MULTIVARIATE INTEGRATION FOR ANALYTIC FUNCTIONS WITH GAUSSIAN KERNELS

FRANCES Y. KUO, IAN H. SLOAN, AND HENRYK WOŹNIAKOWSKI

Abstract. We study multivariate integration of analytic functions defined on $\mathbb{R}^d$. These functions are assumed to belong to a reproducing kernel Hilbert space whose kernel is Gaussian, with nonincreasing shape parameters. We prove that a tensor product algorithm based on the univariate Gauss-Hermite quadrature rules enjoys exponential convergence and computes an ε-approximation for the d-variate integration using an order of $(\ln \varepsilon^{-1})^d$ function values as ε goes to zero. We prove that the exponent d is sharp by proving a lower bound on the minimal (worst case) error of any algorithm based on finitely many function values. We also consider four notions of tractability describing how the minimal number $n(\varepsilon, d)$ of function values needed to find an ε-approximation in the d-variate case behaves as a function of d and $\ln \varepsilon^{-1}$. One of these notions is new. In particular, we prove that for all positive shape parameters, the minimal number $n(\varepsilon, d)$ is larger than any polynomial in d and $\ln \varepsilon^{-1}$ as d and ε^{-1} go to infinity. However, it is not exponential in d^t and $\ln \varepsilon^{-1}$ whenever $t > 1$.

1. Introduction

We study d-variate integration with (product) Gaussian functions. We assume that integrands are analytic functions defined over $\mathbb{R}^d$ and they are from the reproducing kernel Hilbert space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ whose reproducing kernel $\mathcal{K}_{d,\boldsymbol{\gamma}}$ is Gaussian and depends on the so-called shape parameters. We stress that the number of variables, d, can be arbitrary. We prove that a tensor product algorithm based on the univariate *Gauss-Hermite quadratures* enjoys exponential convergence and computes an ε-approximation for the d-variate integration using an order of $(\ln \varepsilon^{-1})^d$ function values as ε goes to zero. We prove that the exponent d is sharp by proving a lower bound on the minimal worst case error of any algorithm based on finitely many function values. We also consider four notions of *tractability* describing how the minimal number $n(\varepsilon, d)$ of function values needed to find an ε-approximation in the d-variate case behaves as a function of d and $\ln \varepsilon^{-1}$.

The function space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ has been used in numerous applications, see [3, 7, 8, 9, 17, 23, 34] for numerical computation, and [2, 4, 12, 22, 24, 28, 29, 32] for statistical learning, as well as [11] for engineering. More precisely, let

$$\mathcal{I}_d f := \int_{\mathbb{R}^d} f(\boldsymbol{x}) \prod_{j=1}^{d} \frac{\exp(-x_j^2)}{\sqrt{\pi}} \, \mathrm{d}\boldsymbol{x} \qquad \text{for } f \in H(\mathcal{K}_{d,\boldsymbol{\gamma}}), \tag{1.1}$$

2010 *Mathematics Subject Classification.* Primary 41A63, 41A99. Secondary 65D30.

where the reproducing kernel $\mathcal{K}_{d,\boldsymbol{\gamma}}$ of the Hilbert space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ is defined by weighted Gaussian kernels,

$$\mathcal{K}_{d,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y}) := \prod_{j=1}^{d} K_{\gamma_j}(x_j,y_j)\,, \qquad K_\gamma(x,y) := \exp(-\gamma^2(x-y)^2)\,, \tag{1.2}$$

for $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$, where $\boldsymbol{\gamma}=\{\gamma_j\}$ is a sequence of *shape parameters* satisfying

$$1 > \gamma_1^2 \ge \gamma_2^2 \ge \cdots > 0 \tag{1.3}$$

with the shape parameters not depending on d. The case of shape parameters greater or equal to one as well as depending on d is left open for future research.

The space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ defined this way is a tensor product of the univariate spaces $H(K_{\gamma_j})$ with reproducing kernels K_{γ_j} for $j=1,2,\ldots,d$. More details about the function spaces $H(K_{\gamma_j})$ and $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ will be given in the next section. Note that throughout this paper we use calligraphic notation to denote functions or operators in the multivariate case, while non-calligraphic notation is used for the univariate case.

Gauss-Hermite quadratures were analyzed in [17] for univariate integration in the space $H(K_\gamma)$. They are classical quadrature rules which sample the integrand at zeros of the orthogonal Hermite polynomials and with quadrature coefficients chosen to achieve the maximal order of exactness for polynomials. Intuitively, Gauss-Hermite quadratures may not be the sensible choice for this function space since the space does not contain any nonzero polynomials. Despite this, our results show that the choice of Gauss-Hermite quadratures is quite reasonable and nearly optimal. We add in passing that Gauss-Hermite quadratures are studied for another space of analytic functions containing all polynomials in the recent paper [14].

For the space $H(K_\gamma)$ with norm denoted by $\|\cdot\|_{H(K_\gamma)}$, it was proved in [17] that the (worst case) error for Gauss-Hermite quadratures Q_n which uses n function values is bounded by[1] (taking $\alpha=1/2$ in [17])

$$e(Q_n) := \sup_{f\in H(K_\gamma)} \frac{|\mathcal{I}_1 f - Q_n f|}{\|f\|_{H(K_\gamma)}} \le \beta_n\,\gamma^{2n}\,, \quad \text{with} \quad \beta_n := \frac{n!}{2^n\,(2n)!}\sqrt{\frac{(4n)!}{(2n)!}}\,. \tag{1.4}$$

Later we will show in Lemma 3.1 that $\{\beta_n\}$ is an almost constant sequence since

$$0.84\ldots = 2^{-1/4} < \beta_n \le \frac{\sqrt{3}}{2} = 0.86\ldots \qquad \text{for all} \quad n\in\mathbb{N} := \{1,2,\ldots\}\,.$$

Hence we have exponential convergence if $\gamma^2<1$, and this is why we assumed that all shape parameters satisfy $\gamma_j^2<1$ in (1.3).

The above result from [17] serves as an upper bound on the *nth minimal error* $e_{1,n}^{\min}$ i.e., the smallest error among all algorithms using n function evaluations for univariate integration in $H(K_\gamma)$. The precise definition of the nth minimal error $e_{d,n}^{\min}$ for the d-variate case (covering also the univariate case $e_{1,n}^{\min}$) is given in (3.1) and (3.2).

The main problem studied in this paper is multivariate integration with an arbitrary number d of variables. We obtain three new results:

[1] It was mistakenly claimed in [17] that the upper bound in (1.4) is the exact error, due to a false step in the lower bound argument.

(1) An upper bound for the nth minimal error for the d-variate case $e^{\rm min}_{d,n}$ based on the error of the tensor product algorithm of univariate Gauss-Hermite quadratures.
(2) A lower bound for the nth minimal error for the univariate case $e^{\rm min}_{1,n}$ which behaves similarly to the error of the Gauss-Hermite quadrature Q_n so that the Gauss-Hermite quadrature is nearly optimal.
(3) A lower bound for the nth minimal error for the d-variate case $e^{\rm min}_{d,n}$ which behaves similarly to the error of the tensor product algorithm of univariate Gaussian-Hermite quadratures so that this algorithm is nearly optimal.

Moreover, we study four notions of *tractability with exponential convergence*. The first three are as in [5, 6, 14, 15, 16], see also [18, 19, 20] for general tractability study. The fourth notion seems to be new and is related to the notions studied in [21, 25]. These notions are defined as follows.

Given $\varepsilon \in (0,1)$ and $d \ge 1$, the information complexity $n(\varepsilon, d)$ is defined as the minimal number of function values n required for the d-variate case so that the nth minimal error $e^{\rm min}_{d,n}$ is at most $\varepsilon\,{\rm CRI}_d$. For the absolute error criterion, we take ${\rm CRI}_d = 1$, whereas for the normalized error criterion, we take ${\rm CRI}_d = \|\mathcal{I}_d\|$, where $\|\mathcal{I}_d\|$ denotes the operator norm of the functional $\mathcal{I}_d$ in $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$. Observe that $\|\mathcal{I}_d\|$ is precisely the error of the zero algorithm, and therefore $\varepsilon\|\mathcal{I}_d\|$ means that we want to reduce the error of the zero algorithm by a factor ε. It will be shown that

$$\|\mathcal{I}_d\| = \prod_{j=1}^{d} (1 + 2\gamma_j^2)^{-1/4}. \tag{1.5}$$

Hence $\|\mathcal{I}_d\| < 1$ for all $\boldsymbol{\gamma}$. Furthermore $\|\mathcal{I}_d\|$ can be exponentially small in d for some $\boldsymbol{\gamma}$. This holds, in particular, for constant weights $\gamma_j^2 = \gamma_1^2$ for all j. Therefore, the normalized error criterion is harder (and sometimes much harder) than the absolute error criterion. Indeed, observe that for the absolute error criterion we have $n(\varepsilon, d) = 0$ if $\varepsilon \ge \|\mathcal{I}_d\|$ since the zero algorithm does the job. When $\|\mathcal{I}_d\|$ is exponentially small in d, we have $n(\varepsilon, d) \ge 1$ only if ε is exponentially small in d. On the other hand, for the normalized error criterion we have $n(\varepsilon, d) \ge 1$ for all $\varepsilon \in (0,1)$.

We specify what we mean by various notions of tractability with exponential convergence. We begin with the notions of exponential and uniform exponential convergence.

- *EXP*: *Exponential convergence* means that there exists a number $q \in (0,1)$ such that for all $d = 1, 2, \dots,$ there are positive numbers $C_{1,d}, C_{2,d}$ and p_d for which

$$e^{\rm min}_{d,n} \le C_{1,d}\, q^{(n/C_{2,d})^{p_d}} \qquad \text{for all } n \in \mathbb{N}. \tag{1.6}$$

  By the *exponent* of EXP we mean

$$p_d^* = \sup\{p_d \mid \text{for } p_d \text{ satisfying (1.6)}\}.$$

- *UEXP*: *Uniform exponential convergence* means that (1.6) holds for p_d independent of d, i.e., $p_d = p > 0$ for all d. In this case, the *exponent* of UEXP is

$$p^* = \sup\{p \mid \text{for } p_d = p \text{ satisfying (1.6) for all } d\}.$$

Obviously if (1.6) holds for some $q \in (0,1)$ it also holds for any $\tilde{q} \in (0,1)$ with appropriately redefined $C_{2,d}$.

The three known notions of tractability for the absolute and normalized error criteria are then defined as follows.

- *EC-SPT*: *Exponential convergence strong polynomial tractability* holds iff there are positive numbers C and τ such that

$$n(\varepsilon,d) \le C(1+\ln \varepsilon^{-1})^{\tau} \quad \text{for all } \varepsilon \in (0,1) \text{ and } d \in \mathbb{N}. \tag{1.7}$$

  The *exponent of EC-SPT* is $\tau^* = \sup\{\tau \mid \text{for } \tau \text{ satisfying (1.7)}\}$.
- *EC-PT*: *Exponential convergence polynomial tractability* holds iff there are positive numbers C and τ_1, τ_2 such that

$$n(\varepsilon,d) \le C\, d^{\,\tau_1}\, (1+\ln \varepsilon^{-1})^{\tau_2} \quad \text{for all } \varepsilon \in (0,1) \text{ and } d \in \mathbb{N}.$$

- *EC-WT*: *Exponential convergence weak tractability* holds iff

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln\, n(\varepsilon,d)}{d+\ln \varepsilon^{-1}} \;=\; 0\,.$$

  This means that the limit is zero for all possible ways in which $d+\varepsilon^{-1}$ goes to infinity.

