



Institut für Numerische Simulation

Rheinische Friedrich-Wilhelms-Universität Bonn

Friedrich-Hirzebruch-Allee 7 • 53115 Bonn • Germany
phone +49 228 73-69828 • fax +49 228 73-69847
www.ins.uni-bonn.de

F. Bartel, A.D. Gilbert, M. Griebel, F.Y. Kuo, I.H. Sloan

**A higher order lattice-based method for
high-dimensional numerical integration using
periodization**

INS Preprint No. 2603

October 2026

A higher order lattice-based method for high-dimensional numerical integration using periodization

Felix Bartel^a, Alexander D. Gilbert^{b,*}, Michael Griebel^{c,d}, Frances Y. Kuo^b and Ian H. Sloan^b

^aMathematical Institute for Machine Learning and Data Science (MIDS), Catholic University of Eichstätt–Ingolstadt, Ingolstadt 85049, Germany

^bSchool of Mathematics and Statistics, UNSW Sydney, Sydney, NSW, 2052, Australia

^cInstitute for Numerical Simulation, University of Bonn, Bonn 53115, Germany

^dFraunhofer Institute for Algorithms and Scientific Computing (SCAI), Sankt Augustin 53757, Germany

ARTICLE INFO

Keywords:

high dimensional integration

lattice rules

periodization

kernel interpolation

ABSTRACT

We present a novel numerical integration method for non-periodic functions over the high-dimensional unit cube $[0, 1]^s$ by a specially crafted (unequally-)weighted quadrature rule using *transformed lattice* points, with the additional option of *subsampling*. The method is designed for integrands whose mixed derivatives up to order α are square-integrable with respect to a product Chebyshev density. The ingredients are (i) periodizing the integrand, (ii) approximating the smooth periodic function by a kernel interpolant at lattice points, (iii) integrating exactly the product of the kernel interpolant and the nonsmooth density, together with subsampling alternatives for (ii) and (iii). With $|J|$ corresponding to the number of function evaluations, we achieve the convergence rate close to the order $|J|^{-(\alpha-1/2)}$ and $|J|^{-\alpha} (1.110721)^s$, where the implied constants are independent of s under favorable conditions. We can interpolate these two results to trade between the convergence rate and the growth in s . We provide numerical experiments to demonstrate our theory.

1. Introduction

In this paper we present a novel approach to the numerical computation of integrals of non-periodic functions over the s -dimensional unit cube $[0, 1]^s$, with s large, by a *specially crafted weighted quadrature rule using transformed (and optionally subsampled) lattice points*:

$$I(f) := \int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x} \approx \sum_{k \in J} w_k f(\boldsymbol{\psi}(\mathbf{t}_k)) =: Q_{\boldsymbol{\psi}}^{\mathbf{X}_J}(f).$$

The ingredients are firstly a simple coordinate transformation $\mathbf{x} = \boldsymbol{\psi}(\mathbf{y})$ that transforms the integrand to a smooth periodic function $F(\mathbf{y}) := f(\boldsymbol{\psi}(\mathbf{y}))$ multiplied by a non-smooth density; then approximation of the smooth function F by a kernel approximation that uses the points $\mathbf{t}_k$ of a carefully designed lattice; then exact integration of kernel interpolant times density to determine the weights w_k ; with the optional alternative of replacing the full lattice by a randomly chosen subset $\mathbf{X}_J$.

The method is designed for functions f that are somewhat smooth, having mixed derivatives up to order $\alpha \in \mathbb{N}$ that are well behaved at the cube boundaries, in the sense that

$$\int_{[0,1]^s} \left| \frac{\partial^{\mathbf{v}}}{\partial \mathbf{x}^{\mathbf{v}}} f(\mathbf{x}) \right|^2 \prod_{j=1}^s \frac{1}{\pi \sqrt{x_j(1-x_j)}} \, d\mathbf{x} < \infty \quad \text{for all } \mathbf{v} \in \{0, \dots, \alpha\}^s,$$

which is satisfied if for instance all the partial derivatives of f are bounded on the unit cube.

1.1. Background

It is well known that, for functions in the classical Sobolev space $H_{\alpha}([0, 1]^s)$ with smoothness $\alpha \geq 1$, a deterministic product quadrature with N points has an error of the order $\mathcal{O}(N^{-\alpha/s})$, thus one encounters the *curse of dimensionality*.

 felix.bartel@ku.de (F. Bartel); alexander.gilbert@unsw.edu.au (A.D. Gilbert); griebel@ins.uni-bonn.de (M. Griebel); f.kuo@unsw.edu.au (F.Y. Kuo); i.sloan@unsw.edu.au (I.H. Sloan)
ORCID(s):

Monte Carlo quadrature has a root-mean-square error of $\mathcal{O}(N^{-1/2})$, which is independent of s but suffers a slow rate regardless of α .

To obtain improved convergence, one should consider the smaller Sobolev space $H_{\alpha}^{\text{mix}}([0, 1]^s)$ of functions with bounded *mixed* derivatives of order α . Conventional Quasi-Monte Carlo (QMC) methods [38, 45] use well chosen deterministic points to improve the rate close to $\mathcal{O}(N^{-1})$. Lattice rules [13, 12] can achieve close to $\mathcal{O}(N^{-\alpha})$ for periodic integrands, while the same rate can be achieved by higher order digital nets [14] for general non-periodic integrands. The same main rate is also achieved by sparse grid quadrature [4]. However, all these bounds include a logarithmic factor in N that grows exponentially with s , and often also exponentially with α . For example, for sparse grids we encounter $(\log N)^{(\alpha+1)(s-1)}$, and for higher order digital nets we encounter $(\log N)^{\alpha \cdot s}$ or $(\log N)^{\alpha+s}$. The *optimal* exponent here is $(\log N)^{(s-1)/2}$, which is achieved by Frolov quadrature [50], but the curse of dimensionality is still present in the $\log N$ term. Moreover, the implied constants all depend on α and exponentially on s , limiting these methods to moderate dimensions s in practice.

To lift this curse, one may work with a *weighted* Sobolev space $H_{\alpha, \gamma}^{\text{mix}}([0, 1]^s)$ and associated weighted norms, see [48, 13], where the underlying weight sequence (γ introduced below) satisfies a certain summability condition. Then the aforementioned lattice rules and higher order QMC will converge with rate $\mathcal{O}(N^{-\alpha+\delta})$ for arbitrarily small $\delta > 0$, without any $\log N$ term, and with the implied constant *independent of s* . (This is the essence of “*strong tractability*” [39].) Similar results hold for anisotropically adapted sparse grid quadrature.

There are other related strategies such as periodization, tent transform, and cosine transform, and we will discuss them briefly at the end of this paper. For now it is worth pointing out that the implied constants in the error bounds for different methods with the same convergence rate can depend on α in very different ways; some may grow super exponentially in α , thus requiring a huge number of points in practice to resolve the preasymptotic regime for large α .

Our approach in this paper offers a different alternative to higher order QMC for the weighted Sobolev space. We construct, in both theory and practice, non-QMC, i.e., unequally weighted, quadrature rules which aim to have at the same time higher order convergence with respect to the number of function evaluations, and little or no growth with respect to s . Specifically, with $|J|$ denoting the number of subsampled lattice points, we achieve a convergence rate close to the order

$$|J|^{-(\alpha-1/2)} \quad \text{and} \quad |J|^{-\alpha} (1.110721)^s,$$

where the implied constants are independent of s under favorable conditions. Both results hold and we can interpolate between them to trade between the convergence rate and the growth in s . This will be explained more precisely later in the introduction. Here it makes sense to state the convergence rate with respect to $|J|$, which corresponds to the number of function evaluations.

1.2. Our approach

Our approach with the full lattice $\mathbf{X} = \{t_0, \dots, t_{N-1}\}$ can be described in three steps as follows:

$$I(f) = \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x} = \int_{[0,1]^s} f(\boldsymbol{\psi}(\mathbf{y})) \mu(\mathbf{y}) d\mathbf{y} = \int_{[0,1]^s} F(\mathbf{y}) \mu(\mathbf{y}) d\mathbf{y} \quad (1)$$

$$\approx \int_{[0,1]^s} (A_{\text{ker}}^{\mathbf{X}} F)(\mathbf{y}) \mu(\mathbf{y}) d\mathbf{y} = \int_{[0,1]^s} \left(\sum_{k=0}^{N-1} a_k(F) K(t_k, \mathbf{y}) \right) \mu(\mathbf{y}) d\mathbf{y} \quad (2)$$

$$= \sum_{k=0}^{N-1} a_k(F) \kappa(t_k) = \sum_{k=0}^{N-1} w_k F(t_k) = \sum_{k=0}^{N-1} w_k f(\boldsymbol{\psi}(t_k)) = Q_{\boldsymbol{\psi}}^{\mathbf{X}}(f), \quad (3)$$

where all of the ingredients F , μ , $A_{\text{ker}}^{\mathbf{X}}$, K , a_k , t_k , κ , w_k will be defined and explained step by step below. Later we may replace $\mathbf{X}$ by a randomly subsampled lattice $\mathbf{X}_J = \{t_k : k \in J\}$.

Step 1. Periodising the integrand

The first step is to make a change of variable,

$$(x_1, \dots, x_s) =: \mathbf{x} = \boldsymbol{\psi}(\mathbf{y}) := (\psi(y_1), \dots, \psi(y_s)),$$

$$\psi(y) := \sin^2(\pi y) = \frac{1}{2}(1 - \cos(2\pi y)) \text{ for } y \in [0, 1], \quad (4)$$

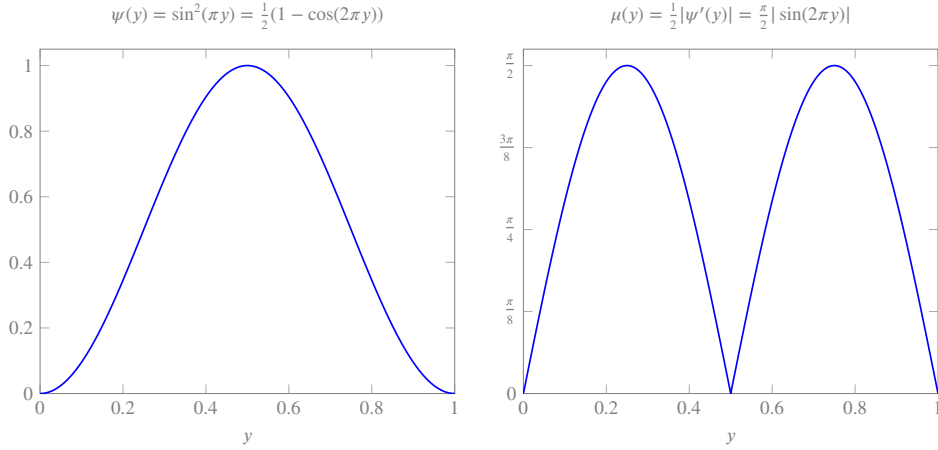


Figure 1: Univariate transformation $x = \psi(y) = \sin^2(\pi y) = \frac{1}{2}(1 - \cos(2\pi y))$ and associated density $\mu(y) = \frac{1}{2}|\psi'(y)| = \frac{\pi}{2}|\sin(2\pi y)|$.

