Construction of Optimal Algorithms for Function Approximation in Gaussian Sobolev Spaces

Yuya Suzuki[1] and Toni Karvonen[2]

[1]*Department of Mathematics and Systems Analysis, Aalto University, P.O. Box 11100, 00076, Helsinki, Finland*

[2]*School of Engineering Science, Lappeenranta–Lahti University of Technology, Yliopistonkatu 34, 53850 Lappeenranta, Finland*

March 20, 2026

Abstract

This paper studies function approximation in Gaussian Sobolev spaces over the real line and measures the error in a Gaussian-weighted L^p-norm. We construct two linear approximation algorithms using n function evaluations that achieve the optimal or almost optimal rate of worst-case convergence in a Gaussian Sobolev space of order α. The first algorithm is based on scaled trigonometric interpolation and achieves the optimal rate $n^{-\alpha}$ up to a logarithmic factor. This algorithm can be constructed in almost-linear time with the fast Fourier transform. The second algorithm is more complicated, being based on spline smoothing, but attains the optimal rate $n^{-\alpha}$.

1 Introduction

This paper is concerned with approximating functions in Gaussian-weighted Sobolev spaces over the real line. We are interested in the *sampling recovery problem*, constructing linear approximation algorithms $A_n(f)$ that recover a function $f \colon \mathbb{R} \to \mathbb{R}$ using n function evaluations. Let $\rho(x) := (2\pi)^{-1/2} \, \mathrm{e}^{-x^2/2}$ be the Gaussian density function and $p \in [1, \infty)$. We measure the approximation error in the weighted $L^p(\mathbb{R})$ norm

$$\|f - A_n(f)\|_{L_\rho^p} := \left(\int_{\mathbb{R}} |f(x) - [A_n(f)](x)|^p \, \rho(x) \, \mathrm{d}x \right)^{1/p} \tag{1.1}$$

and assume that f is an element of the Gaussian Sobolev space

$$W_\rho^{\alpha,q} := \left\{ f \in L_\rho^q \;\middle|\; \|f\|_{W_\rho^{\alpha,q}} := \left(\sum_{\tau=0}^{\alpha} \int_{\mathbb{R}} |f^{(\tau)}(x)|^q \rho(x) \, \mathrm{d}x \right)^{1/q} < \infty \right\} \tag{1.2}$$

for some $\alpha \in \mathbb{N}$ and $p < q < \infty$. This space consists of functions whose weak derivatives up to order α are in L_ρ^q. Since every element of this space admits a continuous representative, in what follows we always consider the continuous representative so that the function evaluation makes sense.

Function approximation in high dimension is an important task. The sampling recovery problem in Gaussian Sobolev spaces has been recently studied by a number of authors [17, 3, 7]. In this paper, we study sampling recovery in the one-dimensional Gaussian Sobolev spaces $W_\rho^{\alpha,q}$ and construct explicit linear algorithms that attain the optimal or almost optimal rate of worst-case convergence. It has been shown by Kuo, Plaskota, and Wasilkowski [15, Theorem 3], and later independently studied by Dũng and Nguyen [3, Theorem 3.3], that the best possible convergence rate of sampling recovery in $W_\rho^{\alpha,q}$ for $p < q$ is of order $n^{-\alpha}$, and we show two alternative methods achieving this rate with explicit computational cost. As one-dimensional algorithms often serve as a foundation for their high-dimensional counterparts, our results motivate further study of such optimal algorithms in a high-dimensional setting.

The first algorithm we propose in Section 2, scaled trigonometric interpolation, achieves the optimal convergence rate up to a logarithmic factor. This algorithm is nothing but a trigonometric interpolation on a suitably truncated interval, and thus can be constructed by Fast Fourier Transform (FFT) with $\mathcal{O}(n \log_2 n)$ computational cost and $\mathcal{O}(n)$ memory usage. Trigonometric interpolation is a popular numerical tool due to the applicability of FFT. In spectral methods, trigonometric bases are often used even if the original partial differential equation is defined on the real line (e.g., [13, 33]). This requires suitable treatment of boundary conditions and truncation of the real line. The truncation interval is chosen to be *large enough*, but analysis of the error that this treatment causes is often omitted. Our results on trigonometric interpolation give theoretical foundation for such methods, and suggest how to choose the truncation interval depending on the number of evaluation points.

In Section 3 we propose the second algorithm, which is based on spline smoothing. The algorithm converges with the optimal rate $n^{-\alpha}$ and is similar to the numerical integration algorithms in [10, 3]. This is achieved by partitioning a truncated real line into unit intervals and constructing an independent spline smoother on each interval. As spline smoothers are known to attain the optimal rate of convergence in classical Sobolev spaces on bounded sets, having the number of evaluation points allocated to an interval decrease exponentially fast as a function of the distance of the interval to the origin ensures that the algorithm attains the optimal rate. The slightly improved convergence rate of this algorithm in comparison to scaled trigonometric interpolation is offset by its higher computational complexity. Figures 1 and 2 compare the two algorithms that we propose.

Numerical integration is closely related to function approximation. Here we mention results related to numerical integration over Gaussian Sobolev spaces for $q = 2$. For the deterministic worst-case error, it has been shown that the rate $n^{-\alpha}$ is optimal by Dick et al. [4, Theorem 1], meaning that any deterministic linear quadrature cannot achieve a convergence rate faster than $n^{-\alpha}$. We note that their result is shown for so-called Hermite spaces, but the norm equivalence between Hermite spaces and Gaussian Sobolev spaces with $q = 2$ is shown in [4, Lemma 6] and [11, Lemma 2.1]. Recently, Dũng and Nguyen [3, Theorem 2.3] showed that this rate is unchanged for general $q \in (1, \infty)$. The measure being Gaussian, Gauss–Hermite quadrature is a natural choice. However, Kazashi, Goda and one of the present authors [11] have shown that Gauss–Hermite quadrature achieves the deterministic worst-case rate of only $n^{-\alpha/2}$. An upper bound of order $n^{-\alpha/2}$ was essentially obtained already by Mastroianni and Monegato [19] and the matching lower bound was proved in [11, Theorem 3.2]. In contrast, a suitably truncated trapezoidal rule is shown to achieve the optimal rate $n^{-\alpha}$ up to a logarithmic factor in [11, Theorem 4.5]. The subsequent work of these authors [8] considers the randomized setting, i.e., algorithms are allowed to be random and their quality is measured in the worst-case root-mean-squared error (RMSE). Therein it is proved that no nonlinear adaptive algorithm can converge with a worst-case RMSE rate faster than $n^{-\alpha-1/2}$, and a randomized trapezoidal rule is shown to attain this rate. The strategy used for obtaining upper bounds of trapezoidal rules both in [11, 8] is based on

$$n = 30$$

$$n = 62$$
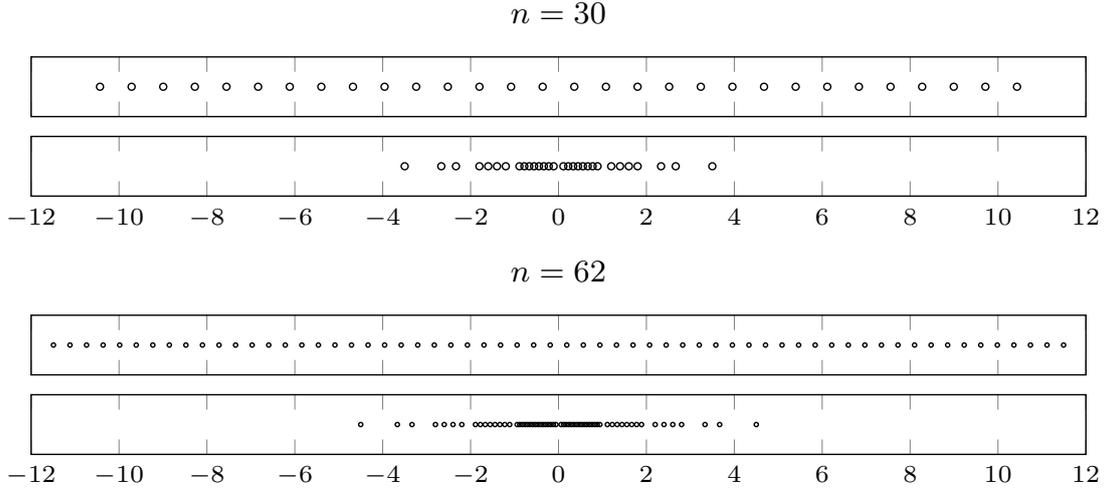
Figure 1: Comparison of the evaluation points used by the algorithms $A_n^\dagger$ and A_n^* from Sections 2 and 3, respectively, for $n = 30$ and $n = 62$. The uniform points used by $A_n^\dagger$ in (2.1), with T given in (2.10) for $\alpha = 4$, $q = 2$, $p = 1$ and $\varepsilon = 0.25$, are displayed in the top panels. The bottom panels display one possible set of points that can be used to construct the algorithm A_n^* of Corollary 3.5. The density of these points decreases exponentially fast when moving away from the origin. The algorithm A_n^* converges with the optimal rate $n^{-\alpha}$, while the rate of convergence of $A_n^\dagger$ is optimal up to a logarithmic factor.

introducing an auxiliary periodic function. This strategy is first used in [25], and we adapt this strategy for function approximation in Section 2.

Upon completing this research in the paper, we learned the paper [5] by Dũng. The methods presented therein achieve the optimal rate of convergence, however they depend on various indices. We emphasize that our methods presented in this paper offer computationally cheap alternatives, especially the trigonometric interpolation. As an additional feature, our methods can be constructed without knowledge of the exact smoothness of the target function. Section 4 concludes this paper by briefly discussing known lower bounds, optimality, and other features of our algorithms.