We also propose a new notion of tractability which depends on two parameters (t,κ) such that $t,\kappa \ge 1$ and $\max(t,\kappa) > 1$, i.e., they are both at least one and at least one of them is larger than one. We have

- *EC-(t,κ)-WT*: *Exponential convergence (t,κ)-weak tractability* holds iff

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln\, n(\varepsilon,d)}{d^{\,t}+(\ln \varepsilon^{-1})^{\kappa}} \;=\; 0\,. \tag{1.8}$$

For $t=1$ and $\kappa>1$ this notion was defined and studied in [21]. For arbitrary (t,κ) and with $\ln \varepsilon^{-1}$ replaced by ε^{-1} this notion was defined and studied in [25].

Note that the case $t=\kappa=1$ is the same as EC-WT. That is why we assume that at least one of t and κ is larger than one to be sure that EC-(t,κ)-WT does not coincide with EC-WT.

EC-(t,κ)-WT means that the information complexity $n(\varepsilon,d)$ is *not* exponential in $d^{\,t}$ and $(\ln \varepsilon^{-1})^{\kappa}$. However, it may happen that $n(\varepsilon,d)$ is exponential in $d^{\,t_1}$ or $(\ln \varepsilon^{-1})^{\kappa_1}$ for $t_1 < t$ or $\kappa_1 < \kappa$. In particular, for $t>1$ we may have an exponential dependence on d which is called the curse of dimensionality. To illustrate this point, let us assume that $n(\varepsilon,d) = (2\,\varepsilon^{-1})^d$ so that the curse of dimensionality holds. It is easy to check that we have EC-(t,κ)-WT iff $t^{-1}+\kappa^{-1} < 1$. On the other hand, if $n(\varepsilon,d) = (2\,\ln\,\varepsilon^{-1})^d$ then the curse still holds but we have EC-(t,κ)-WT iff $t>1$. This means that if we insist that the curse of dimensionality does not hold then we have to take $t=1$ and we have the case studied in [21].

We list these notions of tractability from the most to the least demanding, i.e.,

$$\text{EC-SPT} \implies \text{EC-PT} \implies \text{EC-WT} \implies \text{EC-}(t,\kappa)\text{-WT}.$$

It is easy to see that EC-PT for the absolute or normalized error criterion implies UEXP and $p^* \ge 1/\tau_2$. Hence, the lack of UEXP implies the lack of EC-PT in the absolute and normalized error criteria.

The notions of tractability depend on the error criterion. Indeed, let $n^{\rm ABS}(\varepsilon,d)$ and $n^{\rm NOR}(\varepsilon,d)$ denote the information complexity $n(\varepsilon,d)$ for the absolute and normalized error criteria, respectively. Obviously,

$$n^{\rm NOR}(\varepsilon,d)=n^{\rm ABS}(\varepsilon\,\|\mathcal{I}_d\|,d),$$

where $\|\mathcal{I}_d\|$ is given by (1.5). Clearly, $\|\mathcal{I}_d\|$ is decreasing in d and therefore

$$\|\mathcal{I}_\infty\|:=\lim_{n\to\infty}\|\mathcal{I}_n\|<\|\mathcal{I}_d\|\le\|\mathcal{I}_1\|\qquad\text{for all } d\in\mathbb{N}.$$

Therefore

$$n^{\rm ABS}(\varepsilon\,\|\mathcal{I}_1\|,d)\,\le\,n^{\rm NOR}(\varepsilon,d)\,\le\,n^{\rm ABS}(\varepsilon\,\|\mathcal{I}_\infty\|,d)\qquad\text{for all } d\in\mathbb{N}.$$

It is easy to check that

$$\|\mathcal{I}_\infty\|>0\qquad\text{iff}\qquad\sum_{j=1}^\infty\gamma_j^2<\infty.$$

If $\|\mathcal{I}_\infty\|>0$ then it is also easy to verify that each tractability notion

EC-SPT, EC-PT, EC-WT and EC-(t,κ)-WT

is equivalent for the absolute and normalized error criteria. In this case, it is enough to study only one error criterion.

We study these notions of tractability and prove the following theorem.

Theorem 1.1. *Consider multivariate integration in the space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ with shape parameters $\boldsymbol{\gamma}=\{\gamma_j\}$ satisfying* (1.3)*. Then*

(a) *For all shape parameters, EXP holds.*
(b) *The exponent of EXP is $p_d^*=1/d$ for all $d\in\mathbb{N}$.*
(c) *For all shape parameters, UEXP does* not *hold. This implies the lack of EC-SPT and EC-PT for the absolute and normalized error criteria.*
(d) *For every $d\in\mathbb{N}$ and for the absolute and normalized error criteria, there are positive c_d, C_d and ε_d such that*

$$c_d\left(\frac{\ln\,\varepsilon^{-1}}{\ln\,\ln\,\varepsilon^{-1}}\right)^d\,\le\,n(\varepsilon,d)\,\le\,C_d\,\left(\ln\,\varepsilon^{-1}\right)^d\qquad\textit{for all}\ \ \varepsilon\in(0,\varepsilon_d).$$

The upper bound on $n(\varepsilon,d)$ is attained by the tensor product algorithm of univariate Gaussian-Hermite quadratures.
(e) *For some shape parameters, EC-WT holds for the absolute and normalized error criteria. In particular, this is the case for $\gamma_j^2=\mathcal{O}(\exp(-j^{\,\alpha}))$ with $\alpha>1$.*
(f) *For some shape parameters, EC-WT does not hold for either the absolute or normalized error criterion. In particular, this is the case when $\lim_{j\to\infty}\gamma_j^2>0$.*
(g) *For some shape parameters, EC-$(1,\kappa)$-WT holds for the absolute error criterion. In particular, this is the case when $\lim_{j\to\infty}\gamma_j^2>0$.*
(h) *For all shape parameters, EC-(t,κ)-WT holds for both the absolute and normalized error criteria whenever $t>1$.*

We now discuss the implications of our results presented in Theorem 1.1. First of all, they show that the tensor product of univariate Gauss-Hermite algorithms is nearly optimal for all d and all shape parameters. However, the information complexity, which is the minimal number of function values needed to compute an ε-approximation for the d-variate case, very much depends on d and the shape parameters, Modulo a double logarithm of ε^{-1}, the information complexity as a function of ε^{-1} is proportional to $(\ln \varepsilon^{-1})^{\,d}$ for all shape parameters. We stress that the exponent d of the last bound is sharp. For relatively small d, this is a positive result and we can solve the multivariate integration problem with a relatively small cost. The situation is quite different if d becomes large. Then, although exponential convergence is still present, its exponent decays to zero. This has many negative consequences. We do not have uniform exponential convergence as well as the information complexity is larger that any polynomial in d. This means that EC-PT never holds even for pathologically small shape parameters. If we relax EC-PT and switch to EC-WT then EC-WT holds for some shape parameters, In fact, this holds for exponentially decaying shape parameters. If we further relax the tractability notion then we have EC-(t,κ)-WT for all shape parameters whenever $t>1$. This means, however, that we accept an exponential dependence on $d^{\,t_1}$ for $t_1\in[1,t)$.

We stress that our results for EC-WT and EC-$(1,\kappa)$-WT are not complete. We do not know matching necessary and sufficient conditions on $\boldsymbol{\gamma}$ for which EC-WT and EC-$(1,\kappa)$-WT hold for the absolute and normalized error criteria.

2. Function space with the Gaussian kernel

We first introduce the univariate reproducing kernel Hilbert space $H(K_\gamma)$ with the Gaussian kernel K_γ given by (1.2), where $\gamma^2\in(0,1)$ is the shape parameter. Let $\langle\cdot,\cdot\rangle_{L_2}$ denote the inner product in the L_2 space with weight function $e^{-x^2}/\sqrt{\pi}$, i.e.,

$$\langle f,g\rangle_{L_2} \;:=\; \int_{-\infty}^{\infty} f(x)\,g(x)\,\frac{\exp(-x^2)}{\sqrt{\pi}}\,\mathrm{d}x \qquad \text{for all} \qquad f,g\in L_2.$$

The space $H(K_\gamma)$ consists of analytic functions with the inner product

$$\langle f,g\rangle_{H(K_\gamma)} \;:=\; \sum_{\ell=1}^{\infty}\frac{1}{\lambda_{\ell,\gamma}}\,\langle f,\varphi_{\ell,\gamma}\rangle_{L_2}\,\langle g,\varphi_{\ell,\gamma}\rangle_{L_2} \qquad \text{for all} \qquad f,g\in H(K_\gamma).$$

The norm is given as usual by $\|\cdot\|_{H(K_\gamma)} := \langle\cdot,\cdot\rangle^{1/2}_{H(K_\gamma)}$. Here, $\lambda_{\ell,\gamma}$ and $\varphi_{\ell,\gamma}$ are, respectively, non-increasing eigenvalues and L_2-orthonormal eigenfunctions of the integral operator,

$$\int_{-\infty}^{\infty} K_\gamma(x,y)\,\varphi_{\ell,\gamma}(x)\,\frac{\exp(-x^2)}{\sqrt{\pi}}\,\mathrm{d}x \;=\; \lambda_{\ell,\gamma}\,\varphi_{\ell,\gamma}(y), \qquad \ell=1,2,\dots\,.$$

Specifically we have, see e.g., [22, Section 4.3.1] and [9],

$$\lambda_{\ell,\gamma} \;:=\; (1-\omega_\gamma)\,\omega_\gamma^{\ell-1}, \qquad \text{with} \qquad \omega_\gamma \;:=\; \frac{2\gamma^2}{1+2\gamma^2+\sqrt{1+4\gamma^2}}\,, \tag{2.1}$$

and

$$\varphi_{\ell,\gamma}(x) \;:=\; \sqrt{\frac{(1+4\gamma^2)^{1/4}}{2^{\ell-1}(\ell-1)!}}\,\exp\left(-\frac{2\gamma^2x^2}{1+\sqrt{1+4\gamma^2}}\right)H_{\ell-1}\left((1+4\gamma^2)^{1/4}x\right), \tag{2.2}$$

where $H_{\ell-1}$ is the standard Hermite polynomial of degree $\ell-1$, see e.g., [13],

$$H_{\ell-1}(x) = (-1)^{\ell-1}\,e^{x^2}\,\frac{\mathrm{d}^{\,\ell-1}}{\mathrm{d}x^{\,\ell-1}}\,e^{-x^2} \quad \text{for all } x\in\mathbb{R}.$$

Clearly, $\omega_\gamma < 1$. It is easy to check that ω_γ is an increasing function of $\gamma \in (0,1)$ and $\omega_\gamma = \gamma^2(1+o(1))$ as γ goes to zero. Moreover, $\varphi_{\ell,\gamma}$'s are $H(K_\gamma)$-orthogonal and $\|\varphi_{\ell,\gamma}\|_{H(K_\gamma)} = \lambda_{\ell,\gamma}^{-1/2}$, where the eigenvalue $\lambda_{1,\gamma}$ decreases with increasing γ while for $\ell \ge 2$ the eigenvalue $\lambda_{\ell,\gamma}$ increases with increasing γ.

More details about the characterization of the space $H(K_\gamma)$ can be found in [30].