to convert the problem to a different integration problem over the unit cube with respect to the $\mathbf{y}$ variable as shown in (1). Our choice of the univariate mapping function $\psi : y \in [0, 1] \rightarrow x \in [0, 1]$ traverses the interval $[0, 1]$ twice, see Figure 1 (left), and has the effect of making the transformed function

$$F(\mathbf{y}) := f(\psi(\mathbf{y})), \quad \mathbf{y} \in [0, 1]^s, \quad (5)$$

a smooth 1-periodic function in each coordinate direction. By the usual change of variable rules we then obtain a new expression for the integral in (1), with the product density

$$\mu(\mathbf{y}) := \prod_{j=1}^s \frac{|\psi'(y_j)|}{2} = \prod_{j=1}^s \frac{\pi |\sin(2\pi y_j)|}{2}. \quad (6)$$

Note that the univariate density is 1/2 times the Jacobian, see Figure 1 (right). This possibly unexpected result is a consequence of the double traverse of ψ over the interval $[0, 1]$. To verify this for $s = 1$, we transform the integral of f by making two duplicates, adding half of each, and then applying the one-to-one mappings $\psi : [0, \frac{1}{2}] \rightarrow [0, 1]$ and $\psi : [\frac{1}{2}, 1] \rightarrow [0, 1]$ to each piece, noting that ψ increases from 0 to 1 on the first piece (so ψ' is positive), and decreases from 1 to 0 on the second piece (so ψ' is negative): this gives

$$\begin{aligned} \int_0^1 f(x) dx &= \frac{1}{2} \int_0^1 f(x) dx + \frac{1}{2} \int_0^1 f(x) dx = \frac{1}{2} \int_0^{\frac{1}{2}} f(\psi(y)) \psi'(y) dy + \frac{1}{2} \int_{\frac{1}{2}}^1 f(\psi(y)) \psi'(y) dy \\ &= \frac{1}{2} \int_0^{\frac{1}{2}} f(\psi(y)) \psi'(y) dy - \frac{1}{2} \int_{\frac{1}{2}}^1 f(\psi(y)) \psi'(y) dy = \int_0^1 f(\psi(y)) \frac{|\psi'(y)|}{2} dy. \end{aligned}$$

Step 2. Kernel approximation of the function F

The second step is to approximate the smooth periodic function F in (5) by a *kernel interpolant* using the values of F at the points $\mathbf{X} = \{\mathbf{t}_0, \dots, \mathbf{t}_{N-1}\} \subset [0, 1]^s$, as shown in (2). This approximation takes the form

$$(A_{\text{ker}}^{\mathbf{X}} F)(\mathbf{y}) := \sum_{k=0}^{N-1} a_k(F) K(\mathbf{t}_k, \mathbf{y}), \quad \mathbf{y} \in [0, 1]^s, \quad (7)$$

for a known kernel function $K : [0, 1]^s \times [0, 1]^s \rightarrow \mathbb{R}$, where the coefficients $a_k(F) \in \mathbb{R}$ are determined by the interpolation condition $(A_{\text{ker}}^{\mathbf{X}} F)(\mathbf{t}_{k'}) = F(\mathbf{t}_{k'})$ for all $k' = 0, \dots, N-1$. Defining the vectors $\mathbf{a} = [a_k]_{k=0}^{N-1}$,

$F = [F(\mathbf{t}_k)]_{k=0}^{N-1}$ and the matrix $\mathcal{K} = [K(\mathbf{t}_k, \mathbf{t}_{k'})]_{k,k'=0}^{N-1}$, the interpolation condition can be written as the linear system in matrix-vector form

$$\mathcal{K}\mathbf{a} = F \implies \mathbf{a} = \mathcal{K}^{-1}F. \quad (8)$$

We assume that the transformed function F belongs to the *weighted Korobov space* which is a reproducing kernel Hilbert space (see Subsection 2.1), and we will use its kernel K in (7). For now it suffices to say that the kernel is *shift-invariant* so that $K(\mathbf{t}_k, \mathbf{t}_{k'}) = K(\mathbf{t}_k - \mathbf{t}_{k'}, \mathbf{0})$, i.e., it depends only on the difference between the two points. For the points $\mathbf{t}_k$ in (7) we will use lattice points

$$\mathbf{t}_k = \frac{k\mathbf{z}}{N} \bmod 1, \quad k = 0, \dots, N-1, \quad (9)$$

with generating vector $\mathbf{z} \in \{1, \dots, N-1\}^s$. This pairing of a shift-invariant kernel and a lattice pointset yields a linear system with a circulant matrix $\mathcal{K}$, which can be solved to obtain $\mathbf{a} = \mathcal{K}^{-1}F$ very efficiently using the Fast Fourier transform (FFT) in $\mathcal{O}(N \log N)$ operations. We also have well established theory on the convergence of the approximation error, see Section 4.

Step 3. Exact integration of the approximate integrand

The next step of our method is to integrate the product of the kernel interpolant and the density $A_{\text{ker}}^{\mathbf{X}} F(\mathbf{y}) \mu(\mathbf{y})$ exactly to arrive at (3), with

$$\kappa(\mathbf{t}) := \int_{[0,1]^s} K(\mathbf{t}, \mathbf{y}) \mu(\mathbf{y}) d\mathbf{y}, \quad \mathbf{t} \in [0, 1]^s. \quad (10)$$

If the kernel is a product of univariate kernels (when we have product weights for the weighted Korobov space, see Subsection 2.1) then this last step requires the accurate computation of only s one-dimensional integrals.

Combining $\boldsymbol{\kappa} := [\kappa(\mathbf{t}_k)]_{k=0}^{N-1}$ with $\mathbf{a} = \mathcal{K}^{-1}F$ in (3), we will show in Section 3 that the quadrature weights w_k are given collectively by $\mathbf{w} = \mathcal{K}^{-1} \boldsymbol{\kappa}$, which depends on the kernel and the points, but conveniently is independent of F . Hence, in practice we can compute the weights w_k once and then use the quadrature rule (3) to approximate the integral, without the need to actually compute the kernel approximation (7). In fact, we will explain in Section 3 that the weights given by $\mathbf{w} = \mathcal{K}^{-1} \boldsymbol{\kappa}$ are “optimal” for the reproducing kernel Hilbert space corresponding to the kernel K , in the sense that they minimise the worst case error for weighted integration with respect to the density μ in that space.

Step 2' and Step 3'. The subsampling alternative

Steps 1–3 explained our method for a full lattice pointset $\mathbf{X}$. However, it is well-known [5] that for kernel interpolation with lattice points and Korobov kernels the worst case approximation error achieves only half the optimal convergence rate, i.e., close to $\mathcal{O}(N^{-\alpha/2})$ instead of $\mathcal{O}(N^{-\alpha})$.

This issue can be addressed by using only a small subset of the lattice points. Instead of the full lattice pointset $\mathbf{X}$, we could work with a randomly subsampled lattice $\mathbf{X}_J$ for $J \subset \{0, \dots, N-1\}$ of size $|J| \sim \sqrt{N} \log N$ (see Section 4). To this end, we also need to replace the kernel matrix $\mathcal{K}$ by the $|J| \times |J|$ sub-matrix $\mathcal{K}_J$ and solve instead the linear system

$$\mathcal{K}_J \mathbf{a}_J = F_J.$$

Again, combining $\boldsymbol{\kappa}_J := [\kappa(\mathbf{t}_k)]_{k \in J}$ with $\mathbf{a}_J = \mathcal{K}_J^{-1} F_J$, it can be shown that the quadrature weights $w_{k,J}$ are given collectively by $\mathbf{w}_J = \mathcal{K}_J^{-1} \boldsymbol{\kappa}_J$, which depends on the kernel and the subsampled points, but conveniently is again independent of F . Full details will be explained in Section 3. Subsampling allows us to use fewer function values, and in turn improves the overall convergence rate of our method to be close to $\mathcal{O}(|J|^{-\alpha})$, see below.

Error analysis

The only step to incur an error in our method (1)–(3) is the kernel interpolation, and thus

$$|I(f) - Q_{\psi}^{\mathbf{X}}(f)| = \left| \int_{[0,1]^s} (F(\mathbf{y}) - (A_{\text{ker}}^{\mathbf{X}} F)(\mathbf{y})) \mu(\mathbf{y}) d\mathbf{y} \right|$$

$$\leq \inf_{p \in [2, \infty]} \left\{ \|F - A_{\ker}^{\mathbf{X}} F\|_{L_p([0,1]^s)} \|\mu\|_{L_{p'}([0,1]^s)} \right\}, \quad (11)$$

where we used Hölder's inequality with the conjugate pair $1/p + 1/p' = 1$. An analogous error bound holds if we replace the full lattice $\mathbf{X}$ by a subsampled lattice $\mathbf{X}_J$.

Both L_2 and L_∞ approximation have been analyzed in the literature for lattice-based kernel interpolant in weighted Korobov spaces [7, 8, 34]. The corresponding result for L_2 approximation with subsampled lattice was recently proved in [1], and we extend this to L_∞ approximation in this paper. We then interpolate the L_2 and L_∞ results to obtain the L_p approximation result for all $p \in [2, \infty]$. Here we need *one single lattice generating vector* which is suited for all values of p , thus we need some adaptations of existing results. For the same lattice we obtain upper bounds on $\|F - A_{\ker}^{\mathbf{X}} F\|_{L_p([0,1]^s)}$ and $\|F - A_{\ker}^{\mathbf{X}_J} F\|_{L_p([0,1]^s)}$ in the first factor of (11) for all p , and they depend on the norm of F in the weighted Korobov space.

We then relate this Korobov norm of the transformed F to a special mixed Sobolev norm of the original integrand f :

$$\|F\|_{\text{Kor}, s, \alpha, \gamma} \leq \|f\|_{\text{Sob}, s, \alpha, \beta, \omega}.$$

Details of our two function spaces will be given in Section 2. For now it suffices to know that γ is a rescaling of the weights β (see (21) ahead), and ω is a weight function that appears in the Sobolev norm (see (17) ahead).

For the second factor in (11) we have (see (56) ahead)

$$\|\mu\|_{L_{p'}([0,1]^s)} = \xi_p^s, \quad \text{with} \quad 1 = \xi_\infty \leq \xi_p \leq \xi_2 = \frac{\pi}{\sqrt{8}} \leq 1.110721.$$

Thus, if $p = \infty$ then there is no dependence on s ; otherwise there is a mild exponential growth of at worst $(1.110721)^s$, which occurs when $p = 2$.

In summary, we obtain the error bounds

$$\begin{aligned} |I(f) - Q_{\psi}^{\mathbf{X}}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_p N^{-\theta_p} \xi_p^s \right\} \|f\|_{\text{Sob}, s, \alpha, \beta, \omega}, \quad \theta_p \rightarrow \frac{\alpha}{2} - \frac{1}{4} + \frac{1}{2p}, \\ |I(f) - Q_{\psi}^{\mathbf{X}_J}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_p |J|^{-\theta_p} \xi_p^s \right\} \|f\|_{\text{Sob}, s, \alpha, \beta, \omega}, \quad \theta_p \rightarrow \alpha - \frac{1}{2} + \frac{1}{p}, \end{aligned}$$

with different constants C_p and rates θ_p for the two cases, and they all depend on some arbitrarily small parameter $\epsilon > 0$. We have shown the limits of the rates as $\epsilon \rightarrow 0$, but the constants go to infinity as $\epsilon \rightarrow 0$. The constants C_p and the norm of f can be bounded independently of s under appropriate conditions on the weights γ and β . The only error dependence on s is then through ξ_p^s . The value of p which yields the infimum will change depending on the size of N or $|J|$ relative to s .