Notation. Throughout the paper, we denote the set of all positive integers by $\mathbb{N}$. Unless otherwise specified, in what follows we always assume that p and q in (1.1) and (1.2) satisfy $1 \leq p < q < \infty$. Over a given interval $[a, b]$, we denote unweighted and weighted L^p spaces by $L^p([a, b])$ and $L_\rho^p([a, b])$, respectively, and likewise for unweighted and weighted Sobolev spaces. We use subscripts to denote the dependency of non-negative constants on various parameters. For example, $C_{\alpha, \beta}$ would be a constant that depends *only* on α and β. Otherwise we always specify the parameters that a constant depends or does not depend on.

2   Trigonometric interpolation

This section introduces our first algorithm based on trigonometric interpolation and analyses its L_ρ^p-error in terms of the number n of function evaluation. The algorithm is denoted by $A_n^\dagger : W_\rho^{\alpha, q} \to L_\rho^p$.

3
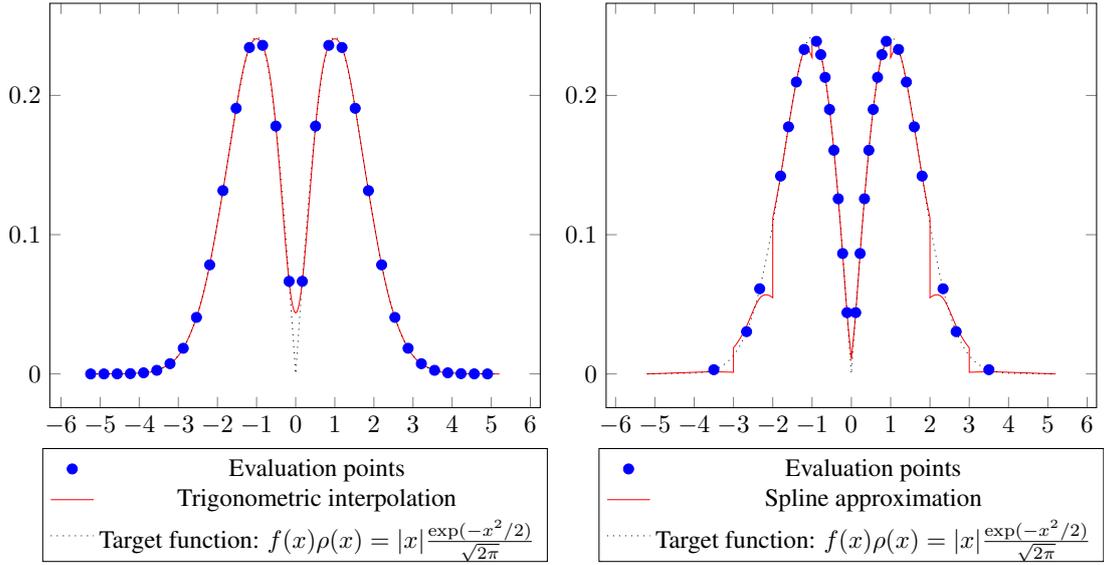
Figure 2: A comparison of trigonometric interpolation from Section 2 and spline approximation from Section 3 when $f(x) = |x|$, $p = 1$ and $q = 2$. We plot the weighted approximations $\rho A_n^\dagger(f)$ with $n = 31$ and $\varepsilon = 0.25$ (left) and $\rho A_n^*(f)$ with $n = 30$ (right). The spline approximation has been constructed using a Matérn kernel of order $\gamma = 3/2$ [see (3.6)], which is a reproducing kernel for a Sobolev space of order two. Observe how the spline approximation is more accurate close to the origin than the trigonometric interpolant.

Definition 2.1 (Trigonometric interpolation with cutoff $A_n^\dagger$)**.** Let $g \coloneqq f\rho^{1/p}$. We define the approximation algorithm

$$A_n^\dagger(f) \coloneqq \rho^{-1/p} \mathcal{I}_n(f\rho^{1/p}) = \rho^{-1/p} \mathcal{I}_n(g), \tag{2.1}$$

with $\mathcal{I}_n$ being the trigonometric interpolation on $[-T, T]$:

$$\mathcal{I}_n(g) \coloneqq \begin{cases} \sum_{k=-\lfloor n/2 \rfloor}^{\lfloor n/2 \rfloor} \widehat{g}_a(k) \phi_k^{[-T,T]}(x) & \text{if } x \in [-T, T], \\ 0 & \text{otherwise,} \end{cases}$$

where $\phi_k^{[-T,T]}(x) \coloneqq \exp(2\pi i k(x + T)/(2T))/\sqrt{2T}$ are orthonormal Fourier basis on $L^2([-T, T])$. The coefficients $\widehat{g}_a(k)$ are calculated using equidistant points as

$$\widehat{g}_a(k) \coloneqq \frac{1}{n} \sum_{j=0}^{n-1} g(\xi_j) \overline{\phi_k^{[-T,T]}}(\xi_j), \quad \text{with} \quad \xi_j \coloneqq \frac{2T}{n} j - T, \ j = 0, \ldots, n-1.$$

We will bound the L_ρ^p error of the above algorithm by decomposing it into two parts:

$$\begin{aligned} \|f - A_n^\dagger(f)\|_{L_\rho^p} &\leq \left(\int_{\mathbb{R} \setminus [-T,T]} \left| f(x) - [A_n^\dagger(f)](x) \right|^p \rho(x)\, \mathrm{d}x \right)^{1/p} \\ &\quad + \left(\int_{-T}^T \left| f(x) - [A_n^\dagger(f)](x) \right|^p \rho(x)\, \mathrm{d}x \right)^{1/p}. \end{aligned} \tag{2.2}$$

For the first term, since we have $A_n^\dagger(f) = 0$ outside the interval $[-T, T]$, we simply bound $(\int_{\mathbb{R}\setminus[-T,T]} |f(x)|^p \rho(x)\,\mathrm{d}x)^{1/p}$ using the decay of $g = f\rho^{1/p}$. In the following lemma, we obtain this decay of g and its derivatives.

Lemma 2.2 (Decay of the function $g = f\rho^{1/p}$). *Let $1 \le p < q < \infty$, $\alpha \in \mathbb{N}$ and $f \in W_\rho^{\alpha,q}$. Then for arbitrary $\varepsilon \in (0, \frac{q-p}{pq})$ and $g = f\rho^{1/p}$, the following quantity is bounded:*

$$\|g\|_{\mathrm{decay}} := \sup_{\substack{x \in \mathbb{R} \\ \tau \in \{0,\dots,\alpha-1\}}} \left| \rho^{1/q-1/p+\varepsilon}(x) g^{(\tau)}(x) \right| \le C_{p,q,\alpha,\varepsilon} \|f\|_{W_\rho^{\alpha,q}} < \infty.$$

Proof. First we note that $1/q - 1/p + \varepsilon < 0$. In the following, we use the Sobolev inequality for $F_\tau(x) := \rho^{1/q-1/p+\varepsilon}(x) g^{(\tau)}(x)$. That is, from the boundedness of $\|F_\tau\|_{L^q(\mathbb{R})}$ and $\|F_\tau'\|_{L^q(\mathbb{R})}$ we deduce the boundedness of $\|F_\tau\|_{L^\infty(\mathbb{R})}$. Let $H_\ell(x)$ be the ℓ-th degree probabilist's Hermite polynomial,

$$H_\ell(x) = \frac{(-1)^\ell}{\sqrt{\ell!}} e^{x^2/2} \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell} e^{-x^2/2}.$$

First notice from the definition of Hermite polynomials that

$$\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell} \rho(x)^{1/p} = p^{-\ell/2} (-1)^\ell \sqrt{\ell!} H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho(x)^{1/p}.$$

Hence, for $\tau = 0, \dots, \alpha - 1$, by applying the product rule to $g^{(\tau)}$ we have

$$
\begin{aligned}
&\|F_\tau(x)\|_{L^q(\mathbb{R})} \\
&\le \sum_{\ell=0}^{\tau} \binom{\tau}{\ell} \left\| p^{-\ell/2}(-1)^\ell \sqrt{\ell!} H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho(x)^{1/p} \rho^{1/q-1/p+\varepsilon}(x) f^{(\tau-\ell)}(x) \right\|_{L^q(\mathbb{R})} \\
&\le \sum_{\ell=0}^{\tau} \binom{\tau}{\ell} \sqrt{\ell!}\, p^{-\ell/2} \left\| H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho(x)^{1/p} \rho^{1/q-1/p+\varepsilon}(x) f^{(\tau-\ell)}(x) \right\|_{L^q(\mathbb{R})} \\
&\le \sum_{\ell=0}^{\tau} \binom{\tau}{\ell} \sqrt{\ell!}\, p^{-\ell/2} \sup_{x \in \mathbb{R}} \left| H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho^\varepsilon(x) \right| \left(\int_{\mathbb{R}} |f^{(\tau-\ell)}(x)|^q \rho(x)\,\mathrm{d}x \right)^{1/q} \\
&= \sum_{\ell=0}^{\tau} \binom{\tau}{\ell} \sqrt{\ell!}\, p^{-\ell/2} \sup_{x \in \mathbb{R}} \left| H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho^\varepsilon(x) \right| \|f^{(\tau-\ell)}(x)\|_{L_\rho^q} < \infty,
\end{aligned}
$$

where in the last line, $\|f^{(\tau-\ell)}(x)\|_{L_\rho^q}$ is bounded because $f \in W_\rho^{\alpha,q}$, and the supremum term is also bounded because H_ℓ has at most only polynomial growth. At the same time, noting that $1/q - 1/p + \varepsilon < 0$, we have