As we mentioned in the introduction, the d-variate function space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ is the tensor product of the univariate spaces $H(K_{\gamma_j})$ with different shape parameters γ_j satisfying (1.3), i.e.,

$$H(\mathcal{K}_{d,\boldsymbol{\gamma}}) \,:=\, H(K_{\gamma_1}) \otimes H(K_{\gamma_2}) \otimes \cdots \otimes H(K_{\gamma_d})\,,$$

where the kernel $\mathcal{K}_{d,\boldsymbol{\gamma}}$ is the product of the kernels K_{γ_j}, see (1.2). For $\boldsymbol{\ell} = (\ell_1,\dots,\ell_d) \in \mathbb{N}^d$ we define

$$\lambda_{\boldsymbol{\ell},\boldsymbol{\gamma}} \,:=\, \prod_{j=1}^d \lambda_{\ell_j,\gamma_j} \,=\, \prod_{j=1}^d [(1-\omega_{\gamma_j})\,\omega_{\gamma_j}^{\ell_j-1}] \quad \text{and} \quad \varphi_{\boldsymbol{\ell},\boldsymbol{\gamma}}(\boldsymbol{x}) \,:=\, \prod_{j=1}^d \varphi_{\ell_j,\gamma_j}(x_j)\,. \tag{2.3}$$

Then we have $\varphi_{\boldsymbol{\ell},\boldsymbol{\gamma}} \in H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ and $\langle \varphi_{\boldsymbol{\ell},\boldsymbol{\gamma}}, \varphi_{\boldsymbol{k},\boldsymbol{\gamma}}\rangle_{L_2} = \delta_{\boldsymbol{\ell},\boldsymbol{k}}$, and the inner product in $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ is given by

$$\langle f,g\rangle_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})} \,:=\, \sum_{\boldsymbol{\ell}\in\mathbb{N}^d} \frac{1}{\lambda_{\boldsymbol{\ell},\boldsymbol{\gamma}}}\, \langle f,\varphi_{\boldsymbol{\ell},\boldsymbol{\gamma}}\rangle_{L_2}\, \langle g,\varphi_{\boldsymbol{\ell},\boldsymbol{\gamma}}\rangle_{L_2} \qquad \text{for all} \qquad f,g \in H(\mathcal{K}_{d,\boldsymbol{\gamma}}),$$

where for square-integrable d-variate functions we define the L_2 norm to be

$$\langle f,g\rangle_{L_2} \,:=\, \int_{\mathbb{R}^d} f(\boldsymbol{x})\,g(\boldsymbol{x}) \prod_{j=1}^d \frac{\exp(-x_j^2)}{\sqrt{\pi}}\,\mathrm{d}\boldsymbol{x} \qquad \text{for all} \qquad f,g \in L_2.$$

3. Upper bound for arbitrary d

In this section we obtain an upper bound on the Nth minimal error $e_{d,N}^{\min}$ in the d-variate case by considering an algorithm that is the tensor product of univariate Gauss-Hermite quadratures. Formally, $e_{d,N}^{\min}$ is defined as

$$e_{d,N}^{\min} = \inf_{\mathcal{A}_{d,N}} \sup_{\|f\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})}\le 1} |\mathcal{I}_d f - \mathcal{A}_{d,N} f|, \tag{3.1}$$

where the infimum is taken over all linear or nonlinear algorithms $\mathcal{A}_{d,N}$ that use at most N function values at non-adaptively or adaptively chosen sample points. It is well known that without loss of generality we can restrict ourselves to linear algorithms (quadratures) using non-adaptive choice of sample points; see [27] for the proof that nonlinear algorithms are not better than linear ones, and [1] that adaption does not help, see also e.g., [18, 31]. That is, we can assume that $\mathcal{A}_{d,N}$ is of the form

$$\mathcal{A}_{d,N} f = \sum_{i=1}^N w_i f(\boldsymbol{t}_i) \quad \text{for all } f \in H(\mathcal{K}_{d,\boldsymbol{\gamma}}),$$

and for some $w_i \in \mathbb{R}$ and $\boldsymbol{t}_i \in \mathbb{R}^d$. This allows us to simplify the definition of $e_{d,N}^{\min}$ to

$$e_{d,N}^{\min} = \inf_{\substack{w_1,\dots,w_N\in\mathbb{R}\\ \boldsymbol{t}_1,\dots,\boldsymbol{t}_N\in\mathbb{R}^d}} \sup_{\|f\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})}\le 1} \left| \mathcal{I}_d f - \sum_{i=1}^N w_i f(\boldsymbol{t}_i)\right|. \tag{3.2}$$

We express the d-variate integration operator (1.1) as

$$\mathcal{I}_d := \bigotimes_{j=1}^{d} I_j \,,$$

where the univariate integration operator $I_j : H(K_{\gamma_j}) \to \mathbb{R}$ is defined as

$$I_j f := \int_{-\infty}^{\infty} f(x) \, \frac{\exp(-x^2)}{\sqrt{\pi}} \, \mathrm{d}x \qquad \text{for} \quad f \in H(K_{\gamma_j}) \,.$$

The dependence of I_j on j is only through the domain $H(K_{\gamma_j})$.

Although the univariate integration operators I_j appear to be the same for all j, the operator norms of I_j's are, in general, different and depend on the shape parameters γ_j. Indeed, it is well known that

$$\|I_j\| := \sup_{f \in H(K_{\gamma_j})} \frac{|I_j f|}{\|f\|_{H(K_{\gamma_j})}} = \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_{\gamma_j}(x,y) \frac{\exp(-x^2)}{\sqrt{\pi}} \frac{\exp(-y^2)}{\sqrt{\pi}} \, \mathrm{d}x \, \mathrm{d}y \right)^{1/2} .$$

By integrating the last formula we obtain

$$\|I_j\| = \frac{1}{(1+2\gamma_j^2)^{1/4}} \,.$$

Hence

$$\|\mathcal{I}_d\| = \prod_{j=1}^{d} \|I_j\| = \prod_{j=1}^{d} \frac{1}{(1+2\gamma_j^2)^{1/4}},$$

which proves (1.5).

The univariate integration operators I_j are approximated by Gauss-Hermite quadratures $Q_{j,n}$ defined as follows. Let $Q_{j,0} := 0$, and for $n \in \mathbb{N}$ let

$$Q_{j,n} f = \sum_{k=1}^{n} w_{n,k} f(x_{n,k}) \quad \text{for } f \in H(K_{\gamma_j})$$

denote the Gauss-Hermite quadrature rule with n points for the space $H(K_{\gamma_j})$. Here, the sample points $x_{n,k}$'s are zeros of the Hermite polynomial H_n and coefficients $w_{n,k}$ are chosen to make the maximal order of exactness, i.e., $I_j p = Q_{j,n} p$ for all polynomials p of of degree up to $2n-1$. The coefficients $w_{n,k}$ are given by

$$w_{n,k} = \frac{2^{n-1} \, n!}{n^2 \, [H_{n-1}(x_{n,k})]^2}.$$

Note that the algorithms $Q_{j,n}$ depend on j only through its domain $H(K_{\gamma_j})$.

For $d \in \mathbb{N}$ and $\boldsymbol{n} = (n_1, n_2, \ldots)$ with $n_j \in \mathbb{N}$, we define

$$\mathcal{Q}_{d,\boldsymbol{n}} := \bigotimes_{j=1}^{d} Q_{j,n_j} \tag{3.3}$$

as the tensor product of the univariate Gauss-Hermite algorithms Q_{j,n_j}. More explicitly,

$$\mathcal{Q}_{d,\boldsymbol{n}} f = \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \cdots \sum_{k_d=1}^{n_d} \left(\prod_{j=1}^{d} w_{n_j,k_j} \right) f(x_{n_1,k_1}, x_{n_2,k_2}, \ldots, x_{n_d,k_d}) \qquad \text{for } f \in H(\mathcal{K}_{d,\boldsymbol{\gamma}}). \tag{3.4}$$

Clearly, the algorithm $\mathcal{Q}_{d,\boldsymbol{n}}$ uses $N = \prod_{j=1}^{d} n_j$ function values.

We know from (1.4) that

$$e(Q_{j,n}) := \sup_{f \in H(K_{\gamma_j})} \frac{|I_j f - Q_{j,n} f|}{\|f\|_{H(K_{\gamma_j})}} = \|I_j - Q_{j,n}\| \le \beta_n \, \gamma_j^{2n} \le \frac{\sqrt{3}}{2} \gamma_j^{2n} \,, \tag{3.5}$$

where the last inequality follows from $\beta_n \le \sqrt{3}/2$ which is shown in Lemma 3.1 below.

Lemma 3.1. *The sequence* $\{\beta_n\}_{n \in \mathbb{N}}$ *defined by*

$$\beta_n := \frac{n!}{2^n \, (2n)!} \sqrt{\frac{(4n)!}{(2n)!}} \qquad \forall \, n \in \mathbb{N}$$

is strictly decreasing, and satisfies

$$0.840896\ldots = \frac{1}{2^{1/4}} < \beta_n \le \frac{\sqrt{3}}{2} = 0.866025\ldots \qquad \textit{for all } n \in \mathbb{N}.$$

Proof. We can write

$$\begin{aligned}
\beta_n &= \sqrt{\frac{(2n+1)(2n+2)\cdots(4n-1)(4n)}{2^{2n}(n+1)^2(n+2)^2\cdots(2n-1)^2(2n)^2}} \\
&= \sqrt{\frac{(n+\frac{1}{2})(n+\frac{3}{2})\cdots(2n-\frac{1}{2})}{(n+1)(n+2)\cdots(2n)}} = \sqrt{\prod_{j=1}^{n} \frac{n+j-\frac{1}{2}}{n+j}} \,,
\end{aligned}$$

which leads to

$$\begin{aligned}
\frac{\beta_n^2}{\beta_{n+1}^2} &= \Bigg(\prod_{j=1}^{n} \frac{n+j-\frac{1}{2}}{n+j} \Bigg) \Bigg(\prod_{j=1}^{n+1} \frac{n+1+j}{n+1+j-\frac{1}{2}} \Bigg) \\
&= \Bigg(\prod_{j=1}^{n} \frac{n+j-\frac{1}{2}}{n+j} \Bigg) \Bigg(\prod_{j=2}^{n+2} \frac{n+j}{n+j-\frac{1}{2}} \Bigg) \\
&= \frac{(n+\frac{1}{2})(2n+1)(2n+2)}{(n+1)(2n+\frac{1}{2})(2n+\frac{3}{2})} = \frac{(n+\frac{1}{2})^2}{(n+\frac{1}{4})(n+\frac{3}{4})} > 1 \,,
\end{aligned}$$

since $(n+\frac{1}{2})^2 - (n+\frac{1}{4})(n+\frac{3}{4}) = \frac{1}{16} > 0$. This indicates that the sequence $\{\beta_n\}$ is strictly decreasing, and so $\beta_n \le \beta_1 = \sqrt{3}/2$ for all $n \in \mathbb{N}$. For the lower bound, Stirling's formula for factorials $n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}$ yields $\beta_n > \lim_{i \to \infty} \beta_i = 2^{-1/4}$ for all $n \in \mathbb{N}$. □

We are ready to estimate the (worst case) error of the algorithm $\mathcal{Q}_{d,\boldsymbol{n}}$,

$$e(\mathcal{Q}_{d,\boldsymbol{n}}) := \sup_{f \in H(\mathcal{K}_{d,\boldsymbol{\gamma}})} \frac{|\mathcal{I}_d f - \mathcal{Q}_{d,\boldsymbol{n}} f|}{\|f\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})}} = \|\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}}\|.$$

Theorem 3.2. *The error of the algorithm* $\mathcal{Q}_{d,\boldsymbol{n}}$ *defined by* (3.3) *satisfies*

$$e(\mathcal{Q}_{d,\boldsymbol{n}}) \le \frac{\sqrt{3}}{2} \sum_{s=1}^{d} \gamma_s^{2\,n_s} \prod_{j=1}^{s-1} \frac{1}{(1+2\gamma_j^2)^{1/4}}. \tag{3.6}$$

Proof. We derive an error bound by induction on d. For $d = 1$ the error estimate clearly holds due to (3.5). For $d > 1$, we have

$$\begin{aligned}\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}} &= \mathcal{I}_{d-1} \otimes I_d \; - \; \mathcal{Q}_{d-1,\boldsymbol{n}} \otimes Q_{d,n_d} \\ &= \mathcal{I}_{d-1} \otimes (I_d - Q_{d,n_d}) \; + \; (\mathcal{I}_{d-1} - \mathcal{Q}_{d-1,\boldsymbol{n}}) \otimes Q_{d,n_d}.\end{aligned}$$

