + 1}\right)^{1/2}$$

as well as

$$\int_{\mathbb{R}} u_j^2 d\mu_0 = \left(\frac{1}{4\sigma_j^2 + 1}\right)^{1/2}.$$

We recall from Chapter 1 the definition of the appropriate domain

$$\mathfrak{X} = \{\boldsymbol{x} \in \mathbb{R}^J : \sum_{j \in J} |\ell_{\sigma_j}(x_j, x_j) - 1| < \infty \text{ and } \sum_{j \in J} |u_j(x_j) - 1| < \infty\},$$

which here simplifies to

$$\mathfrak{X} = \{\boldsymbol{x} \in \mathbb{R}^J : \sum_{j \in J} |u_j(x_j) - 1| < \infty\}.$$

Since $u_j(x_j) > 0$ always holds, the sum $\sum_{j \in J} |u_j(x_j) - 1|$ converges if and only if the product $\prod_{j\in\mathbb{N}} u_j(x_j)$ converges to a value other than zero, and we obtain

$$\mathfrak{X} = \ell^2(\boldsymbol{\sigma}).$$

We note that Theorem 1.4 is also applicable here if (3.3) holds, since

$$\sigma_j^2, \qquad \left|\int_{\mathbb{R}} u_j d\mu_0 - 1\right| \qquad \text{and} \qquad \left|\int_{\mathbb{R}} u_j^2 d\mu_0 - 1\right|$$

all are asymptotically equivalent in j and we have $\mu(\mathfrak{X}) = 1$ already by Lemma 3.2(i). Of course, in this case the L^2-embedding is easier obtained by Lemma 3.2(ii).

In subsequent sections, we consider integration and L^2-approximation, which are trivial if (3.3) does not hold, see Remark 3.8 and Section 3.3.1. See Gnewuch et al. (2022, Sec. 4.2) for further results on different domains for $\widetilde{L}_{\boldsymbol{\sigma}}$ in the case $d = \infty$.

For every d the mapping

$$L_{\boldsymbol{\sigma}} := \widetilde{L}_{\boldsymbol{\sigma}}|_{\ell^2(\boldsymbol{\sigma}^2)\times\ell^2(\boldsymbol{\sigma}^2)}$$

and the Hilbert space $H(L_{\boldsymbol{\sigma}})$ are called a Gaussian kernel and, despite a different use in stochastic analysis, a Gaussian space, respectively. Accordingly, $H(L_{\boldsymbol{\sigma}})$ consists of real-valued functions on the domain

$$\mathfrak{X}(L_{\boldsymbol{\sigma}}) := \ell^2(\boldsymbol{\sigma}^2).$$

3.1.3 Hermite Kernels

We introduce a new notation for Hermite spaces to make the dependence on the Fourier weights clearer. This is necessary in this chapter, since we often manipulate the weights or consider multiple scales of weights at the same time. We consider specific Hermite spaces of infinite smoothness, see Section 2.4.1 and 2.1.2.

Let $0 < \beta < 1$. The univariate Hermite spaces we consider are determined by the sequence of Fourier weights $(\alpha_\nu)_{\nu\in\mathbb{N}_0} = (\beta^{-\nu})_{\nu\in\mathbb{N}_0}$, so the kernel k is given by

$$k_\beta(x, y) := \sum_{\nu\in\mathbb{N}_0} \beta^\nu \cdot h_\nu(x) \cdot h_\nu(y)$$

for $x, y \in \mathbb{R}$. We call β the base parameter.

For infinite-variate Hermite spaces, let $\boldsymbol{\beta} := (\beta_j)_{j\in J}$ be a sequence of base parameters, i.e., $0 < \beta_j < 1$.

For every sequence $\boldsymbol{\beta}$ with

$$\sum_{j\in J} \beta_j < \infty \tag{3.5}$$

we define the Hermite kernel $K_{\boldsymbol{\beta}}$ by

$$K_{\boldsymbol{\beta}}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j\in J} k_{\beta_j}(x_j, y_j)$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}(K_{\boldsymbol{\beta}})$ with the maximal domain

$$\mathfrak{X}(K_{\boldsymbol{\beta}}) := \ell^2(\boldsymbol{\beta}),$$

cf. (2.54). If $d < \infty$ then (3.5) is trivially satisfied and $\mathfrak{X}(K_{\boldsymbol{\beta}}) = \mathbb{R}^d$. Recall that $\mu(\mathfrak{X}(K_{\boldsymbol{\beta}})) = 1$ and $H(K_{\boldsymbol{\beta}}) \subseteq L^2(\mu)$ with a compact identical embedding, cf. Theorem 2.2.

3.1.4 The Isometric Isomorphism

As we will see, the relation

$$1 - \beta_j = \frac{2}{1 + \left(1 + 8\sigma_j^2\right)^{1/2}} \tag{3.6}$$

for every $j \in J$ is of key importance. First of all, we note that (3.6) defines a bijection between the set of shape parameters $\sigma_j > 0$ and the set of base parameters $0 < \beta_j < 1$.

Remark 3.4. Consider the case $d = \infty$, and assume that (3.6) is satisfied for every $j \in \mathbb{N}$. If $\lim_{j\to\infty} \sigma_j = 0$ or $\lim_{j\to\infty} \beta_j = 0$, then $\beta_j \asymp \sigma_j^2$ and therefore $\ell^2(\boldsymbol{\sigma}^2) = \ell^2(\boldsymbol{\beta})$ as vector spaces. Moreover, (3.3) and (3.5) are equivalent. We conclude that (3.6) defines a bijection between the set of square-summable sequences of shape parameters $\sigma_j > 0$ and the set of summable sequences of base parameters $0 < \beta_j < 1$.

Let $\boldsymbol{c} := (c_j)_{j\in J}$ be a positive sequence with $\sum_{j\in J} \omega_j < \infty$ for $\omega_j := |c_j - 1|$. We define $\varphi_{\boldsymbol{c}} \colon \mathbb{R}^J \to [0, \infty[$ by

$$\varphi_{\boldsymbol{c}}(\boldsymbol{x}) := \begin{cases} \exp\Big(-\sum_{j\in J} \frac{c_j^2-1}{4} \cdot x_j^2\Big) & \text{if } \boldsymbol{x} \in \ell^2(\boldsymbol{\omega}), \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\boldsymbol{x} \mapsto \boldsymbol{c}\boldsymbol{x}$ and $\boldsymbol{x} \mapsto \boldsymbol{c}^{-1}\boldsymbol{x}$ define bijections on $\ell^2(\boldsymbol{\sigma}^2)$. We define a linear mapping $Q_{\boldsymbol{c}}$ on the space of all functions $f \colon \ell^2(\boldsymbol{\sigma}^2) \to \mathbb{R}$ by

$$Q_{\boldsymbol{c}} f(\boldsymbol{x}) := \boldsymbol{c}_*^{1/2} \cdot \varphi_{\boldsymbol{c}}(\boldsymbol{x}) \cdot f(\boldsymbol{c}\boldsymbol{x})$$

for $\boldsymbol{x} \in \ell^2(\boldsymbol{\sigma}^2)$, reminding the reader of the notation $\boldsymbol{c}_* = \prod_{j\in J} c_j$.

In Gnewuch et al. (2022, Thm. 5.8) the following result was established. Note that in that paper, instead of the univariate kernels k_{β_j}, kernels $(1-\beta_j)k_{\beta_j}$ were considered.

Theorem 3.5. *Assume that* (3.3) *is satisfied. Moreover, assume that* (3.6) *and*

$$c_j = \big(1 + 8\sigma_j^2\big)^{1/4} \tag{3.7}$$

are satisfied for every $j \in J$. Then $Q_{\boldsymbol{c}}$ defines an isometric isomorphism on $L^2(\mu)$ and between $H((\mathbf{1}-\boldsymbol{\beta})_* K_{\boldsymbol{\beta}})$ and $H(L_{\boldsymbol{\sigma}})$.

Regarding the proof, we mention that at first the theorem is established in the univariate case utilizing Mehler's formula that makes it possible to write univariate Gaussian kernels in terms of the Hermite polynomials. The multi- and infinite-variate case then follow using tensorisation arguments similar to Theorem A.3.

By definition, the value of $Q_{\boldsymbol{c}} f$ at $\boldsymbol{x}$ is determined by the value of f at $\boldsymbol{c}\boldsymbol{x}$. The analogous property for $Q_{\boldsymbol{c}}^{-1}$ reads as follows.

Lemma 3.6. *Assume that* (3.3) *is satisfied. Moreover, assume that* (3.7) *is satisfied for every $j \in J$. Then we have $\varphi_{\boldsymbol{c}} > 0$ on $\ell^2(\boldsymbol{\sigma}^2)$ and*

$$Q_{\boldsymbol{c}}^{-1} f(\boldsymbol{x}) = \frac{1}{\boldsymbol{c}_*^{1/2} \varphi_{\boldsymbol{c}}(\boldsymbol{c}^{-1}\boldsymbol{x})} \cdot f(\boldsymbol{c}^{-1}\boldsymbol{x})$$

for all $f \colon \ell^2(\boldsymbol{\sigma}^2) \to \mathbb{R}$ and $\boldsymbol{x} \in \ell^2(\boldsymbol{\sigma}^2)$.

Proof. We only have to verify that $\varphi_{\boldsymbol{c}} > 0$ on $\ell^2(\boldsymbol{\sigma}^2)$. In the non-trivial case $d = \infty$ the latter follows from $c_j^2 - 1 \asymp \sigma_j^2$, which holds due to (3.3) and (3.7). □

3.2 Integration

In what follows, we derive new results for integration and L^2-approximation on Gaussian and Hermite spaces. This will be done by transferring known results from Hermite spaces with the help of the isometric isomorphism from Section 3.1.4 to Gaussian spaces and vice versa. Since the actual transference mechanisms for integration and L^2-approximation differ significantly, we discuss both applications separately in Sections 3.2 and 3.3, respectively.

Throughout this section, $\boldsymbol{\sigma} := (\sigma_j)_{j\in J}$ and $\boldsymbol{\beta} := (\beta)_{j\in J}$ denote sequences of shape parameters and base parameters, respectively. Initially, we do not impose any summability requirements on $\boldsymbol{\sigma}$ or $\boldsymbol{\beta}$. For the integration problem, an appropriate relation between σ_j and β_j, which is different from (3.6), will be determined later, see (3.11).

3.2.1 The Norm of the Integration Functional

Before our main discussion of the integration problem, we briefly discuss its norm, or equivalently the 0-th minimal error

$$e_0(M) = \sup_{\|f\|_{H(M)}\leqslant 1} \left| \int_{\mathfrak{X}} f \, d\mu \right|$$

for $M = L_{\boldsymbol{\sigma}}$ and $M = K_{\boldsymbol{\beta}}$. We remark that $e_0(M)$ will appear as a constant multiple times throughout the rest of the section.

Lemma 3.7. *For the integration problem, if* (3.3) *is satisfied then*

$$e_0(L_{\boldsymbol{\sigma}}) = \prod_{j\in J} \frac{1}{\left(1 + 4\sigma_j^2\right)^{1/4}}.$$

If (3.5) *is satisfied then*

$$e_0(K_{\boldsymbol{\beta}}) = 1.$$

Proof. For $d = 1$, we refer to the known results

$$e_0^4(\ell_{\sigma_j}) = \frac{1}{1 + 4\sigma_j^2} < 1,$$

see Kuo et al. (2017b, Eqn. (1.5)), as well as $e_0(k_{\beta_j}) = 1$ see Irrgeher et al. (2015, p. 385).

The case $d \in \mathbb{N}$ follows immediately.

For $d = \infty$, we make use of the tensor product structure of the integration functional, see Theorem A.3, noting that the d-fold tensor product of $\mathbb{R}$ (as a Hilbert space) is isometrically isomorphic to $\mathbb{R}$. Cf. Gnewuch et al. (2022, Sec. A.4–A.6), which may also be used here. We obtain

$$e_0(M) = \prod_{j \in J} e_0(m_j)$$

with $m_j := \ell_{\sigma_j}$ or with $m_j := k_{\beta_j}$. □

Remark 3.8. Consider the reproducing kernel $\widetilde{L}_{\boldsymbol{\sigma}}$ on the domain $\mathbb{R}^{\mathbb{N}}$, as in (3.4). If (3.3) is not satisfied, we still have $H(\widetilde{L}_{\boldsymbol{\sigma}}) \subseteq L^2(\mu)$ due to Lemma 3.2. However, in this case, integration is trivial, since for the norm z of the integration functional we obtain

$$\begin{aligned} z^2 &= \int_{\mathbb{R}^{\mathbb{N}}} \int_{\mathbb{R}^{\mathbb{N}}} \widetilde{L}_{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y}) \, d\mu(\boldsymbol{x}) \, d\mu(\boldsymbol{y}) \\ &= \prod_{j \in \mathbb{N}} \int_{\mathbb{R}} \int_{\mathbb{R}} \ell_{\sigma_j}(x_j, y_j) \, d\mu_0(x_j) \, d\mu_0(y_j) \\ &= \prod_{j \in \mathbb{N}} \frac{1}{(1 + 4\sigma_j^2)^{1/2}} = 0. \end{aligned}$$

3.2.2 The Transference Result

For any positive sequence $\boldsymbol{\tau} \in \mathbb{R}^J$ we define $t_{\boldsymbol{\tau}} \colon \mathbb{R}^J \to \mathbb{R}^J$ by $t_{\boldsymbol{\tau}}(\boldsymbol{x}) := \boldsymbol{\tau}^{-1}\boldsymbol{x}$.