$$
\begin{aligned}
&\|F_\tau'\|_{L^q(\mathbb{R})} \\
&\le \left\| \rho^{1/q-1/p+\varepsilon}(x) g^{(\tau+1)} \right\|_{L^q(\mathbb{R})} + \left\| (1/p - 1/q - \varepsilon) x \rho^{1/q-1/p+\varepsilon}(x) g^{(\tau)} \right\|_{L^q(\mathbb{R})} \\
&\le \sum_{\ell=0}^{\tau+1} \binom{\tau+1}{\ell} \sqrt{\ell!}\, p^{-\ell/2} \sup_{x \in \mathbb{R}} \left| H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho^\varepsilon(x) \right| \|f^{(\tau-\ell+1)}(x)\|_{L_\rho^q} \\
&\quad + (1/p - 1/q - \varepsilon) \sum_{\ell=0}^{\tau} \binom{\tau}{\ell} \sqrt{\ell!}\, p^{-\ell/2} \sup_{x \in \mathbb{R}} \left| x H_\ell\left(\frac{x}{\sqrt{p}}\right) \rho^\varepsilon(x) \right| \|f^{(\tau-\ell)}(x)\|_{L_\rho^q} < \infty,
\end{aligned}
$$

where in the last line, again we used $\|f^{(\tau-\ell+1)}(x)\|_{L_\rho^q} \le \|f\|_{W_\rho^{\alpha,q}} < \infty$. Thus we have $\|g\|_{\mathrm{decay}} < \infty$. $\qquad\square$

For the second term in (2.2), we introduce a suitable auxiliary function G and consider the bound

$$
\left(\int_{-T}^{T} |f(x) - [A_n^\dagger(f)](x)|^p \rho(x)\, \mathrm{d}x \right)^{1/p}
$$

$$
= \left(\int_{-T}^{T} |g(x) - [\mathcal{I}_n(g)](x)|^p\, \mathrm{d}x \right)^{1/p}
$$

$$
= \|g - \mathcal{I}_n(g)\|_{L^p([-T,T])}
$$

$$
\leq \|g - G\|_{L^p([-T,T])} + \|G - \mathcal{I}_n(G)\|_{L^p([-T,T])} + \|\mathcal{I}_n(G) - \mathcal{I}_n(g)\|_{L^p([-T,T])}. \tag{2.3}
$$

Lemma 2.7 will upper bound the three terms in (2.3). For that, we need the following Lemmas 2.3 to 2.6.

Lemma 2.3 (Auxiliary periodic function and its properties)**.** *Let $1 \leq p < q < \infty$, $\alpha \in \mathbb{N}$ and $T > 0$. For $f \in W_\rho^{\alpha,q}$, let $g = f\rho^{1/p}$ and define $G : [-T - d, T + d] \to \mathbb{R}$ by*

$$
G(x) := g(x) - \sum_{\tau=1}^{\alpha} \frac{B_\tau^{[-T,T]}(x)}{\tau!} \left(\int_{-T}^{T} g^{(\tau)}(y)\, \mathrm{d}y \right), \tag{2.4}
$$

with an arbitrarily small $d > 0$, where $B_\tau^{[-T,T]}$ denotes the scaled Bernoulli polynomial of degree τ on $[-T,T]$, i.e.,

$$
B_\tau^{[-T,T]}(x) = (2T)^{\tau-1} B_\tau\left(\frac{x+T}{2T} \right)
$$

with B_τ being the standard Bernoulli polynomial of degree τ. Then, the auxiliary function G is $(\alpha - 1)$-times continuously differentiable with $G^{(\alpha-1)}$ being absolutely continuous on $[-T,T]$, and satisfies $G^{(\tau)}(-T) = G^{(\tau)}(T)$ for all $\tau = 0, \ldots, \alpha - 1$. Furthermore, the auxiliary function G has its α-th weak derivative $G^{(\alpha)}$ in $L^q([-T,T])$ which can be identified as a function over a $2T$-periodic torus $\mathbb{T}([-T,T))$ with

$$
\fint_{\mathbb{T}([-T,T))} |G^{(\alpha)}(x)|^q\, \mathrm{d}x < \infty.
$$

Proof. The proof follows the strategy of [11, Lemma 4.1], where the case with $p = 1$ and $q = 2$ was proved. Recall that for $\tau, \tau' \in \mathbb{N}_0$ scaled Bernoulli polynomials satisfy

$$
\frac{\mathrm{d}^{\tau'}}{\mathrm{d}x^{\tau'}} \frac{B_\tau^{[-T,T]}(x)}{\tau!} = \begin{cases} 0 & \text{if } \tau' > \tau, \\ \frac{B_{\tau-\tau'}^{[-T,T]}(x)}{(\tau-\tau')!} & \text{if } \tau' \leq \tau \end{cases}
$$

and

$$
\int_{-T}^{T} B_\tau^{[-T,T]}(x)\, \mathrm{d}x = \begin{cases} 1 & \text{if } \tau = 0, \\ 0 & \text{if } \tau \neq 0. \end{cases}
$$

From these properties we have

$$
\int_{-T}^{T} G^{(\tau)}(x)\, \mathrm{d}x = \int_{-T}^{T} g^{(\tau)}(x)\, \mathrm{d}x - \left(\int_{-T}^{T} B_0^{[-T,T]}(x)\, \mathrm{d}x \right) \left(\int_{-T}^{T} g^{(\tau)}(s)\, \mathrm{d}s \right) = 0,
$$

for $\tau = 1, \ldots, \alpha$. Since $f \in W_\rho^{\alpha,q}$, $g^{(\tau)}$ and $G^{(\tau)}$ are absolutely continuous on $(-T - d, T + d)$ for $\tau = 0, 1, \ldots, \alpha - 1$. By using the fundamental theorem of calculus, we now obtain

$$
G^{(\tau)}(-T) = G^{(\tau)}(T) \quad \text{for} \quad \tau = 0, \ldots, \alpha - 1. \tag{2.5}
$$

6

Next, we note that $G^{(\alpha)}$ and $g^{(\alpha)}$ only differ by the constant $\frac{1}{2T}\int_{-T}^{T}g^{(\alpha)}(y)\,\mathrm{d}y$ on $[-T,T]$. Since $f^{(\alpha)}\in L^q_\rho$, we know $g^{(\alpha)}$ is in $L^q([-T,T])$. Due to matching boundary values $G^{(\alpha-1)}(-T)=G^{(\alpha-1)}(T)$ and absolute continuity of $G^{(\alpha-1)}$ on $(-T-\delta,T+\delta)$, we have $\|G^{(\alpha)}\|_{L^p(\mathbb{T}([-T,T]))}<\infty$. $\square$ $\square$

Lemma 2.4 (Norm estimate on bounded intervals)**.** *Let $1\le p<q<\infty$, $\alpha\in\mathbb{N}$, $a<b$ and $\varepsilon\in[0,q-p)$. Then*

$$\|g\|_{W^{\alpha,q}([a,b])}\le C_{p,q,\alpha,\varepsilon}\,\rho^{\varepsilon/p}(c)\|f\|_{W^{\alpha,q}_\rho},$$

where $c=\min_{x\in[a,b]}|x|$.

Proof. We have

$$\|g\|_{W^{\alpha,q}([a,b])}=\left(\sum_{\tau=0}^{\alpha}\int_{a}^{b}|g^{(\tau)}(x)|^q\,\mathrm{d}x\right)^{1/q}=\left(\sum_{\tau=0}^{\alpha}\int_{a}^{b}|(f\rho^{1/p}))^{(\tau)}(x)|^q\,\mathrm{d}x\right)^{1/q}$$

$$\le\sum_{\tau=0}^{\alpha}\left(\int_{a}^{b}\left|\sum_{\ell=0}^{\tau}\binom{\tau}{\ell}(-1)^\ell\sqrt{\ell!}f^{(\tau-\ell)}(x)p^{-\ell/2}H_\ell\left(\frac{x}{\sqrt{p}}\right)\rho^{1/p}(x)\right|^q\,\mathrm{d}x\right)^{1/q}$$

$$\le\sum_{\tau=0}^{\alpha}\sum_{\ell=0}^{\tau}\binom{\tau}{\ell}\left(\int_{a}^{b}\left|f^{(\tau-\ell)}(x)\sqrt{\ell!}p^{-\ell/2}H_\ell\left(\frac{x}{\sqrt{p}}\right)\rho^{1/p}(x)\right|^q\,\mathrm{d}x\right)^{1/q}$$

$$\le\sum_{\tau=0}^{\alpha}\sum_{\ell=0}^{\tau}\binom{\tau}{\ell}\left(\sup_{x\in[a,b]}\left|\sqrt{\ell!}p^{-\ell/2}H_\ell\left(\frac{x}{\sqrt{p}}\right)\rho^{(q-p)/p}(x)\right|^q\int_{a}^{b}|f^{(\tau-\ell)}(x)|^q\rho(x)\,\mathrm{d}x\right)^{1/q}$$

$$\le\sum_{\tau=0}^{\alpha}\sum_{\ell=0}^{\tau}\binom{\tau}{\ell}\left(\sup_{x\in[a,b]}\left|\sqrt{\ell!}p^{-\ell/2}H_\ell\left(\frac{x}{\sqrt{p}}\right)\rho^{(q-p-\varepsilon)/p}(x)\right|^q\right.$$

$$\left.\times\rho^{\varepsilon q/p}(c)\int_{a}^{b}|f^{(\tau-\ell)}(x)|^q\rho(x)\,\mathrm{d}x\right)^{1/q}$$

$$\le C_{p,q,\alpha,\varepsilon}\,\rho^{\varepsilon/p}(c)\|f\|_{W^{\alpha,q}_\rho},$$

where we used the facts that $H_\ell(\cdot/\sqrt{p})\rho^{(q-p-\varepsilon)/p}$ is bounded on the real line and that $\rho(x)$ is decreasing in $|x|$. $\square$ $\square$

For obtaining trigonometric interpolation error on $[-T,T]$, we make use of results from trigonometric interpolation on $[0,2\pi]$. For that sake, we need the following estimate for the effect of linear scaling $S_T:[0,2\pi]\to[-T,T]$ given by $S_T(x):=Tx/\pi-T$.