Hence

$$\|\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}}\| \;\le\; \|\mathcal{I}_{d-1}\| \, \|I_d - Q_{d,n_d}\| \;+\; \|\mathcal{I}_{d-1} - \mathcal{Q}_{d-1,\boldsymbol{n}}\| \, \|Q_{d,n_d}\|,$$

where the operator norm $\|Q_{d,n_d}\|$ is defined analogously to $\|\mathcal{I}_d\|$ and $\|\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}}\|$. We now show that $\|Q_{d,n_d}\| \le 1$. We have from the reproducing property of the reproducing kernel K_{γ_d} that

$$Q_{d,n_d} f \;=\; \left\langle f, \sum_{k=1}^{n_d} w_{n_d,k} K_{\gamma_d}(\cdot, x_{n_d,k}) \right\rangle_{H(K_{\gamma_d})}.$$

Therefore

$$\begin{aligned}\|Q_{d,n_d}\| &= \left\| \sum_{k=1}^{n_d} w_{n_d,k} K_{\gamma_d}(\cdot, x_{n_d,k}) \right\|_{H(K_{\gamma_d})} \\ &= \left(\sum_{k_1=1}^{n_d} \sum_{k_2=1}^{n_d} w_{n_d,k_1} w_{n_d,k_2} K_{\gamma_d}(x_{n_d,k_1}, x_{n_d,k_2}) \right)^{1/2} \\ &= \left(\sum_{k_1=1}^{n_d} \sum_{k_2=1}^{n_d} w_{n_d,k_1} w_{n_d,k_2} \, e^{-\gamma_d^2 (x_{n_d,k_1} - x_{n_d,k_2})^2} \right)^{1/2}.\end{aligned}$$

The coefficients w_{n_d,k_1} and w_{n_d,k_2} of the Gauss-Hermite quadratures are positive and they sum up to one. This yields

$$\|Q_{d,n_d}\| \le \left(\sum_{k_1=1}^{n_d} \sum_{k_2=1}^{n_d} w_{n_d,k_1} w_{n_d,k_2} \right)^{1/2} \;=\; \sum_{k=1}^{n_d} w_{n_d,k} = 1,$$

as claimed. Using this, the formula for $\|\mathcal{I}_{d-1}\|$ and (3.5), we conclude

$$\|\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}}\| \;\le\; \frac{\sqrt{3}}{2} \, \gamma_d^{2n_d} \prod_{j=1}^{d-1} \frac{1}{(1 + 2\gamma_j^2)^{1/4}} + \|\mathcal{I}_{d-1} - \mathcal{Q}_{d-1,\boldsymbol{n}}\|.$$

By induction on d and remembering that $e(\mathcal{Q}_{d,\boldsymbol{n}}) = \|\mathcal{I}_d - \mathcal{Q}_{d,\boldsymbol{n}}\|$, we complete the proof. $\square$

The error of the algorithm $\mathcal{Q}_{d,\boldsymbol{n}}$ which uses $N = \prod_{j=1}^d n_j$ function values is, of course, an upper bound on the minimal error $e_{d,N}^{\min}$. From Theorem 3.2 we know that for any $n_j \in \mathbb{N}$ for which $N = \prod_{j=1}^d n_j$ we have

$$e_{d,N}^{\min} \le \frac{\sqrt{3}}{2} \sum_{s=1}^{d} \gamma_s^{2\,n_s} \prod_{j=1}^{s-1} \frac{1}{(1 + 2\gamma_j^2)^{1/4}}. \tag{3.7}$$

This estimate allows us to show that exponential convergence in the sense of (1.6) holds for all shape parameters with $p_d = 1/d$.

Indeed, for $N \in \mathbb{N}$ take $n_s = n$ for all $s = 1, 2, \dots, d$ with $n = \lfloor N^{1/d} \rfloor$. Then $n^d \le N \le (n+1)^d$. We now simplify the bound on the minimal error by replacing by 1 the factors in (3.7) which are smaller than 1, and by replacing all γ_j^2 by its largest value γ_1^2, i.e., we use

$$e_{d,N}^{\min} \le e_{d,n^d}^{\min} \le d\,\gamma_1^{2n} \le \frac{d}{\gamma_1^2}\,\gamma_1^{2N^{1/d}}.$$

This means that (1.6) holds with

$$C_{1,d} = \frac{d}{\gamma_1^2}, \quad q = \gamma_1^2, \quad C_{2,d} = 1, \quad p_d = \frac{1}{d}. \tag{3.8}$$

This proves Theorem 1.1(a) and part of Theorem 1.1(b), in that it proves $p_d^* \ge 1/d$.

We now turn to the information complexity $n(\varepsilon, d)$ which is formally defined as

$$n(\varepsilon, d) = \min\{N \,:\, e_{d,N}^{\min} \le \varepsilon\, \mathrm{CRI}_d\},$$

where $\mathrm{CRI}_d = 1$ for the absolute error criterion, and $\mathrm{CRI}_d = \|\mathcal{I}_d\|$ for the normalized error criterion.

We are ready to show an upper bound on $n(\varepsilon, d)$ by defining n_s to be the minimal integer for which each term in the sum in (3.6) is at most $\varepsilon\,\mathrm{CRI}_d/d$, which will imply $e(Q_{d,\boldsymbol{n}}) \le \varepsilon\,\mathrm{CRI}_d$. That is,

$$n_s \;=\; \max\left(1, \left\lceil \frac{\ln\left(\frac{\sqrt{3}\,d}{2\varepsilon\,\mathrm{CRI}_d}\;\prod_{j=1}^{s-1}(1+2\gamma_j^2)^{-1/4}\right)}{\ln\frac{1}{\gamma_s^2}} \right\rceil\right).$$

Then

$$N = \prod_{s=1}^{d} n_s = \prod_{s=1}^{d} \max\left(1, \left\lceil \frac{\ln\frac{d}{\varepsilon\,\mathrm{CRI}_d} \;+\; \ln\left(\frac{\sqrt{3}}{2}\;\prod_{j=1}^{s-1}(1+2\gamma_j^2)^{-1/4}\right)}{\ln\frac{1}{\gamma_s^2}} \right\rceil\right).$$

The value of the second logarithm in the numerator is negative since its argument is less than 1. Therefore

$$N \le \prod_{s=1}^{d} \left\lceil \frac{\ln\frac{d}{\varepsilon\,\mathrm{CRI}_d}}{\ln\frac{1}{\gamma_s^2}} \right\rceil = \left(\ln\frac{1}{\varepsilon}\right)^d \frac{1+o(1)}{\prod_{s=1}^{d}\ln\frac{1}{\gamma_s^2}} \qquad \text{as } \varepsilon \to 0.$$

We can further estimate

$$\frac{1}{\mathrm{CRI}_d} \;\le\; \prod_{j=1}^{d}(1+2\gamma_j^2)^{1/4} \;\le\; (1+2\gamma_1^2)^{d/4}.$$

Therefore

$$N \;\le\; \left(1 + \frac{\ln\frac{d}{\varepsilon} \;+\; d\,\ln(1+2\gamma_1^2)^{1/4}}{\ln\frac{1}{\gamma_1^2}}\right)^d.$$

Since $n(\varepsilon, d) \le N$, this proves the following corollary.

Corollary 3.3. *For all shape parameters $\boldsymbol{\gamma} = \{\gamma_j\}$ and for both the absolute and normalized error criteria, we have the following estimates*

$$n(\varepsilon, d) \;\le\; \prod_{s=1}^{d} \left\lceil \frac{\ln \frac{d}{\varepsilon\,\mathrm{CRI}_d}}{\ln \frac{1}{\gamma_s^2}} \right\rceil \tag{3.9}$$

$$\le\; \left(1 + \frac{\ln \frac{d}{\varepsilon} \;+\; d\,\ln(1+2\gamma_1^2)^{1/4}}{\ln \frac{1}{\gamma_1^2}} \right)^{d} \tag{3.10}$$

for all $\varepsilon \in (0,1)$ *and* $d \in \mathbb{N}$*. That is, for every* $d \in \mathbb{N}$ *there are positive* C_d *and* ε_d *such that*

$$n(\varepsilon, d) \;\le\; C_d \left(\ln \frac{1}{\varepsilon} \right)^{d} \qquad \textit{for all } \varepsilon \in (0, \varepsilon_d). \tag{3.11}$$

Note that (3.11) in Corollary 3.3 proves the upper bound in Theorem 1.1(d).

We now study EC-WT. For $\delta \in (0,1)$ define

$$s^*(\delta) = \min\{ s \;:\; \gamma_s^2 \le \delta \}$$

with the convention that $\min \emptyset = \infty$. Observe that

$$s \ge s^*(\varepsilon\,\mathrm{CRI}_d/d) \qquad \Longrightarrow \qquad \left\lceil \frac{\ln \frac{d}{\varepsilon\,\mathrm{CRI}_d}}{\ln \frac{1}{\gamma_s^2}} \right\rceil = 1.$$

From (3.9) in Corollary 3.3 we conclude

$$\ln n(\varepsilon, d) \;\le\; \min\left(d, s^*\left(\varepsilon\,\mathrm{CRI}_d/d\right)\right) \ln \left(1 + \frac{\ln \frac{d}{\varepsilon\,\mathrm{CRI}_d}}{\ln \frac{1}{\gamma_1^2}} \right).$$

Let $x = \max(d, \ln \varepsilon^{-1})$. Then $d \le x$ and $1/\varepsilon \le e^x$ and therefore $d/\varepsilon \le x\,e^x$. Since the function s^* is non-increasing, we have

$$\frac{\ln n(\varepsilon, d)}{d + \ln \varepsilon^{-1}} \;\le\; \min\left(1, \frac{s^*\left(\mathrm{CRI}_d/(x\,e^x)\right)}{x}\right) \ln \left(1 + \frac{\ln \frac{x\,e^x}{\mathrm{CRI}_d}}{\ln \frac{1}{\gamma_1^2}} \right).$$

Consider now the special case when $\gamma_s^2 = \mathcal{O}(\exp(-s^{\,\alpha}))$ with $\alpha > 1$, which holds iff

$$s^*(\delta) = \mathcal{O}\left(\left(\ln \delta^{-1}\right)^{1/\alpha} \right). \tag{3.12}$$

For such shape parameters $\sum_{j=1}^{\infty} \gamma_j^2 < \infty$ and CRI_d is of order 1 also for the normalized error criterion. From (3.12) we conclude that

$$\min\left(1, \frac{s^*\left(\mathrm{CRI}_d/(x\,e^x)\right)}{x}\right) \ln \left(1 + \frac{\ln \frac{x\,e^x}{\mathrm{CRI}_d}}{\ln \frac{1}{\gamma_1^2}} \right) \;=\; \mathcal{O}\left(\frac{\ln x}{x^{1-1/\alpha}} \right)$$

which obviously goes to zero as x approaches infinity. This proves EC-WT for the absolute and normalized error criteria and proves Theorem 1.1(e).

We now show that EC-$(1,\kappa)$-WT in the sense of (1.8) holds for the absolute error criterion and for shape parameters such that $\gamma_\infty^2 := \lim_{j\to\infty} \gamma_j^2 > 0$. This will be shown based on the error bound of $\mathcal{Q}_{d,\boldsymbol{n}}$ given by (3.6). Note that $\gamma_j^2 \ge \gamma_\infty^2$

and $\prod_{j=1}^{s-1}(1+2\gamma_j^2)^{-1/4} \le (1+2\gamma_\infty^2)^{-(s-1)/4}$. The bound on $e(\mathcal{Q}_{d,\boldsymbol{n}})$ can now be simplified to

$$e(\mathcal{Q}_{d,\boldsymbol{n}}) \;\le\; \sum_{s=1}^{d} \gamma_1^{2n_s} \frac{1}{(1+2\gamma_\infty^2)^{(s-1)/4}}.$$

Define $n_s = n_{s,d}$ for $s = 1, \dots, d$ such that

$$\gamma_1^{2n_s} \frac{1}{(1+2\gamma_\infty^2)^{(s-1)/4}} \;\le\; \frac{\varepsilon}{d}$$

so that $e(\mathcal{Q}_{d,\boldsymbol{n}}) \le \varepsilon$.