In this section, our goal is, roughly speaking, to utilize the isomorphism $Q_{\boldsymbol{c}}$ to translate algorithms for integration on $H(K_{\boldsymbol{\beta}})$ into algorithms for integration on $H(L_{\boldsymbol{\sigma}})$. The main challenge is that integration of f with respect to μ does not translate to integration of Qf with respect to μ, but with respect to the image measure $t_{\boldsymbol{\tau}}\mu$, where $\boldsymbol{c}$ and $\boldsymbol{\tau}$ are suitably related, see Lemma 3.10 for details. To offset this, we will establish and use the fact that, for a sequence $\boldsymbol{\tau}'$, the integration problem on $H(L_{\boldsymbol{\sigma\tau}'})$ with respect to μ is equivalent to the integration problem on $H(L_{\boldsymbol{\sigma}})$ with respect to $t_{\boldsymbol{\tau}'}\mu$. Then, we only need to find fitting parameters so that both changes of measure cancel each other out, leading to the different relation between $\boldsymbol{\sigma}$ and $\boldsymbol{\beta}$ given by (3.11), as opposed to (3.6).

Lemma 3.9. *Let $\boldsymbol{\tau} := (\tau_j)_{j\in J}$ denote a positive sequence with $\sum_{j\in J}|\tau_j - 1| < \infty$. The image measure $t_{\boldsymbol{\tau}}\mu$ of μ with respect to $t_{\boldsymbol{\tau}}$ has the density $\boldsymbol{\tau}_* \cdot \varphi_{\boldsymbol{\tau}}^2$ with respect to μ.*

Proof. To establish the statement of the lemma, it suffices to show for an arbitrary cylinder set $A \subseteq \mathbb{R}^J$ that $t_{\boldsymbol{\tau}}\mu(A) = \int_A \boldsymbol{\tau}_* \cdot \varphi_{\boldsymbol{\tau}}^2 d\mu$.

In the case $d < \infty$ any cylinder set is of the form $A := A_1 \times \cdots \times A_d$ with measurable sets $A_j \subseteq \mathbb{R}$. Since

$$\begin{aligned}\mu_0(\tau_j A_j) &= (2\pi)^{-1/2} \cdot \int_{\tau_j A_j} \exp(-x_j^2/2)dx_j \\ &= \tau_j \cdot \int_{A_j} \exp\left(-\frac{\tau_j^2 - 1}{2} \cdot x_j^2\right) d\mu_0(x_j),\end{aligned}$$

we obtain

$$t_{\boldsymbol{\tau}}\mu(A) = \mu\left(t_{\boldsymbol{\tau}}^{-1}(A)\right) = \prod_{j=1}^{d} \mu_0(\tau_j A_j) = \boldsymbol{\tau}_* \cdot \int_A \varphi_{\boldsymbol{\tau}}^2(\boldsymbol{x}) d\mu(\boldsymbol{x}),$$

which in turn yields the claim.

In the case $d = \infty$ we put $J_0 := \{j \in \mathbb{N} : \tau_j < 1\}$. Using the monotone convergence theorem we obtain

$$\int_{\mathbb{R}^{\mathbb{N}}} \varphi_{\boldsymbol{\tau}}^2(\boldsymbol{x}) d\mu(\boldsymbol{x}) \leqslant \int_{\mathbb{R}^{\mathbb{N}}} \exp\left(-\sum_{j\in J_0} \frac{\tau_j^2 - 1}{2} \cdot x_j^2\right) d\mu(\boldsymbol{x}) = \prod_{j\in J_0} \tau_j^{-1} < \infty. \tag{3.8}$$

For $d = \infty$ any cylinder set is of the form $A := A_1 \times \ldots$ with measurable sets $A_j \subseteq \mathbb{R}$, where $A_j = \mathbb{R}$ for $j > j_0$. For every $j_1 > j_0$ we obtain

$$\begin{aligned} t_{\boldsymbol{\tau}}\mu(A) &= \prod_{j=1}^{j_1} \mu_0(\tau_j A_j) = \prod_{j=1}^{j_1} \left(\tau_j \cdot \int_{A_j} \exp\left(-\frac{\tau_j^2 - 1}{2} \cdot x_j^2 \right) d\mu_0(x_j) \right) \\ &= \prod_{j=1}^{j_1} \tau_j \cdot \int_A \exp\left(-\sum_{j=1}^{j_1} \frac{\tau_j^2 - 1}{2} \cdot x_j^2 \right) d\mu(\boldsymbol{x}). \end{aligned}$$

Clearly, we have

$$\exp\left(-\sum_{j=1}^{j_1} \frac{\tau_j^2 - 1}{2} \cdot x_j^2 \right) \leqslant \exp\left(-\sum_{j \in J_0} \frac{\tau_j^2 - 1}{2} \cdot x_j^2 \right)$$

for all $\boldsymbol{x} \in \mathbb{R}^{\mathbb{N}}$. Thus, due to (3.8), the dominated convergence theorem yields

$$t_{\boldsymbol{\tau}}\mu(A) = \boldsymbol{\tau}_* \cdot \int_A \varphi_{\boldsymbol{\tau}}^2(\boldsymbol{x}) d\mu(\boldsymbol{x}),$$

which in turn establishes the claim. □

Lemma 3.10. *Let* $\boldsymbol{c} := (c_j)_{j \in J}$ *and* $\boldsymbol{\tau} := (\tau_j)_{j \in J}$ *denote positive sequences with* $\sum_{j \in J} |c_j - 1| < \infty$ *and* $\sum_{j \in J} |\tau_j - 1| < \infty$. *If*

$$\tau_j^2 = \frac{c_j^2 + 1}{2} \tag{3.9}$$

for every $j \in J$ *then we have*

$$I(Q_{\boldsymbol{c}} f \circ t_{\boldsymbol{\tau}}) = \frac{\boldsymbol{\tau}_*}{\boldsymbol{c}_*^{1/2}} \cdot I(f) \tag{3.10}$$

for every $f \in L^1(\mu)$.

Proof. By definition,

$$Q_{\boldsymbol{c}} f \circ t_{\boldsymbol{\tau}} = \boldsymbol{c}_*^{1/2} \cdot (\varphi_{\boldsymbol{c}} \circ t_{\boldsymbol{\tau}}) \cdot (f \circ t_{\boldsymbol{c}^{-1}\boldsymbol{\tau}}) = \left(\boldsymbol{c}_*^{1/2} \cdot (\varphi_{\boldsymbol{c}} \circ t_{\boldsymbol{c}}) \cdot f \right) \circ t_{\boldsymbol{c}^{-1}\boldsymbol{\tau}}.$$

Note that $\sum_{j\in J}\omega_j < \infty$ for $\omega_j := |c_j^{-1}\tau_j - 1|$. Lemma 3.9 implies

$$I(Q_{\boldsymbol{c}} f \circ t_{\boldsymbol{\tau}}) = \frac{\boldsymbol{\tau}_*}{\boldsymbol{c}_*^{1/2}} \cdot \int_{\mathbb{R}^J} (\varphi_{\boldsymbol{c}} \circ t_{\boldsymbol{c}}) \cdot \varphi^2_{\boldsymbol{c}^{-1}\boldsymbol{\tau}} \cdot f d\mu$$

for every $f\colon \mathbb{R}^J \to \mathbb{R}$ such that $(\varphi_{\boldsymbol{c}} \circ t_{\boldsymbol{c}}) \cdot \varphi^2_{\boldsymbol{c}^{-1}\boldsymbol{\tau}} \cdot f \in L^1(\mu)$.

Lemma 3.1 yields $\mu(\ell^2(\boldsymbol{\omega})) = 1$. Let $\boldsymbol{x} \in \ell^2(\boldsymbol{\omega})$. We have $\varphi_{\boldsymbol{c}^{-1}\boldsymbol{\tau}}(\boldsymbol{x}) > 0$ and, since $\omega_j \asymp |c_j - 1|$, we also have $\varphi_{\boldsymbol{c}}(t_{\boldsymbol{c}}(\boldsymbol{x})) > 0$. Since

$$-\ln\left(\varphi_{\boldsymbol{c}}(t_{\boldsymbol{c}}(\boldsymbol{x})) \cdot \varphi^2_{\boldsymbol{c}^{-1}\boldsymbol{\tau}}(\boldsymbol{x})\right) = \sum_{j\in J} \frac{x_j^2}{4c_j^2} \cdot \left(c_j^2 - 1 + 2(\tau_j^2 - c_j^2)\right),$$

we conclude that (3.10) is satisfied for every $f \in L^1(\mu)$ if $c_j^2 - 1 = 2(c_j^2 - \tau_j^2)$ for every $j \in J$. The latter is equivalent to (3.9) for every $j \in J$. □

We are now able to give the following main result, cf. Gnewuch et al. (2026, Thm. 4.3).

Theorem 3.11. *Assume that* (3.3) *is satisfied. Moreover, assume that*

$$1 - \beta_j = \frac{1}{1 + 2\sigma_j^2}, \tag{3.11}$$

$$c_j = (1 + 4\sigma_j^2)^{1/2}, \tag{3.12}$$

$$\tau_j = (1 + 2\sigma_j^2)^{1/2} \tag{3.13}$$

are satisfied for every $j \in J$. For every quadrature formula A with nodes from $\ell^2(\boldsymbol{\sigma}^2)$ and

$$B(f) := \frac{\boldsymbol{c}_*^{1/2}}{\boldsymbol{\tau}_*} \cdot A(Q_{\boldsymbol{c}} f \circ t_{\boldsymbol{\tau}})$$

we have

$$e(A, L_{\boldsymbol{\sigma}}) = (\boldsymbol{1} + 4\boldsymbol{\sigma}^2)_*^{-1/4} \cdot e(B, K_{\boldsymbol{\beta}}).$$

Proof. At first, we only assume that (3.3) is satisfied. We alter the shape parameters of the Gaussian kernel $L_{\boldsymbol{\sigma}}$. To this end, let $\boldsymbol{\tau} := (\tau_j)_{j\in J}$ denote any positive sequence with $\sum_{j\in J}|\tau_j - 1| < \infty$. Obviously, $\sum_{j\in J}(\tau_j\sigma_j)^2 < \infty$. Recall that $t_{\boldsymbol{\tau}}$ defines a bijection on $\ell^2(\boldsymbol{\sigma}^2)$. Since

$$L_{\boldsymbol{\tau\sigma}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j \in J} \ell_{\tau_j \sigma_j}(x_j, y_j) = \prod_{j \in J} \ell_{\sigma_j}(\tau_j x_j, \tau_j y_j) = L_{\boldsymbol{\sigma}}(\boldsymbol{\tau x}, \boldsymbol{\tau y})$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \ell^2(\boldsymbol{\sigma}^2)$, the spaces $H(L_{\boldsymbol{\tau\sigma}})$ and $H(L_{\boldsymbol{\sigma}})$ are isometrically isomorphic via $f \mapsto f \circ t_{\boldsymbol{\tau}}$. Consequently,

$$e(A, L_{\boldsymbol{\sigma}}) = \sup_{\|f\|_{H(L_{\boldsymbol{\tau\sigma}})} \leqslant 1} |I(f \circ t_{\boldsymbol{\tau}}) - A(f \circ t_{\boldsymbol{\tau}})|.$$

Next, we apply Theorem 3.5 with $\boldsymbol{\tau\sigma}$ instead of $\boldsymbol{\sigma}$, which leads to

$$1 - \beta_j = \frac{2}{1 + (1 + 8(\tau_j \sigma_j)^2)^{1/2}} \tag{3.14}$$

and

$$c_j = \left(1 + 8(\tau_j \sigma_j)^2\right)^{1/4}, \tag{3.15}$$

instead of (3.6) and (3.7). Assume that (3.14) and (3.15) are satisfied for every $j \in J$. Since $\sum_{j \in J} (\tau_j \sigma_j)^2 < \infty$, we obtain $\sum_{j \in J} \beta_j < \infty$ by Remark 3.4, which in turn ensures that $K_{\boldsymbol{\beta}}$ is well-defined. Let $Q := Q_{\boldsymbol{c}}$. Then $Q_{|H(K_{\boldsymbol{\beta}})}$ is an isomorphism from $H(K_{\boldsymbol{\beta}})$ to $H(L_{\boldsymbol{\tau\sigma}})$ with

$$\|Qf\|^2_{H(L_{\boldsymbol{\tau\sigma}})} = \|f\|^2_{H((\boldsymbol{1}-\boldsymbol{\beta})_* K_{\boldsymbol{\beta}})} = (\boldsymbol{1} - \boldsymbol{\beta})_*^{-1} \cdot \|f\|^2_{H(K_{\boldsymbol{\beta}})}$$

for every $f \in H(K_{\boldsymbol{\beta}})$. We conclude that

$$e(A, L_{\boldsymbol{\sigma}}) = (1 - \boldsymbol{\beta})_*^{1/2} \cdot \sup_{\|f\|_{H(K_{\boldsymbol{\beta}})} \leqslant 1} |I(Qf \circ t_{\boldsymbol{\tau}}) - A(Qf \circ t_{\boldsymbol{\tau}})|.$$