Lemma 2.5 (Boundedness of $\|g\circ S_T\|_{W^{\alpha,q}([0,2\pi])}$)**.** *Let $1\le p<q<\infty$, $\alpha\in\mathbb{N}$, and $f\in W^{\alpha,q}_\rho$. For $g=f\rho^{1/p}$ we define $h:=g\circ S_T$. Then, for $T\ge1$ we have*

$$\|h\|_{W^{\alpha,q}([0,2\pi])}\le C_{p,q,\alpha}T^{\alpha-1/q}\|f\|_{W^{\alpha,q}_\rho}.$$

Proof. We have

$$\|h\|_{W^{\alpha,q}([0,2\pi])} = \left(\sum_{\tau=0}^{\alpha}\int_0^{2\pi}|h^{(\tau)}(x)|^q\,\mathrm{d}x\right)^{1/q}$$

$$= \left(\sum_{\tau=0}^{\alpha}\int_0^{2\pi}|(g\circ S_T)^{(\tau)}(x)|^q\,\mathrm{d}x\right)^{1/q}$$

$$= \left(\sum_{\tau=0}^{\alpha}\int_0^{2\pi}|(T/\pi)^\tau(g^{(\tau)}\circ S_T)(x)|^q\,\mathrm{d}x\right)^{1/q}$$

$$= \left(\sum_{\tau=0}^{\alpha}(T/\pi)^{\tau q-1}\int_{-T}^T|g^{(\tau)}(x)|^q\,\mathrm{d}x\right)^{1/q}$$

$$\le C_{q,\alpha}T^{\alpha-1/q}\|g\|_{W^{\alpha,q}([-T,T])}.$$

The claim then follows from Lemma 2.4 with $a = -T$, $b = T$, and $\varepsilon = 0$. $\qquad\square\qquad\square$

Lemma 2.6 (L^p interpolation error for periodic functions)**.** *Let $1 \le p < q < \infty$, $\alpha \in \mathbb{N}$, and $f \in W_\rho^{\alpha,q}$. Define $g = f\rho^{1/p}$ and an auxiliary periodic function G as in (2.4). Then*

$$\|G - \mathcal{I}_n(G)\|_{L^p([-T,T])} \le C_{p,q,\alpha}T^{\alpha-1/q+1/p}\|f\|_{W_\rho^{\alpha,q}}\frac{1}{n^\alpha}.$$

Proof. The result by Temlyakov [32, Theorem 2.4.4], where the interpolation operator is denoted by I_m, applies to our setting. Therein, the function space considered is defined via convolution kernels, but the norm equivalence to periodic Sobolev spaces

$$W^{\alpha,q}(\mathbb{T}[0,2\pi]) := \left\{f \mid f \in W^{\alpha,q}([0,2\pi]), f^{(\tau)}(0) = f^{(\tau)}(2\pi) \text{ for } \tau = 0,\dots,\alpha-1\right\}$$

is shown in [34, Theorem 2.7]. Hence, by letting $S_T(x) := Tx/\pi - T$, [32, Theorem 2.4.4] leads to

$$\|G\circ S_T - \mathcal{I}_n(G\circ S_T)\|_{W^{\alpha,q}(\mathbb{T}(0,2\pi))} \le C'_{p,q,\alpha}\|G\circ S_T\|_{W^{\alpha,q}(\mathbb{T}(0,2\pi))}\frac{1}{n^\alpha}.$$

Further, by Lemma 2.5, the above error is bounded by

$$\le C''_{p,q,\alpha}T^{\alpha-1/q}\|f\|_{W_\rho^{\alpha,q}}\frac{1}{n^\alpha}.$$

At the same time, since S_T is a plain linear scaling, we have

$$\|G\circ S_T - \mathcal{I}_n(G\circ S_T)\|_{L^p(\mathbb{T}(0,2\pi))} = \left(\frac{\pi}{T}\right)^{1/p}\|G - \mathcal{I}_n(G)\|_{L^p(\mathbb{T}(-T,T))}.$$

Thus we have proved the claim. $\qquad\square\qquad\square$

Lemma 2.7 (Error bound on $[-T,T]$ for $A_n^\dagger$)**.** *Let $1 \le p < q < \infty$, $\alpha \in \mathbb{N}$, and $f \in W_\rho^{\alpha,q}$. Choose $\varepsilon \in (0, \frac{q-p}{pq})$ arbitrarily. Then the approximation error on $[-T,T]$ by the algorithm $A_n^\dagger$ is bounded by*

$$\|f - A_n^\dagger(f)\|_{L_\rho^p([-T,T])} = \left(\int_{-T}^T\left|f(x)\rho^{1/p}(x) - \mathcal{I}_n(f\rho^{1/p})\right|^p\,\mathrm{d}x\right)^{1/p}$$

$$\le C_1\|f\|_{W_\rho^{\alpha,q}}T^{\alpha-1/q+1/p}n^{-\alpha}$$

$$+ C_2\|f\|_{W_\rho^{\alpha,q}}(\ln n)\max\{1,(2T)^{\alpha-1+1/p}\}\mathrm{e}^{-\frac{(q-p)/pq-\varepsilon}{2}T^2},$$

where the constants C_1 and C_2 do not depend on n or T.

Proof. From Lemma 2.2, we know that for $g := f\rho^{1/p}$ the quantity

$$\|g\|_{\text{decay}} = \sup_{\substack{x \in \mathbb{R} \\ \tau \in \{0,\dots,\alpha-1\}}} \left| \rho^{1/q-1/p+\varepsilon}(x) g^{(\tau)}(x) \right|$$

is bounded. As mentioned above, we consider the following bound:

$$
\begin{aligned}
\|f - A_n^\dagger(f)\|_{L_\rho^p([-T,T])} &= \|g - \mathcal{I}_n g\|_{L^p([-T,T])} \\
&\leq \|g - G\|_{L^p([-T,T])} + \|G - \mathcal{I}_n(G)\|_{L^p([-T,T])} + \|\mathcal{I}_n(G) - \mathcal{I}_n(g)\|_{L^p([-T,T])}.
\end{aligned} \tag{2.6}
$$

For the first term in the last line of (2.6), we have

$$
\begin{aligned}
\|g - G\|_{L^p([-T,T])} &\leq \sum_{\tau=1}^{\alpha} \frac{\|B_\tau^{[-T,T]}(x)\|_{L^p([-T,T])}}{\tau!} \left| \int_{-T}^{T} g^{(\tau)}(s)\,\mathrm{d}s \right| \\
&\leq \sum_{\tau=1}^{\alpha} \frac{\|B_\tau^{[-T,T]}(x)\|_{L^\infty([-T,T])} (2T)^{1/p}}{\tau!} \left| \int_{-T}^{T} g^{(\tau)}(s)\,\mathrm{d}s \right| \\
&\leq \sum_{\tau=1}^{\alpha} (2T)^{\tau-1+1/p} \left| \int_{-T}^{T} g^{(\tau)}(s)\,\mathrm{d}s \right| \\
&\leq \alpha \max\{1, (2T)^{\alpha-1+1/p}\} \|g\|_{\text{decay}}\, C_{p,q,\varepsilon}\, \mathrm{e}^{-\frac{(q-p)/pq-\varepsilon}{2}T^2},
\end{aligned} \tag{2.7}
$$

where in the penultimate line, we used $\left| \frac{B_\tau^{[-T,T]}(x)}{\tau!} \right| \leq \frac{(2T)^{\tau-1}}{2}$ for $x \in [-T,T]$; see [25, Equation (6)] or [16].

We proceed to the third term. We have

$$
\begin{aligned}
&\|\mathcal{I}_n(g - G)\|_{L^p([-T,T])} \\
&\leq \|\mathcal{I}_n\|_{C([-T,T]) \to L^p([-T,T])} \|g - G\|_{L^p([-T,T])} \\
&\leq \|\mathcal{I}_n\|_{C([-T,T]) \to L^p([-T,T])} \alpha \max\{1, (2T)^{\alpha-1+1/p}\}\, \|g\|_{\text{decay}}\, C_{p,q,\varepsilon}\, \mathrm{e}^{-\frac{(q-p)/pq-\varepsilon}{2}T^2}.
\end{aligned} \tag{2.8}
$$

Now the operator norm of the interpolation, from continuous functions to L^p, can be bounded using the Lebesgue constant Λ_n for trigonometric interpolation by

$$\|\mathcal{I}_n\|_{C([-T,T]) \to L^p([-T,T])} \leq 2T\Lambda_n \leq 2T \left(\frac{\pi+4}{\pi} + \frac{2\ln n}{\pi} \right),$$

where this explicit constant here can be found, e.g., in [28, Theorem 2.1], or we refer to [6] as an earlier result.