Let

$$m^*(\varepsilon, d) := 1 + \left\lceil 4\, \frac{\ln(d\,\gamma_1^2) \;+\; \ln\,\varepsilon^{-1}}{\ln\,(1+2\gamma_\infty^2)} \right\rceil.$$

If $m^*(\varepsilon,d) \le d$, then for $s \in [m^*(\varepsilon,d), d]$ we have $\gamma_1^2/(1+2\gamma_\infty^2)^{(s-1)/4} \le \varepsilon/d$, and we can take $n_{s,d} = 1$. For $s \in [1, \min(d, m^*(\varepsilon,d)-1)]$, we can take

$$n_{s,d} = \left\lceil \frac{\ln\,d \;+\; \ln\,\varepsilon^{-1} \;-\; (s-1)\,\ln(1+2\gamma_\infty^2)^{1/4}}{\ln\,\frac{1}{\gamma_1^2}} \right\rceil \;\le\; 1 + \frac{\ln\,d \;+\; \ln\,\varepsilon^{-1}}{\ln\,\frac{1}{\gamma_1^2}}.$$

Therefore

$$N = \prod_{s=1}^{d} n_{s,d} \;\le\; \left(1 + \frac{\ln\,d \;+\; \ln\,\varepsilon^{-1}}{\ln\,\frac{1}{\gamma_1^2}}\right)^{\min(d, m^*(\varepsilon,d)-1)}.$$

Clearly,

$$m^*(\varepsilon,d) = \mathcal{O}\left(1 + \ln\,d \;+\; \ln\,\varepsilon^{-1}\right)$$

with the factor in the big $\mathcal{O}$ notation independent of both d and ε^{-1}. Moreover, $n(\varepsilon,d) \le N$ and

$$\ln\,n(\varepsilon,d) \;\le\; \ln\,N \;=\; \mathcal{O}\left(\min(d, 1+\ln\,d \;+\; \ln\,\varepsilon^{-1})\,\ln(1+\ln\,d \;+\; \ln\,\varepsilon^{-1})\right).$$

Let $x = \max(d, (\ln\,\varepsilon^{-1})^\kappa)$ for $\kappa > 1$. Then x goes to infinity with $d + \varepsilon^{-1}$. For large x we have

$$\frac{\ln\,n(\varepsilon,d)}{d + [\ln\,\varepsilon^{-1}]^\kappa} = \mathcal{O}\left(\frac{\min(x, 1+\ln\,x + x^{1/\kappa})\,\ln(1+\ln\,x + x^{1/\kappa})}{x}\right) = \mathcal{O}\left(\frac{\ln\,x}{x^{1-1/\kappa}}\right),$$

which goes to zero as x approaches infinity since $\kappa > 1$. This means that EC-$(1,\kappa)$-WT holds and proves Theorem 1.1(g).

We now consider EC-(t,κ)-WT with $t > 1$. From (3.10) in Corollary 3.3 we have

$$\ln\,n(\varepsilon,d) \;=\; \mathcal{O}\left(d\,\ln\left(\ln\,\varepsilon^{-1} \;+\; d\right)\right),$$

with the factor in the big $\mathcal{O}$ notation independent of d and ε^{-1}. With $x = \max(d^t, (\ln\,\varepsilon^{-1})^\kappa)$ we can rewrite the last formula as

$$\ln\,n(\varepsilon,d) \;=\; \mathcal{O}\left(x^{1/t}\,\ln\,x\right).$$

Hence,

$$\frac{\ln\,n(\varepsilon,d)}{d^t + (\ln\,\varepsilon^{-1})^\kappa} \;=\; \mathcal{O}\left(\frac{\ln\,x}{x^{1-1/t}}\right),$$

which goes to zero as x approaches infinity since $t > 1$. This proves Theorem 1.1(h).

4. Lower bound for $d = 1$

We provide a lower bound on the nth minimal error $e_{1,n}^{\min}$ for the univariate case. This will be done as follows. For any quadrature rule A_n which uses sample points t_i for $i = 1, 2, \ldots, n$, we construct a univariate function $f^* \in H(K_\gamma)$ for which the integral is 1 and f^* vanishes at all sample points t_i. Then the quadrature approximation is 0 and the error of A_n is lower bounded by $|If^* - A_n f|/\|f^*\|_{H(K_\gamma)} = 1/\|f^*\|_{H(K_\gamma)}$. We then estimate the norm of f^* from above and independently of the sample points t_i. This in turn yields a lower bound on $e_{1,n}^{\min}$.

Theorem 4.1. *The nth minimal error of univariate integration in the function space $H(K_\gamma)$ is bounded from below by*

$$e_{1,n}^{\min} \ \geq \ \sqrt{\frac{2\,(1+4\gamma^2)^{1/4}}{(1+2\gamma^2+\sqrt{1+4\gamma^2})e}\,\frac{\omega_\gamma^n n!}{(n+1)(2n)!}}\,, \tag{4.1}$$

where $\omega_\gamma < 1$ is given by (2.1).

Proof. First we observe from (2.2) that for any $\gamma > 0$ and any polynomial p, the function

$$f(x) \ = \ \exp\left(-\frac{2\gamma^2 x^2}{1+\sqrt{1+4\gamma^2}}\right) p(x) \qquad \text{for all} \quad x \in \mathbb{R} \tag{4.2}$$

is a linear combination of eigenfunctions $\varphi_{\ell,\gamma}$ from (2.2), and hence it belongs to $H(K_\gamma)$.

In the following, we will choose $\tilde\gamma > 0$ so that

$$\frac{4\tilde\gamma^2}{1+\sqrt{1+4\tilde\gamma^2}} \ = \ \frac{2\gamma^2}{1+\sqrt{1+4\gamma^2}}\,. \tag{4.3}$$

The existence of such a $\tilde\gamma$ follows from the fact that the expression on the left is a continuous function of $\tilde\gamma \in (0,\infty)$ with range $(0,\infty)$.

Let $\tilde\gamma > 0$ be arbitrary for the moment. Given $n \in \mathbb{N}$ and arbitrary points $t_1, \ldots, t_n$ from $\mathbb{R}$, we choose real numbers $a_1, \ldots, a_{n+1}$ such that

$$\sum_{\ell=1}^{n+1} a_\ell\, \varphi_{\ell,\tilde\gamma}(t_i) = 0 \quad \text{for all} \quad i = 1, 2, \ldots, n\,,$$

where $\varphi_{\ell,\tilde\gamma}$ is given by (2.2), but with γ replaced by $\tilde\gamma$. Since the linear system has more unknowns than the number of conditions, there exists a solution such that

$$\max_{1\le \ell\le n+1} |a_\ell| = 1 = a_k \ \text{ for some index } k.$$

Following the proof idea from [26], we now define

$$f^*(x) \ := \ \varphi_{k,\tilde\gamma}(x) \sum_{\ell=1}^{n+1} a_\ell\, \varphi_{\ell,\tilde\gamma}(x)\,.$$

Then $f^*(t_i) = 0$ for all $i = 1, 2, \ldots, n$ so that any quadrature rule based on linear combinations of these function values is 0, that is, for any choice of coefficients $w_1, \ldots, w_n$ we have

$$A_n f^* \ = \ \sum_{i=1}^{n} w_i\, f^*(t_i) \ = \ 0\,.$$

On the other hand, we have

$$I f^* \;=\; \int_{-\infty}^{\infty} f^*(x)\, \frac{\exp(-x^2)}{\sqrt{\pi}}\, \mathrm{d}x \;=\; \sum_{\ell=1}^{n+1} a_\ell\, \langle \varphi_{k,\tilde\gamma}, \varphi_{\ell,\tilde\gamma}\rangle_{L_2} \;=\; a_k \;=\; 1\,.$$

We write

$$f^*(x) \;=\; \sum_{\ell=1}^{n+1} a_\ell\, g_{k,\ell}(x)\,, \qquad \text{with} \qquad g_{k,\ell}(x) \;:=\; \varphi_{k,\tilde\gamma}(x)\, \varphi_{\ell,\tilde\gamma}(x)\,. \tag{4.4}$$

Then from (2.2) we have

$$\begin{aligned} g_{k,\ell}(x) \;&=\; \sqrt{\frac{(1+4\tilde\gamma^2)^{1/2}}{2^{k-1}(k-1)!\, 2^{\ell-1}(\ell-1)!}}\, \exp\left(-\frac{4\tilde\gamma^2 x^2}{1+\sqrt{1+4\tilde\gamma^2}}\right) \\ &\quad\times H_{k-1}\left((1+4\tilde\gamma^2)^{1/4} x\right) H_{\ell-1}\left((1+4\tilde\gamma^2)^{1/4} x\right)\,. \end{aligned} \tag{4.5}$$

The choice of $\tilde\gamma$ satisfying (4.3) and the fact that f in (4.2) belongs to $H(K_\gamma)$ then ensure that $g_{k,\ell} \in H(K_\gamma)$, and therefore $f^* \in H(K_\gamma)$. Thus we can estimate

$$\|f^*\|_{H(K_\gamma)} \;\le\; \sum_{\ell=1}^{n+1} |a_\ell|\, \|g_{k,\ell}\|_{H(K_\gamma)} \;\le\; \sum_{\ell=1}^{n+1} \|g_{k,\ell}\|_{H(K_\gamma)}\,.$$

The worst case error in $H(K_\gamma)$ is bounded from below by

$$e(A_n) \;\ge\; \frac{|I f^* - A_n f^*|}{\|f^*\|_{H(K_\gamma)}} \;\ge\; \frac{1}{\sum_{\ell=1}^{n+1} \|g_{k,\ell}\|_{H(K_\gamma)}}\,.$$

Since the functions $g_{k,\ell}$ do not depend on A_n, i.e., on w_i and t_i for $i = 1, 2 \ldots, n$, we have

$$e_{1,n}^{\min} \;\ge\; \frac{1}{\sum_{\ell=1}^{n+1} \|g_{k,\ell}\|_{H(K_\gamma)}}\,.$$

The proof is completed by inserting the bound on $\|g_{k,\ell}\|_{H(K_\gamma)}$ from Lemma 4.2 below. □

Lemma 4.2. *For integers $k, \ell \in [1, n+1]$ and parameter $\tilde\gamma$ which is the unique solution of* (4.3)*, the $H(K_\gamma)$-norm of the function $g_{k,\ell}$ given by* (4.4) *satisfies*

$$\|g_{k,\ell}\|_{H(K_\gamma)} \;\le\; \sqrt{\frac{(1+2\gamma^2+\sqrt{1+4\gamma^2})e}{2\,(1+4\gamma^2)^{1/4}}\, \frac{(2n)!}{\omega_\gamma^n\, n!}}\,,$$

where $\omega_\gamma < 1$ is given by (2.1)*.*

Proof. By rationalizing the denominators on both sides of (4.3), we obtain

$$\sqrt{1+4\tilde\gamma^2} - 1 \;=\; \frac{4\tilde\gamma^2}{1+\sqrt{1+4\tilde\gamma^2}} \;=\; \frac{2\gamma^2}{1+\sqrt{1+4\gamma^2}} \;=\; \frac{\sqrt{1+4\gamma^2}-1}{2}\,,$$

which yields

$$\sqrt{1+4\tilde\gamma^2} \;=\; \frac{1+\sqrt{1+4\gamma^2}}{2}\,, \tag{4.6}$$

and therefore $\tilde\gamma$ is given uniquely by

$$\tilde\gamma \;:=\; \frac{1}{2}\sqrt{\left(\frac{1+\sqrt{1+4\gamma^2}}{2}\right)^2 - 1} \;<\; \gamma\,. \tag{4.7}$$