Moreover, if the function f is *twice* as smooth compared to the reproducing kernel with smoothness α , then a generalization of the recent work [46] yields *double the rates*:

$$\begin{aligned} |I(f) - Q_{\psi}^{\mathbf{X}}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_p N^{-\theta_p} \xi_p^s \right\} \|f\|_{\text{Sob}, s, 2\alpha, \eta^2, \omega}, \quad \theta_p \rightarrow \alpha - \frac{1}{4} + \frac{1}{2p}, \\ |I(f) - Q_{\psi}^{\mathbf{X}_J}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_p |J|^{-\theta_p} \xi_p^s \right\} \|f\|_{\text{Sob}, s, 2\alpha, \eta^2, \omega}, \quad \theta_p \rightarrow 2\alpha - \frac{1}{2} + \frac{1}{p}, \end{aligned}$$

with new constants C_p , noting that the norm is now with smoothness 2α and squared and rescaled weights η^2 . Full details are given in Section 5.

The different convergence rates for $p = 2$ and $p = \infty$ are summarized in Table 1, with interpolated results for the values of $p \in (2, \infty)$. We stress that our method does not depend on the value of p and we do not need to specify a value of p for the error bound to hold. We will automatically get the smallest error bound over all possible values of $p \in [2, \infty]$. There is a trade-off between the polynomial rate in N or $|J|$ and the exponential dependence on s .

1.3. Structure of this paper

In Section 2 we define a weighted Korobov space and a special weighted Sobolev space, then relate the Korobov norm of F to the Sobolev norm of f . In Section 3 we provide further details on kernel interpolation and associated

	f with smoothness α		f with smoothness 2α	
	$p = 2$	$p = \infty$	$p = 2$	$p = \infty$
full lattice $\mathbf{X}$	$N^{-\alpha/2} (1.110721)^s$	$N^{-(\alpha/2-1/4)}$	$N^{-\alpha} (1.110721)^s$	$N^{-(\alpha-1/4)}$
subsampled lattice $\mathbf{X}_J$	$ J ^{-\alpha} (1.110721)^s$	$ J ^{-(\alpha-1/2)}$	$ J ^{-2\alpha} (1.110721)^s$	$ J ^{-(2\alpha-1/2)}$

Table 1

Summary of convergence results for quadrature rule using full and subsampled lattices.

cost, as well as the exact integration to obtain our final quadrature weights w_k . In Section 4 we provide error analysis to construct one single lattice generating vector for kernel interpolation, with a full lattice and a subsampled lattice, under both the L_2 and L_∞ approximation settings, and with results interpolated for L_p approximation for all $p \in [2, \infty]$. In Section 5 we apply these approximation error bounds to obtain the final error bound for our method, given in our main result, Theorem 8, and we also obtain doubling the rates in Corollary 10 if the integrand f is twice as smooth compared to the kernel. In Section 6 we present numerical results. In Section 7 we compare our method with related strategies and give some concluding remarks.

2. Function spaces

In Step 1 of our method (see (1)) we introduced a transformation function $F = f \circ \psi$ in (5). We now define two Hilbert spaces needed for our analysis: a weighted Korobov space of periodic functions for our transformed function F , and a special weighted mixed Sobolev space for our original integrand f . We then relate the Korobov norm of F to the Sobolev norm of f .

2.1. Weighted Korobov space of periodic functions

Consider all periodic functions defined on $[0, 1]^s$ with absolutely convergent Fourier series

$$F(\mathbf{y}) = \sum_{\mathbf{h} \in \mathbb{Z}^s} \hat{F}_{\mathbf{h}} e^{2\pi i \mathbf{h} \cdot \mathbf{y}}, \quad \text{where} \quad \hat{F}_{\mathbf{h}} = \int_{[0,1]^s} F(\mathbf{y}) e^{-2\pi i \mathbf{h} \cdot \mathbf{y}} d\mathbf{y}. \quad (12)$$

We define the *weighted Korobov space* $H_{s,\alpha,\gamma}^{\text{Kor}}$ with dimension $s \in \mathbb{N}$, smoothness parameter $\alpha > 1/2$ and positive weights $\gamma = \{\gamma_{\mathbf{u}}\}_{\mathbf{u} \in \mathbb{N}}$ to be the space of such functions with finite norm defined by

$$\|F\|_{\text{Kor},s,\alpha,\gamma}^2 := \sum_{\mathbf{h} \in \mathbb{Z}^s} r_{s,\alpha,\gamma}(\mathbf{h}) |\hat{F}_{\mathbf{h}}|^2, \quad r_{s,\alpha,\gamma}(\mathbf{h}) := \gamma_{\text{supp}(\mathbf{h})}^{-1} \prod_{j \in \text{supp}(\mathbf{h})} |h_j|^{2\alpha}, \quad (13)$$

where $\text{supp}(\mathbf{h}) := \{j \in \{1 : s\} : h_j \neq 0\}$ and $\{1 : s\} := \{1, \dots, s\}$. The user-defined weight $\gamma_{\mathbf{u}}$ controls the importance of the part of the function f that depends only on the variables $\mathbf{y}_{\mathbf{u}} := (y_j)_{j \in \mathbf{u}}$. We fix $\gamma_{\emptyset} := 1$. For integer $\alpha \in \mathbb{N}$ it is known that

$$\|F\|_{\text{Kor},s,\alpha,\gamma}^2 = \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{(2\pi)^{2\alpha|\mathbf{u}|}} \frac{1}{\gamma_{\mathbf{u}}} \int_{[0,1]^{|\mathbf{u}|}} \left| \int_{[0,1]^{s-|\mathbf{u}|}} \left(\prod_{j \in \mathbf{u}} \frac{\partial^\alpha}{\partial y_j^\alpha} \right) F(\mathbf{y}) d\mathbf{y}_{\{1:s\} \setminus \mathbf{u}} \right|^2 d\mathbf{y}_{\mathbf{u}}. \quad (14)$$

The space $H_{s,\alpha,\gamma}^{\text{Kor}}$ is a reproducing kernel Hilbert space, with inner product

$$\langle F, G \rangle_{\text{Kor},s,\alpha,\gamma} := \sum_{\mathbf{h} \in \mathbb{Z}^s} r_{s,\alpha,\gamma}(\mathbf{h}) \hat{F}_{\mathbf{h}} \overline{\hat{G}_{\mathbf{h}}}$$

and the reproducing kernel

$$K_{s,\alpha,\gamma}^{\text{Kor}}(\mathbf{y}, \mathbf{y}') := \sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \ker_{\alpha}(y_j, y'_j), \quad \ker_{\alpha}(y, y') := \sum_{h \in \mathbb{Z} \setminus \{0\}} \frac{e^{2\pi i h(y-y')}}{|h|^{2\alpha}}. \quad (15)$$

In general, a reproducing kernel $K(\cdot, \cdot)$ for a Hilbert space H over a domain D satisfies $K(\mathbf{y}, \cdot) \in H$ for all $\mathbf{y} \in D$, and $\langle K(\mathbf{y}, \cdot), F \rangle_H = F(\mathbf{y})$ for all $F \in H$ and $\mathbf{y} \in D$; the latter is known as the reproducing property.

For integer α we can write

$$\ker_\alpha(y, y') = \frac{(2\pi)^{2\alpha}}{(-1)^{\alpha+1}(2\alpha)!} B_{2\alpha}(\{y - y'\}), \quad y, y' \in [0, 1], \quad (16)$$

where the braces indicate that $y - y'$ is to be replaced by its fractional part in $[0, 1]$, and $B_{2\alpha}(y)$ is the Bernoulli polynomial of even degree 2α . For example, for $\alpha = 1, 2, 3$ we have

$$B_2(y) = y^2 - y + \frac{1}{6}, \quad B_4(y) = y^4 - 2y^3 + y^2 - \frac{1}{30}, \quad B_6(y) = y^6 - 3y^5 + \frac{5y^4}{2} - \frac{y^2}{2} + \frac{1}{42}.$$

In the special case of *product weights*, i.e., $\gamma_u = \prod_{j \in u} \gamma_j$ for some positive sequence $(\gamma_j)_{j \geq 1}$, the kernel takes a particularly simple form

$$K_{s,\alpha,\gamma}^{\text{Kor}}(\mathbf{y}, \mathbf{y}') = \prod_{j=1}^s (1 + \gamma_j \ker_\alpha(y_j, y'_j)).$$

2.2. Weighted Sobolev space of non-periodic functions

For an integer $\alpha \geq 1$, we define a special weighted Sobolev space of $H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ to be the class of functions on $[0, 1]^s$ with mixed square-integrable derivatives up to order α in each coordinate, equipped with the norm

$$\|f\|_{\text{Sob},s,\alpha,\beta,\omega}^2 := \sum_{u \subseteq \{1:s\}} \frac{1}{\beta_u} \sum_{\substack{\mathbf{v} \in \{0:\alpha\}^s \\ \text{supp}(\mathbf{v})=u}} \int_{[0,1]^s} \left| \frac{\partial^{\mathbf{v}}}{\partial \mathbf{x}^{\mathbf{v}}} f(\mathbf{x}) \right|^2 \prod_{j=1}^s \omega(x_j) d\mathbf{x}, \quad \omega(x) := \frac{1}{\pi \sqrt{x(1-x)}}. \quad (17)$$

Analogously to the weighted Korobov space, here the positive weight β_u controls the importance of the part of f with respect to the variables $\mathbf{x}_u = (x_j)_{j \in u}$, with $\beta_\emptyset := 1$. In addition, there is now a univariate weight function ω , i.e., the Chebyshev density on $[0, 1]$, which controls the behaviour of f at the boundaries of the cube. This special space is a subspace of the Sobolev space without weight function. The particular choice of the weight function will be seen to arise naturally in the next subsection, see Lemma 3.

With weights $\beta_u = 1$ and weight function $\omega \equiv 1$, the norm (17) is exactly the classical mixed Sobolev norm. In other words, our special norm (17) is closer to the mixed Sobolev norm than the “anchored” norms and the “unanchored” or “ANOVA” norms typically appearing the QMC literature.

2.3. Relating the Korobov norm of F to the Sobolev norm of f

In order to utilise error bounds in the weighted Korobov space, we need to relate the Korobov norm of the transformed periodic function $F = f \circ \boldsymbol{\psi}$ to the Sobolev norm of the original (non-periodic) function f , where we recall that $\boldsymbol{\psi}(\mathbf{y}) = (\psi(y_1), \dots, \psi(y_s))$ is defined in (4).

First, we show how the derivatives are related by using the Faà di Bruno formula for computing the multivariate chain rule. In what follows, we write the order $\mathbf{v}$ mixed derivative of a function using the shorthand notation $f^{(\mathbf{v})}$.

Lemma 1. *Let $f \in H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ for integer $\alpha \geq 1$ and weights β . For any multiindex $\mathbf{v} \in \{0 : \alpha\}^s$ the mixed derivative of the transformed function $F = f \circ \boldsymbol{\psi}$ satisfies*

$$|F^{(\mathbf{v})}(\mathbf{y})| \leq (2\pi)^{|\mathbf{v}|} \sum_{\substack{\lambda \leq \mathbf{v} \\ \text{supp}(\lambda) = \text{supp}(\mathbf{v})}} |f^{(\lambda)}(\boldsymbol{\psi}(\mathbf{y}))| \prod_{j=1}^s S(v_j, \lambda_j), \quad (18)$$

where $S(v, \lambda)$ denotes the Stirling number of the second kind.