For the second term in (2.6), we know from Lemma 2.6 that

$$\|G - \mathcal{I}_n(G)\|_{L^p([-T,T])} \leq C_{p,q,\alpha} T^{\alpha-1/q+1/p} \|f\|_{W_\rho^{\alpha,q}} \frac{1}{n^\alpha}. \tag{2.9}$$

Summing up all three terms, (2.7),(2.8), and (2.9), we obtain the error bound

$$
\begin{aligned}
&\|f - A_n^\dagger(f)\|_{L_\rho^p([-T,T])} \\
&\leq C_1 T^{\alpha-1/q+1/p} n^{-\alpha} + C_2 (\ln n) \max\{1, (2T)^{\alpha-1+1/p}\} \mathrm{e}^{-\frac{(q-p)/pq-\varepsilon}{2}T^2}.
\end{aligned}
$$

Hence we proved the claim. $\qquad\qquad \square \qquad\qquad\qquad\qquad\qquad\qquad \square$

9

Lemma 2.8 (Error bound on tails)**.** *Let $1 \le p < q < \infty$, $\alpha \in \mathbb{N}$, and $f \in W_\rho^{\alpha,q}$. Set $g = f\rho^{1/p}$. Choose $\varepsilon \in (0, \frac{q-p}{pq})$ arbitrarily. For the error outside the interval $[-T,T]$, we have*

$$\left(\int_{\mathbb{R} \setminus [-T,T]} |f(x)|^p \, \rho(x) \, \mathrm{d}x \right)^{1/p} \le C_{p,q,\varepsilon} \, \|g\|_{\text{decay}} \, T^{-1/p} \, \mathrm{e}^{-\frac{(q-p)/pq - \varepsilon}{2} T^2}.$$

Proof. Let $r := 1/p - 1/q - \varepsilon > 0$. Using Lemma 2.2 we have

$$\left(\int_{\mathbb{R} \setminus [-T,T]} |g(x)|^p \, \mathrm{d}x \right)^{1/p} \le \left(\int_{\mathbb{R} \setminus [-T,T]} \left(\|g\|_{\text{decay}} \frac{\mathrm{e}^{-rx^2/2}}{\sqrt{2\pi}^r} \right)^p \mathrm{d}x \right)^{1/p}$$

$$\le \|g\|_{\text{decay}} \left(\frac{2}{T} \int_T^\infty \frac{x \mathrm{e}^{-rpx^2/2}}{\sqrt{2\pi}^{rp}} \, \mathrm{d}x \right)^{1/p}$$

$$= \|g\|_{\text{decay}} \left(\frac{2}{rpT\sqrt{2\pi}^{rp}} \, \mathrm{e}^{-rpT^2/2} \right)^{1/p} = \frac{2^{1/p}}{(rpT)^{1/p}\sqrt{2\pi}^r} \, \mathrm{e}^{-rT^2/2}.$$

This proves the claim. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$ $\hfill\square$

Theorem 2.9 (Error bound on $\mathbb{R}$ for $A_n^\dagger$)**.** *Let $1 \le p < q < \infty$, $\alpha \in \mathbb{N}$, and $f \in W_\rho^{\alpha,q}$. Choose $\varepsilon \in (0, \frac{q-p}{pq})$ arbitrarily. Set the cut-off interval $[-T,T]$ with*

$$T = \sqrt{2\alpha \left(\frac{q-p}{pq} - \varepsilon \right)^{-1} \ln n}. \tag{2.10}$$

Then, for any integer $n \ge 2$, the approximation error on $\mathbb{R}$ by the algorithm $A_n^\dagger$ is bounded by

$$\|f - A_n^\dagger(f)\|_{L_\rho^p} = \left(\int_{\mathbb{R}} \left| g(x)\rho^{1/p}(x) - \mathcal{I}_n(g\rho^{1/p}) \right|^p \mathrm{d}x \right)^{1/p}$$

$$\le C_{p,q,\alpha,\varepsilon} \|f\|_{W_\rho^{\alpha,q}} \frac{(\ln n)^{\frac{\alpha}{2} + \frac{1}{2} + \frac{1}{2p}}}{n^\alpha}.$$

Proof. We consider the error bound (2.2). With our choice of the cutoff interval, Lemma 2.7 tells us that

$$\left(\int_{-T}^T \left| f(x) - [A_n^\dagger(f)](x) \right|^p \rho(x) \, \mathrm{d}x \right)^{1/p}$$

$$\le C_1 (\ln n)^{\frac{\alpha}{2} - \frac{1}{2q} + \frac{1}{2p}} n^{-\alpha} + C_2 (\ln n)^{\frac{\alpha}{2} + \frac{1}{2} + \frac{1}{2p}} n^{-\alpha},$$

here we used the fact that T is always bigger than 1 for $n \ge 2$, since $0 < \frac{q-p}{pq} - \varepsilon < 1$. For the error outside the interval, using Lemma 2.8, we obtain

$$\left(\int_{\mathbb{R} \setminus [-T,T]} |f(x)|^p \, \rho(x) \, \mathrm{d}x \right)^{1/p}$$

$$\le C_{p,q,\varepsilon} \, \|g\|_{\text{decay}} \, \mathrm{e}^{-\frac{(q-p)/pq - \varepsilon}{2} T^2} = C_{p,q,\varepsilon} \, \|g\|_{\text{decay}} \, n^{-\alpha}.$$

Therefore, the total error is bounded by

$$\|f - A_n^\dagger(f)\|_{L_\rho^p} \le C_{p,q,\alpha,\varepsilon} \|f\|_{W_\rho^{\alpha,q}} (\ln n)^{\frac{\alpha}{2} + \frac{1}{2} + \frac{1}{2p}} n^{-\alpha}. \qquad\square$$

Thus we have proved the claim. $\hfill\square$

Remark 2.10 (Periodic auxiliary function G). This proof strategy of introducing the auxiliary function G can be applied to other algorithms or problems. Let us elaborate this part here. Consider an abstract problem of approximating a linear operator or a linear functional $S : \mathcal{H}_1 \to \mathcal{H}_2$ by a linear algorithm $A_n : \mathcal{H}_1 \to \mathcal{H}_2$. This is a typical setting in *information-based complexity*, see, e.g., [24, Section 4.4]. Then, the error has the following upper bound for any $g \in \mathcal{H}_1$:

$$\|S(g) - A_n(g)\|_{\mathcal{H}_2} \leq \|S(g) - S(G)\|_{\mathcal{H}_2} + \|S(G) - A_n(G)\|_{\mathcal{H}_2} + \|A_n(G) - A_n(g)\|_{\mathcal{H}_2}$$
$$= \|S(g - G)\|_{\mathcal{H}_2} + \|S(G) - A_n(G)\|_{\mathcal{H}_2} + \|A_n(G - g)\|_{\mathcal{H}_2}. \qquad (2.11)$$

This strategy is first used in [25] for a multivariate integration problem in which $S(g) = \int_{\mathbb{R}^d} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ and A_n is a specific quadrature rule, called a scaled rank-1 lattice rule. Therein, the first error term in (2.11), $\|S(g - G)\|_{\mathcal{H}_2}$, happens to vanish due to properties of the integration functional. However, as we have seen, this term can be bounded just like the third term $\|A_n(G - g)\|_{\mathcal{H}_2}$ for general linear algorithms.

In Theorem 2.9, the choice of the cut-off parameter T depends on the smoothness α of the approximand. However, this information may not be available in practice. Following [11, Corollary 4.4] and [8, Remark 3.11], one can replace α by a slowly increasing function $\gamma(n)$ such as $\max(\ln(\ln n), 0)$, and still achieve the convergence rate $n^{-\alpha}$, up to a factor $(\gamma(n) \ln n)^{\frac{\alpha}{2} + \frac{1}{2} + \frac{1}{2p}}$.

Corollary 2.11 (α-free construction). *Let assumptions of Theorem 2.9 be satisfied. Choose $\varepsilon \in (0, \frac{q-p}{pq})$ arbitrarily. Choose a non-decreasing function $\gamma(n) : \mathbb{N} \to \mathbb{R}$ satisfying $\lim_{n \to \infty} \gamma(n) = \infty$, and set the cut-off interval $[-T, T]$ with*

$$T = \sqrt{2\gamma(n) \left(\frac{q - p}{pq} - \varepsilon \right)^{-1} \ln n}.$$

Then, for any integer $n \geq \gamma^{-1}(\alpha) := \min\{m \in \mathbb{N} \mid \gamma(m) \geq \alpha\}$, the approximation error on $\mathbb{R}$ by the algorithm $A_n^{\dagger}$ is bounded by

$$\|f - A_n^{\dagger}(f)\|_{L_{\rho}^p} = \left(\int_{\mathbb{R}} \left| g(x) \rho^{1/p}(x) - \mathcal{I}_n(g\rho^{1/p}) \right|^p \mathrm{d}x \right)^{1/p}$$
$$\leq C_{p,q,\alpha,\varepsilon} \|f\|_{W_{\rho}^{\alpha,q}} \frac{(\gamma(n) \ln n)^{\frac{\alpha}{2} + \frac{1}{2} + \frac{1}{2p}}}{n^{\alpha}}.$$

Remark 2.12 (Computational cost). As mentioned in Introduction, the computational cost and memory usage of an FFT implementation of algorithm $A_n^{\dagger}$ are only $\mathcal{O}(n \log_2 n)$ and $\mathcal{O}(n)$, respectively. The advantage of the method is mainly this cheap cost of construction. However, when the evaluation of the approximand function f is much more expensive, one may wish to have the optimal convergence rate including the logarithmic factor, at the sacrifice of cheap construction. In such a case, we propose the optimal algorithm described in Section 3.

Remark 2.13 (Higher dimensions). For multidimensional settings, where the tensor product of one-dimensional Gaussian Sobolev spaces or the dominating-mixed smoothness is considered, it is possible to extend our algorithm in such a way that the construction cost is still $\mathcal{O}(n \log_2 n)$. One possibility is using rank-1 lattices as interpolation nodes, e.g. [18, 20, 31, 30]. However, the convergence rate of approximation algorithms based on rank-1 lattice points deteriorates at least to $n^{-\alpha/2}$ for periodic functions [2], and thus unfortunately it is not possible to obtain the same rate $n^{-\alpha}$. Another possibility is the fast Smolyak construction [9] combined with our algorithm $A_n^{\dagger}$, which may attain the optimal rate up to a logarithmic factor.