Using the identity for the product of Hermite polynomials, see [33],

$$H_m(x)\,H_p(x) \;=\; 2^p p! \sum_{r=0}^{p} \binom{m}{p-r} \frac{1}{2^r r!}\, H_{m-p+2r}(x)\,, \qquad m \ge p\,,$$

we can rewrite (4.5) as

$$\begin{aligned} g_{k,\ell}(x) \;=\; & \sum_{r=0}^{\min(k,\ell)-1} \sqrt{\frac{(1+4\tilde\gamma^2)^{1/2}}{2^{k-1}(k-1)!\,2^{\ell-1}(\ell-1)!}}\, \exp\left(-\frac{4\tilde\gamma^2 x^2}{1+\sqrt{1+4\tilde\gamma^2}}\right) \\ & \times \frac{2^{\min(k,\ell)-1}(\min(k,\ell)-1)!}{2^r r!} \binom{\max(k,\ell)-1}{\min(k,\ell)-1-r} \\ & \times H_{|k-\ell|+2r}\left((1+4\tilde\gamma^2)^{1/4}x\right)\,. \end{aligned}$$

Substituting (4.3) and (4.6) into the above expression yields

$$\begin{aligned} g_{k,\ell}(x) \;=\; & \sum_{r=0}^{\min(k,\ell)-1} \sqrt{\frac{1+\sqrt{1+4\gamma^2}}{2\cdot 2^{k-1}(k-1)!\,2^{\ell-1}(\ell-1)!}}\, \exp\left(-\frac{2\gamma^2 x^2}{1+\sqrt{1+4\gamma^2}}\right) \\ & \times \frac{2^{\min(k,\ell)-1}(\min(k,\ell)-1)!}{2^r r!} \binom{\max(k,\ell)-1}{\min(k,\ell)-1-r} \\ & \times H_{|k-\ell|+2r}\left(t_\gamma\,(1+4\gamma^2)^{1/4}x\right)\,, \end{aligned}$$

where

$$t_\gamma := \sqrt{\frac{1+\sqrt{1+4\gamma^2}}{2}\,\frac{1}{(1+4\gamma^2)^{1/4}}} = \sqrt{\frac{1}{2}\left(1+\frac{1}{\sqrt{1+4\gamma^2}}\right)} \;\to\; 1^- \text{ as } \gamma \to 0\,.$$

Next we use the multiplicative theorem of the Hermite polynomials, see [10],

$$H_m(ty) \;=\; \sum_{i=0}^{\lfloor m/2\rfloor} t^{m-2i}(t^2-1)^i \binom{m}{2i} \frac{(2i)!}{i!} H_{m-2i}(y)\,.$$

This yields

$$\begin{aligned} g_{k,\ell}(x) \;=\; & \sum_{r=0}^{\min(k,\ell)-1} \sum_{i=0}^{\left\lfloor\frac{|k-\ell|+2r}{2}\right\rfloor} \sqrt{\frac{1+\sqrt{1+4\gamma^2}}{2\cdot 2^{k-1}(k-1)!\,2^{\ell-1}(\ell-1)!}} \\ & \times \frac{2^{\min(k,\ell)-1}(\min(k,\ell)-1)!}{2^r r!} \binom{\max(k,\ell)-1}{\min(k,\ell)-1-r} \\ & \times t_\gamma^{|k-\ell|+2r-2i}(t_\gamma^2-1)^i \binom{|k-\ell|+2r}{2i} \frac{(2i)!}{i!} \\ & \times H_{|k-\ell|+2r-2i}\left((1+4\gamma^2)^{1/4}x\right) \exp\left(-\frac{2\gamma^2 x^2}{1+\sqrt{1+4\gamma^2}}\right)\,. \end{aligned}$$

Using (2.2), we can rewrite this expression as

$$
\begin{aligned}
g_{k,\ell}(x) \;=\; \sum_{r=0}^{\min(k,\ell)-1} \sum_{i=0}^{\lfloor \frac{|k-\ell|+2r}{2} \rfloor} & \sqrt{\frac{1+\sqrt{1+4\gamma^2}}{2\cdot 2^{k-1}(k-1)!\,2^{\ell-1}(\ell-1)!}} \\
& \times \frac{2^{\min(k,\ell)-1}(\min(k,\ell)-1)!}{2^r r!} \binom{\max(k,\ell)-1}{\min(k,\ell)-1-r} \\
& \times t_\gamma^{|k-\ell|+2r-2i}(t_\gamma^2-1)^i \binom{|k-\ell|+2r}{2i} \frac{(2i)!}{i!} \\
& \times \sqrt{\frac{2^{|k-\ell|+2r-2i}(|k-\ell|+2r-2i)!}{(1+4\gamma^2)^{1/4}}}\; \varphi_{|k-\ell|+2r-2i+1,\gamma}(x)\,.
\end{aligned}
$$

After some elementary but tedious manipulations, this simplifies to

$$
\begin{aligned}
g_{k,\ell}(x) \;=\; \sum_{r=0}^{\min(k,\ell)-1} \sum_{i=0}^{\lfloor \frac{|k-\ell|}{2} \rfloor + r} & \sqrt{\mu_\gamma}\; t_\gamma^{|k-\ell|+2r-2i} \left(\frac{t_\gamma^2-1}{2}\right)^i \sqrt{\frac{(k-1)!\,(\ell-1)!}{(|k-\ell|+2r-2i)!}} \\
& \times \frac{1}{i!\,(\min(k,\ell)-1-r)!} \binom{|k-\ell|+2r}{r} \varphi_{|k-\ell|+2r-2i+1,\gamma}(x)\,,
\end{aligned}
$$

where

$$
\mu_\gamma \;:=\; \frac{1+\sqrt{1+4\gamma^2}}{2(1+4\gamma^2)^{1/4}} \;>\; 1\,.
$$

We change the summation variable i to $p = r - i$, and then interchange the summation variables p and r. Thus we obtain

$$
\begin{aligned}
g_{k,\ell}(x) \;&=\; \sum_{p=-\lfloor \frac{|k-\ell|}{2} \rfloor}^{\min(k,\ell)-1} \sum_{r=\max(p,0)}^{\min(k,\ell)-1} \sqrt{\mu_\gamma}\; t_\gamma^{|k-\ell|+2p} \left(\frac{t_\gamma^2-1}{2}\right)^{r-p} \sqrt{\frac{(k-1)!\,(\ell-1)!}{(|k-\ell|+2p)!}} \\
&\qquad\qquad \times \frac{1}{(r-p)!\,(\min(k,\ell)-1-r)!} \binom{|k-\ell|+2r}{r} \varphi_{|k-\ell|+2p+1,\gamma}(x) \\
&=\; \sum_{p=-\lfloor \frac{|k-\ell|}{2} \rfloor}^{\min(k,\ell)-1} c_{k,\ell,p}\, \varphi_{|k-\ell|+2p+1,\gamma}(x)\,,
\end{aligned}
$$

where

$$
\begin{aligned}
c_{k,\ell,p} \;:=\; & \sqrt{\mu_\gamma}\; t_\gamma^{|k-\ell|+2p} \sqrt{\frac{(k-1)!\,(\ell-1)!}{(|k-\ell|+2p)!}} \\
& \times \sum_{r=\max(p,0)}^{\min(k,\ell)-1} \left(\frac{t_\gamma^2-1}{2}\right)^{r-p} \frac{1}{(r-p)!\,(\min(k,\ell)-1-r)!} \binom{|k-\ell|+2r}{r}\,.
\end{aligned}
$$

Since $|k-\ell|+2p+1$ corresponds to a unique index for every p, we conclude from the $H(K_\gamma)$-orthogonality of the eigenfunctions and (2.1) that

$$\|g_{k,\ell}\|^2_{H(K_\gamma)} = \sum_{p=-\lfloor\frac{|k-\ell|}{2}\rfloor}^{\min(k,\ell)-1} c^2_{k,\ell,p}\,\|\varphi_{|k-j|+2p+1,\gamma}\|^2_{H(K_\gamma)}$$

$$= \sum_{p=-\lfloor\frac{|k-\ell|}{2}\rfloor}^{\min(k,\ell)-1} \frac{c^2_{k,\ell,p}}{\lambda_{|k-j|+2p+1,\gamma}} = \sum_{p=-\lfloor\frac{|k-\ell|}{2}\rfloor}^{\min(k,\ell)-1} \frac{c^2_{k,\ell,p}}{(1-\omega_\gamma)\,\omega_\gamma^{|k-\ell|+2p}}\,. \tag{4.8}$$

Note that $\binom{|k-\ell|+2r}{r}$ is an increasing function of r so that

$$\binom{|k-\ell|+2r}{r} \le \binom{k+\ell-2}{\min(k,\ell)-1},$$

which in turn can be upper bounded by the "middle" binomial coefficient so that

$$\binom{|k-\ell|+2r}{r} \le \binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil}.$$

We can also bound $1/(\min(k,\ell)-1-r)!$ by 1, and therefore

$$|c_{k,\ell,p}| \le \sqrt{\mu_\gamma}\; t_\gamma^{|k-\ell|+2p}\sqrt{\frac{(k-1)!\,(\ell-1)!}{(|k-\ell|+2p)!}}\binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil} \times \sum_{r=\max(p,0)}^{\min(k,\ell)-1}\left|\frac{t_\gamma^2-1}{2}\right|^{r-p}\frac{1}{(r-p)!}$$

$$\le \sqrt{\mu_\gamma}\; t_\gamma^{|k-\ell|+2p}\sqrt{\frac{(k-1)!\,(\ell-1)!}{(|k-\ell|+2p)!}}\binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil}\exp\left(\left|\frac{t_\gamma^2-1}{2}\right|\right). \tag{4.9}$$

Note that we have $\omega_\gamma<1$ and

$$\frac{t_\gamma^2-1}{2} = \frac{1-\sqrt{1+4\gamma^2}}{4\sqrt{1+4\gamma^2}} = \frac{-\gamma^2}{1+4\gamma^2+\sqrt{1+4\gamma^2}} < 0\,.$$

Thus from (4.8) and (4.9) we obtain

$$\|g_{k,\ell}\|^2_{H(K_\gamma)} \le \frac{1}{(1-\omega_\gamma)\,\omega_\gamma^{k+\ell-2}}\sum_{p=-\lfloor\frac{|k-\ell|}{2}\rfloor}^{\min(k,\ell)-1} c^2_{k,\ell,p}$$

$$\le \frac{\mu_\gamma(k-1)!\,(\ell-1)!}{(1-\omega_\gamma)\,\omega_\gamma^{k+\ell-2}}\binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil}^2\exp(1-t_\gamma^2)\sum_{p=-\lfloor\frac{|k-\ell|}{2}\rfloor}^{\min(k,\ell)-1}\frac{(t_\gamma^2)^{|k-\ell|+2p}}{(|k-\ell|+2p)!}$$

$$\le \frac{\mu_\gamma(k-1)!\,(\ell-1)!}{(1-\omega_\gamma)\,\omega_\gamma^{k+\ell-2}}\binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil}^2\exp(1-t_\gamma^2)\,\exp(t_\gamma^2)\,.$$

It is easy to check that the sequence

$$a_m = \binom{m}{\lceil\frac{m}{2}\rceil} \quad \text{for } m\in\mathbb{N}$$

is non-decreasing. Therefore for $k, \ell \in [1, n+1]$ we have

$$(k-1)!\,(\ell-1)!\binom{k+\ell-2}{\lceil\frac{k+\ell-2}{2}\rceil}^2 \le (n!)^2\binom{2n}{n}^2 = \left(\frac{(2n)!}{n!}\right)^2.$$

Hence, we may further estimate the norm of $g_{k,\ell}$ by

$$\|g_{k,\ell}\|_{H(K_\gamma)} \le \sqrt{\frac{\mu_\gamma e}{1-\omega_\gamma}\frac{(2n)!}{\omega_\gamma^n\, n!}}\,.$$

The proof is completed by substituting the expressions for μ_γ and ω_γ. □

We can simplify the lower bound (4.1) on $e_{1,n}^{\min}$ in Theorem 4.1 by using Stirling's formula $n!/(2n)! \sim [e/(4n)]^n/\sqrt{2}$, and by noting that $\omega_\gamma \ge 2\gamma^2/(3+\sqrt{5})$. Then for some $c_1, c_2 \in (0,1)$ we know that (4.1) is lower bounded by $c_1(c_2\,\gamma^2/n)^n/(n+1)$ for all $n \in \mathbb{N}$. This and the upper bound on $e_{1,n}^{\min}$ supplied by the error of the Gauss-Hermite quadrature (1.4) yield

$$\frac{c_1}{n+1}\left(\frac{c_2\,\gamma^2}{n}\right)^n \le e_{1,n}^{\min} \le \gamma^{2n}.$$

Hence, both the lower and upper bounds on the minimal error $e_{1,n}^{\min}$ are exponentially small in n. We claim that the difference between lower and upper bounds is not really essential although we have an extra factor $(c_2/n)^n/(n+1)$ in the lower bound. Indeed, let us compare what is the minimal $n = n(\varepsilon, 1)$ for which $e_{1,n}^{\min}$ is at most ε. It is easy to check, and we will do it for the minimal $n(\varepsilon, d)$ for general d in the next section, that there are positive numbers c_3, c_4 and $\varepsilon_1 \in (0,1)$ such that

$$c_3\,\frac{\ln \varepsilon^{-1}}{\ln\ln \varepsilon^{-1}} \le n(\varepsilon,1) \le c_4 \ln \varepsilon^{-1} \qquad \text{for all } \varepsilon \in (0, \varepsilon_1).$$

As we see, there is not much difference between these bounds. That is why we say that the Gauss-Hermite quadrature is nearly optimal.