Finally, we apply Lemma 3.10, noting that the unique solution of (3.9) and (3.15) with $\tau_j > 0$ is given by (3.12) and (3.13). In the sequel, we assume that (3.11), (3.12), and (3.13) are satisfied, in addition to (3.3). We use

$$(\tau_j \sigma_j)^2 = (1 + 2\sigma_j^2)\sigma_j^2$$

to obtain

$$1 + 8(\tau_j \sigma_j)^2 = (1 + 4\sigma_j^2)^2.$$

It follows that (3.14) is satisfied, too, and obviously, $\sum_{j \in J} |\tau_j - 1| < \infty$ as well as $\sum_{j \in J} |c_j - 1| < \infty$. Lemma 3.10 yields

$$e(A, L_{\boldsymbol{\sigma}}) = (1 - \boldsymbol{\beta})_*^{1/2} \cdot \frac{\boldsymbol{\tau}_*}{\boldsymbol{c}_*^{1/2}} \cdot \sup_{\|f\|_{H(K_{\boldsymbol{\beta}})} \leqslant 1} \left| I(f) - \frac{\boldsymbol{c}_*^{1/2}}{\boldsymbol{\tau}_*} \cdot A(Qf \circ t_{\boldsymbol{\tau}}) \right|.$$

We use

$$\frac{\tau_j}{c_j^{1/2}} = \frac{\left(1 + 2\sigma_j^2\right)^{1/2}}{\left(1 + 4\sigma_j^2\right)^{1/4}},$$

to obtain

$$(1 - \beta_j)^{1/2} \frac{\tau_j}{c_j^{1/2}} = \frac{1}{(1 + 4\sigma_j^2)^{1/4}}$$

and therefore

$$(1 - \boldsymbol{\beta})_*^{1/2} \cdot \frac{\boldsymbol{\tau}_*}{\boldsymbol{c}_*^{1/2}} = (\mathbf{1} + 4\boldsymbol{\sigma}^2)_*^{-1/4}. \qquad \square$$

Remark 3.12. Let the assumptions from Theorem 3.11 be satisfied. For a quadrature formula A on the Gaussian space $H(L_{\boldsymbol{\sigma}})$ with nodes from $\ell^2(\boldsymbol{\sigma}^2)$ the quadrature formula B on the Hermite space $H(K_{\boldsymbol{\beta}})$ according to Theorem 3.11 is easily determined explicitly. Let

$$e_j := \left(\frac{1 + 4\sigma_j^2}{1 + 2\sigma_j^2}\right)^{1/2}$$

for every $j \in J$. Then, for $f \colon \ell^2(\boldsymbol{\sigma}^2) \to \mathbb{R}$ we have

$$Bf = \boldsymbol{e}_* \cdot \sum_{i=1}^{n} f(\boldsymbol{y}_i) \cdot b_i,$$

with $\boldsymbol{y}_i \in \ell^2(\boldsymbol{\sigma}^2)$ and $b_i \in \mathbb{R}$ given by

$$\boldsymbol{y}_i := \boldsymbol{e}\boldsymbol{x}_i \qquad \text{and} \qquad b_i := \varphi_{\boldsymbol{c}}(\boldsymbol{\tau}^{-1}\boldsymbol{x}_i) \cdot a_i.$$

It is easily verified that $(\boldsymbol{x}_i, a_i) \mapsto (\boldsymbol{y}_i, b_i)$ defines a bijection on the set $\ell^2(\boldsymbol{\sigma}^2) \times \mathbb{R}$ of pairs of nodes and coefficients.

We discuss the cost of the algorithm B. In the case $d \in \mathbb{N}$, we assume $\mathrm{cost}(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \mathbb{R}^J$, which leads to

$$\mathrm{cost}(A) = \mathrm{cost}(B) = n.$$

In the case $d = \infty$, we use the unrestricted subspace sampling model, see Section 1.3, specifically with the default value $a = 0$. Since we have $e_j \neq 0$ for all $j \in J$, it follows that $\mathrm{cost}(\boldsymbol{x}_i) = \mathrm{cost}(\boldsymbol{y}_i)$ for all i, and therefore

$$\mathrm{cost}(A) = \mathrm{cost}(B).$$

Corollary 3.13. *Assume that* (3.3) *and* (3.11) *are satisfied. In the cost model cases as given in Remark 3.12, for every $n \in \mathbb{N}$ we have*

$$\frac{e_n(L_{\boldsymbol{\sigma}})}{e_0(L_{\boldsymbol{\sigma}})} = \frac{e_n(K_{\boldsymbol{\beta}})}{e_0(K_{\boldsymbol{\beta}})} = e_n(K_{\boldsymbol{\beta}}).$$

i.e., the normalized n-th minimal errors for integration on Gaussian spaces and on the corresponding Hermite spaces, related by (3.11), *coincide.*

Proof. We combine Theorem 3.11 and Remark 3.12 to obtain

$$e_n(L_{\boldsymbol{\sigma}}) = (\mathbf{1} + 4\boldsymbol{\sigma}^2)_*^{-1/4} \cdot e_n(K_{\boldsymbol{\beta}})$$

for every $n \in \mathbb{N}_0$. Use Lemma 3.7 to establish the desired result. □

For the remainder of this section, we utilize Corollary 3.13 to study the asymptotic behavior of the n-th minimal errors, which, by Corollary 3.13, is the same on $H(L_{\boldsymbol{\sigma}})$ and $H(K_{\boldsymbol{\beta}})$. We remark, however, that Corollary 3.13 actually even gives us the exact relation between $e_n(L_{\boldsymbol{\sigma}})$ and $e(K_{\boldsymbol{\beta}})$ for every $n \in \mathbb{N}$. In particular, integration on $H(L_{\boldsymbol{\sigma}})$ is actually easier than on $H(K_{\boldsymbol{\beta}})$ and, at least for large shape parameters, the corresponding constant factor might be significant for moderately large n.

3.2.3 Univariate Integration

Let us mention prior work concerning integration on $H(\ell_\sigma)$ and on $H(k_\beta)$, where explicit upper bounds for the worst-case error $e(A, M)$ of particular quadrature formulas A and explicit lower bounds for $e_n(M)$ are obtained. Of course, these upper bounds immediately yield upper bounds for minimal errors.

First, we mention Gauss-Hermite rules, which are analyzed in Irrgeher et al. (2015) on Hermite spaces and in Kuo and Woźniakowski (2012) as well as Kuo et al. (2017b) on Gaussian spaces. These are Gaussian quadrature rules based on exactly integrating the Hermite polynomials. More specifically, for $n \in \mathbb{N}$, the n-point Gauss-Hermite rule is given by

$$A_n f = \sum_{i=1}^{n} f(\boldsymbol{x}_i) \cdot a_i,$$

where the x_i are the zeros of the Hermite polynomial of degree n, and the weights a_i are chosen in such a way to ensure that A_n integrates all polynomials of degree less or equal than $2n - 1$ exactly. While of course the set of all polynomials is dense in $H(k_\beta)$, the space $H(\ell_\sigma)$ contains no nontrivial polynomials, so intuitively Gauss-Hermite rules are better fitted for integration on $H(k_\beta)$ than on $H(\ell_{\boldsymbol{\sigma}})$. Nevertheless, Kuo and Woźniakowski (2012) and Kuo et al. (2017b) establish positive results.

In Karvonen et al. (2021), scaled Gauss-Hermite rules were studied on Gaussian spaces. For $n \in \mathbb{N}$, and $\boldsymbol{x}_i, a_i$ given as for the n-point Gauss-Hermite rule, the n-point scaled Gauss-Hermite rule is given by

$$B_n(f) = \frac{1}{(1 + 2\sigma^2)^{1/2}} \sum_{i=1}^{n} f\left(\frac{\boldsymbol{x}_i}{(1 + 2\sigma^2)^{1/2}}\right) \cdot a_i \cdot \exp\left(\frac{\boldsymbol{x}_i^2 \cdot \sigma^2}{1 + 2\sigma^2}\right),$$

which is motivated by integrating $2n - 1$ members of a specific orthonormal basis of $H(\ell_{\boldsymbol{\sigma}})$ exactly.

Lower bounds for the n-minimal errors have been established in Irrgeher et al. (2015) for Hermite spaces and in Kuo et al. (2017b) for Gaussian spaces.

In the following we combine these results in the best possible way regarding the convergence rate of the n-th minimal errors with the transference result established in Theorem 3.11 and Corollary 3.13.

We remark that in Kuo and Woźniakowski (2012) and Kuo et al. (2017b), integration with respect to the normal distribution $N(0, 1/2)$ was studied, rather than with respect to μ_0. Similarly, in Karvonen et al. (2021), integration with respect to $N(0, \alpha^2)$ was studied for an arbitrary positive variance parameter α. However, as discussed in Section 3.2.2, the integration problem on $H(\ell_{\sigma\alpha})$ with respect to μ_0 is equivalent to the integration problem on $H(\ell_\sigma)$ with respect to $N(0, \alpha^2)$, so that results depend only on the product of α^2 and σ, rather than both variables separately. This is also reflected in the results of Karvonen et al. (2021). We are able to make use of all three papers by simply substituting the parameters appropriately.

Theorem 3.14. *Let*

$$\begin{aligned} C_1(\sigma) &:= 2^{-1} \cdot (1 + 4\sigma^2)^{-1/4}, \\ C_2(\sigma) &:= \pi^{-1/4} \cdot (1 + 2\sigma^2)^{-1/2}, \\ C(\beta) &:= \pi^{-1/4} \cdot (1 - \beta^2)^{1/4}. \end{aligned}$$

For every shape parameter $\sigma > 0$ and every $n \in \mathbb{N}$ the n-th minimal error $e_n(\ell_\sigma)$ for integration on the Gaussian space $H(\ell_\sigma)$ satisfies

$$C_1(\sigma) \cdot \left(\frac{\sigma^2}{1 + 2\sigma^2}\right)^{2n} \cdot (n+1)^{-2} \leqslant e_n(\ell_\sigma) \leqslant C_2(\sigma) \cdot \left(\frac{2\sigma^2}{1 + 2\sigma^2}\right)^{n} \cdot n^{-1/4}.$$

For every base parameter $0 < \beta < 1$ and every $n \in \mathbb{N}$ the n-th minimal error $e_n(k_\beta)$ for integration on the Hermite space $H(k_\beta)$ satisfies

$$\frac{1}{2} \cdot \left(\frac{\beta}{2}\right)^{2n} \cdot (n+1)^{-2} \leqslant e_n(k_\beta) \leqslant C(\beta) \cdot \beta^n \cdot n^{-1/4}.$$

Proof. The upper bound for $e_n(\ell_\sigma)$ follows immediately from Karvonen et al. (2021, Thm. 2.5), and the lower bound for $e_n(k_\beta)$ is established in Irrgeher et al. (2015, Thm. 2).

Assume that

$$1 - \beta = \frac{1}{1 + 2\sigma^2}.$$

First of all, Corollary 3.13 and Lemma 3.7 yield

$$e_n(\ell_\sigma) = (1 + 4\sigma^2)^{-1/4} \cdot e_n(k_\beta).$$

Moreover, we have

$$\beta = \frac{2\sigma^2}{1 + 2\sigma^2}$$

and

$$1 - \beta^2 = \frac{1 + 4\sigma^2}{(1 + 2\sigma^2)^2}.$$

Consequently, the upper bound for $e_n(k_\beta)$ follows from the upper bound for $e_n(\ell_\sigma)$ and the lower bound for $e_n(\ell_\sigma)$ follows from the lower bound for $e_n(k_\beta)$. □

The lower bound for $e_n(\ell_\sigma)$ from Theorem 3.14 improves the lower bound from Kuo et al. (2017b, Thm. 4.1), which is super-exponentially small in n, while the upper bound for $e_n(k_\beta)$ from Theorem 3.14 only slightly improves the upper bound from Irrgeher et al. (2015, Prop. 1). Obviously, the upper and lower bounds from Theorem 3.14 are not tight at all. On the level of algorithms, on $H(k_\beta)$, the scaled Gauss-Hermite rules, further transformed according to our transference principle and specifically to Remark 3.12, achieve a slightly better error rate than the Gauss-Hermite rules previously considered in Irrgeher et al. (2015).

3.2.4 Multivariate Integration

In this section we study the multivariate case $d \in \mathbb{N}$, where upper bounds for the minimal errors $e_n(M)$ have been established in Kuo et al. (2017b) and Karvonen et al. (2021) for Gaussian spaces as well as in Irrgeher et al. (2015) and Irrgeher et al. (2016b) for Hermite spaces. In all three papers, the upper bounds are obtained by quadrature formulas that are full tensor products of suitable univariate quadrature formulas, which have already been mentioned in Section 3.2.3. For lower bounds and for the study of tractability concepts we refer to Kuo et al. (2017b) and Irrgeher et al. (2015).

Exponential Convergence

At first, we consider the case of $d \in \mathbb{N}$ being fixed, and thus (3.3) and (3.5) are trivially satisfied for any choice of parameters σ_j and β_j, respectively.

Let either $M := L_{\boldsymbol{\sigma}}$ or $M := K_{\boldsymbol{\beta}}$. Motivated by Theorem 3.14, we study whether there exist constants $C_1, C_2, p > 0$ fulfilling

$$e_n(M) \leqslant C_1 \cdot \exp(-C_2 \cdot n^p) \tag{3.16}$$

for all $n \in \mathbb{N}$. In this case, we say that exponential convergence is achieved for the integration problem on $H(M)$ and

$$p^*(M) := \sup\{p \geqslant 0 \colon \exists\, C_1, C_2 > 0\ \forall\, n \in \mathbb{N} \colon e_n(M) \leqslant C_1 \cdot \exp(-C_2 \cdot n^p)\}$$

is the most relevant quantity.