11

3 Spline algorithms

In this section we consider a more refined decomposition of the real line. By partitioning the line into intervals of unit length and having the number of points placed within each interval decrease exponentially fast as one moves away from the origin, we construct an algorithm that attains the optimal rate $n^{-\alpha}$ in $W_\rho^{\alpha,q}$ when $q = 2$. Naive version of this construction based on solving a linear system [see (3.7)] requires $\mathcal{O}(n^3)$ computational cost, and is similar to (and inspired by) those used in [3] and [10, Section 3]. Faster computation is possible in some special cases that we do not discuss here.

3.1 Generic results

Let $W^{\alpha,q}(\mathbb{I})$ stand for the standard unweighted Sobolev space on an interval $\mathbb{I}$. That is, each element of this space is α times weakly differentiable on $\mathbb{I}$ and the weak derivatives are in $L^q(\mathbb{I})$. For each $\nu \in \mathbb{N}$, let $A_\nu^{\mathbb{I},1} \colon W^{\alpha,q}(\mathbb{I}_1) \to L^p(\mathbb{I}_1)$ be an approximation to $f \in W^{\alpha,q}(\mathbb{I}_1)$ based on ν function values on the interval $\mathbb{I}_1 = [0,1]$. We may interpret $A_\nu^{\mathbb{I},1}(f)$ as an element of $L^p(\mathbb{R})$ that vanishes on $\mathbb{R} \setminus \mathbb{I}_1$. We assume that there are positive constants C_0 and β independent of f or ν such that

$$\|f - A_\nu^{\mathbb{I},1}(f)\|_{L^p(\mathbb{I}_1)} \leq C_0 \nu^{-\beta} \|f\|_{W^{\alpha,q}(\mathbb{I}_1)} \tag{3.1}$$

for all $f \in W^{\alpha,q}(\mathbb{I}_1)$ and $\nu \in \mathbb{N}$. Let $\mathbb{I}_k = [k-1, k]$ for $k \in \mathbb{Z}$. By straightforward translation we obtain algorithms $A_\nu^{\mathbb{I},k} \colon W^{\alpha,q}(\mathbb{I}_k) \to L^p(\mathbb{I}_k)$ such that

$$\|f - A_\nu^{\mathbb{I},k}(f)\|_{L^p(\mathbb{I}_k)} \leq C_0 \nu^{-\beta} \|f\|_{W^{\alpha,q}(\mathbb{I}_k)} \tag{3.2}$$

for all $f \in W^{\alpha,q}(\mathbb{I}_k)$ and $\nu \in \mathbb{N}$. For $m \in \mathbb{N}$ and $\nu_1, \ldots, \nu_m \in \mathbb{N}$ we set $n = 2(\nu_1 + \cdots + \nu_m)$. Moreover, set $\nu_{-k} = \nu_k$. Define the algorithm $A_n^* \colon W_\rho^{\alpha,q} \to L_\rho^p$ that uses n function values to approximate $f \in W_\rho^{\alpha,q}$ via

$$A_n^*(f) = \rho^{-1/p} \sum_{0 < |k| \leq m} A_{\nu_k}^{\mathbb{I},k}(g_k), \quad \text{where } g_k = f|_{\mathbb{I}_k} \rho^{1/p}. \tag{3.3}$$

Note that $A_n^*(f)$ vanishes on $\mathbb{R} \setminus [-m, m]$. We shall show that A_n^* inherits the convergence rate (3.1) of $A_\nu^{\mathbb{I},1}$ if m and $\nu_1, \ldots, \nu_m$ are selected properly.

Proposition 3.1 (Generic error bound). *Let $1 \leq p < q < \infty$, $\alpha \in \mathbb{N}$, and $f \in W_\rho^{\alpha,q}$. Then there are positive constants C and δ, which do not depend on m or $\nu_1, \ldots, \nu_m$, such that*

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C \left(\sum_{k=1}^m \mathrm{e}^{-\delta k^2} \nu_k^{-\beta} + \mathrm{e}^{-\delta m^2} \right).$$

Proof. The proof is similar to those in Section 2. The error decomposes as

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq \sum_{0 < |k| \leq m} \|g_k - A_{\nu_k}^{\mathbb{I},k}(g_k)\|_{L^p(\mathbb{I}_k)} + \left(\int_{\mathbb{R} \setminus [-m,m]} |f(x)|^p \rho(x) \, \mathrm{d}x \right)^{1/p}.$$

Let $\varepsilon \in (0, \frac{q-p}{pq})$. Then Lemma 2.8 bounds the tail term as

$$\int_{\mathbb{R} \setminus [-m,m]} |f(x)|^p \rho(x) \, \mathrm{d}x = \int_{\mathbb{R} \setminus [-m,m]} |g(x)|^p \, \mathrm{d}x \leq C_{p,q,\varepsilon} \|g\|_{\mathrm{decay}} \, \mathrm{e}^{-\frac{(q-p)/pq-\varepsilon}{2} m^2},$$

12

where $g = f\rho^{1/p}$ and $\|g\|_{\mathrm{decay}}$ is finite by Lemma 2.2. The error bound (3.2) and Lemma 2.4 yield

$$\|g_k - A_{\nu_k}^{\mathbb{I},k}(g_k)\|_{L^p(\mathbb{I}_k)} \leq C_0 \nu_k^{-\beta} \|g_k\|_{W^{\alpha,q}(\mathbb{I}_k)} \leq C_0 C_{p,q,\alpha,\varepsilon}\, \rho^{\varepsilon/p}(|k|-1)\nu_k^{-\beta}\|f\|_{W_\rho^{\alpha,q}}.$$

The above bound does not depend on the sign of k because we require that $\nu_{-k} = \nu_k$. Hence there are positive constants C and δ such that

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C\left(\sum_{k=1}^m e^{-\delta k^2}\nu_k^{-\beta} + e^{-\delta m^2}\right),$$

which completes the proof. $\qquad\qquad\square\qquad\qquad\qquad\qquad\square$

Theorem 3.2 (Error bound for exponentially decaying points)**.** *Let $1 \leq p < q < \infty$, $\alpha \in \mathbb{N}$, and $m \in \mathbb{N}$. Set $\nu_k = 2^{m-k}$ for $k = 1, \ldots, m$. Then the algorithm A_n^* defined in (3.3) uses $n = 2(2^m - 1)$ function values and satisfies*

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C n^{-\beta}$$

for any $f \in W_\rho^{\alpha,q}$, where the constant C does not depend on n.

Proof. The algorithm uses $n = 2\sum_{k=1}^m \nu_k = 2\sum_{k=1}^m 2^{m-k} = 2(2^m - 1)$ function values. Proposition 3.1 yields

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C\left(\sum_{k=1}^m e^{-\delta k^2}\nu_k^{-\beta} + e^{-\delta m^2}\right) = C\left(2^{-\beta m}\sum_{k=1}^m e^{-\delta k^2}2^{\beta k} + e^{-\delta m^2}\right).$$

Since $\sum_{k=1}^\infty e^{-\delta k^2}2^{\beta k} < \infty$ and $e^{-\delta m^2}$ decays faster than $2^{-\beta m}$ as $m \to \infty$, the claim follows from

$$2^{-\beta m} \leq (2^m - 1)^{-\beta} = (n/2)^{-\beta}.$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

To attain the optimal rate $n^{-\alpha}$ in $W_\rho^{\alpha,q}$ we need to exhibit a linear algorithm $A_\nu^{\mathbb{I},1}$ that satisfies (3.1) for $\beta = \alpha$. It is well known that for any $p, q \in [1, \infty]$ and $\nu \in \mathbb{N}$ one may construct an algorithm $A_\nu^{\mathbb{I},1}$ that uses ν function values and has the error bound

$$\|f - A_\nu^{\mathbb{I},1}(f)\|_{L^p(\mathbb{I}_1)} \leq C_0 \nu^{-\alpha+(1/q-1/p)_+}\|f\|_{W^{\alpha,q}(\mathbb{I}_1)} \qquad (3.4)$$

for all $f \in W^{\alpha,q}(\mathbb{I}_1)$, where $(x)_+ = \max\{0, x\}$. See, for example, Section 1.3.11 in [22] or [14, 23]. Because we assume that $1 \leq p < q < \infty$, Theorem 3.2 and (3.4) yield the following corollary, which gives the known upper bound on approximation error in $W_\rho^{\alpha,q}$ [3, Theorem 3.3].

Corollary 3.3 (Convergence rate for sampling recovery)**.** *Let $1 \leq p < q < \infty$ and $\alpha \in \mathbb{N}$. For each $n \in \mathbb{N}$ there is a linear algorithm of the form (3.3) that uses n function values and satisfies*

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C n^{-\alpha},$$

for any $f \in W_\rho^{\alpha,q}$, where the constant C does not depend on n.

3.2 Spline smoothing

The construction of an algorithm $A_\nu^{\mathbb{I},1}$ satisfying (3.4) is particularly simple when $q = 2$, in which case $W^{\alpha,q}(\mathbb{I}_1)$ is a Hilbert space.

Definition 3.4 (Spline smoother). Let $W(\mathbb{I}_1)$ be any Hilbert space that is norm-equivalent to $W^{\alpha,2}(\mathbb{I}_1)$ and let $0 \leq x_1 < \cdots < x_\nu \leq 1$ be points on $\mathbb{I}_1$. For any $\lambda > 0$, the minimiser $s_{f,\nu,\lambda}$ of

$$\sum_{i=1}^{\nu} [f(x_i) - s(x_i)]^2 + \lambda \|s\|_{W(\mathbb{I}_1)}^2 \tag{3.5}$$

among $s \in W^{\alpha,2}(\mathbb{I}_1)$ is unique. This minimiser is called the *spline smoother* to f at $x_1, \ldots, x_\nu$.