PROOF. We follow [27, Appendix B], which analysed similar periodising transformations in the context of PDEs with random coefficients. The recursive Faà di Bruno formula from [43] gives

$$F^{(\mathbf{v})}(\mathbf{y}) = \sum_{\substack{\lambda \in \mathbb{N}_0^s \\ 1 \leq |\lambda| \leq |\mathbf{v}|}} f^{(\lambda)}(\boldsymbol{\psi}(\mathbf{y})) \kappa_{\mathbf{v},\lambda}(\mathbf{y}), \quad (19)$$

where, since $\boldsymbol{\psi}$ is a componentwise map, the sequence $\kappa_{\mathbf{v},\lambda}(\mathbf{y})$ can be defined recursively as $\kappa_{\mathbf{v},\mathbf{0}}(\mathbf{y}) \equiv \delta_{\mathbf{v},\mathbf{0}}$, $\kappa_{\mathbf{v},\lambda}(\mathbf{y}) \equiv 0$ if $|\mathbf{v}| < |\lambda|$ or $\lambda \not\geq \mathbf{0}$, and otherwise

$$\kappa_{\mathbf{v}+\mathbf{e}_j,\lambda}(\mathbf{y}) = \sum_{m=0}^{v_j} \psi^{(m+1)}(y_j) \kappa_{\mathbf{v}-m\mathbf{e}_j,\lambda-\mathbf{e}_j}(\mathbf{y}).$$

Here $\mathbf{e}_j \in \mathbb{N}_0^s$ is the j th unit basic vector. Note that the recursive formula also implies that $\kappa_{\mathbf{v},\lambda} \equiv 0$ if $\text{supp}(\lambda) \neq \text{supp}(\mathbf{v})$.

For $m \geq 1$, the derivative of ψ defined in (4) can be bounded by

$$|\psi^{(m)}(y)| = \left| \frac{d^m}{dy^m} (\sin^2(\pi y)) \right| = \left| \frac{d^m}{dy^m} \left(\frac{1 - \cos(2\pi y)}{2} \right) \right| = \left| (2\pi)^m \cos\left(2\pi y + m\frac{\pi}{2}\right) \right| \leq (2\pi)^m,$$

which is the same bound as satisfied by the derivatives of the transformation in [27, Theorem B.1]. Hence, the sequence $\kappa_{\mathbf{v},\lambda}$ satisfies the same bounds

$$|\kappa_{\mathbf{v},\lambda}(\mathbf{y})| \leq (2\pi)^{|\mathbf{v}|} \prod_{j=1}^s S(v_j, \lambda_j).$$

Substituting this into (19) yields (18), where we have simplified the limits in the sum by using the property that $S(v, \lambda) = 0$ if $v < \lambda$ and $S(v, 0) = \delta_{v,0}$.

Next, we need to know under what conditions the map $\mathbf{x} = \boldsymbol{\psi}(\mathbf{y})$ can be inverted. To see this in one dimension, let $g : [0, 1] \rightarrow \mathbb{R}$ be integrable and consider the periodised function $g \circ \psi$. Since $g \circ \psi$ is even periodic, i.e., $g(\psi(y)) = g(\psi(1 - y))$, the integral of $g \circ \psi$ can be written as

$$\int_0^1 g(\psi(y)) dy = \int_0^{1/2} g(\psi(y)) dy + \int_{1/2}^1 g(\psi(y)) dy = 2 \int_0^{1/2} g(\psi(y)) dy. \quad (20)$$

The mapping ψ is now monotone on the smaller interval $[0, \frac{1}{2}]$, and so we can make the change of variables

$$x = \psi(y) = \sin^2(\pi y) = \frac{1}{2}(1 - \cos(2\pi y)), \quad dx = \pi \sin(2\pi y) dy, \quad y = \frac{1}{2\pi} \cos^{-1}(1 - 2x),$$

to obtain

$$\int_0^{1/2} g(\psi(y)) dy = \int_0^1 \frac{g(x)}{\pi \sin(\cos^{-1}(1 - 2x))} dx = \frac{1}{2\pi} \int_0^1 \frac{g(x)}{\sqrt{x(1-x)}} dx,$$

where we used the identity $\sin(\cos^{-1}(z)) = \sqrt{1 - z^2}$ for $z \in (-1, 1)$. Substituting this into (20) leads to

$$\int_0^1 g(\psi(y)) dy = \int_0^1 g(x) \omega(x) dx,$$

where ω is the weight function defined in (17). Thus the inverse mapping is defined provided that g is integrable with respect to the weight function ω .

Since $\boldsymbol{\psi}$ is a componentwise mapping, this can be generalised to higher dimensions to give the following lemma. Here we define the weighted L^1 space by

$$L_{\omega}^1([0, 1]^s) := \left\{ g : [0, 1]^s \rightarrow \mathbb{R} : \int_{[0,1]^s} |g(\mathbf{x})| \prod_{j=1}^s \omega(x_j) d\mathbf{x} < \infty \right\}.$$

Lemma 2. *Let $g \in L_{\omega}^1([0, 1]^s)$. Then the mapping $\boldsymbol{\psi}$ can be “inverted” in the following sense*

$$\int_{[0,1]^s} g(\boldsymbol{\psi}(\mathbf{y})) d\mathbf{y} = \int_{[0,1]^s} g(\mathbf{x}) \prod_{j=1}^s \omega(x_j) d\mathbf{x},$$

and the integral on the right is finite.

Now we can bound the Korobov norm of $F = f \circ \psi$ by the Sobolev norm of f .

Lemma 3. *Let $f \in H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ for integer $\alpha \geq 1$ and positive weights $\beta = \{\beta_u\}_{u \in \mathbb{N}}$. Then the Korobov norm of $F = f \circ \psi$ satisfies*

$$\|F\|_{\text{Kor},s,\alpha,\gamma} \leq \|f\|_{\text{Sob},s,\alpha,\beta,\omega},$$

where $\gamma = \{\gamma_u\}_{u \in \mathbb{N}}$ are scaled weights given by

$$\gamma_u := c_\alpha^{|\mathbf{u}|} \beta_u, \quad \text{with} \quad c_\alpha := \sum_{\lambda=1}^{\alpha} S(\alpha, \lambda)^2. \quad (21)$$

PROOF. Applying the Cauchy–Schwarz inequality to the (squared) Korobov norm (14) then substituting in the bound (18) for each $\mathbf{v} = \alpha \sum_{j \in \mathbf{u}} \mathbf{e}_j$ gives

$$\begin{aligned} \|F\|_{\text{Kor},s,\alpha,\gamma}^2 &\leq \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{(2\pi)^{2\alpha|\mathbf{u}|}} \frac{1}{\gamma_u} \int_{[0,1]^s} \left| \left(\prod_{j \in \mathbf{u}} \frac{\partial^\alpha}{\partial y_j^\alpha} \right) F(\mathbf{y}) \right|^2 d\mathbf{y} \\ &\leq \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{\gamma_u} \int_{[0,1]^s} \left| \sum_{\substack{\lambda \in \{1:\alpha\}^s \\ \text{supp}(\lambda) = \mathbf{u}}} f^{(\lambda)}(\psi(\mathbf{y})) \prod_{j \in \mathbf{u}} S(\alpha, \lambda_j) \right|^2 d\mathbf{y}, \end{aligned}$$

where we have simplified the inner sum using $S(\alpha, 0) = 1$.

Applying the Cauchy–Schwarz inequality to the sum over λ , we can bound this by

$$\begin{aligned} \|F\|_{\text{Kor},s,\alpha,\gamma}^2 &\leq \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{\gamma_u} \int_{[0,1]^s} \left(\sum_{\substack{\lambda \in \{1:\alpha\}^s \\ \text{supp}(\lambda) = \mathbf{u}}} |f^{(\lambda)}(\psi(\mathbf{y}))|^2 \right) \left(\sum_{\substack{\lambda \in \{1:\alpha\}^s \\ \text{supp}(\lambda) = \mathbf{u}}} \prod_{j \in \mathbf{u}} S(\alpha, \lambda_j)^2 \right) d\mathbf{y} \\ &= \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{\gamma_u} \left(\sum_{\lambda=1}^{\alpha} S(\alpha, \lambda)^2 \right)^{|\mathbf{u}|} \sum_{\substack{\lambda \in \{1:\alpha\}^s \\ \text{supp}(\lambda) = \mathbf{u}}} \int_{[0,1]^s} |f^{(\lambda)}(\psi(\mathbf{y}))|^2 d\mathbf{y} \\ &= \sum_{\mathbf{u} \subseteq \{1:s\}} \frac{1}{\beta_u} \sum_{\substack{\lambda \in \{1:\alpha\}^s \\ \text{supp}(\lambda) = \mathbf{u}}} \int_{[0,1]^s} |f^{(\lambda)}(\mathbf{x})|^2 \prod_{j=1}^s \omega(x_j) d\mathbf{x} = \|f\|_{\text{Sob},s,\alpha,\beta,\omega}^2, \end{aligned}$$

where we used the connection between weights given in (21), as well as Lemma 2, which is valid since $f \in H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ implies that $|f^{(\lambda)}(\cdot)|^2 \in L_\omega^1([0,1]^s)$ for all $\lambda \in \{0:\alpha\}^s$.

Note that the norm inequality in Lemma 3 is sharp for the constant function.

3. Kernel interpolation and integration

Here we provide further details for Steps 2 and 3 of our method.

3.1. Kernel interpolation

In Step 2 of our method (see (2)) we approximated the transformed function F by its kernel interpolant

$$(A_{\text{ker}}^X F)(\mathbf{y}) := \sum_{k=0}^{N-1} a_k(F) K(\mathbf{t}_k, \mathbf{y}).$$

Defining $\mathbf{a} = [a_k(F)]_{k=0}^{N-1}$, $\mathbf{F} = [F(\mathbf{t}_k)]_{k=0}^{N-1}$, and $\mathcal{K} = [K(\mathbf{t}_k, \mathbf{t}_{k'})]_{k,k'=0}^{N-1}$, we already explained that the coefficients $a_k(F)$ are given by solving the linear system $\mathcal{K}\mathbf{a} = \mathbf{F}$ to obtain $\mathbf{a} = \mathcal{K}^{-1}\mathbf{F}$.

In the case of a Korobov kernel and a full lattice $\mathbf{X}$, we have $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y}, \mathbf{0})$ and thus $\mathcal{K}$ is a circulant matrix

$$\mathcal{K} = [K(\mathbf{t}_k, \mathbf{t}_{k'})]_{k,k'=0}^{N-1} = [K(\mathbf{t}_{(k-k') \bmod N}, \mathbf{0})]_{k,k'=0}^{N-1},$$

i.e., each entry only depends on the difference of the row and column indices. Thus, it is possible to diagonalize it by a discrete Fourier transform, which can be performed in almost linear cost using the fast Fourier transform (FFT). The kernel approximation with lattices is therefore computable using three FFTs in $\mathcal{O}(N \log N)$ arithmetic operations,

$$\mathbf{a} = \mathcal{K}^{-1} \mathbf{F} = \text{ifft}(\text{fft}(\mathbf{F}) ./ \text{fft}(\mathbf{c})),$$

where $\mathbf{c} = [K(\mathbf{t}_k, \mathbf{0})]_{k=0}^{N-1}$ is the first column of $\mathcal{K}$, and $./$ denotes componentwise division, and where fft and ifft denote the FFT and inverse FFT, respectively.

The cost for evaluating the kernel will depend on the structure of the weights γ_u , ranging from linear in s for product weights to quadratic in s for POD and SPOD weights, see [25, Table 1].

Kernel methods have a long history, tracing back to Wahba's seminal work on splines [51]. They have an important role in radial basis functions [52], and have recently been used in high dimensional approximation, see, e.g., [56, 55, 22, 25, 26]. We will provide error analysis for lattice-based kernel methods in Section 4.