Exponential convergence has been studied by Kuo et al. (2017b) and Karvonen et al. (2021) for Gaussian spaces, as well as by Irrgeher et al. (2015) for Hermite spaces. We present a moderate generalization of the result for Gaussian spaces, together with a proof that employs the transference mechanism.

Theorem 3.15. *Exponential convergence is achieved for the integration problem on* $H(L_{\boldsymbol{\sigma}})$ *with*

$$p = 1/d = p^*(L_{\boldsymbol{\sigma}}).$$

Proof. After a possibly necessary reordering of the univariate kernels, the shape parameters satisfy $\sigma_1 \geqslant \cdots \geqslant \sigma_d$, which implies $\beta_1 \geqslant \cdots \geqslant \beta_d$ for the base parameters given by (3.11). Thus we may employ Irrgeher et al. (2015, Thm. 1.1) in the specific case $\omega := 1/\mathrm{e}$ as well as $a_j := \ln(1/\beta_j)$ and $b_j := 1$ for all $j \in \mathbb{N}$, which yields exponential convergence on $H(K_{\boldsymbol{\beta}})$ with $p^*(K_{\boldsymbol{\beta}}) = 1/d$. Furthermore, (3.16) is satisfied for $M = K_{\boldsymbol{\beta}}$ in the extremal case $p = 1/d$, see Irrgeher et al. (2015, Thm. 4).

Corollary 3.13 implies that (3.16) holds for $M = K_{\boldsymbol{\beta}}$ if and only if it holds for $M = L_{\boldsymbol{\sigma}}$ with C_1 being replaced by $C_1 \cdot e_0(L_{\boldsymbol{\sigma}})$, but with C_2 and p unchanged. Thus, we obtain the desired result. □

Theorem 3.15 has been established in Kuo et al. (2017b, Thm. 1.1) under the additional assumption

$$\max_{1 \leqslant j \leqslant d} \sigma_j^2 < 1/2,$$

while the general case has been left open for future research. In the case

$$\sigma_1 = \cdots = \sigma_d > 0$$

the exponential convergence on $H(L_{\boldsymbol{\sigma}})$ with $p = 1/d$ follows directly from Karvonen et al. (2021, Cor. 2.11).

Alternatively to the transference mechanism, the following monotonicity property allows to establish Theorem 3.15 in full generality and working exclusively with Gaussian spaces. Suppose that $\boldsymbol{\sigma}^{(1)} \leqslant \boldsymbol{\sigma}^{(2)}$, i.e., $\sigma_j^{(1)} \leqslant \sigma_j^{(2)}$ for $j = 1, \ldots, d$. Using (3.1) we conclude that $H(L_{\boldsymbol{\sigma}^{(1)}}) \subseteq H(L_{\boldsymbol{\sigma}^{(2)}})$ with a continuous identical embedding. Therefore we obtain $p^*(L_{\boldsymbol{\sigma}}) \leqslant 1/d$ for every $\boldsymbol{\sigma}$ from Kuo et al. (2017b, Thm. 1.1.(b)), and exponential convergence with $p = 1/d$ for every $\boldsymbol{\sigma}$ from Karvonen et al. (2021, Cor. 2.11).

Exponential Convergence Weak Tractability

Exponential convergence has been studied alongside uniform exponential convergence and several notions of tractability in Irrgeher et al. (2015), Irrgeher et al. (2016b), and Kuo et al. (2017b). We specify the setting by arbitrary sequences $\boldsymbol{\sigma} := (\sigma_j)_{j \in \mathbb{N}}$ and $\boldsymbol{\beta} := (\beta_j)_{j \in \mathbb{N}}$ of shape parameters $\sigma_j > 0$ and base parameters $0 < \beta_j < 1$. For every $d \in \mathbb{N}$ we put $\boldsymbol{\sigma}_d := (\sigma_1, \ldots, \sigma_d)$ and $\boldsymbol{\beta}_d := (\beta_1, \ldots, \beta_d)$, and we consider either $M_d := L_{\boldsymbol{\sigma}_d}$ for all $d \in \mathbb{N}$ or $M_d := K_{\boldsymbol{\beta}_d}$ for all $d \in \mathbb{N}$.

By definition, we have uniform exponential convergence if there exists a constant $p > 0$ with the following property: For every $d \in \mathbb{N}$ there exist $C_1, C_2 > 0$ such that

$$e_n(M_d) \leqslant C_1 \cdot \exp(-C_2 \cdot n^p)$$

for all $n \in \mathbb{N}$. For all sequences $\boldsymbol{\sigma}$, Theorem 3.15 excludes uniform exponential convergence, and the lack of this property excludes exponential convergence polynomial tractability, see Kuo et al. (2017b, p. 832). We are thus led to the study of a weaker tractability concept, which is discussed in the following.

For $0 < \varepsilon < 1$ the information complexity for the absolute and the normalized error criterion are defined by

$$n^{\text{abs}}(\varepsilon, M_d) := \inf\{n \in \mathbb{N} \colon e_n(M_d) \leqslant \varepsilon\}$$

and

$$n^{\text{norm}}(\varepsilon, M_d) := \inf\{n \in \mathbb{N} \colon e_n(M_d)/e_0(M_d) \leqslant \varepsilon\},$$

respectively. Lemma 3.7 yields $n^{\text{abs}}(\varepsilon, M_d) \leqslant n^{\text{norm}}(\varepsilon, M_d)$, with equality if M_d is a Hermite kernel.

Let $t \geqslant 1$ and $\kappa > 0$. By definition, exponential convergence (t, κ)-weak tractability, for short EC-(t, κ)-WT, holds for $(M_d)_{d\in\mathbb{N}}$ and the normalized error criterion, if

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln(n^{\text{norm}}(\varepsilon, M_d))}{d^t + (\ln(\varepsilon^{-1}))^\kappa} = 0.$$

For Hermite spaces, the case $t = \kappa = 1$ is studied in Irrgeher et al. (2015) as well as Irrgeher et al. (2016b), while for Gaussian spaces, the case $t \geqslant 1$ and $\kappa \geqslant 1$ is studied in Kuo et al. (2017b).

First, we give a necessary condition and a sufficient condition for EC-$(1, 1)$-WT in the case of Gaussian spaces and the normalized error criterion. For non-increasing shape parameters the two conditions coincide and hence we even obtain a characterization of this tractability property.

Theorem 3.16. *Let $\boldsymbol{\sigma}$ be a sequence of shape parameters. If EC-$(1,1)$-WT holds for $(L_{\boldsymbol{\sigma}_d})_{d\in\mathbb{N}}$ and the normalized error criterion, then we have*

$$\inf_{j\in\mathbb{N}} \sigma_j = 0.$$

If $\boldsymbol{\sigma}$ is non-increasing and

$$\lim_{j\to\infty} \sigma_j = 0,$$

then EC-$(1,1)$-WT holds for $(L_{\boldsymbol{\sigma}_d})_{d\in\mathbb{N}}$ and the normalized error criterion.

Proof. For both statements, we make use of corresponding results for Hermite spaces, similarly to the proof of Theorem 3.15. Therefore, let $\boldsymbol{\beta}$ be the sequence of base parameters given by (3.11).

For the first statement, suppose that $\inf_{j\in\mathbb{N}} \sigma_j > 0$ holds. Then, we also have $\inf_{j\in\mathbb{N}} \beta_j > 0$, which means there exists a constant sequence $\widetilde{\boldsymbol{\beta}}$ of base parameters such that $\boldsymbol{\beta}$ is bounded from below by $\widetilde{\boldsymbol{\beta}}$. Thus, we may use Irrgeher et al. (2015, Thm. 1.3) in the specific case $\omega := 1/e$ as well as $a_j := \ln(1/\beta_j)$ and $b_j := 1$, which shows that EC-$(1,1)$-WT does not hold for $(K_{\tilde{\boldsymbol{\beta}}_d})_{d\in\mathbb{N}}$. Further, for all $d \in \mathbb{N}$ we have identical embeddings of $H(K_{\tilde{\boldsymbol{\beta}}_d})$ into $H(K_{\boldsymbol{\beta}_d})$ of norm one, cf. (2.47), which implies $e_n(K_{\tilde{\boldsymbol{\beta}}_d}) \leqslant e_n(K_{\boldsymbol{\beta}_d})$ for all $d, n \in \mathbb{N}$. Thus, EC-$(1,1)$-WT does not hold for $(K_{\boldsymbol{\beta}_d})_{d\in\mathbb{N}}$. Together with Corollary 3.13, this shows the claim.

To show the second statement, we assume that $\boldsymbol{\sigma}$ is non-increasing and converges to 0. Then, $\boldsymbol{\beta}$ is also non-increasing and converges to 0. Thus, by Irrgeher et al. (2016b, Thm. 3), again in the specific case $\omega := 1/e$ as well as $a_j := \ln(1/\beta_j)$ and $b_j := 1$, we have that EC-$(1,1)$-WT holds for $H(K_{\boldsymbol{\beta}_d})_{d\in\mathbb{N}}$. Again by Corollary 3.13, this shows the claim. □

Theorem 3.16 improves the previously best known sufficient condition for EC-$(1,1)$-WT for Gaussian spaces, namely, that the sequence of shape parameters is non-increasing and converges to 0 exponentially fast, see Kuo et al. (2017b, Thm. 1.1(e)). On the other hand, the necessity of $\boldsymbol{\sigma}$ converging to 0 was already known under the additional assumptions that the boundedness condition

$$\sup_{j\in\mathbb{N}} \sigma_j^2 < 1/2 \tag{3.17}$$

is satisfied and $\boldsymbol{\sigma}$ is non-increasing, see Kuo et al. (2017b, Thm. 1.1(f)). We remark that the first result for Hermite spaces we have used in the proof, Irrgeher et al. (2016b, Thm. 3), is fully constructive. Therefore, via the transference principle, the sufficient condition in Theorem 3.16 is also obtained constructively.

Our next goal is to study EC-(t,k)-WT in the case $t > 1$. As an intermediary step, we derive a new upper bound for $\ln(n^{\mathrm{norm}}(\varepsilon, L_{\boldsymbol{\sigma}_d}))$ with an explicit dependence on ε, d, and $\boldsymbol{\sigma}$. For $d \in \mathbb{N}$ we put

$$h(d) := \begin{cases} 0 & \text{if } d \leqslant 2, \\ \ln(\ln(d)) & \text{otherwise.} \end{cases}$$

Theorem 3.17. *There exists a constant $C > 0$ such that*

$$\ln(n^{\text{norm}}(\varepsilon, L_{\boldsymbol{\sigma}_d})) \leqslant C \cdot \left(d\,h(d) + d\ln(\ln(\varepsilon^{-1})) + \sum_{j=1}^{d} (\ln(\sigma_j))^{+} \right)$$

for every $d \in \mathbb{N}$, every $\varepsilon \in \,]0, 1/3]$, and every sequence $\boldsymbol{\sigma} := (\sigma_j)_{j\in\mathbb{N}}$ of shape parameters.

Proof. Put

$$\zeta_j := \ln\left(1 + \frac{1}{2\sigma_j^2}\right).$$

From Karvonen et al. (2021, Thm. 2.10) we get

$$e_N(L_{\boldsymbol{\sigma}_d}) \leqslant \sum_{j=1}^{d} (1 + 2\sigma_j^2)^{-1/2} \cdot \prod_{\substack{i=1 \\ i\neq j}}^{d} (1 + 4\sigma_i^2)^{-1/4} \cdot \exp(-n_j\zeta_j)$$

for all $d \in \mathbb{N}$ and $(n_1, \ldots, n_d) \in \mathbb{N}^d$ with $N := \prod_{j=1}^{d} n_j$. Together with Lemma 3.7 this implies

$$\frac{e_N(L_{\boldsymbol{\sigma}_d})}{e_0(L_{\boldsymbol{\sigma}_d})} \leqslant \sum_{j=1}^{d} (1 + 2\sigma_j^2)^{-1/2} \cdot (1 + 4\sigma_j^2)^{1/4} \cdot \exp(-n_j\zeta_j) \leqslant \sum_{j=1}^{d} \exp(-n_j\zeta_j).$$

Choosing

$$n_j := \left\lceil \frac{\ln(d/\varepsilon)}{\zeta_j} \right\rceil$$

we obtain $\exp(-n_j\zeta_j) \leqslant \varepsilon/d$ and therefore $n^{\text{norm}}(\varepsilon, L_{\boldsymbol{\sigma}_d}) \leqslant N$, i.e.,

$$n(\varepsilon, d) := n^{\text{norm}}(\varepsilon, L_{\boldsymbol{\sigma}_d}) \leqslant \prod_{j=1}^{d} \left(\frac{\ln(d/\varepsilon)}{\zeta_j} + 1 \right), \tag{3.18}$$

cf. Kuo et al. (2017b, Eqn. (3.9)).