Because $\alpha \geq 1$, the Sobolev embedding theorem ensures that $W^{\alpha,2}(\mathbb{I}_1)$, and thus also $W(\mathbb{I}_1)$, is continuously embedded in the space of continuous functions on $\mathbb{I}_1$. It follows that for each $x \in \mathbb{I}_1$ the point evaluation functional $f \mapsto f(x)$ is continuous on $W(\mathbb{I}_1)$. By the Riesz representation theorem, for each $x \in \mathbb{I}_1$ there is a *representer* $\ell_x \in W(\mathbb{I}_1)$ such that $f(x) = \langle f, \ell_x \rangle_{W(\mathbb{I}_1)}$ for every $f \in W(\mathbb{I}_1)$. Using these representers we may define the *reproducing kernel* of $W(\mathbb{I}_1)$ as

$$K(x,y) = \langle \ell_x, \ell_y \rangle_{W(\mathbb{I}_1)}.$$

This kernel is positive-semidefinite and satisfies $f(x) = \langle f, K(\cdot, x) \rangle_{W(\mathbb{I}_1)}$.

Some reproducing kernels are available in closed form. The full scale of Sobolev spaces (also those of fractional order) is reproduced by the class of *Matérn kernels* popular in machine learning and kriging [26, 29]. Let l be a positive scaling parameter and $\mathcal{K}_\gamma$ the modified Bessel function of the second kind of order γ. The Matérn kernel K_γ of order $\gamma > 0$ given by

$$K_\gamma(x,y) = \frac{2^{1-\gamma}}{\Gamma(\gamma)} \left(\frac{\sqrt{2\gamma}\,|x-y|}{l} \right)^\gamma \mathcal{K}_\gamma \left(\frac{\sqrt{2\gamma}\,|x-y|}{l} \right) =: \Phi_\gamma(|x-y|), \tag{3.6}$$

is a reproducing kernel for a space $W_\gamma(\mathbb{I}_1)$ that is norm-equivalent to $W^{\alpha,2}(\mathbb{I}_1)$ for $\alpha = \gamma + 1/2$. The curious coefficients $\sqrt{2\gamma}$ and $2^{1-\gamma}/\Gamma(\gamma)$ ensure that K_γ tends pointwise to the *Gaussian kernel* $\exp(-(x-y)^2/(2l^2))$ as $\gamma \to \infty$. The norm-equivalence of $W_\gamma(\mathbb{I}_1)$ and $W^{\alpha,2}(\mathbb{I}_1)$ can be verified as follows. It goes back at least to the work of Kimeldorf and Wahba [12] that a positive-semidefinite kernel of the form $K(x,y) = \Phi(|x-y|)$ for $\Phi \colon \mathbb{R} \to \mathbb{R}$ that is continuous and integrable is a reproducing kernel of a Hilbert space whose squared norm is proportional to

$$\int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \, \widehat{\Phi}(\omega)^{-1} \, \mathrm{d}\omega,$$

where $\widehat{f}$ and $\widehat{\Phi}$ are the Fourier transforms of f and Φ (see also Theorem 10.12 in [36]). Suppose that $\alpha = \gamma + 1/2$ is an integer. Because the Fourier transform of the function Φ_γ in (3.6) is proportional to $(c_1 + c_2\omega^2)^{-(\gamma+1/2)}$ for positive c_1 and c_2 [36, Theorem 6.13], the space $W_\gamma(\mathbb{R})$ in which K_γ is reproducing on $\mathbb{R}$ has the squared norm

$$\|f\|_{W_\gamma(\mathbb{R})}^2 \propto \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 (c_1 + c_2\omega^2)^{\gamma+1/2} \, \mathrm{d}\omega = \sum_{\tau=0}^{\alpha} \binom{\alpha}{\tau} c_1^{\alpha-\tau} c_2^\tau \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \omega^{2\tau} \, \mathrm{d}\omega$$

$$= \sum_{\tau=0}^{\alpha} \binom{\alpha}{\tau} c_1^{\alpha-\tau} c_2^\tau \int_{\mathbb{R}} |f^{(\tau)}(x)|^2 \, \mathrm{d}x,$$

where $f^{(\tau)}$ denotes the τ-th order weak derivative and we have used the binomial theorem and Parseval's identity. This norm is equivalent to the norm of the standard Sobolev space $W^{\alpha,2}(\mathbb{R})$

and by taking a restriction of $W_\gamma(\mathbb{R})$ on $\mathbb{I}_1$ we obtain a space $W_\gamma(\mathbb{I}_1)$ that is norm-equivalent to $W^{\alpha,2}(\mathbb{I}_1)$.

When one has access to the reproducing kernel K of $W(\mathbb{I}_n)$, constructing the spline smoother is straightforward, though not necessarily computationally convenient. It is a standard result [35, Section 1.3] that the minimiser of (3.5) takes the form

$$s_{f,\nu,\lambda}(x) = \sum_{i=1}^\nu a_i K(x, x_i). \tag{3.7}$$

The coefficients $\boldsymbol{a} = (a_1, \ldots, a_\nu)$ are the solution to the linear system $(\boldsymbol{K}_\nu + \lambda \boldsymbol{I}_\nu)\boldsymbol{a} = \boldsymbol{f}_\nu$ of ν equations, where $\boldsymbol{K}_\nu = (K(x_i, x_j))_{i,j=1}^\nu$, $\boldsymbol{I}_\nu$ is the $\nu \times \nu$ identity matrix and $\boldsymbol{f}_\nu = (f(x_1), \ldots, f(x_\nu))$. In learning theory this result is known as the *representer theorem* [27]. Note that this linear system has a unique solution when $\lambda > 0$ because the matrix $\boldsymbol{K}_\nu$ is positive-semidefinite. Let $h_\nu = \max_{i=0,\ldots,\nu}|x_{i+1} - x_i|$, where we use the convention $x_0 = 0$ and $x_{\nu+1} = 1$. By Proposition 3.6 in [37] (see [1] for additional results), there is a positive constant C_0 independent of $f \in W^{\alpha,2}(\mathbb{I}_1)$ and the points $x_1, \ldots, x_\nu$ such that

$$\|f - s_{f,\nu,\lambda}\|_{L^p(\mathbb{I}_1)} \leq C_0 h_\nu^\alpha \|f\|_{W^{\alpha,2}(\mathbb{I}_1)} \tag{3.8}$$

if $1 \leq p < 2$ and $\lambda \leq \nu^{-2\alpha}$. This estimate provides an optimal rate of convergence for the following spline smoother based algorithm. Namely, suppose that $0 = x_{\nu,0} < x_{\nu,1} < \cdots < x_{\nu,\nu} \leq x_{\nu,\nu+1} = 1$ for each $\nu \in \mathbb{N}$ and

$$\sup_{\nu \in \mathbb{N}} \max_{i=0,\ldots,\nu} \nu \, |x_{\nu,i+1} - x_{\nu,i}| < \infty. \tag{3.9}$$

For example, the equispaced points

$$x_{\nu,i} = \frac{i}{\nu+1}$$

satisfy (3.9) since $\nu \, |x_{\nu,i+1} - x_{\nu,i}| = \nu/(\nu+1) < 1$. Set $\lambda_\nu = \nu^{-2\alpha}$ and let s_{f,ν,λ_ν} be the spline smoother to f at the points $x_{\nu,1}, \ldots, x_{\nu,\nu}$. When

$$A_\nu^{\mathbb{I},1}(f) = s_{f,\nu,\lambda_\nu}, \tag{3.10}$$

the algorithm A_n^* in (3.3) is a sum of $2m$ spline smoothers and uses n function values.

Corollary 3.5 (Convergence of a spline smoother algorithm)**.** *Let $1 \leq p < 2 \leq q$ and $\alpha \in \mathbb{N}$. Let m and $\nu_1, \ldots, \nu_m$ be as in Theorem 3.2 and $A_n^{\mathbb{I},1}$ as in (3.10). Then the algorithm A_n^* defined in (3.3) uses $n = 2(2^m - 1)$ function values and satisfies*

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq C n^{-\alpha}$$

for any $f \in W_\rho^{\alpha,q}$, where the constant C does not depend on n.

Proof. Because the Gaussian density function ρ defines a finite measure on $\mathbb{R}$, we have $L_\rho^q \subseteq L_\rho^2$ and consequently $W_\rho^{\alpha,q} \subseteq W_\rho^{\alpha,2}$. The claim then follows from (3.8) and the fact that for points that satisfy (3.9) we have $h_\nu \leq C_h n^{-1}$ for a positive constant C_h independent of ν. $\qquad\square$ $\qquad\square$

In fact, the smoothness of the Sobolev space that is used to construct the spline smoothers s_{f,ν,λ_ν} does *not* have to coincide with the smoothness of f. It is a consequence of the escape theorem of Narcowich, Ward and Wendland [21, Theorem 4.2] that the estimate (3.8) remains valid even when the Hilbert space $W(\mathbb{I}_1)$ is norm-equivalent to $W^{\tilde{\alpha},2}(\mathbb{I}_1)$ for $\tilde{\alpha} \geq \alpha$. From this we obtain the following corollary similar in spirit to Corollary 2.11.

Corollary 3.6 (Misspecified smoothness). *Let $\tilde{\alpha} \geq \alpha$ and consider the setting of Corollary 3.5 but suppose that the spline smoothers s_{f,ν,λ_ν} are constructed using a Hilbert space $W(\mathbb{I}_1)$ norm-equivalent to $W^{\tilde{\alpha},2}(\mathbb{I}_n)$. Then*

$$\|f - A_n^*(f)\|_{L_\rho^p} \leq Cn^{-\alpha}$$

for any $f \in W_\rho^{\alpha,q}$, where the constant C does not depend on n.