3.2. The numerical integration step

In Step 3 of our method (see (3)) we integrated the product of the kernel interpolant $A_{\text{ker}}^{\mathbf{X}} F$ and the density μ exactly. Thus our initial integral $I(f)$ has been approximated by

$$I_{\mu}(F) := \int_{[0,1]^s} F(\mathbf{y}) \mu(\mathbf{y}) d\mathbf{y} \approx \sum_{k=0}^{N-1} a_k(F) \kappa(\mathbf{t}_k), \quad \kappa(\mathbf{t}_k) = \int_{[0,1]^s} K(\mathbf{t}_k, \mathbf{y}) \mu(\mathbf{y}) d\mathbf{y}. \quad (22)$$

In our case, since the Korobov kernel $K = K_{\text{Kor},s,\alpha,\gamma}$ can be written explicitly in terms of Bernoulli polynomials (see (16)), the integrals $\kappa(\mathbf{t}_k)$ can be evaluated analytically at essentially no cost.

Since the interpolant coefficients are linear in the function evaluations, by defining $\boldsymbol{\kappa} = [\kappa(\mathbf{t}_k)]_{k=0}^{N-1}$, we can write

$$\sum_{k=0}^{N-1} a_k(F) \kappa(\mathbf{t}_k) = \mathbf{a}^{\top} \boldsymbol{\kappa} = \mathbf{F}^{\top} \mathcal{K}^{-1} \boldsymbol{\kappa} = \sum_{k=0}^{N-1} [\mathcal{K}^{-1} \boldsymbol{\kappa}]_k F(\mathbf{t}_k).$$

Hence the approximation (22) is equivalent to a weighted quadrature rule

$$Q_{\psi}^{\mathbf{X}}(f) := \sum_{k=0}^{N-1} w_k f(\psi(\mathbf{t}_k)) = \sum_{k=0}^{N-1} w_k F(\mathbf{t}_k), \quad \text{with } w_k := [\mathcal{K}^{-1} \boldsymbol{\kappa}]_k. \quad (23)$$

Since the points $\mathbf{t}_k$ are lattice points (see (9)), the transformed points $\psi(\mathbf{t}_k)$ of the weighted quadrature rule are related to a *Chebyshev lattice* [10], see Section 7.4. The effect of the transformation is to concentrate the points near the boundary and since ψ traverses the interval twice there is a duplication of points. A 2-dimensional example of a lattice and the corresponding transformed lattice points are shown in Figure 2.

The weights $\mathbf{w} = [w_k]_{k=0}^{N-1}$ given by (23) are in fact *optimal weights* for μ -weighted integration in the Korobov space $H_{s,\alpha,\gamma}^{\text{Kor}}$ with the pointset $\mathbf{X} := \{\mathbf{t}_0, \dots, \mathbf{t}_{N-1}\}$ as studied in [23, 40]. By “optimal” we refer here to the weights that minimise the *worst case integration error* of the quadrature rule for weighted integration with respect to μ in the space $H_{s,\alpha,\gamma}^{\text{Kor}}$,

$$e^{\text{int},\mu}(Q_{\mu}^{\mathbf{X}}, H_{s,\alpha,\gamma}^{\text{Kor}}) := \sup_{F \in H_{s,\alpha,\gamma}^{\text{Kor}}, \|F\|_{\text{Kor},s,\alpha,\gamma} \leq 1} \left| I_{\mu}(F) - \sum_{k=0}^{N-1} w_k F(\mathbf{t}_k) \right|.$$

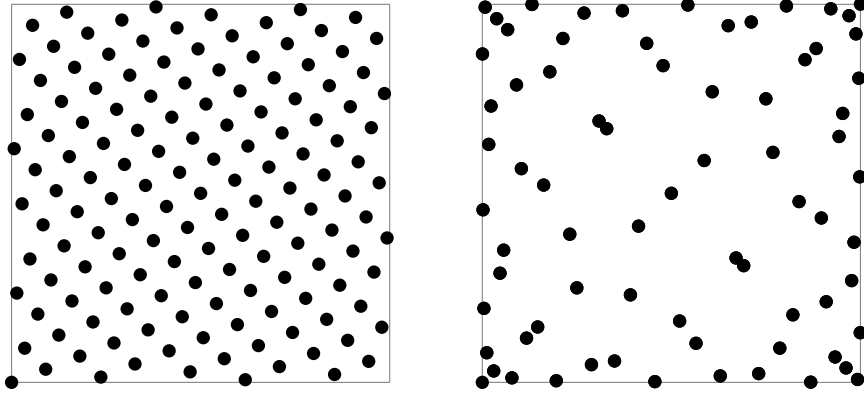


Figure 2: Two-dimensional lattice rule with $N = 144$ and $\mathbf{z} = (1, 89)^\top$ on the left, and the same lattice points after applying the transformation ψ on the right, showing $\lfloor 144/2 + 1 \rfloor = 73$ distinct points.

Step \ Weights γ	product	POD	SPOD
CBC	$sN \log(N)$	$sN \log(N) + s^2 \log(s)N$	$sN \log(N) + s^3 \alpha^2 N$
Compute $K(t_k, \mathbf{0})$, $\kappa(t_k)$ for all k	sN	$sN + s^2 \log(s)N$	$s^2 \alpha^2 N$
Compute weights $\mathbf{w}$	$N \log(N)$	$N \log(N)$	$N \log(N)$
Total	$sN \log(N)$	$sN \log(N) + s^2 \log(s)N$	$sN \log(N) + s^3 \alpha^2 N$

Table 2

Cost of steps to construct quadrature rule for full lattice (23) for different weights γ .

3.3. Implementation for full lattice

To construct the quadrature rule (23) for the full lattice $\mathbf{X} = \{t_0, \dots, t_{N-1}\}$, we must choose the generating vector $\mathbf{z}$ to obtain the lattice points $\{t_k\}$ by (9) and then compute the weights $\{w_k\}$. The error analysis and construction for $\mathbf{z}$ will be given in Section 4 below. For now it suffices to say that $\mathbf{z}$ can be constructed efficiently using the fast *component-by-component* (CBC) algorithm for approximation, which for product, POD and SPOD weights is outlined in [7]. The cost of the CBC construction depends on the weights and is summarised in Table 2.

Then, given lattice points, to compute the corresponding weights $\{w_k\}$ we first compute the kernel integrals $\kappa(t_k)$ from (22). Expanding the kernel as in (15) and using the fact that $\int_0^1 \mu(y) dy = 1$, we have

$$\kappa(t_k) = \sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \int_0^1 \ker_{\alpha}(t_{k,j}, y_j) \mu(y_j) dy_j = \sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}} \prod_{j \in \mathbf{u}} \Lambda_{\alpha}(t_{k,j}), \quad (24)$$

where for $t \in [0, 1)$ we define the one-dimensional integral

$$\Lambda_{\alpha}(t) := \int_0^1 \ker_{\alpha}(t, y) \mu(y) dy = \frac{\pi(2\pi)^{2\alpha}}{2(-1)^{\alpha+1}(2\alpha)!} \int_0^1 B_{2\alpha}(\{t - y\}) |\sin(2\pi y)| dy, \quad (25)$$

which can be computed exactly. In particular,

$$\begin{aligned} & \int_0^1 B_{2\alpha}(\{t - y\}) |\sin(2\pi y)| dy \\ &= \begin{cases} \int_0^t B_{2\alpha}(t - y) \sin(2\pi y) dy + \int_t^{1/2} B_{2\alpha}(t - y + 1) \sin(2\pi y) dy - \int_{1/2}^1 B_{2\alpha}(t - y + 1) \sin(2\pi y) dy & \text{if } t \in [0, 1/2], \\ \int_0^{1/2} B_{2\alpha}(t - y) \sin(2\pi y) dy - \int_{1/2}^t B_{2\alpha}(t - y) \sin(2\pi y) dy - \int_t^1 B_{2\alpha}(t - y + 1) \sin(2\pi y) dy & \text{if } t \in [1/2, 1). \end{cases} \end{aligned}$$

For $\alpha = 1, 2, 3$ and $t \in [0, 1)$ these integrals are given by

$$\Lambda_1(t) = \frac{\pi^2}{3} + \frac{\pi}{2} |\sin(2\pi t)| + \frac{\pi^2}{2} \left(4t^2 - 4t - 2 \left| t - \frac{1}{2} \right| + \frac{1}{2} - \frac{2}{\pi^2} \right),$$

Algorithm 1 Constructing the quadrature points and weights for the full (or subsampled) lattice

Input: α, γ, N (and J for the subsampled variant)

Output: $\mathbf{z}, \mathbf{X} = \{\mathbf{t}_k\}_{k=0}^{N-1}, \mathbf{w} = [w_k]_{k=0}^{N-1}$

- 1: construct $\mathbf{z}$ using (fast) CBC algorithm
 - 2: **for** $t \in \{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}\}$ **do** compute $\Lambda_\alpha(t)$ in (25)
 - 3: **for** $k = 0, 1, 2, \dots, N-1$ **do**
 - 4: compute $\mathbf{t}_k = \frac{k\mathbf{z}}{N} \bmod 1$
 - 5: compute $K(\mathbf{t}_k, \mathbf{0})$ using (15)
 - 6: compute $\kappa(\mathbf{t}_k)$ using (24)
 - 7: **end for**
 - 8: for the full lattice compute the weights $\mathbf{w}$ via fast FFT using (26) (for the subsample variant compute the weights $\mathbf{w}_J$ in (28) via one of three methods)
-

$$\Lambda_2(t) = \frac{-\pi^4}{3} \left[2t^4 - 4t^3 + \left(\frac{7}{2} - \frac{6}{\pi^2} \right) t^2 - \left(\frac{3}{2} - \frac{6}{\pi^2} \right) t - 2 \left| t - \frac{1}{2} \right|^3 + \frac{3}{\pi^2} \left| t - \frac{1}{2} \right| + \frac{59}{240} - \frac{7}{4\pi^2} + \frac{3}{\pi^4} - \frac{3}{2\pi^3} |\sin(2\pi t)| \right],$$

and, denoting the indicator function by $\mathbb{1}(\cdot)$,

$$\begin{aligned} \Lambda_3(t) = & \frac{2}{45} \left(2\pi^6 t^6 - 6\pi^6 t^5 + \left(\frac{65}{4}\pi^6 - 15\pi^4 \right) t^4 + \left(\frac{-45}{2}\pi^6 + 30\pi^4 \right) t^3 + \left(\frac{299}{16}\pi^6 - \frac{195}{4}\pi^4 + 45\pi^2 \right) t^2 \right. \\ & + \left(\frac{-135}{16}\pi^6 + \frac{135}{4}\pi^4 - 45\pi^2 \right) t + \frac{2143}{1344}\pi^6 - \frac{299}{32}\pi^4 + \frac{195}{8}\pi^2 - \frac{45}{2} \\ & - 3\pi \left(\mathbb{1}(t \leq \tfrac{1}{2}) \left[\frac{-15}{4} \sin(2\pi t) - \pi^5 t^5 + 5\pi^5 t^4 + \left(5\pi^3 - \frac{15}{2}\pi^5 \right) t^3 + \left(-15\pi^3 + \frac{25}{4}\pi^5 \right) t^2 \right. \right. \\ & + \left. \left(\frac{-15}{2}\pi + \frac{45}{4}\pi^3 - \frac{45}{16}\pi^5 \right) t + \frac{15}{2}\pi - \frac{25}{8}\pi^3 + \frac{17}{32}\pi^5 \right] \\ & + \mathbb{1}(t > \tfrac{1}{2}) \left[\frac{15}{4} \sin(2\pi t) + \pi^5 t^5 + \left(-5\pi^3 - \frac{5}{2}\pi^5 \right) t^3 + \frac{15}{4}\pi^5 t^2 \right. \\ & + \left. \left. \left(\frac{15}{2}\pi + \frac{15}{4}\pi^3 - \frac{35}{16}\pi^5 \right) t - \frac{15}{8}\pi^3 + \frac{15}{32}\pi^5 \right] \right) \Bigg). \end{aligned}$$

Thus the integrals $\kappa(\mathbf{t}_k)$ can be computed via (24), where depending on the type of weights γ the sum over $\mathbf{u}$ can be computed efficiently in the same way as evaluating the kernel (see Table 2 for the cost). Note that all components of the lattice points $\mathbf{t}_k$ are multiples of $1/N$, so we only need to compute $\Lambda_\alpha(t)$ for $t \in \{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}\}$.