For $z \geqslant 0$ we define

$$\varphi(z) := \ln(1 + z),$$

and we put $\tau := \ln(d/\varepsilon) \geqslant 0$ as well as

$$J_1 := \{j \in \{1,\ldots,d\} \colon \tau \leqslant 2\zeta_j\} \qquad \text{and} \qquad J_2 := \{1,\ldots,d\} \backslash J_1.$$

Clearly, $\varphi(z) \leqslant z$ for every $z \geqslant 0$ and $\varphi(z) \leqslant 3\ln(z)$ if $z > 2$, and therefore

$$\begin{aligned}\ln(n(\varepsilon,d)) &\leqslant \sum_{j=1}^{d} \ln(\tau/\zeta_j + 1) \leqslant \sum_{j\in J_1} \tau/\zeta_j + 3\sum_{j\in J_2} \ln(\tau/\zeta_j)\\ &\leqslant 2d + 3\sum_{j\in J_2} \ln(\tau/\zeta_j).\end{aligned}$$

In the sequel, we assume that $\varepsilon \in \,]0,1/3]$, so that $\ln(\tau) \geqslant \ln(\ln(3)) > 0$. With $c_1 := 2/\ln(\ln(3)) + 3$ we obtain

$$\ln(n(\varepsilon,d)) \leqslant c_1 d \ln(\tau) + 3\sum_{j\in J_2} \ln(1/\zeta_j). \tag{3.19}$$

Let us consider the particular case that

$$\inf_{j\in\mathbb{N}} \sigma_j \geqslant 2.$$

We put

$$\psi(y) := \ln\big(1 + 1/(2y^2)\big)$$

for $y \geqslant 2$, so that $\zeta_j = \psi(\sigma_j)$. Since $\lim_{y\to\infty} y^2 \cdot \psi(y) = 1/2$, we have

$$c_2 := \inf_{y\geqslant 2} y^2 \cdot \psi(y) \in \,]0,1/2].$$

With $c_3 := 2 + \ln(1/c_2)/\ln(2)$ we obtain

$$\ln(1/\psi(y)) \leqslant \ln(y^2/c_2) = \ln(y)\cdot(2 + \ln(1/c_2)/\ln(y)) \leqslant c_3 \ln(y)$$

for every $y \geqslant 2$. Consequently,

$$\sum_{j\in J_2} \ln(1/\zeta_j) = \sum_{j\in J_2} \ln\left((1/\psi(\sigma_j)\right) \leqslant c_3 \sum_{j=1}^{d} \ln(\sigma_j).$$

Together with (3.19), this implies

$$\ln(n(\varepsilon, d)) \leqslant c_1 d \ln(\tau) + 3c_3 \sum_{j=1}^{d} \ln(\sigma_j).$$

Observe that the upper bound in (3.18) is a monotonically increasing function in each of the variables σ_j. For any sequence of shape parameters we therefore have

$$\ln(n(\varepsilon, d)) \leqslant c_1 d \ln(\tau) + 3c_3 \sum_{j=1}^{d} \ln(\max(\sigma_j, 2)).$$

Since

$$\ln(\max(\sigma_j, 2)) \leqslant (\ln(\sigma_j))^+ + 1,$$

we obtain

$$\ln(n(\varepsilon, d)) \leqslant c_1 d \ln(\tau) + 3c_3 d + 3c_3 \sum_{j=1}^{d} (\ln(\sigma_j))^+.$$

Finally, there exists a constant $c_4 > 0$ such that

$$\begin{aligned}\ln(\tau) &= \ln\left(\ln(d) + \ln(\varepsilon^{-1})\right) \leqslant \ln\left(2\max(\ln(d), \ln(\varepsilon^{-1}))\right) \\ &\leqslant c_4 \ln\left(\max(\ln(d), \ln(\varepsilon^{-1}))\right) \leqslant c_4\left(h(d) + \ln(\ln(\varepsilon^{-1}))\right)\end{aligned}$$

for all $d \in \mathbb{N}$ and $\varepsilon \in \,]0, 1/3]$. □

Next, we present a sufficient condition for EC-(t,k)-WT to hold for some $t > 1$ and every $\kappa > 0$.

Theorem 3.18. *If there exist $c, \alpha \geqslant 0$ such that the shape parameters satisfy*

$$\sigma_j \leqslant \exp(c j^{\alpha})$$

for all $j \in \mathbb{N}$, then we have EC-(t,κ)-WT for $(L_{\boldsymbol{\sigma}_d})_{d \in \mathbb{N}}$ and the normalized and the absolute error criterion for every $t > \alpha + 1$ and every $\kappa > 0$.

Proof. It suffices to consider the normalized error criterion. We apply the upper bound for

$$n(\varepsilon, d) := n^{\mathrm{norm}}(\min(\varepsilon, 1/3), L_{\boldsymbol{\sigma}_d}) \geqslant n^{\mathrm{norm}}(\varepsilon, L_{\boldsymbol{\sigma}_d}).$$

from Theorem 3.17. Clearly, we have

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{d\,h(d) + d\ln(\ln(\max(\varepsilon^{-1}, 3)))}{d^t + \ln(\varepsilon^{-1})^\kappa} = 0$$

even for $t > 1$, so it suffices to show

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\sum_{j=1}^d (\ln(\sigma_j))^+}{d^t + \ln(\varepsilon^{-1})^\kappa} = 0$$

for $t > \alpha + 1$. This is indeed true, since

$$\sum_{j=1}^d (\ln(\sigma_j))^+ \leqslant \sum_{j=1}^d c j^\alpha \leqslant c d^{\alpha+1}. \qquad \square$$

Theorem 3.18 extends the result established in Kuo et al. (2017b, Thm. 1.1.(h)), where the boundedness condition (3.17) is shown to imply EC-(t,κ)-WT for every $t > 1$ and every $\kappa \geqslant 1$ under the additional assumption that $\boldsymbol{\sigma}$ is non-increasing. Theorem 3.18 yields, in particular, that the latter already holds if $(\sigma_j)_{j\in\mathbb{N}}$ is polynomially bounded.

Next, we apply the transference result to obtain new results for Hermite spaces, where the normalized and the absolute error criterion coincide.

At first, corresponding to Theorem 3.17, we provide a new upper bound for $\ln(n^{\mathrm{norm}}(\varepsilon, K_{\boldsymbol{\beta}_d}))$ with an explicit dependence on ε, d, and $\boldsymbol{\beta}$.

Theorem 3.19. *There exists a constant $C > 0$ such that*

$$\ln(n^{\mathrm{norm}}(\varepsilon, K_{\boldsymbol{\beta}_d})) \leqslant C \cdot \left(d\,h(d) + d\ln(\ln(\varepsilon^{-1})) - \sum_{j=1}^d \ln(1-\beta_j) \right)$$

for every $d \in \mathbb{N}$, every $\varepsilon \in \,]0, 1/3]$, and every sequence $\boldsymbol{\beta} := (\beta_j)_{j\in\mathbb{N}}$ of base parameters in $(0,1)$.

Proof. Since (3.11) implies $2\sigma_j^2 < 1/(1-\beta_j)$, the statement follows from Corollary 3.13

and Theorem 3.17. □

Next, corresponding to Theorem 3.18, we give a sufficient condition for EC-$(t,\kappa)-WT$ to hold for some $t > 1$ and every $\kappa > 0$ in the case of Hermite spaces.

Theorem 3.20. *If there exist $c, \alpha \geqslant 0$ such that the base parameters satisfy*

$$\beta_j \leqslant 1 - \exp(-cj^{\alpha})$$

for all $j \in \mathbb{N}$, then we have EC-(t,κ)-WT for $(K_{\boldsymbol{\beta}_d})_{d\in\mathbb{N}}$ and the normalized error criterion for every $t > \alpha + 1$ and every $\kappa \geqslant 1$.

Proof. The assumption on the shape parameters β_j implies $-\ln(1-\beta_j) \leqslant cj^{\alpha}$. Thus we may proceed as in the proof of Theorem 3.18 to verify the claim. □

Remark 3.21. On the one hand, rather mild assumptions on the shape parameters or the base parameters imply EC-(t,κ)-WT for every $t > 1$ regardless of choice of κ, see Theorems 3.18 and 3.20. On the other hand, in the case $t = 1$, much stronger properties are required even for EC-$(t,1)$-WT to hold, see Theorem 3.16 and Irrgeher et al. (2015, Thm. 1.3). We add that necessary conditions for EC-(t,κ)-WT are only known in the case $t = 1$ and $0 < \kappa \leqslant 1$.

3.2.5 Infinite-Variate Integration

In this section, we consider the case $d = \infty$. For $M = L_{\boldsymbol{\sigma}}$ or $M = K_{\boldsymbol{\beta}}$, we study the decay $\mathrm{dec}(M)$ of the n-th minimal worst-case errors, see (1.18) and (1.19). Note that a lower bound for $\mathrm{dec}(M)$ corresponds to an upper bound for the n-th minimal errors $e_n(M)$ and vice versa. Our goal is to determine $\mathrm{dec}(L_{\boldsymbol{\sigma}})$, and to this end we put

$$\rho := \liminf_{j\to\infty} \frac{\ln(1/\sigma_j^2)}{\ln(j)}. \tag{3.20}$$

Note that (3.3) implies $\rho \geqslant 1$, see, e.g., Gnewuch et al. (2019, Lemma B.3).

For the cost of an algorithm, we employ the unrestricted subspace sampling model, cf. Section 1.3. Specifically, we choose the default value $a = 0$.

For the following main result on infinite-variate integration, cf. Gnewuch et al. (2026, Thm. 4.14).

Theorem 3.22. *Assume that* (3.3) *is satisfied. The polynomial decay rate of* n*-th minimal errors of integration on the Gaussian space* $H(L_{\boldsymbol{\sigma}})$ *is*

$$\operatorname{dec}(L_{\boldsymbol{\sigma}}) = \frac{1}{2}(\rho - 1).$$

Proof. Consider the sequence $\boldsymbol{\beta}$ of base parameters given by (3.11). Corollary 3.13 immediately yields

$$\operatorname{dec}(L_{\boldsymbol{\sigma}}) = \operatorname{dec}(K_{\boldsymbol{\beta}}) \tag{3.21}$$

and thus it suffices to show $\operatorname{dec}(K_{\boldsymbol{\beta}}) = \frac{1}{2}(\rho - 1)$. This follows from Corollary 2.22 in the case (EG).

Indeed, we already established in Section 2.1.2 that $\boldsymbol{\beta}$ gives rise to Fourier weights $\alpha_{\nu,j} = \beta_j^{-\nu}$ fulfilling (EG), since $\beta_j^{-\nu} = 2^{r_j \cdot \nu}$ for $\nu, j \in \mathbb{N}$, where we set $r_j := -\log_2(\beta_j)$, and possibly reordering the kernels so that $r_1 = \inf_{j \in \mathbb{N}} r_j$.

Hence we obtain from Corollary 2.22 that the polynomial decay rate of n-th minimal errors of integration on the Hermite space $H(K_{\boldsymbol{\beta}})$ is

$$\operatorname{dec}(K_{\boldsymbol{\beta}}) = \frac{1}{2}(\widetilde{\rho} - 1),$$

where

$$\widetilde{\rho} := \liminf_{j \to \infty} \frac{r_j \cdot \ln(2)}{\ln(j)}.$$

Given (3.11) we have $\beta_j \asymp \sigma_j^2$, and therefore

$$\rho = \liminf_{j \to \infty} \frac{\ln(1/\beta_j)}{\ln(j)} = \widetilde{\rho}. \qquad \square$$

Algorithms that yield $\operatorname{dec}(L_{\boldsymbol{\sigma}}) \geqslant (\rho - 1)/2$ are obtained by applying the transference result to properly chosen multivariate decomposition methods (MDMs) on the space $H(K_{\boldsymbol{\beta}})$. We roughly sketched the construction of MDMs in a more general setting in Section 2.3.2. See Gnewuch et al. (2024, Sec. 4.3) for a more detailed sketch in the case of Hermite spaces.

Comparing the result of Theorem 3.22 to Remark 3.8, we observe that if a sequence $\boldsymbol{\sigma}$ fulfilling (3.3) but decaying relatively slowly, so that ρ approaches one, lead to $\operatorname{dec}(L_{\boldsymbol{\sigma}})$ being close to zero, which corresponds to the integration problem being harder than if $\boldsymbol{\sigma}$ would decay faster. On the other hand, if $\boldsymbol{\sigma}$ decays even slower, so that (3.3) is not fulfilled anymore, the integration problem becomes trivial. This might seem counterintuitive. However, recall Corollary 3.13, which states that for fixed n, we actually have that $e_n(L_{\boldsymbol{\sigma}})$ is smaller than $e_n(K_{\boldsymbol{\beta}})$ by a factor of $e_0(L_{\boldsymbol{\sigma}})$. If ρ approaches one, this constant factor becomes very large, which corresponds to the integration problem becoming easier.

3.3 L^2-Approximation

As before, we consider sequences $\boldsymbol{\sigma} := (\sigma_j)_{j\in J}$ and $\boldsymbol{\beta} := (\beta)_{j\in J}$ of shape parameters and base parameters, respectively.

3.3.1 The Norm of the L^2-Approximation Operator

As in the case of the integration problem, we first discuss the norm of the L^2-approximation problem, or equivalently $e_0(M)$ for $M = L_{\boldsymbol{\sigma}}$ and $M = K_{\boldsymbol{\beta}}$, since it will appear as a constant later.