5. Lower bound for $d \ge 1$

In this section we extend the lower bound to arbitrary d by making use of the tensor product structure.

Theorem 5.1. *Let N be given by*

$$N+1 := \prod_{j=1}^{d}(m_j+1) \quad \text{with } m_j \in \mathbb{N}. \tag{5.1}$$

The Nth minimal error of multivariate integration in the function space $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ is bounded from below by

$$e_{d,N}^{\min} \ge \frac{1}{N+1}\prod_{j=1}^{d}\left(\sqrt{\frac{2\,(1+4\gamma_j^2)^{1/4}}{(1+2\gamma_j^2+\sqrt{1+4\gamma_j^2})e}\,\frac{\omega_{\gamma_j}^{m_j}\, m_j!}{(2m_j)!}}\right), \tag{5.2}$$

where $\omega_{\gamma_j} < 1$ is given by (2.1), but with γ replaced by γ_j.

Proof. We follow closely the argument in the proof of Theorem 4.1 for the univariate case. In fact, for $d = 1$ we have $N = m_1$ and the lower bound in (5.2) is the same as the lower bound (4.1) with $\gamma = \gamma_1$.

Let $\tilde{\boldsymbol{\gamma}} = \{\tilde{\gamma}_j\}$ be defined by (see (4.7))

$$\tilde{\gamma}_j \ := \ \frac{1}{2}\sqrt{\left(\frac{1+\sqrt{1+4\gamma_j^2}}{2}\right)^2 - 1} \ < \ \gamma_j \quad \text{for all} \quad j = 1, 2, \ldots .$$

Given N and positive integers $m_1, \ldots, m_d$ satisfying (5.1), we define the index set

$$\mathcal{L} \ := \ \{\boldsymbol{\ell} = (\ell_1, \ell_2, \ldots, \ell_d) \in \mathbb{N}^d \ : \ 1 \le \ell_j \le m_j + 1 \quad \text{for all} \quad j = 1, 2, \ldots, d\} \, .$$

Clearly the cardinality of $\mathcal{L}$ is $|\mathcal{L}| = \prod_{j=1}^d (m_j + 1) = N + 1$. Given arbitrary points $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N$ from $\mathbb{R}^d$, we choose real numbers $a_{\boldsymbol{\ell}}$ for $\boldsymbol{\ell} \in \mathcal{L}$ such that

$$\sum_{\boldsymbol{\ell} \in \mathcal{L}} a_{\boldsymbol{\ell}} \, \varphi_{\boldsymbol{\ell}, \tilde{\boldsymbol{\gamma}}}(\boldsymbol{t}_i) = 0 \quad \text{for all} \quad i = 1, 2, \ldots, N \, ,$$

where $\varphi_{\boldsymbol{\ell}, \tilde{\boldsymbol{\gamma}}}(\boldsymbol{x}) = \prod_{j=1}^d \varphi_{\ell_j, \tilde{\gamma}_j}(x_j)$ is as defined in (2.3) but with each γ_j replaced by $\tilde{\gamma}_j$.

Since the linear system has more unknowns than the number of conditions, there exists a solution such that $\max_{\boldsymbol{\ell} \in \mathcal{L}} |a_{\boldsymbol{\ell}}| = 1 = a_{\boldsymbol{k}}$ for some multiindex $\boldsymbol{k} \in \mathcal{L}$. We now define

$$f^*(\boldsymbol{x}) \ := \ \varphi_{\boldsymbol{k}, \tilde{\boldsymbol{\gamma}}}(\boldsymbol{x}) \sum_{\boldsymbol{\ell} \in \mathcal{L}} a_{\boldsymbol{\ell}} \, \varphi_{\boldsymbol{\ell}, \tilde{\boldsymbol{\gamma}}}(\boldsymbol{x}) \, .$$

Then $f^*(\boldsymbol{t}_i) = 0$ for all $i = 1, 2, \ldots, N$ so that any quadrature rule $\mathcal{A}_{d,N}$ based on linear combinations of these function values is 0, that is, for any choice of weights $w_1, \ldots, w_N$ we have

$$\mathcal{A}_{d,N} f^* \ = \ \sum_{i=1}^N w_i \, f^*(\boldsymbol{t}_i) \ = \ 0 \, .$$

On the other hand, we have

$$\mathcal{I}_d f^* \ = \ \int_{\mathbb{R}^d} f^*(\boldsymbol{x}) \prod_{j=1}^d \frac{\exp(-x_j^2)}{\sqrt{\pi}} \, \mathrm{d}\boldsymbol{x} \ = \ \sum_{\boldsymbol{\ell} \in \mathcal{L}} a_{\boldsymbol{\ell}} \, \langle \varphi_{\boldsymbol{k}, \tilde{\boldsymbol{\gamma}}}, \varphi_{\boldsymbol{\ell}, \tilde{\boldsymbol{\gamma}}} \rangle_{L_2} \ = \ a_{\boldsymbol{k}} \ = \ 1 \, .$$

Thus the worst case error in $H(\mathcal{K}_{d,\boldsymbol{\gamma}})$ is bounded from below by

$$e(\mathcal{A}_{d,N}) \ \ge \ \frac{|\mathcal{I}_d f^* - \mathcal{A}_{d,N} f^*|}{\|f^*\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})}} \ = \ \frac{1}{\|f^*\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})}} \, . \tag{5.3}$$

Due to the tensor product structure, we can write

$$f^*(\boldsymbol{x}) \ = \ \sum_{\boldsymbol{\ell} \in \mathcal{L}} a_{\boldsymbol{\ell}} \prod_{j=1}^d \varphi_{k_j, \tilde{\gamma}_j}(x_j) \, \varphi_{\ell_j, \tilde{\gamma}_j}(x_j) \ = \ \sum_{\boldsymbol{\ell} \in \mathcal{L}} a_{\boldsymbol{\ell}} \prod_{j=1}^d g_{k_j, \ell_j}(x_j) \, ,$$

where g_{k_j, ℓ_j} is as defined in (4.4), but with k, ℓ and $\tilde{\gamma}$ replaced by k_j, ℓ_j, and $\tilde{\gamma}_j$, respectively. Using $|a_{\boldsymbol{\ell}}| \le a_{\boldsymbol{k}} = 1$ and Lemma 4.2, but with n replaced by m_j for each j, we obtain

$$\begin{aligned}\|f^*\|_{H(\mathcal{K}_{d,\boldsymbol{\gamma}})} &\le \sum_{\boldsymbol{\ell}\in\mathcal{L}} \prod_{j=1}^d \|g_{k_j,\ell_j}\|_{H(K_{\gamma_j})} \\ &= \prod_{j=1}^d \sum_{\ell_j=1}^{m_j+1} \|g_{k_j,\ell_j}\|_{H(K_{\gamma_j})} \\ &\le \prod_{j=1}^d \left[(m_j+1)\sqrt{\frac{(1+2\gamma_j^2+\sqrt{1+4\gamma_j^2})e}{2\,(1+4\gamma_j^2)^{1/4}}\,\frac{(2m_j)!}{\omega_{\gamma_j}^{m_j}\,m_j!}}\right].\end{aligned}$$

Applying this estimate in (5.3) completes the proof. □

Based on the lower bound (5.2) we now show that the exponent of EXP is at least $1/d$. Let

$$g(x) := \sqrt{\frac{2\,(1+4x)^{1/4}}{(1+2x+\sqrt{1+4x})e}} \qquad \text{for } x\in[0,1].$$

Since $g(x)>0$ for all $x\in[0,1]$ and g is continuous, we have

$$g_{\min} := \min_{x\in[0,1]} g(x) \;>\; 0.$$

Applying this to (5.2) we have for $N+1=\prod_{j=1}^d (m_j+1)$,

$$e^{\min}_{d,N} \;\ge\; \frac{g_{\min}^d}{N+1}\prod_{j=1}^d\left(\omega_{\gamma_j}^{m_j}\,\frac{m_j!}{(2m_j)!}\right). \tag{5.4}$$

For any integer $k\ge 2$, take $m_j=k-1$ for all $j=1,2,\dots,d$. Then $N=k^d-1$ and $k=(N+1)^{1/d}$. Clearly,

$$\frac{m_j!}{(2m_j)!} = \frac{(k-1)!}{[2(k-1)]!} = \frac{k!}{(2k)!}\,\frac{(2k-1)\,2k}{k} \;=\; \frac{k!}{(2k)!}\,(4k-2).$$

Since $\sqrt{2\pi}\,k^{k+1/2}e^{-k}\le k!\le e\,k^{k+1/2}\,e^{-k}$ and $4k-2\ge 3k$ we obtain from (5.4),

$$e^{\min}_{d,N} \;\ge\; g_{\min}^d\left[\frac{3\sqrt{\pi}}{e}\,\left(\frac{e}{4k}\right)^k\right]^d \prod_{j=1}^d \omega_{\gamma_j}^{k-1}. \tag{5.5}$$

Since $\omega_{\gamma_j}\ge\omega_{\gamma_d}$, we rewrite the last estimate in terms of N as

$$e^{\min}_{d,N} \;\ge\; \left(\frac{3\sqrt{\pi}}{e}\,g_{\min}\right)^d \left(\frac{e}{4}\right)^{d\,(N+1)^{1/d}} \left(\frac{1}{N+1}\right)^{(N+1)^{1/d}} \omega_{\gamma_d}^{d((N+1)^{1/d}-1)}. \tag{5.6}$$

Take arbitrary $q\in(0,1)$. Using $a^x=q^{x\,\ln a^{-1}/\ln q^{-1}}$ for any positive a and real x, we rewrite (5.6) as

$$e^{\min}_{d,N} \;\ge\; \left(\frac{3\sqrt{\pi}}{e}\,g_{\min}\right)^d q^{\,p_{d,N}}$$

with

$$p_{d,N} \;=\; d\left[((N+1)^{1/d}-1)\,\frac{\ln\,\omega_d^{-1}}{\ln\,q^{-1}} + (N+1)^{1/d}\,\frac{\ln(4/e)}{\ln\,q^{-1}}\right] + \frac{(N+1)^{1/d}\,\ln(N+1)}{\ln\,q^{-1}}.$$

For N tending to infinity we have

$$p_{d,N} = \frac{1+o(1)}{\ln\, q^{-1}}\; N^{1/d}\, \ln\, N\,.$$

This shows that (1.6) can only hold if $p_d < 1/d$, and therefore $p_d^* \le 1/d$. From this and (3.8) we conclude that the exponent of EXP is

$$p_d^* = \frac{1}{d}\,,$$

which proves Theorem 1.1(b).