Next we compute the weights $\mathbf{w} = [w_k]_{k=0}^{N-1}$ by solving the system $\mathcal{K}\mathbf{w} = \boldsymbol{\kappa}$, which since $\mathcal{K}$ is circulant can be done in (almost) linear time using the FFT. Let $\mathbf{c} = [K(\mathbf{t}_k, \mathbf{0})]_{k=0}^{N-1}$ denote the first column of the kernel matrix as above, then we can also solve the linear system for the quadrature weights using the FFT,

$$\mathbf{w} = \text{ifft}(\text{fft}(\boldsymbol{\kappa})/\text{fft}(\mathbf{c})), \quad (26)$$

with cost $\mathcal{O}(N \log N)$.

The procedure to compute the points and weights for the quadrature rule (23) is summarised in Algorithm 1. Since the linear system can be solved in $\mathcal{O}(N \log N)$ the cost of constructing the quadrature rule is dominated by the CBC construction cost. The cost breakdown is summarised in Table 2.

3.4. Implementation for subsampled lattice

For a subset $J \subset \{0, 1, \dots, N-1\}$, the quadrature rule for the subsampled lattice $\mathbf{X}_J = \{\mathbf{t}_k\}_{k \in J}$ is

$$Q_{\psi}^{X_J}(f) := \sum_{k \in J} w_{J,k} f(\psi(\mathbf{t}_k)), \quad \text{with} \quad w_{J,k} := [\mathcal{K}_J^{-1} \boldsymbol{\kappa}_J]_k, \quad (27)$$

Method	Cost in $ J $	Cost in N
1. direct	$ J ^3$	$N^{\frac{3}{2}(1+\delta)} \log(N)^3$
2. iterative	$\# \text{ iter.} \times J ^2$	$\# \text{ iter.} \times N^{1+\delta} (\log N)^2$
3. iterative with FFT	N/A	$\# \text{ iter.} \times N \log N$

Table 3

Summary of cost of different methods for solving the subsampled system (28) in terms of size of the subsampled lattice $|J|$ and full lattice N . In practice we take $|J| = \sqrt{N} \log N$ so $\delta = 0$.

where now the weights $\mathbf{w}_J = [w_{J,k}]_{k \in J}$ come from the subsampled kernel approximation and we denote $\mathcal{K}_J = [K(\mathbf{t}_k, \mathbf{t}_{k'})]_{k,k' \in J}$ and $\boldsymbol{\kappa}_J = [\kappa(\mathbf{t}_k)]_{k \in J}$. Here we have followed the same steps as in (23) to obtain the subsampled weights.

Algorithm 1 Steps 1–7 remain the same for the subsampled lattice, however, Step 8 is now replaced by solving

$$\mathcal{K}_J \mathbf{w}_J = \boldsymbol{\kappa}_J, \quad (28)$$

using one of three methods:

1. Solve (28) directly with cost $\mathcal{O}(|J|^3)$.
2. Solve (28) using an iterative method, such as conjugate gradient, with a cost $\mathcal{O}(|J|^2)$ times the number of iterations. Note that for an iterative method the number of iterations will depend on the condition number of the matrix $\mathcal{K}_J$, which can be large in practice.
3. Solve (28) using an iterative method but to perform the matrix vector multiplications at each iteration by embedding $\mathcal{K}_J$ back into the full kernel matrix $\mathcal{K}$ and then exploiting the FFT. For $\mathbf{v} \in \mathbb{R}^{|J|}$, we can perform the matrix-vector multiplication $\mathcal{K}_J \mathbf{v}$ by embedding $\mathcal{K}_J$ into $\mathcal{K}$ and padding $\mathbf{v}$ with zeroes, performing the multiplication with the full matrix $\mathcal{K}$ (using the FFT) then selecting the entries corresponding to the set J . Mathematically, we compute

$$\mathcal{K}_J \mathbf{v} = [\mathcal{K}(\mathbf{v}; \mathbf{0}_{-J})]_J = [\text{ifft}(\text{fft}(\mathbf{c}) * \text{fft}((\mathbf{v}; \mathbf{0}_{-J})))]_J$$

where $*$ denotes componentwise multiplication and $(\mathbf{v}; \mathbf{0}_{-J})_k = v_k$ if $k \in J$ and 0 otherwise. Performing the matrix-vector multiplication in this way costs $\mathcal{O}(N \log N)$ due to the FFT and so the overall cost of solving (28) is the same order as computing the weights for the full lattice via (26). Note again, however, that this cost will still depend on the number of iterations, which can be high in practice.

To compare the cost of the three methods in terms of both the size of the subsampled lattice $|J|$ and the size of the full lattice N , we use the property that the subsampling theory in Section 4 below implies the upper bound $|J| \lesssim N^{\frac{1+\delta}{2}} \log N$ for some $\delta > 0$. Specifically, from Theorem 7 below it follows that $\delta = \epsilon/(\alpha - \epsilon)$ for $\epsilon > 0$ from the convergence rate. Using this upper bound on $|J|$ the cost of each method for solving (28) is summarised in Table 3. Note that in practice, similar to [1] we take $|J| = \lceil \sqrt{N} \log N \rceil$, i.e., we treat δ as 0.

Thus, depending on the method used to solve (28) the cost of constructing the quadrature rule will be the same as the full lattice outlined in Table 2 with Method 3 or possibly higher with Methods 1 and 2. The benefit of using the subsampled rule is that, after the quadrature rule is constructed, only $|J| = \sqrt{N} \log N$ function evaluations are required which is much less.

4. Theory for lattice-based kernel interpolant and subsampling

In Step 2 of our method (see (2)) we approximated the transformed function F by its kernel interpolant $A_{\text{ker}}^X F$, while the alternative Step 2' replaced X by a subsampled lattice X_J . In this section we provide the full error analysis for both.

To obtain error bounds for a kernel interpolant with a full lattice, we make use of error bounds from a ‘‘classical lattice algorithm’’. To obtain error bounds for a kernel interpolant with a subsampled lattice, we need error bounds from a ‘‘least squares approximation’’. We will discuss these in the next two subsections. We will need to adapt the existing theory to have one single lattice generating vector for all values of $p \in [2, \infty]$ in (11).

We will start with a general function space setting, and later specialize to the weighted Korobov space. For functions defined on $[0, 1]^s$ with absolutely convergent Fourier series (12), we consider a general function space H equipped with the norm

$$\|F\|_H^2 := \sum_{\mathbf{h} \in \mathbb{Z}^s} r(\mathbf{h}) |\hat{F}_{\mathbf{h}}|^2, \quad (29)$$

where the “radius” function $r(\mathbf{h})$ enforces a decay in the Fourier coefficients, with $r(\mathbf{h}) \rightarrow \infty$ for $\|\mathbf{h}\|_2 \rightarrow \infty$. The specific choice of $r(\mathbf{h}) = r_{s,\alpha,\gamma}(\mathbf{h})$ defined in (13) yields the weighted Korobov space.

4.1. Classical lattice algorithm and kernel interpolant

For a given frequency index set $\mathcal{A} \subset \mathbb{Z}^s$ and a lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$, a classical lattice algorithm approximates F by

$$(A_{\mathcal{A}}^{\mathbf{X}} F)(\mathbf{y}) := \sum_{\mathbf{h} \in \mathcal{A}} \left(\frac{1}{N} \sum_{k=0}^{N-1} F(\mathbf{t}_k) e^{-2\pi i \mathbf{h} \cdot \mathbf{z}/N} \right) e^{2\pi i \mathbf{h} \cdot \mathbf{y}}, \quad \mathbf{t}_k = \frac{k}{N} \mathbf{z} \bmod 1, \quad (30)$$

see, e.g., [35, 29, 28, 41, 7, 12]. For the same lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$, recall that a kernel interpolant $A_{\text{ker}}^{\mathbf{X}} F$ approximates F by (7), where the coefficients $a_k(F)$ are obtained by solving a linear system $(A_{\text{ker}}^{\mathbf{X}} F)(\mathbf{t}_{k'}) = F(\mathbf{t}_{k'})$ for $k' = 0, \dots, N-1$.

The *worst case L_p approximation error* of any algorithm A in the space H is defined as

$$e^{\text{app}}(A, H, L_p) := \sup_{\|F\|_H \leq 1} \|F - AF\|_{L_p}.$$

Recall that the kernel interpolant is optimal among all algorithms using the same function evaluations, see e.g., [25, Theorem 2.2]. We have the following approximation error bounds.

Theorem 4. *For a frequency index set $\mathcal{A} := \{\mathbf{h} \in \mathbb{Z}^s : r(\mathbf{h}) \leq M\}$ with $M > 0$ and a lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$, we have*

$$e^{\text{app}}(A_{\text{ker}}^{\mathbf{X}}, H, L_2) \leq e^{\text{app}}(A_{\mathcal{A}}^{\mathbf{X}}, H, L_2) < \left(\frac{1}{M} + M S_N(\mathbf{z}) \right)^{1/2}, \quad (31)$$

$$e^{\text{app}}(A_{\text{ker}}^{\mathbf{X}}, H, L_{\infty}) \leq e^{\text{app}}(A_{\mathcal{A}}^{\mathbf{X}}, H, L_{\infty}) < \left(\sum_{\mathbf{h} \notin \mathcal{A}} \frac{1}{r(\mathbf{h})} \right)^{1/2} + \sqrt{|\mathcal{A}|} \sqrt{M S_N(\mathbf{z})}, \quad (32)$$

where

$$S_N(\mathbf{z}) := \sum_{\mathbf{h} \in \mathbb{Z}^s} \frac{1}{r(\mathbf{h})} \sum_{\substack{\boldsymbol{\ell} \in \mathbb{Z}^s \setminus \{0\} \\ \boldsymbol{\ell} \cdot \mathbf{z} \equiv 0 \pmod{N}}} \frac{1}{r(\mathbf{h} + \boldsymbol{\ell})}. \quad (33)$$

In particular, the choice $M = 1/\sqrt{S_N(\mathbf{z})}$ minimizes the right-hand side of (31) and yields

$$e^{\text{app}}(A_{\text{ker}}^{\mathbf{X}}, H, L_2) \leq e^{\text{app}}(A_{\mathcal{A}}^{\mathbf{X}}, H, L_2) < \sqrt{\frac{2}{M}} = \sqrt{2} (S_N(\mathbf{z}))^{1/4}. \quad (34)$$

PROOF. For $F \in H$ we write $F - A_{\mathcal{A}}^{\mathbf{X}} F = (F - P_{\mathcal{A}} F) + (P_{\mathcal{A}} F - A_{\mathcal{A}}^{\mathbf{X}} F)$, where $(P_{\mathcal{A}} F)(\mathbf{y}) = \sum_{\mathbf{h} \in \mathcal{A}} \hat{F}_{\mathbf{h}} e^{2\pi i \mathbf{h} \cdot \mathbf{y}}$ is the L_2 -projection of F onto $\text{span}\{e^{2\pi i \mathbf{h} \cdot \mathbf{y}} : \mathbf{h} \in \mathcal{A}\}$. The L_2 bound (31) was derived in [7, Section 2.4]. We have $\|F - A_{\mathcal{A}}^{\mathbf{X}} F\|_{L_2}^2 = \|F - P_{\mathcal{A}} F\|_{L_2}^2 + \|P_{\mathcal{A}} F - A_{\mathcal{A}}^{\mathbf{X}} F\|_{L_2}^2$, with

$$\|F - P_{\mathcal{A}} F\|_{L_2}^2 = \sum_{\mathbf{h} \notin \mathcal{A}} \frac{r(\mathbf{h})}{r(\mathbf{h})} |\hat{F}_{\mathbf{h}}|^2 \leq \frac{1}{M} \|F\|_H^2 \quad \text{and} \quad \|P_{\mathcal{A}} F - A_{\mathcal{A}}^{\mathbf{X}} F\|_{L_2}^2 < M S_N(\mathbf{z}) \|F\|_H^2. \quad (35)$$

We state the latter as a strict inequality, because in the derivation of the quantity $S_N(\mathbf{z})$, see e.g., [7, Formula (2.7)], infinitely many positive terms with index $\mathbf{h} \notin \mathcal{A}$ were added.