Lemma 3.23. *For the L^2-approximation problem. if* (3.3) *is satisfied then*

$$e_0(L_{\boldsymbol{\sigma}}) = \prod_{j\in J} \frac{\sqrt{2}}{\left(1+\left(1+8\sigma_j^2\right)^{1/2}\right)^{1/2}}.$$

If (3.5) *is satisfied then*

$$e_0(K_{\boldsymbol{\beta}}) = 1.$$

Proof. As in Lemma 3.7, we make use of the tensor product structure of the problem, along with known results in the case $d = 1$, to obtain the result. In the case $d = 1$, we have

$$e_0^2(\ell_{\sigma_j}) = \frac{2}{1+\left(1+8\sigma_j^2\right)^{1/2}} < 1,$$

see, e.g., Sloan and Woźniakowski (2018, Eqn. (10)), and $e_0(k_{\beta_j}) = 1$, see Irrgeher et al. (2016b, p. 104). Consequently,

$$e_0(M) = \prod_{j \in J} e_0(m_j)$$

with $m_j := \ell_{\sigma_j}$ or with $m_j := k_{\beta_j}$. □

Similarly to Remark 3.8, one can show that if (3.3) does not hold, the L^2-approximation problem is trivial.

3.3.2 The Transference Result

In contrast to integration, we may use the relation (3.6) directly to establish a transference result for L^2-approximation in a straightforward way. For this main result, cf. Gnewuch et al. (2026, Thm. 5.1).

Theorem 3.24. *Assume that* (3.3) *is satisfied. Moreover, assume that* (3.6) *and* (3.7) *are satisfied for every* $j \in J$*. For every linear sampling method* A *with nodes from* $\ell^2(\boldsymbol{\sigma}^2)$ *we have*

$$e(A, L_{\boldsymbol{\sigma}}) = (\mathbf{1} - \boldsymbol{\beta})_*^{1/2} \cdot e(Q_{\boldsymbol{c}}^{-1} A Q_{\boldsymbol{c}}, K_{\boldsymbol{\beta}}).$$

Proof. We apply Theorem 3.5. Put $Q := Q_{\boldsymbol{c}}$. Since $Q_{|H(K_{\boldsymbol{\beta}})}$ is an isomorphism from $H(K_{\boldsymbol{\beta}})$ to $H(L_{\boldsymbol{\sigma}})$ with

$$\|Qf\|_{H(L_{\boldsymbol{\sigma}})}^2 = \|f\|_{H((\mathbf{1}-\boldsymbol{\beta})_* K_{\boldsymbol{\beta}})}^2 = (\mathbf{1} - \boldsymbol{\beta})_*^{-1} \cdot \|f\|_{H(K_{\boldsymbol{\beta}})}^2$$

for every $f \in H(K_{\boldsymbol{\beta}})$, we have

$$e(A, L_{\boldsymbol{\sigma}}) = (\mathbf{1} - \boldsymbol{\beta})_*^{1/2} \cdot \sup_{\|f\|_{H(K_{\boldsymbol{\beta}})} \leqslant 1} \|Qf - AQf\|_{L^2(\mu)}.$$

Furthermore,

$$\|Qf - AQf\|_{L^2(\mu)} = \|f - Q^{-1} AQf\|_{L^2(\mu)},$$

since Q is an isometric isomorphism on $L^2(\mu)$. □

Remark 3.25. Let the assumptions from Theorem 3.24 be satisfied. For a linear sampling method A on the Gaussian space $H(L_{\boldsymbol{\sigma}})$ according to (1.15) with nodes from $\ell^2(\boldsymbol{\sigma})$ the linear sampling method $Q_{\boldsymbol{c}}^{-1} A Q_{\boldsymbol{c}}$ on the Hermite space $H(K_{\boldsymbol{\beta}})$ is easily determined explicitly: For $f \colon \ell^2(\boldsymbol{\sigma}^2) \to \mathbb{R}$ we have

$$Q_{\boldsymbol{c}}^{-1} A Q_{\boldsymbol{c}}(f) = \sum_{i=1}^{n} Q_{\boldsymbol{c}} f(\boldsymbol{x}_i) \cdot Q_{\boldsymbol{c}}^{-1} a_i = \sum_{i=1}^{n} f(\boldsymbol{y}_i) \cdot b_i$$

with $\boldsymbol{y}_i \in \ell^2(\boldsymbol{\sigma}^2)$ and $b_i \in L^2(\mu)$ given by

$$\boldsymbol{y}_i := \boldsymbol{c}\boldsymbol{x}_i \qquad \text{and} \qquad b_i(\boldsymbol{z}) := \frac{\varphi_{\boldsymbol{c}}(\boldsymbol{x}_i)}{\varphi_{\boldsymbol{c}}(\boldsymbol{c}^{-1}\boldsymbol{z})} \cdot a_i(\boldsymbol{c}^{-1}\boldsymbol{z}),$$

see Lemma 3.6.

It is easy to see that $(\boldsymbol{x}_i, a_i) \mapsto (\boldsymbol{y}_i, b_i)$ defines a bijection on the set $\ell^2(\boldsymbol{\sigma}^2) \times L^2(\mu)$ of pairs of nodes and coefficients. Furthermore, in the case $d \in \mathbb{N}$, and in the case $d = \infty$ for the unrestricted subspace sampling model with default value $a = 0$, we have

$$\operatorname{cost}(A) = \operatorname{cost}(Q_{\boldsymbol{c}}^{-1} A Q_{\boldsymbol{c}}),$$

cf. Remark 3.12.

Corollary 3.26. *Assume that* (3.3) *and* (3.6) *are satisfied. In the cost model cases as given in Remark 3.25, for every $n \in \mathbb{N}$ we have*

$$\frac{e_n(L_{\boldsymbol{\sigma}})}{e_0(L_{\boldsymbol{\sigma}})} = \frac{e_n(K_{\boldsymbol{\beta}})}{e_0(K_{\boldsymbol{\beta}})} = e_n(K_{\boldsymbol{\beta}}),$$

i.e., the normalized n-th minimal errors for L^2-approximation on Gaussian spaces and on the corresponding Hermite spaces, related by (3.6)*, coincide.*

Proof. Combine Theorem 3.24, Remark 3.25, and Lemma 3.23. □

3.3.3 Multivariate L^2-Approximation

L^2-approximation in the case of finitely many variables has been studied previously by Fasshauer et al. (2012) and Sloan and Woźniakowski (2018) for Gaussian spaces and by Irrgeher et al. (2016a) and Irrgeher et al. (2016b) for Hermite spaces. While the latter paper employs linear sampling methods as described in Section 1.3, Irrgeher et al. (2016a) and Sloan and Woźniakowski (2018) allow a larger class of algorithms, namely linear algorithms A based on linear information, i.e.,

$$A(f) := \sum_{i=1}^{n} \lambda_i(f) \cdot a_i \tag{3.22}$$

with $n \in \mathbb{N}$, bounded linear functionals λ_i on $H(M)$ and coefficients $a_i \in L^2(\mu)$; Fasshauer et al. (2012) consider both classes of algorithms. In the case of general linear algorithms, the same tractability concepts have been studied on the one hand in Fasshauer et al. (2012) and Sloan and Woźniakowski (2018) for Gaussian spaces, and on the other hand in Irrgeher et al. (2016a) for Hermite spaces, with rather complete and matching results for both types of function space.

It was already observed in Irrgeher et al. (2016b) that the linear sampling methods studied there achieve the same results as algorithms of the form (3.22) for several tractability concepts.

Based on the previous work from Dolbeault et al. (2023), more general tractability results were established in Krieg et al. (2023). In particular, general linear information and function evaluation do not lead to different tractability results for weak tractability with exponential convergence or most of the other tractability notions considered in Irrgeher et al. (2016a), Irrgeher et al. (2016b), Fasshauer et al. (2012) and Sloan and Woźniakowski (2018).

This leads to already rather complete results, though in a non constructive way. Therefore, we do not apply our transference result to this case here. Our results for infinitely many variables, however, are new.

3.3.4 Infinite-Variate L^2-Approximation

As in Section 3.2.5, we consider ρ given by (3.20) and consider the unrestricted subspace sampling model with default value $a = 0$.

The next main result, cf. Gnewuch et al. (2026, Thm. 5.4), follows easily from Corollary 2.22 by means of our transference result.

Theorem 3.27. *Assume that* (3.3) *is satisfied. The polynomial decay rate of n-th minimal errors of L^2-approximation on the Gaussian space $H(L_{\boldsymbol{\sigma}})$ is*

$$\operatorname{dec}(L_\sigma) = \frac{1}{2}(\rho - 1).$$

Proof. We proceed as in the proof of Theorem 3.22: By Corollary 3.26, we only have to consider $\operatorname{dec}(K_{\boldsymbol{\beta}})$, and $\operatorname{dec}(K_{\boldsymbol{\beta}})$ is known for L^2-approximation, too. This time, however, the sequence $\boldsymbol{\beta} = (\beta_j)_{j\in\mathbb{N}}$ is defined as in (3.6), but again $\beta_j \asymp \sigma_j^2$ and the weights $\beta_j^{-\nu} = 2^{r_j \cdot \nu}$ with $\nu, j \in \mathbb{N}$ are exponentially growing Fourier weights in the sense of Section 2.1.2, possibly after reordering the kernels. As before,

$$\rho = \liminf_{j\to\infty} \frac{\ln(1/\beta_j)}{\ln(j)} = \liminf_{j\to\infty} \frac{r_j \cdot \ln(2)}{\ln(j)}.$$

Hence we obtain from Corollary 2.22 that the polynomial decay rate of n-th minimal errors of L^2-approximation on the Hermite space $H(K_{\boldsymbol{\beta}})$ is

$$\operatorname{dec}(K_{\boldsymbol{\beta}}) = \frac{1}{2}(\rho - 1). \qquad \square$$

Appendix A

Tensor Products of Hilbert Spaces

Throughout Chapters 1-3, we make use of the countable product of Hilbert spaces in an abstract setting. Here, we give a brief introduction into this setting and give some basic results.

We stress that most of the results found here were already established in von Neumann (1939) and, based on that, in the author's master's thesis Rüßmann (2020). These results are only included here for the reader's convenience and orientation. However, we added Theorem A.3, which is a generalization of a previous result.

A.1 Tensor Products of Arbitrary Hilbert Spaces

In this section we consider Hilbert spaces $H_j \neq \{0\}$ with $j \in \mathbb{N}$ over the same scalar field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Our first goal is to construct the *incomplete tensor product*; see von Neumann (1939) or Rüßmann (2020) for a thorough study. For a more comprehensive overview, see Gnewuch et al. (2022, App. A).

We define

$$\mathcal{H} = \bigtimes_{j \in \mathbb{N}} H_j$$

and fix a sequence $\boldsymbol{u} = (u_j)_{j \in \mathbb{N}} \in \mathcal{H}$ such that $\|u_j\|_{H_j} = 1$ for all $j \in \mathbb{N}$. Let

$$C^{(\boldsymbol{u})} = \left\{ \boldsymbol{f} \in \mathcal{H} \colon \sum_{j \in \mathbb{N}} \big| \|f_j\|_{H_j} - 1 \big| < \infty \text{ and } \sum_{j \in \mathbb{N}} \big| \langle u_j, f_j \rangle_{H_j} - 1 \big| < \infty \right\}.$$

We mention that for a second sequence $\boldsymbol{v} = (v_j)_{j\in\mathbb{N}} \in \mathcal{H}$ of unit vectors, we have

$$C^{(\boldsymbol{u})} = C^{(\boldsymbol{v})} \text{ iff } \sum_{j\in\mathbb{N}} |\langle u_j, v_j\rangle_{H_j} - 1| < \infty, \tag{A.1}$$

with $C^{(\boldsymbol{u})} \cap C^{(\boldsymbol{v})} = \varnothing$ otherwise. Then,

$$\mathcal{K}^{(\boldsymbol{u})}(\boldsymbol{g}, \boldsymbol{f}) = \prod_{j\in\mathbb{N}} \langle f_j, g_j\rangle_{H_j}$$

is well-defined for all $\boldsymbol{f}, \boldsymbol{g} \in C^{(\boldsymbol{u})}$ and

$$\mathcal{K}^{(\boldsymbol{u})} \colon C^{(\boldsymbol{u})} \times C^{(\boldsymbol{u})} \to \mathbb{K}$$

is a reproducing kernel. We define the incomplete tensor product of the spaces H_j by

$$H^{(\boldsymbol{u})} := \bigotimes_{j\in\mathbb{N}} H_j^{(u_j)} = H(\mathcal{K}^{(\boldsymbol{u})}).$$

For any $\boldsymbol{f} \in C^{(\boldsymbol{u})}$ the function $\bigotimes_{j\in\mathbb{N}} f_j := \mathcal{K}^{(\boldsymbol{u})}(\cdot, \boldsymbol{f})$ is called an elementary tensor. Obviously, $\| \bigotimes_{j\in\mathbb{N}} f_j\|_{H^{(\boldsymbol{u})}} = \prod_{j\in\mathbb{N}} \|f_j\|_{H_j}$, and the span of the elementary tensors is dense in $H^{(\boldsymbol{u})}$. We denote the set of all sequences $\boldsymbol{f} \in \mathcal{H}$ with only differ from u_j finitely often by $C^{*(\boldsymbol{u})}$. Perhaps less obvious, even the span of those elementary tensors based on sequences $\boldsymbol{f} \in C^{*(\boldsymbol{u})}$ is dense in $H^{(\boldsymbol{u})}$, see, e.g., Rüßmann (2020, Prop. 2.39).