This also implies that we cannot have UEXP, which in turn implies the lack of EC-PT. Hence, Theorem 1.1(c) is also proved.

We now find a lower estimate on $n(\varepsilon,d)$ for a fixed d and for ε tending to zero for the absolute and normalized error criteria. We know that $e_{d,N}^{\min} \le \varepsilon\,\mathrm{CRI}_d$ for all $N \ge n(\varepsilon,d)$. We already remarked in the introduction that $n(\varepsilon,d)=0$ for $\varepsilon \ge \|\mathcal{I}_d\|$ for the absolute error criterion. That is why we take $\varepsilon < \|\mathcal{I}_d\| = \prod_{j=1}^d (1+2\gamma_j^2)^{-1/4}$ for the absolute error criterion and $\varepsilon \in (0,1)$ for the normalized error criterion. Then $n(\varepsilon,d) \ge 1$. Choose now N as the smallest integer of the form $N = k^d - 1$ with $k \ge 2$ and $N \ge n(\varepsilon,d)$. Clearly, this corresponds to

$$k = k(\varepsilon,d) := \left\lceil (1+n(\varepsilon,d))^{1/d} \right\rceil \tag{5.7}$$

which indeed is at least 2. Since $N \ge n(\varepsilon,d)$, we have $e_{d,N}^{\min} \le \varepsilon\,\mathrm{CRI}_d$ and hence we obtain from (5.5) by taking logarithms

$$k\,\ln\,k \;+\; k\,\alpha_d \;+\; \ln\,\frac{e}{3\sqrt{\pi}\,g_{\min}} \;\ge\; \frac{\ln\,\frac{1}{\varepsilon\,\mathrm{CRI}_d}}{d}, \tag{5.8}$$

where

$$\alpha_d = \ln\frac{4}{e} + \frac{1}{d}\sum_{j=1}^d \ln\,\frac{1}{\omega_{\gamma_j}}. \tag{5.9}$$

Note that we have replaced $\omega_{\gamma_j}^{k-1}$ in the lower bound (5.5) by $\omega_{\gamma_j}^{k}$, which is valid since $\omega_{\gamma_j} < 1$. This proves that $k = k(\varepsilon,d)$ as well as $n(\varepsilon,d)$ goes to infinity as ε goes to zero. Take now an arbitrary $c \in (0,1)$ and let

$$x = c\,\frac{\ln\,\frac{1}{\varepsilon\,\mathrm{CRI_d}}}{d} \quad \text{and} \quad k = \frac{x}{\ln\,x}.$$

Then for ε tending to zero we have x tending to infinity and

$$k\,\ln\,k \;+\; k\,\alpha_d \;+\; \ln\,\frac{e}{3\,\sqrt{\pi}\,g_{\min}} \;=\; c\;(1+o(1))\,\frac{\ln\,\frac{1}{\varepsilon\,\mathrm{CRI}_d}}{d}.$$

This means that the inequality (5.8) is not satisfied for small ε. Hence for small ε, the solution of (5.8) satisfies

$$k \;\ge\; \frac{x}{\ln\,x}.$$

We have from (5.7) that $k-1 < (1+n(\varepsilon,d))^{1/d} \le k$, and so

$$n(\varepsilon,d) \;\ge\; (k-1)^d \;\ge\; \Big(\frac{x}{\ln\,x} - 1\Big)^d.$$

Since x is of order $\ln \varepsilon^{-1}$ for small ε, the lower bound on $n(\varepsilon, d)$ is of order

$$\left(\frac{\ln \varepsilon^{-1}}{\ln \ln \varepsilon^{-1}}\right)^d.$$

This proves the lower estimate in Theorem 1.1(d). Together with (3.10) of Corollary 3.3 this completes the proof of Theorem 1.1(d).

We now turn to EC-WT. We want to show that EC-WT does not hold if $\lim_{j\to\infty} \gamma_j^2 > 0$. It is enough to show this fact for the absolute error criterion since the normalized error criterion is harder.

Observe that α_d in (5.9) is uniformly bounded in d iff $\ln(1/\omega_{\gamma_j})$ does not go to infinity. This, in turn, holds iff γ_j^2 does not go to zero. Hence, for $\gamma_\infty^2 := \lim_{j\to\infty} \gamma_j^2 > 0$ there is a positive α such that $\alpha_d \le \alpha$. Then k also satisfies the inequality

$$k \ln k \;+\; k\,\alpha \;+\; \ln \frac{e}{3\sqrt{\pi}\, g_{\min}} \;\ge\; \frac{\ln \varepsilon^{-1}}{d}.$$

Choose a (large) positive number C such that

$$C > \max\left(\frac{1}{12}\,\ln(1+2\gamma_1^2), \ln \frac{e}{3\sqrt{\pi}\, g_{\min}}\right).$$

For any $d \in \mathbb{N}$, we take $\varepsilon = \varepsilon_d$ such that $\ln \varepsilon_d^{-1} = 3\,C\,d$. Then we have

$$\varepsilon_d \;=\; e^{-3\,C\,d} \;<\; (1+2\gamma_1^2)^{-d/4} \;\le\; \prod_{j=1}^{d}(1+2\gamma_j^2)^{-1/4} \;=\; \|\mathcal{I}_d\|,$$

as needed. Then k satisfies the inequality

$$k \ln k \;+\; k\,\alpha \;\ge\; 2\,C.$$

As before, we can show that $k \ge 1 + C/\ln C$ for large C which is independent of d and ε^{-1}. Then

$$\frac{\ln n(\varepsilon_d, d)}{d + \ln \varepsilon_d^{-1}} \ge \frac{\ln (k-1)^d}{(3C+1)d} \ge \frac{\ln(C/\ln C)}{3C+1} > 0 \quad \text{for all } d.$$

This proves that the left hand side of the last inequality cannot go to zero with d approaching infinity. This contradicts EC-WT and proves Theorem 1.1(f).

6. Concluding remarks

We have completed the proof for our main result, Theorem 1.1. The upper bound of the nth minimal error is obtained by considering the tensor product algorithm of univariate Gauss-Hermite quadratures. We have obtained complete knowledge regarding EXP, UEXP, EC-SPT, EC-PT, and EC-(t,κ)-WT for $t > 1$. However, our results for EC-WT and EC-$(1,\kappa)$-WT are not complete.

Acknowledgements

The authors acknowledge the support of the Australian Research Council under the projects DP110100442, FT130100655, and DP150101770. Henryk Woźniakowski is supported partially by the National Science Foundation and by the National Science Centre, Poland, based on the decision DEC-2013/09/B/ST1/04275.

References

1. N. S. Bakhvalov, *On the optimality of linear methods for operator approximations in convex classes of functions*, *USSR Comp. Math. Mech. Phys.*, 11, 244–249, 1971.
2. A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*, Kluwer Academic Publishers, Boston, 2004.
3. M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, 2003.
4. F. Cucker and D. X. Zhou, *Learning theory: An approximation theory viewpoint*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2007.
5. J. Dick, G. Larcher, F. Pillichshammer, and H. Woźniakowski, *Exponential convergence and tractability of multivariate integration for Korobov spaces*, Math. Comp., 80, 905–930, 2011.
6. J. Dick, P. Kritzer, F. Pillichshammer, and H. Woźniakowski, *Approximation of analytic functions in Korobov spaces* J. Complexity, 30, 2–28, 2014.
7. G. E. Fasshauer, *Meshfree Approximation Methods with* MATLAB, World Scientific Publishers, Singapore, 2007.
8. G. E. Fasshauer, F. J. Hickernell, and H. Woźniakowski, *Average case approximation: convergence and tractability of Gaussian kernels*, *Monte Carlo and Quasi-Monte Carlo 2010*, eds. L. Plaskota and H. Woźniakowski, 329–345, Springer Verlag, 2012.
9. G. E. 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.
10. E. Feldheim, *Relations entre les polynomes de Jacobi, Laguerre et Hermite*, Acta mathematica, 75, 117–138, 1942.
11. A. I. J. Forrester, A. Sóbester, and A. J. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*, Wiley, Chichester, 2008.
12. T. Hastie, R. Tibshirani, and J. Friedman, *Elements of statistical learning: Data mining, inference, and prediction*, Second Edition, Springer Series in Statistics, Springer, New York, 2009.
13. F. B. Hildebrand, *Introduction to Numerical Analysis*, Second Edition, McGraw-Hill, 1974.
14. C. Irrgeher, P. Kritzer, G. Leobacher, and F. Pillichshammer, *Integration in Hermite spaces of analytic functions*, J. Complexity, 31, 380–404, 2015.
15. P. Kritzer, F. Pillichshammer, and H. Woźniakowski, *Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces*, Math. Comp., 83, 1189–1206, 2014.
16. P. Kritzer, F. Pillichshammer, and H. Woźniakowski, *Tractability of multivariate analytic problems*, in *Uniform Distribution and Quasi-Monte Carlo Methods, Discrepancy, Integration and Applications*, eds. P. Kritzer, H. Niederreiter, F. Pillichshammer and A. Winterhof, 147–170, DeGruyter, Berlin, 2014.
17. F. Y. Kuo and H. Woźniakowski, *Gauss-Hermite quadratures for functions from Hilbert spaces with Gaussian reproducing kernel*, BIT Numer. Math., 52, 425–436, 2012.
18. E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, Volume I: Linear Information*, EMS, Zurich, 2008.
19. E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, Volume II: Standard Information for Functionals*, EMS, Zurich, 2010.
20. E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, Volume III: Standard Information for Operators*, EMS, Zurich, 2012.
21. A. Papageorgiou and I. Petras, A new criterion for tractability of multivariate problems, *J. Complexity*, 30, 604–619, 2014.
22. C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006 (online version at `http://www.gaussianprocess.org/gpml/`).
23. R. Schaback and H. Wendland, *Kernel techniques: From machine learning to meshless methods*, *Acta Numerica*, 15, 543–639, 2006.
24. B. Schölkopf and A. J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond*, MIT Press, Cambridge, Massachusetts, 2002.

25. P. Siedlecki and M. Weimar, *Notes on (s,t)-weak tractability: a refined classification of problems with (sub)exponential information-based complexity*, *J. Approx. Theory*, 200, 227–258, 2015.
26. I. H. Sloan and H. Woźniakowski, *An intractability result for multiple integration*, *Math. Comp.*, 66, 1119–1124, 1997.
27. S. A. Smolyak, *On optimal restoration of functions and functionals of them*, (in Russian), Candidate Dissertation, *Moscow State University*, 1965.
28. M. L. Stein, *Interpolation of spatial data. Some theory for Kriging*, Springer-Verlag, New York, 1999.
29. I. Steinwart and A. Christmann, *Support Vector Machines*, Springer Verlag, Berlin, 2008.
30. I. Steinwart, D. Hush, and C. Scovel, *An explicit description of the reproducing kernel Hilbert spaces of Gaussian RBF kernels*, *IEEE Transactions on Information Theory*, vol. 52, no.10, 4635–4663, 2006.
31. J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.
32. G. Wahba, *Spline models for observational data*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 59, SIAM, Philadelphia, 1990.
33. G. N. Watson, *A Note on the Polynomials of Hermite and Laguerre*, J. London Math. Soc. s1-13 (1), 29–32, 1938
34. H. Wendland, *Scattered data approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2005.

School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia

E-mail address: `f.kuo@unsw.edu.au`

School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia

E-mail address: `i.sloan@unsw.edu.au`

Department of Computer Science, Columbia University, New York, NY 10027, USA, and Institute of Applied Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland

E-mail address: `henryk@cs.columbia.edu`