We now proceed to derive the L_∞ bound (32). For $F \in H$, by the triangle inequality we have

$$\|F - A_{\mathcal{A}}^X F\|_{L_\infty} \leq \|F - P_{\mathcal{A}} F\|_{L_\infty} + \|P_{\mathcal{A}} F - A_{\mathcal{A}}^X F\|_{L_\infty}. \quad (36)$$

For the first term on the right-hand side of (36), we use the Cauchy–Schwarz inequality to obtain

$$\|F - P_{\mathcal{A}} F\|_{L_\infty} \leq \sum_{\mathbf{h} \notin \mathcal{A}} \sqrt{\frac{r(\mathbf{h})}{r(\mathbf{h})}} |\hat{F}_{\mathbf{h}}| \leq \left(\sum_{\mathbf{h} \notin \mathcal{A}} \frac{1}{r(\mathbf{h})} \right)^{1/2} \|F\|_H.$$

For the second term on the right-hand side of (36), we write $(P_{\mathcal{A}} F - A_{\mathcal{A}}^X F)(\mathbf{y}) =: \sum_{\mathbf{h} \in \mathcal{A}} \hat{q}_{\mathbf{h}} e^{2\pi i \mathbf{h} \cdot \mathbf{y}}$ and apply the Cauchy–Schwarz inequality again to obtain

$$\|P_{\mathcal{A}} F - A_{\mathcal{A}}^X F\|_{L_\infty} \leq \sum_{\mathbf{h} \in \mathcal{A}} |\hat{q}_{\mathbf{h}}| \leq \sqrt{|\mathcal{A}|} \left(\sum_{\mathbf{h} \in \mathcal{A}} |\hat{q}_{\mathbf{h}}|^2 \right)^{1/2} = \sqrt{|\mathcal{A}|} \|P_{\mathcal{A}} F - A_{\mathcal{A}}^X F\|_{L_2},$$

which can be bounded using (35). Combining the two estimates gives (32) as required.

We remark that [34, Lemma 5.1 and Lemma 5.4] included two alternative upper bounds for (32). The first bound there is similar to (32) (there is an extra factor 3 but the square root is outside the sum which is tighter). The second bound there makes use of the precise definition of $r(\mathbf{h})$ in weighted Korobov spaces and requires a reparametrization of the quantity $S_N(\mathbf{z})$ with respect to the weights γ and the smoothness parameter α , and it yields a slightly better overall convergence rate when $\alpha > 1$. We do not consider this approach in this paper, because later we need both the L_2 and L_∞ bounds to hold for the *same* lattice generating vector $\mathbf{z}$ obtained from the criterion $S_N(\mathbf{z})$, rather than two different ones obtained from different criteria.

A lattice $\mathbf{X} = \{\frac{k}{N} \mathbf{z} \bmod 1\}_{k=0}^{N-1}$ is said to have the *reconstructing property* on a frequency index set $\mathcal{B} \subseteq \mathbb{Z}^s$ if and only if

$$\mathbf{h} \cdot \mathbf{z} \not\equiv \mathbf{h}' \cdot \mathbf{z} \pmod{N} \quad \text{for all } \mathbf{h}, \mathbf{h}' \in \mathcal{B} \quad \text{with } \mathbf{h} \neq \mathbf{h}'. \quad (37)$$

In other words, every index has a unique value of dot product with $\mathbf{z}$ modulo N . We say that there are “no aliasing frequencies” in the index set. When this holds, the classical lattice algorithm $A_{\mathcal{B}}^X$ is a projection onto $\text{span}\{e^{2\pi i \mathbf{h} \cdot \mathbf{y}} : \mathbf{h} \in \mathcal{B}\}$, see e.g., [1, Theorem 3.2]. In the following, we will write $\mathcal{B}$ for an index set when we have the reconstruction property and $\mathcal{A}$ for a generic index set.

It is known from [1, Theorem 3.3] that any lattice $\mathbf{X}$ has the reconstructing property on $\{\mathbf{h} \in \mathbb{Z}^s : r(\mathbf{h}) < e^{\text{app}(A^X, H, L_2)^{-2}}\}$, where A^X can be any linear or nonlinear algorithm using $\mathbf{X}$. In particular, [1, Corollary 3.4] applied this result to the algorithm $A_{\mathcal{A}}^X$ satisfying (34), implying that the lattice $\mathbf{X} = \{\frac{k}{N} \mathbf{z} \bmod 1\}_{k=0}^{N-1}$ has the reconstructing property on

$$\mathcal{B} := \left\{ \mathbf{h} \in \mathbb{Z}^s : r(\mathbf{h}) \leq \frac{1}{2\sqrt{S_N(\mathbf{z})}} \right\}, \quad (38)$$

where $S_N(\mathbf{z})$ is defined by (33). This index set $\mathcal{B}$ has half the radius of the index set $\mathcal{A}$ satisfying (34). From (31) with this half radius $M = 1/(2\sqrt{S_N(\mathbf{z})})$, we obtain for the index set (38),

$$e^{\text{app}(A_{\text{ker}}^X, H, L_2)} \leq e^{\text{app}(A_{\mathcal{B}}^X, H, L_2)} < \sqrt{\frac{1.25}{M}} = \sqrt{2.5} (S_N(\mathbf{z}))^{1/4},$$

which differs from (34) only by a constant scaling factor, as expected.

4.2. Subsampled least squares approximation and subsampled kernel interpolant

Next we use a subset of the lattice points $\mathbf{X}_J = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k \in J}$ for $J \subseteq \{0, \dots, N-1\}$ and consider a *least squares approximation* on a frequency index set $B \subseteq \mathbb{Z}^s$,

$$(S_B^{\mathbf{X}_J} F)(\mathbf{y}) := g(\mathbf{y}) = \sum_{\mathbf{h} \in B} \hat{g}_{\mathbf{h}} e^{2\pi i \mathbf{h} \cdot \mathbf{y}}, \quad (39)$$

where the coefficients $[\hat{g}_{\mathbf{h}}]_{\mathbf{h} \in B}$ are obtained by minimizing $\sum_{k \in J} |g(t_k) - F(t_k)|^2$.

Theorem 5. Let $B \subseteq \mathbb{Z}^s$ be a nonempty frequency index set (not necessarily defined by (38)) for which the lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$ has the reconstructing property. For $J \subseteq \{0, \dots, N-1\}$ a multiset with $|J| \geq |B|$, define the matrices

$$\mathbf{L}_{J,B} := [e^{2\pi i k \mathbf{h} \cdot \mathbf{z}/n}]_{k \in J, \mathbf{h} \in B} \quad \text{and} \quad \Phi_{J,B} := \left[\frac{e^{2\pi i k \mathbf{h} \cdot \mathbf{z}/n}}{\sqrt{r(\mathbf{h})}} \right]_{k \in J, \mathbf{h} \in B}. \quad (40)$$

Suppose $\mathbf{L}_{J,B}$ has full rank $|B|$. Then the least squares approximation $S_B^{\mathbf{X}_J} F$ with the subsampled lattice $\mathbf{X}_J = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k \in J}$ is a projection onto $\text{span}\{e^{2\pi i \mathbf{h} \cdot \mathbf{y}} : \mathbf{h} \in B\}$, with $[\hat{g}_{\mathbf{h}}]_{\mathbf{h} \in B} = (\mathbf{L}_{J,B}^* \mathbf{L}_{J,B})^{-1} \mathbf{L}_{J,B}^* [F(t_k)]_{k \in J}$. Moreover, we have

$$e^{\text{app}}(A_{\ker}^{\mathbf{X}_J}, H, L_2) \leq e^{\text{app}}(S_B^{\mathbf{X}_J}, H, L_2) < \left(\sup_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} + \frac{\|\Phi_{J,B}\|_2^2}{\sigma_{\min}^2(\mathbf{L}_{J,B})} \right)^{1/2}, \quad (41)$$

$$e^{\text{app}}(A_{\ker}^{\mathbf{X}_J}, H, L_{\infty}) \leq e^{\text{app}}(S_B^{\mathbf{X}_J}, H, L_{\infty}) < \left(\sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} \right)^{1/2} + \sqrt{|B|} \frac{\|\Phi_{J,B}\|_2}{\sigma_{\min}(\mathbf{L}_{J,B})}, \quad (42)$$

where $\sigma_{\min}(\mathbf{L}_{J,B})$ denotes the smallest singular value of $\mathbf{L}_{J,B}$.

PROOF. The L_2 bound (41) was derived in [1, Lemma 4.1, Lemma 4.2, and Theorem 4.3]. The L_{∞} bound (42) can be obtained analogously to the proof of Theorem 4.

Theorem 6. For a lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$, define the frequency index set $B \subseteq \mathbb{Z}^s$ by (38) (thus $\mathbf{X}$ has the reconstructing property on B). For $t \geq 4$, let $J \subseteq \{0, \dots, N-1\}$ be a multiset of uniformly i.i.d. drawn integers with

$$|J| := \lceil 12|B|(\log|B| + t) \rceil.$$

Then for the subsampled lattice $\mathbf{X}_J = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k \in J}$, with probability exceeding $1 - 3 \exp(-t)$ we have

$$e^{\text{app}}(A_{\ker}^{\mathbf{X}_J}, H, L_2) \leq e^{\text{app}}(S_B^{\mathbf{X}_J}, H, L_2) \leq \left(\frac{7}{|B|} \sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} + 14 \sqrt{S_N(\mathbf{z})} \right)^{1/2}, \quad (43)$$

$$e^{\text{app}}(A_{\ker}^{\mathbf{X}_J}, H, L_{\infty}) \leq e^{\text{app}}(S_B^{\mathbf{X}_J}, H, L_{\infty}) \leq (1 + \sqrt{7}) \left(\sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} \right)^{1/2} + \sqrt{12} \sqrt{|B|} (S_N(\mathbf{z}))^{1/4}. \quad (44)$$

PROOF. The L_2 bound (43) was derived in [1, Theorem 4.7] using [1, Lemma 4.4, Theorem 4.5, and Lemma 4.6]. We now derive the L_{∞} bound (44). From the proof of [1, Theorem 4.7], we conclude with probability exceeding $1 - 3 \exp(-t)$ that

$$\begin{aligned} \frac{\|\Phi_{J,B}\|_2}{\sigma_{\min}(\mathbf{L}_{J,B})} &\leq \left(\frac{7}{|B|} \sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} + 4 \left(\sup_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} + \sqrt{S_N(\mathbf{z})} \right) \right)^{1/2} \\ &\leq \left(\frac{7}{|B|} \sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} + 12 \sqrt{S_N(\mathbf{z})} \right)^{1/2} \leq \frac{\sqrt{7}}{\sqrt{|B|}} \left(\sum_{\mathbf{h} \notin B} \frac{1}{r(\mathbf{h})} \right)^{1/2} + \sqrt{12} (S_N(\mathbf{z}))^{1/4}. \end{aligned}$$

Substituting this bound into (42) yields (44) as required.