We mention that for any $d \in \mathbb{N}$, the finite tensor product $\bigotimes_{j=1}^d H_j$ is isometrically embedded into $H^{(\boldsymbol{u})}$ by mapping any elementary tensor $\bigotimes_{j=1}^d f_j$ to the elementary tensor $\bigotimes_{j\in\mathbb{N}} g_j$, where $g_j = f_j$ for $j \leqslant d$ and $g_j = u_j$ otherwise.

Remark A.1. Consider any choice of orthonormal bases $(e_{\eta,j})_{\eta\in N_j}$ in each of the spaces H_j. Assuming $0 \in N_j$ for notational convenience, we require that

$$e_{0,j} = u_j$$

for every $j \in \mathbb{N}$. Let $\boldsymbol{N}$ denote the set of all sequences $\boldsymbol{\eta} := (\eta_j)_{j\in\mathbb{N}}$ with $\eta_j \in N_j$ for every $j \in \mathbb{N}$ and with $\{j \in \mathbb{N} \colon \eta_j \neq 0\}$ being finite. Then the elementary tensors

$$e_{\boldsymbol{\eta}} := \bigotimes_{j \in \mathbb{N}} e_{\eta_j, j} \tag{A.2}$$

with $\boldsymbol{\eta} \in \boldsymbol{N}$ form an orthonormal basis of $H^{(\boldsymbol{u})}$.

We denote by $\boldsymbol{U}$ the set of finite subsets of $\mathbb{N}$.

Remark A.2. Based on Remark A.1 we introduce an orthogonal decomposition of $H^{(\boldsymbol{u})}$. For $\boldsymbol{\nu} \in \boldsymbol{U}$, define $H_{\boldsymbol{\nu}}$ as the closed subspace determined by those $e_{\boldsymbol{\eta}}$ for which $\eta_j \neq 0$ if and only if $j \in \boldsymbol{\nu}$. Clearly, we have

$$H^{(\boldsymbol{u})} = \bigoplus_{\boldsymbol{\nu} \in \boldsymbol{U}} H_{\boldsymbol{\nu}}.$$

Further, for $\boldsymbol{\nu} \in \boldsymbol{U}$, the space $H_{\boldsymbol{\nu}}$ is isometrically isomorphic to the finite Hilbert space tensor product

$$\bigotimes_{j \in \boldsymbol{\nu}} \widetilde{H}_j,$$

where $\widetilde{H}_j$ is the orthogonal complement of $\operatorname{span}\{u_j\}$ in H_j.

We now present a new result regarding the tensor product of operators on Hilbert spaces. This slightly generalizes the result from Gnewuch et al. (2022, Sec A.4), making it more immediately useful in the context of this thesis.

Theorem A.3. *For $j \in \mathbb{N}$, let H_j, G_j be nonzero Hilbert spaces and $u_j \in H_j$ with $\|u_j\|_{H_j} = 1$ as well as $v_j \in G_j$ with $\|v_j\|_{G_j} = 1$. Further, for each $j \in \mathbb{N}$, let*

$$T_j \colon H_j \to G_j$$

be a bounded linear operator such that

$$\sum_{j=1}^{\infty} \left| \|T_j\| - 1 \right| < \infty, \tag{A.3}$$

$$\sum_{j=1}^{\infty} \left| \|T_j u_j\|_{G_j} - 1 \right| < \infty, \tag{A.4}$$

and

$$\sum_{j=1}^{\infty} |\|\langle T_j u_j, v_j\rangle_{G_j}\| - 1| < \infty. \tag{A.5}$$

Then, there exists exactly one bounded linear operator $T^{\otimes}\colon H^{(\boldsymbol{u})} \to G^{(\boldsymbol{v})}$ *fulfilling*

$$T^{\otimes} \bigotimes_{j\in\mathbb{N}} f_j = \bigotimes_{j\in\mathbb{N}} T_j f_j \tag{A.6}$$

for all $\boldsymbol{f} \in C^{(\boldsymbol{u})}$.

Proof. We denote the dense linear subspace F of $H^{(\boldsymbol{u})}$ given as the span of those elementary tensors $\bigotimes_{j\in\mathbb{N}} f_j$ for which $f_j \neq u_j$ only finitely many times. Then, (A.4) along with (A.5) ensure that $(T_j u_j)_{j\in\mathbb{N}}$ is a sequence in $C^{(\boldsymbol{v})}$, which is not changed by exchanging finitely many u_j for f_j. Therefore, we can define an operator $T\colon F \to G^{(\boldsymbol{v})}$ by setting

$$T(\bigotimes_{j\in\mathbb{N}} f_j) = \bigotimes_{j\in\mathbb{N}} (T_j f_j)$$

for $\boldsymbol{f} \in C^{*(\boldsymbol{u})}$ and extending linearly. The operator T is bounded, in fact we have

$$\|T\| = \prod_{j\in\mathbb{N}} \|T_j\|$$

which we will infer in the following from the corresponding statement for finite tensor products, see, e.g., Hackbusch (2012, Prop. 4.150).

To show $\|T\| \geqslant \prod_{j\in\mathbb{N}} \|T_j\|$, let $h_j \in H_j$ for each $j \in \mathbb{N}$ such that $\|h_j\|_{H_j} = 1$ and $\|T_j h_j\|_{G_j} = \|T_j\|$. For a fixed $d \in \mathbb{N}$, define the elementary tensor $\bigotimes_{j\in\mathbb{N}} f_j$, given by $f_j = h_j$ for $j \leqslant d$ and $f_j = u_j$ otherwise. Obviously, we have $\bigotimes_{j\in\mathbb{N}} f_j \in F$ with $\|\bigotimes_{j\in\mathbb{N}} f_j\|_{H^{(\boldsymbol{u})}} = 1$. Further, we have

$$\|T \bigotimes_{j\in\mathbb{N}} f_j\|_{G^{(\boldsymbol{v})}} = \prod_{j=1}^{d} \|T_j\| \cdot \prod_{j=d+1}^{\infty} \|T_j u_j\|_{G_j}.$$

By letting d tend to infinity bearing in mind (A.4), the inequality indeed holds.

For the other inequality, let

$$h = \sum_{i=1}^{n} \bigotimes_{j\in\mathbb{N}} f_{j,i} \in F$$

with $\|h\|_{H^{(\boldsymbol{u})}} = 1$. Here, we have $f_{j,i} \neq u_j$ only finitely often for each i. Hence, there exists a $d_0 \in \mathbb{N}$ such that $f_{j,i} = u_j$ for all $i \in \{1, \dots, n\}$ and $j > d_0$. Let $d \geqslant d_0$ and define $\tilde{h} \in \bigotimes_{j=1}^{d} H_j$ by

$$\tilde{h} = \sum_{i=1}^{n} \bigotimes_{j=1}^{d} f_{j,i},$$

and recall that by isometric embedding of $\bigotimes_{j=1}^{d} H_j$ into $H^{(\boldsymbol{u})}$, we have $\|\widetilde{h}\|_{\bigotimes_{j=1}^{d} H_j} = 1$. Thus, we have

$$\|\bigotimes_{j=1}^{d} T_j(\widetilde{h})\|_{\bigotimes_{j=1}^{d} G_j} \leqslant \prod_{j=1}^{d} \|T_j\|.$$

We define a sequence $\boldsymbol{w}$ of unit vectors $w_j = (T_j u_j)/\|T_j u_j\|_{G_j}$ in G_j. Then by (A.4) and (A.5) in connection with (A.1), we have $C^{(\boldsymbol{v})} = C^{(\boldsymbol{w})}$ and thereby $G^{(\boldsymbol{v})} = G^{(\boldsymbol{w})}$. We denote by $\overline{h}$ the image of the isometric embedding with respect to $\boldsymbol{w}$ of $\bigotimes_{j=1}^{d}(T_j)(\widetilde{h})$ into $G^{(\boldsymbol{w})}$, i.e.

$$\overline{h} = \sum_{i=1}^{n} \bigotimes_{j\in\mathbb{N}} g_{j,i}$$

with $g_{j,i} = T_j f_{j,i}$ for $j \leqslant d$ and $g_{j,i} = w_j$ otherwise. Then

$$\|Th\|_{G^{\boldsymbol{u}}} = \|\sum_{i=1}^{n} \bigotimes_{j\in\mathbb{N}} T_j h_j\|_{G^{\boldsymbol{u}}} = (\prod_{j=d+1}^{\infty} \|T_j u_j\|) \cdot \|\overline{h}\|_{G^{\boldsymbol{u}}} \leqslant (\prod_{j=d+1}^{\infty} \|T_j u_j\|) \cdot (\prod_{j=1}^{d} \|T_j\|)$$

and the desired inequality follows by letting d tend to ∞.

We obtain $T^{\otimes}$ by extending T to $H^{(\boldsymbol{u})}$ via the BLT-theorem. □

A.2 Tensor Products of Reproducing Kernel Hilbert Spaces

Subsequently, we consider the special case that $H_j := H(k_j)$ with reproducing kernels k_j on a shared domain D. To $\boldsymbol{x} \in D^{\mathbb{N}}$ we associate

$$\tau(\boldsymbol{x}) := (k_j(\cdot, x_j))_{j\in\mathbb{N}} \in \bigtimes_{j\in\mathbb{N}} H(k_j),$$

and we put

$$\mathfrak{X}^{(\boldsymbol{u})} := \{\boldsymbol{x} \in D^{\mathbb{N}} \colon \tau(\boldsymbol{x}) \in C^{(\boldsymbol{u})}\}.$$

If $\mathfrak{X}^{(\boldsymbol{u})} \neq \varnothing$ then

$$K^{(\boldsymbol{u})}(\boldsymbol{x}, \boldsymbol{y}) := \mathcal{K}^{(\boldsymbol{u})}(\tau(\boldsymbol{x}), \tau(\boldsymbol{y})) = \prod_{j\in\mathbb{N}} k_j(x_j, y_j)$$

with $\boldsymbol{x}, \boldsymbol{y} \in \mathfrak{X}^{(\boldsymbol{u})}$ yields a reproducing kernel

$$K^{(\boldsymbol{u})} \colon \mathfrak{X}^{(\boldsymbol{u})} \times \mathfrak{X}^{(\boldsymbol{u})} \to \mathbb{K}$$

of tensor product form.

In the sense of the following result the incomplete tensor product $H^{(\boldsymbol{u})}$ is the RKHS with tensor product kernel $K^{(\boldsymbol{u})}$. This result was first proven in Rüßmann (2020, Thm. 4.10), see also Gnewuch et al. (2022, Thm. A.6) for a more succinct version of the proof.

Theorem A.4. *If $\mathfrak{X}^{(\boldsymbol{u})} \neq \varnothing$ then*

$$\Phi \colon H^{(\boldsymbol{u})} \to \mathbb{K}^{\mathfrak{X}^{(\boldsymbol{u})}},$$

given by

$$\Phi g(\boldsymbol{x}) := g(\tau(\boldsymbol{x})), \qquad \boldsymbol{x} \in \mathfrak{X}^{(\boldsymbol{u})},$$

is an isometric isomorphism between $H^{(\boldsymbol{u})}$ and $H(K^{(\boldsymbol{u})})$. In particular, for $\boldsymbol{f} \in C^{(\boldsymbol{u})}$

and $\boldsymbol{x} \in \mathfrak{X}^{(\boldsymbol{u})}$ *the product* $\prod_{j\in\mathbb{N}} f_j(x_j)$ *converges and*

$$\left(\Phi \bigotimes_{j\in\mathbb{N}} f_j\right)(\boldsymbol{x}) = \prod_{j\in\mathbb{N}} f_j(x_j).$$

In addition to the reproducing kernels k_j and the unit vectors $u_j \in H(k_j)$ we consider reproducing kernels $\ell_j \colon D \times D \to \mathbb{K}$ such that $H(k_j) \subseteq H(\ell_j)$ and $\|u_j\|_{H(\ell_j)} = 1$ for every $j \in \mathbb{N}$. Moreover, we let

$$T_j \colon H(k_j) \hookrightarrow H(\ell_j)$$

denote the corresponding identical embedding. Analogously to $\mathfrak{X}^{(\boldsymbol{u})}$ and $K^{(\boldsymbol{u})}$ we consider the set $\mathfrak{Y}^{(\boldsymbol{u})}$ of all $\boldsymbol{x} \in D^{\mathbb{N}}$ such that

$$\sum_{j\in\mathbb{N}} |\ell_j(x_j, x_j) - 1| < \infty \quad \text{and} \quad \sum_{j\in\mathbb{N}} |u_j(x_j) - 1| < \infty$$

and the tensor product kernel

$$L^{(\boldsymbol{u})}(\boldsymbol{x}, \boldsymbol{y}) := \prod_{j\in\mathbb{N}} \ell_j(x_j, y_j), \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathfrak{Y}^{(\boldsymbol{u})},$$

assuming that $\mathfrak{Y}^{(\boldsymbol{u})} \neq \varnothing$.

We obtained the following result in Gnewuch et al. (2024, Thm. A.3).