Remark 3.7 (Fractional smoothness). Because the results in [1, 21, 37] that we have used apply also to fractional Sobolev spaces, Corollaries 3.5 and 3.6 generalise to the fractional setting if the weighted Sobolev space $W_\rho^{\alpha,q}$ is defined appropriately for $\alpha \notin \mathbb{N}$.

Remark 3.8 (Higher dimensions). By replacing the unit intervals $\mathbb{I}_k$ with suitably selected unit cubes as in [3], Corollaries 3.5 and 3.6 could be generalised also to the d-dimensional weighted isotropic Sobolev space $W_\rho^{\alpha,q}(\mathbb{R}^d)$ consisting of all functions whose Gaussian-weighted mixed weak derivatives of total order at most α are q-integrable. The estimate (3.8) remains valid when $\mathbb{I}_1$ is replaced with a d-dimensional unit cube $\mathbb{I}_1^d$ and h_ν with the d-dimensional fill-distance $\max_{\boldsymbol{x} \in \mathbb{I}_1^d} \min_{i=1,\dots,\nu} \|\boldsymbol{x} - \boldsymbol{x}_i\|$, where $\|\cdot\|$ is the Euclidean norm.

4 Optimality

We conclude this paper by discussing optimality, i.e., general lower bounds for algorithms which are based on function values. For $p < q$, the optimal rate is derived by Kuo, Plaskota, and Wasilkowski [15, Theorem 3]. They consider a more general class of weights, which is applicable to ours. Their result states that the error of any algorithm A_n that only uses n function evaluation has the lower bound

$$\sup_{\substack{f \in W_\rho^{\alpha,q} \\ \|f\|_{W_\rho^{\alpha,q}} \leq 1}} \|f - A_n(f)\|_{L_\rho^p} \geq C_{p,q,\alpha} n^{-\alpha},$$

for $1 \leq p < q$. In this sense, our methods achieving error convergence $n^{-\alpha}$ are optimal. In the same paper, they propose a piecewise Lagrange approximation that achieves the optimal rate [15, Theorem 2].

Even though our algorithms do not cover the extreme case $p = q$, this has been recently studied by Dũng [5, Theorem 3.4]. In this case the optimal rate degenerates to $n^{-\alpha/2}$. Dũng [5] also covers extreme cases, such as $q = 1$ or $q = \infty$, deriving the convergence rate $n^{-\alpha/2+(1-1/p)/2}$ for $q = 1$ and $n^{-\alpha/2+1/(2p)}$ for $q = \infty$. These rates can be achieved by his B-spline algorithms [5, Theorem 2.2 and 2.4].

In this paper, we focused on implementable and efficient algorithms achieving the optimal approximation rate. In practice, we often do not know the smoothness α of the target function and we have access to only function values. We explored the case when α is not known exactly in Corollaries 2.11 and 3.6. It is an interesting question of whether optimal algorithms can be constructed for the extreme cases of (p, q) without knowledge of α. We leave this to future study.

Acknowledgement

References

[1] R. Arcangéli, M. C. L. de Silanes, and J. J. Torrens. An extension of a bound for functions in Sobolev spaces, with applications to (m, s)-spline interpolation and smoothing. *Numer. Math.*, 107(2):181–211, 2007.

[2] G. Byrenheid, L. Kämmerer, T. Ullrich, and T. Volkmer. Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness. *Numer. Math.*, 136(4):993–1034, 2017.

[3] Dinh Dũng and Van Kien Nguyen. Optimal numerical integration and approximation of functions on $\mathbb{R}^d$ equipped with Gaussian measure. *IMA J. Numer. Anal.*, 44(2):1242–1267, 2024.

[4] 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(2):684–707, 2018.

[5] Dinh Dũng. Weighted approximate sampling recovery and integration based on B-spline interpolation and quasi-interpolation. *arXiv preprint* `arXiv:2501.01167 [math.NA]`, 2025.

[6] H. Ehlich and K. Zeller. Auswertung der Normen von Interpolationsoperatoren. *Math. Ann.*, 164:105–112, 1966.

[7] M. Gnewuch, A. Hinrichs, K. Ritter, and R. Rüßmann. Infinite-dimensional integration and L^2-approximation on Hermite spaces. *J. Approx. Theory*, 300:Paper No. 106027, 32, 2024.

[8] Takashi Goda, Yoshihito Kazashi, and Yuya Suzuki. Randomizing the trapezoidal rule gives the optimal RMSE rate in Gaussian Sobolev spaces. *Math. Comp.*, 93(348):1655–1676, 2024.

[9] M. Griebel and J. Hamaekers. Fast discrete Fourier transform on generalized sparse grids. In *Sparse grids and applications—Munich 2012*, volume 97 of *Lect. Notes Comput. Sci. Eng.*, pages 75–107. Springer, Cham, 2014.

[10] T. Karvonen, C. J. Oates, and M. Girolami. Integration in reproducing kernel Hilbert spaces of Gaussian kernels. *Math. Comp.*, 90(331):2209–2233, 2021.

[11] 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.

[12] G. S. Kimeldorf and G. Wahba. A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Stat.*, 41(2):495–502, 1970.

[13] C. Klein. Fourth order time-stepping for low dispersion Korteweg-de Vries and nonlinear Schrödinger equations. *Electron. Trans. Numer. Anal.*, 29:116–135, 2007/08.

[14] D. Krieg and M. Sonnleitner. Random points are optimal for the approximation of Sobolev functions. *IMA J. Numer. Anal.*, 2023.

[15] F. Y. Kuo, L. Plaskota, and G. W. Wasilkowski. Optimal algorithms for doubly weighted approximation of univariate functions. *J. Approx. Theory*, 201:30–47, 2016.

[16] D. H. Lehmer. On the maxima and minima of Bernoulli polynomials. *Amer. Math. Monthly*, 47:533–538, 1940.

[17] G. Leobacher, F. Pillichshammer, and A. Ebert. Tractability of L_2-approximation and integration in weighted Hermite spaces of finite smoothness. *J. Complexity*, 78:101768, 2023.

[18] D. Li and F. J. Hickernell. Trigonometric spectral collocation methods on lattices. In *Recent advances in scientific computing and partial differential equations (Hong Kong, 2002)*, volume 330 of *Contemp. Math.*, pages 121–132. Amer. Math. Soc., Providence, RI, 2003.

[19] G. Mastroianni and G. Monegato. Error estimates for Gauss-Laguerre and Gauss-Hermite quadrature formulas. In *Approximation and computation (West Lafayette, IN, 1993)*, volume 119 of *Internat. Ser. Numer. Math.*, pages 421–434. Birkhäuser Boston, Boston, MA, 1994.

[20] H. Munthe-Kaas and T. Sørevik. Multidimensional pseudo-spectral methods on lattice grids. *Appl. Numer. Math.*, 62(3):155–165, 2012.

[21] F. J. Narcowich, J. D. Ward, and H. Wendland. Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions. *Constr. Approx.*, 24(2):175–186, 2006.

[22] E. Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*, volume 1349 of *Lecture Notes in Mathematics*. Springer-Verlag, 1988.

[23] E. Novak and H. Triebel. Function spaces in Lipschitz domains and optimal rates of convergence for sampling. *Constr. Approx.*, 23:325–350, 2006.

[24] E. Novak and H. Woźniakowski. *Tractability of Multivariate Problems. Vol. I: Linear Information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society, 2008.

[25] D. Nuyens and Y. Suzuki. Scaled lattice rules for integration on $\mathbb{R}^d$ achieving higher-order convergence with error analysis in terms of orthogonal projections onto periodic spaces. *Math. Comp.*, 92(339):307–347, 2023.

[26] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, 2006.

[27] B. Schölkopf, R. Herbrich, and A. J. Smola. A generalized representer theorem. In *Computational Learning Theory*, volume 2111 of *Lecture Notes in Computer Science*, pages 416–426. Springer, 2001.

[28] T. Sørevik and M. A. Nome. Trigonometric interpolation on lattice grids. *BIT*, 56(1):341–356, 2016.

[29] M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Series in Statistics. Springer, 1999.

[30] Y. Suzuki and D. Nuyens. Rank-1 lattices and higher-order exponential splitting for the time-dependent Schrödinger equation. In Bruno Tuffin and Pierre L'Ecuyer, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2018 (MCQMC 2018). Springer Proceedings in Mathematics & Statistics, vol 324.*, pages 485–502. Springer, Cham, 2020.

[31] Y. Suzuki, G. Suryanarayana, and D. Nuyens. Strang splitting in combination with rank-1 and rank-r lattices for the time-dependent Schrödinger equation. *SIAM J. Sci. Comput.*, 41(6):B1254–B1283., 2019.

[32] V. Temlyakov. *Multivariate Approximation*, volume 32 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2018.

[33] M. Thalhammer, M. Caliari, and C. Neuhauser. High-order time-splitting Hermite and Fourier spectral methods. *J. Comput. Phys.*, 228(3):822–832, 2009.

[34] T. Ullrich. *Smolyak's Algorithm, Sparse Grid Approximation and Periodic Function Spaces with Dominating Mixed Smoothness*. PhD thesis, Friedrich-Schiller-Universität Jena, 2007.

[35] G. Wahba. *Spline Models for Observational Data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, 1990.

[36] H. Wendland. *Scattered Data Approximation*. Number 17 in Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2005.

[37] H. Wendland and C. Rieger. Approximate interpolation with applications to selecting smoothing parameters. *Numer. Math.*, 101(4):729–748, 2005.