Theorem A.5. *If* $\varnothing \neq \mathfrak{Y}^{(\boldsymbol{u})} \subseteq \mathfrak{X}^{(\boldsymbol{u})}$ *and* $\sum_{j\in\mathbb{N}} |\|T_j\| - 1| < \infty$ *then*

$$\{f|_{\mathfrak{Y}^{(\boldsymbol{u})}} \colon f \in H(K^{(\boldsymbol{u})})\} \subseteq H(L^{(\boldsymbol{u})}),$$

and the operator norm of the restriction $T \colon H(K^{(\boldsymbol{u})}) \to H(L^{(\boldsymbol{u})})$, $f \mapsto f|_{\mathfrak{Y}^{(\boldsymbol{u})}}$ *is given by*

$$\|T\| = \prod_{j\in\mathbb{N}} \|T_j\|.$$

Proof. The statement of the theorem is obtained via tensorization and application of

Theorem A.4. More precisely, let

$$H^{(\boldsymbol{u})} := \bigotimes_{j\in\mathbb{N}} (H(k_j))^{(u_j)}, \qquad G^{(\boldsymbol{u})} := \bigotimes_{j\in\mathbb{N}} (H(\ell_j))^{(u_j)},$$

and let $T^{\otimes}\colon H^{(\boldsymbol{u})} \to G^{(\boldsymbol{u})}$ denote the tensor product of the identical embeddings T_j. Note that

$$\|T^{\otimes}\| = \prod_{j\in\mathbb{N}} \|T_j\|, \tag{A.7}$$

which can easily be inferred from the corresponding statement for finite tensor products, see, e.g., Hackbusch (2012, Prop. 4.150) and cf. the proof of Theorem A.3. Moreover, let $\Phi\colon H^{(\boldsymbol{u})} \to H(K^{(\boldsymbol{u})})$ and $\Upsilon\colon G^{(\boldsymbol{u})} \to H(L^{(\boldsymbol{u})})$ denote the isometric isomorphisms according to Theorem A.4. It suffices to show that

$$\Upsilon \circ T^{\otimes} \circ \Phi^{-1} f = f|_{\mathfrak{Y}^{(\boldsymbol{u})}}$$

for every $f \in H(K^{(\boldsymbol{u})})$.

Consider any orthonormal bases $(e_{\boldsymbol{\eta}})_{\boldsymbol{\eta}\in\boldsymbol{N}}$ according to Remark A.1. Let

$$f := \sum_{\boldsymbol{\eta}\in\boldsymbol{N}} c_{\boldsymbol{\eta}} \cdot \Phi e_{\boldsymbol{\eta}} \in H(K^{(\boldsymbol{u})})$$

with $c_{\boldsymbol{\eta}} \in \mathbb{R}$ such that $\sum_{\boldsymbol{\eta}\in\boldsymbol{N}} |c_{\boldsymbol{\eta}}|^2 < \infty$. We obtain

$$\Upsilon \circ T^{\otimes} \circ \Phi^{-1} f(\boldsymbol{x}) = \sum_{\boldsymbol{\eta}\in\boldsymbol{N}} c_{\boldsymbol{\eta}} \cdot \Upsilon \circ T^{\otimes} e_{\boldsymbol{\eta}}(\boldsymbol{x})$$

for every $\boldsymbol{x} \in \mathfrak{Y}^{(\boldsymbol{u})}$. Furthermore,

$$T^{\otimes} e_{\boldsymbol{\eta}} = \bigotimes_{j\in\mathbb{N}} T_j e_{\nu_j,j} \in G^{(\boldsymbol{u})},$$

and therefore

$$\Upsilon \circ T^{\otimes} e_{\boldsymbol{\eta}}(\boldsymbol{x}) = \prod_{j\in\mathbb{N}} e_{\nu_j,j}(x_j) = \Phi e_{\boldsymbol{\eta}}(\boldsymbol{x})$$

for every $\boldsymbol{x} \in \mathfrak{Y}^{(\boldsymbol{u})}$. It follows that

$$\Upsilon \circ T^{\otimes} \circ \Phi^{-1} f(\boldsymbol{x}) = f(\boldsymbol{x})$$

for every $\boldsymbol{x} \in \mathfrak{Y}^{(\boldsymbol{u})}$, as claimed.

The statement on the operator norm of T follows from (A.7) as well as the fact that Φ and Υ are isometries. $\square$

Bibliography

N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

S. N. Chandler-Wilde, D. P. Hewett, and A. Moiola. Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples. *Mathematika*, 61:414–443, 2015.

J. Dick and M. Gnewuch. Infinite-dimensional integration in weighted Hilbert spaces: anchored decompositions, optimal deterministic algorithms, and higher order convergence. *Found. Comput. Math.*, 14:1027–1077, 2014.

J. Dick, C. Irrgeher, G. Leobacher, and F. Pillichshammer. On the optimal order of integration in Hermite spaces with finite smoothness. *SIAM J. Numer. Anal.*, 56:684–707, 2018.

M. Dolbeault, D. Krieg, and M. Ullrich. A sharp upper bound for sampling numbers in L_2. *Appl. and Comput. Harm. Ana.*, 63:113–134, 2023.

D. Dũng and V. K. Nguyen. Optimal numerical integration and approximation of functions on $\mathbb{R}^d$ equipped with Gaussian measure. *IMA J. Numer. Anal.*, 2023. drad051.

G. Fasshauer, F. J. Hickernell, and H. Woźniakowski. On dimension-independent rates of convergence for function approximation with Gaussian kernels. *SIAM J. Numer. Anal.*, 50:247–271, 2012.

M. Gnewuch. Lower error bounds for randomized multilevel and changing dimension algorithms. In J. Dick, F. Y. Kuo, G. W. Peters, and I. H. Sloan, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2013*, pages 399–415. Springer, Heidelberg, 2013.

M. Gnewuch and M. Wnuk. Explicit error bounds for randomized Smolyak algorithms and an application to infinite-dimensional integration. *J. Approx. Theory*, 251:105342, 2020.

M. Gnewuch, M. Hefter, A. Hinrichs, and K. Ritter. Embeddings of weighted Hilbert spaces and applications to multivariate and infinite-dimensional integration. *J. Approx. Theory*, 222:8–39, 2017.

M. Gnewuch, M. Hefter, A. Hinrichs, K. Ritter, and G. W. Wasilkowski. Embeddings for infinite-dimensional integration and L_2-approximation with increasing smoothness. *J. Complexity*, 54:101406, 1–32, 2019.

M. Gnewuch, M. Hefter, A. Hinrichs, and K. Ritter. Countable tensor products of Hermite spaces and spaces of Gaussian kernels. *J. Complexity*, 71:101654, 2022.

M. Gnewuch, A. Hinrichs, K. Ritter, and R. Rüßmann. Infinite-dimensional integration and L^2-approximation on Hermite spaces. *J. Approx. Theory*, 300:106027, 2024.

M. Gnewuch, K. Ritter, and R. Rüßmann. Multi- and infinite-variate integration and L^2-approximation on Hilbert spaces with Gaussian kernels. arXiv:2412.05368, to appear in Math. Comp., 2026.

W. Hackbusch. *Tensor spaces and numerical tensor calculus*, volume 42 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2012.

M. Hefter and K. Ritter. On embeddings of weighted tensor product Hilbert spaces. *J. Complexity*, 31(3):405–423, 2015.

F. J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$. *J. Complexity*, 26(3):229–254, 2010.

C. Irrgeher and G. Leobacher. High-dimensional integration on $\mathbb{R}^d$, weighted Hermite spaces, and orthogonal transforms. *J. Complexity*, 31:174–205, 2015.

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

C. Irrgeher, P. Kritzer, F. Pillichshammer, and H. Woźniakowski. Tractability of multivariate approximimation defined over Hilbert spaces with exponential weights. *J. Approx. Theory*, 207:301–338, 2016a.

C. Irrgeher, P. Kritzer, F. Pillichshammer, and H. Woźniakowski. Approximation in Hermite spaces of smooth functions. *J. Approx. Theory*, 207:98–126, 2016b.

T. Jahn and T. Ullrich. On the optimal constants in the two-sided Stechkin inequalities. *J. Approx. Theory*, 269:Paper No. 105607, 25, 2021.

T. Karvonen and S. Särkkä. Gaussian kernel quadrature at scaled Gauss-Hermite nodes. *BIT Numer. Math.*, 59:877–902, 2019.

T. Karvonen, C. J. Oates, and M. Girolami. Integration in reproducing kernel Hilbert spaces of Gaussian kernels. *Math. Comp.*, 90:2209–2223, 2021.

Y. Kazashi, Y. Suzuki, and T. Goda. Sub-optimality of Gauss-Hermite quadrature and optimality of trapezoidal rule for functions with finite smoothness. *SIAM J. Numer. Anal.*, 61 (3):1426–1448, 2023.

D. Krieg, P. Siedlecki, M. Ullrich, and H. Woźniakowski. Exponential tractability of l 2-approximation with function values. *Advances in Computational Mathematics*, 49(2):18, 2023.

F. Y. Kuo and H. Woźniakowski. Gauss-Hermite quadratures for functions from Hilbert spaces with Gaussian reproducing kernels. *BIT Numer. Math.*, 52:425–436, 2012.

F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski. On decompositions of multivariate functions. *Math. Comp.*, 79:953–966, 2010a.

F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski, and H. Woźniakowski. Liberating the dimension. *J. Complexity*, 26:422–454, 2010b.

F. Y. Kuo, D. Nuyens, L. Plaskota, I. H. Sloan, and G. W. Wasilkowski. Infinite-dimensional integration and the multivariate decomposition method. *J. Comput. Appl. Math.*, 326: 317–234, 2017a.

F. Y. Kuo, I. H. Sloan, and H. Woźniakowski. Multivariate integration for analytic functions with Gaussian kernels. *Math. Comp.*, 86:829–853, 2017b.

G. Leobacher, F. Pillichshammer, and A. Ebert. Tractability of L_2-approximation and integration in weighted Hermite spases of finite smoothness. *J. Complexity*, 78, 2023. 101768.

H. Q. Minh. Some properties of Gaussian reproducing kernel Hilbert spaces and their implications for function approximation and learning theory. *Constr. Approx.*, 32:307–338, 2010.

B. Niu, F. J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic multi-level algorithms for infinite-dimensional integration on $\mathbb{R}^{\mathbb{N}}$. *J. Complexity*, 27:331–351, 2011.

E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Vol. 1: Linear Information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.

E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 2: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.

E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 3: Standard information for operators*, volume 18 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2012.

V. I. Paulsen and M. Raghupathi. *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*. Cambridge Univ. Press, Cambridge, 2016.

L. Plaskota and G. W. Wasilkowski. Tractability of infinite-dimensional integration in the worst case and randomized settings. *J. Complexity*, 27:505–518, 2011.

H. Rabitz and Ö. F. Alış. General foundations of high-dimensional model representations. *J. Math. Chemistry*, 25:197–233, 1999.

R. Rüßmann. Tensor products of Hilbert spaces. Master's thesis, Department of Mathematics, TU Kaiserslautern, 2020.

I. H. Sloan and H. Woźniakowski. Multivariate approximation for analytic functions with Gaussian kernels. *J. Complexity*, 45:1–21, 2018.

I. Steinwart and A. Christmann. *Support vector machines*. Information Science and Statistics. Springer-Verlag, New York, 2008.

I. Steinwart and C. Scovel. Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr. Approx.*, 35:363–417, 2012.

I. Steinwart, D. Hush, and C. Scovel. An explicit description of the reproducing kernel Hilbert space of Gaussian RBF kernels. *IEEE Trans. Inform. Theory*, 52:4635–4663, 2006.

G. Szegö. *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

J. von Neumann. On infinite direct products. *Compositio Math.*, 6:1–77, 1939.

G. Wasilkowski and H. Woźniakowski. Liberating the dimension for function approximation. *Journal of Complexity*, 27(1):86–110, 2011.

G. W. Wasilkowski. Liberating the dimension for L_2-approximation. *J. Complexity*, 28: 304–319, 2012.

G. W. Wasilkowski. On tractability of linear tensor product problems for ∞-variate classes of functions. *J. Complexity*, 29:351–369, 2013.

G. W. Wasilkowski and H. Woźniakowski. Explicit cost bounds for algorithms for multivariate tensor product problems. *J. Complexity*, 11:1–56, 1995.

G. W. Wasilkowski and H. Woźniakowski. Liberating the dimension for function approximation: standard information. *J. Complexity*, 27:417–440, 2011.

Wissenschaftlicher Werdegang

2013	Abitur am Städtischen Gymnasium Schmallenberg
2013	Beginn des Studiums der Mathematik an der TU Kaiserslautern
2016	Bachelor of Science in Mathematik mit Anwendungsfach Informatik an der TU Kaiserslautern
2020	Master of Science in Mathematik mit Vertiefungsgebiet Stochastische Analysis und Anwendungsfach Informatik an der TU Kaiserslautern
2020	Wissenschaftlicher Mitarbeiter an der TU Kaiserslautern bzw. RPTU Kaiserslautern-Landau
2022	Beginn der Promotion in der Gruppe von Prof. Dr. Klaus Ritter

Scientific Career

2013	Abitur at the Städtisches Gymnasium Schmallenberg
2013	Starting to study Mathematics at TU Kaiserslautern
2016	Bachelor of Science in Mathematics with subsidiary subject Computer Science at TU Kaiserslautern
2020	Master of Science in Mathematics with specialization in Stochastic Analysis and subsidiary subject Computer Science at TU Kaiserslautern
2020	Scientific assistant at TU Kaiserslautern / RPTU Kaiserslautern-Landau
2022	Starting Ph.D.-program in the group of Prof. Dr. Klaus Ritter