Remark 1. The random techniques used in Theorem 6 enforce a logarithmic oversampling, i.e., the number of points $|J|$ is larger than the number of frequencies $|B|$ by a logarithmic factor. This logarithmic gap was decreased in [37] in a non-constructive way and later in [2] in a constructive way. Shortly after it was shown that the logarithmic gap can be eliminated altogether in [15], again in a non-constructive way. Recently, this was also made constructive in [6]. Note that all mentioned constructive approaches rely on the so-called BSS algorithm, which has arithmetic complexity of $\mathcal{O}(|B|^3 N)$. For this reason we favor the random approach and concede the loss in error decay.

4.3. Specializing to weighted Korobov space

We now make use of the specific definition of $r(\mathbf{h}) = r_{s,\alpha,\gamma}(\mathbf{h})$ in (13) for the weighted Korobov space and apply known results from [34] for all the required quantities. Note that the smoothness parameter α in [34] corresponds to our 2α here.

Given $N \geq 2$, $s \in \mathbb{N}$, $\alpha > 1/2$ and weights $\gamma = \{\gamma_{\mathbf{u}}\}_{\mathbf{u} \subseteq \mathbb{N}}$, the generating vector $\mathbf{z} \in \mathbb{Z}^s$ obtained from the CBC construction following [34, Algorithm 3.2] satisfies [34, Theorem 3.3]:

$$\frac{2\zeta(2\alpha)\gamma_{\{1\}}}{N^{2\alpha}} \leq S_N(\mathbf{z}) \leq \left(\frac{\Phi_{1,\lambda}}{\varphi(N)} \right)^{\frac{1}{\lambda}} \quad \text{for all } \lambda \in \left(\frac{1}{2\alpha}, 1 \right], \quad (45)$$

where $\varphi(N) = |\{z \in \{1, \dots, N\} : \gcd(z, N) = 1\}|$ is the Euler totient function (e.g., $\varphi(N) = N - 1$ for N prime), $\zeta(x) = \sum_{h=1}^{\infty} h^{-x}$ is the Riemann zeta function, and

$$\Phi_{1,\lambda} := (2^{4\alpha\lambda+1} + 1) \left(\sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} |\mathbf{u}| \gamma_{\mathbf{u}}^{\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right) \left(\sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right). \quad (46)$$

For an index set defined by $\mathcal{A} := \{\mathbf{h} \in \mathbb{Z}^s : r_{s,\alpha,\gamma}(\mathbf{h}) \leq M_{\mathcal{A}}\}$ with $M_{\mathcal{A}} \geq 1$, it follows from [34, Lemma 5.2] and [34, Lemma 5.3] that for all $\lambda \in (\frac{1}{2\alpha}, 1)$ we have

$$(\gamma_{\{1\}} M_{\mathcal{A}})^{\frac{1}{2\alpha}} \leq |\mathcal{A}| \leq (\Phi_{2,\lambda} M_{\mathcal{A}})^{\lambda}, \quad \Phi_{2,\lambda} := \left(\sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right)^{\frac{1}{\lambda}}, \quad (47)$$

$$\sum_{\mathbf{h} \notin \mathcal{A}} \frac{1}{r_{s,\alpha,\gamma}(\mathbf{h})} \leq \left(\frac{\Phi_{3,\lambda}}{M_{\mathcal{A}}} \right)^{\frac{1-\lambda}{2\alpha\lambda}}, \quad \Phi_{3,\lambda} := \frac{1}{\gamma_{\{1\}}^{\lambda}} \left(\left(\frac{\lambda}{1-\lambda} \right)^{\lambda} \sum_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{\lambda} (2\zeta(2\alpha\lambda))^{|u|} \right)^{\frac{2\alpha}{1-\lambda}}. \quad (48)$$

We combine all results in a unified theorem for one single lattice generating vector $\mathbf{z} \in \mathbb{Z}^s$. For the full lattice in L_2 and L_{∞} approximations we recover the convergence rates from [34]. For the subsampled lattice in L_2 approximation we recover the convergence rate from [1]. The constants here may differ from [34, 1] due to our need to ensure that all bounds hold for the same $\mathbf{z}$. The results for the subsampled lattice in L_{∞} approximation, and all results for L_p approximation for $p \in (2, \infty)$ are new.

In comparison, it is known [32] that the optimal polynomial rate for any algorithm with N function evaluations is $N^{-\alpha}$ for L_2 approximation and $N^{-(\alpha-1/2)}$ for L_{∞} approximation.

Theorem 7. Let $H_{s,\alpha,\gamma}^{\text{Kor}}$ be the weighted Korobov space defined in Section 2.1 with $\alpha > 1/2$ and positive weights $\gamma = (\gamma_{\mathbf{u}})_{\mathbf{u} \subseteq \mathbb{N}}$. For $N \geq 7$ sufficiently large, obtain $\mathbf{z} \in \mathbb{Z}^s$ by the component-by-component construction in [34, Algorithm 3.2] so that $S_N(\mathbf{z}) = S_{s,\alpha,\gamma,N}(\mathbf{z}) \leq 1/4$ satisfies (45). Define

$$\mathcal{A} := \left\{ \mathbf{h} \in \mathbb{Z}^s : r_{s,\alpha,\gamma}(\mathbf{h}) \leq \frac{1}{\sqrt{S_N(\mathbf{z})}} \right\} \quad \text{and} \quad \mathcal{B} := \left\{ \mathbf{h} \in \mathbb{Z}^s : r_{s,\alpha,\gamma}(\mathbf{h}) \leq \frac{1}{2\sqrt{S_N(\mathbf{z})}} \right\}. \quad (49)$$

Then, for all $p \in [2, \infty]$ and arbitrarily small $\epsilon > 0$, we have the following results:

(a) The full lattice $\mathbf{X} = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k=0}^{N-1}$ satisfies

$$e^{\text{app}}(A_{\text{ker}}^{\mathbf{X}}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq e^{\text{app}}(A_{\mathcal{A}}^{\mathbf{X}}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq C_{p,\epsilon}^{\text{full}} \left(\frac{1}{\varphi(N)} \right)^{\theta_{p,\epsilon}^{\text{full}}},$$

$$\theta_{p,\epsilon}^{\text{full}} = \begin{cases} \frac{1}{2}(\alpha - \epsilon) & \text{if } p = 2, \\ \frac{1}{2}(\alpha - \frac{1}{2} - \epsilon)(1 - \frac{\epsilon}{\alpha}) & \text{if } p = \infty, \\ \frac{1}{2}(\alpha - \frac{1}{2} + \frac{1}{p} - \epsilon(1 - \frac{2}{p}))(1 - \frac{\epsilon}{\alpha}) & \text{if } p \in [2, \infty]. \end{cases} \quad (50)$$

(b) For $t \geq 4$, let $J \subseteq \{0, \dots, N-1\}$ be a multiset of uniformly i.i.d. drawn integers with

$$|J| := \lceil 12 |B| (\log |B| + t) \rceil.$$

The subsampled lattice $\mathbf{X}_J = \{\frac{k}{N}\mathbf{z} \bmod 1\}_{k \in J}$ satisfies, with probability exceeding $1 - 3 \exp(-t)$,

$$e^{\text{app}}(A_{\text{ker}}^{\mathbf{X}_J}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq e^{\text{app}}(S_B^{\mathbf{X}_J}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq C_{p,\epsilon}^{\text{sub}} \left(\frac{\log |J| + t}{|J|} \right)^{\theta_{p,\epsilon}^{\text{sub}}},$$

$$\theta_{p,\epsilon}^{\text{sub}} = \begin{cases} (\alpha - \epsilon)(1 - \frac{\epsilon}{\alpha}) & \text{if } p = 2, \\ (\alpha - \frac{1}{2} - \epsilon)(1 - \frac{\epsilon}{\alpha}) & \text{if } p = \infty, \\ (\alpha - \frac{1}{2} + \frac{1}{p} - \epsilon)(1 - \frac{\epsilon}{\alpha}) & \text{if } p \in [2, \infty]. \end{cases} \quad (51)$$

If N is a prime power, then we have

$$N^{\frac{1-\epsilon/\alpha}{2}} \log N \lesssim |J| \lesssim N^{\frac{1}{2(1-\epsilon/\alpha)}} \log N \quad \text{and} \quad \left(\frac{|J|}{\log |J|} \right)^{2(1-\epsilon/\alpha)} \lesssim N \lesssim \left(\frac{|J|}{\log |J|} \right)^{\frac{2}{1-\epsilon/\alpha}}.$$

All the constants $C_{p,\epsilon}^{\text{full}}$ and $C_{p,\epsilon}^{\text{sub}}$ are bounded independently of s if the weights γ satisfy

$$\sum_{\mathbf{u} \in \mathbb{N}_s, |\mathbf{u}| < \infty} \gamma_{\mathbf{u}}^{\frac{1}{2(\alpha-\epsilon)}} \left[2 e^{1/e} \zeta\left(\frac{2\alpha}{2(\alpha-\epsilon)}\right) \right]^{|\mathbf{u}|} < \infty, \quad (52)$$

and the constants go to infinity as $\epsilon \rightarrow 0$.

PROOF. For both the full lattice and the subsampled lattice, we consider the cases $p = 2$ and $p = \infty$ separately and then use interpolation to obtain the result for general $p \in [2, \infty]$. In each case below, for the parameter λ in (45)–(48) we substitute $\lambda = \frac{1}{2(\alpha-\epsilon)}$ with arbitrarily small $\epsilon > 0$.

- FULL LATTICE WITH $p = 2$. We substitute the upper bound in (45) into (34) to obtain (50), with

$$\sqrt{2} \left(\frac{\Phi_{1,\lambda}}{\varphi(N)} \right)^{\frac{1}{4\lambda}} =: C_{2,\epsilon}^{\text{full}} \left(\frac{1}{\varphi(N)} \right)^{\theta_{2,\epsilon}^{\text{full}}}, \quad \theta_{2,\epsilon}^{\text{full}} := \frac{1}{4\lambda} = \frac{1}{2}(\alpha - \epsilon).$$

- FULL LATTICE WITH $p = \infty$. We use $M_{\mathcal{A}} = 1/\sqrt{S_N(\mathbf{z})} \geq 2$, substitute (48) and the upper bound in (47) into (32), and apply the upper bound in (45) to obtain (50), with

$$\begin{aligned} \left(\frac{\Phi_{3,\lambda}}{M_{\mathcal{A}}} \right)^{\frac{1-\lambda}{4\alpha\lambda}} + (\Phi_{2,\lambda} M_{\mathcal{A}})^{\frac{\lambda}{2}} \left(\frac{M_{\mathcal{A}}}{M_{\mathcal{A}}^2} \right)^{\frac{1}{2}} &\leq (\Phi_{3,\lambda}^{\frac{1-\lambda}{4\alpha\lambda}} + \Phi_{2,\lambda}^{\frac{\lambda}{2}}) \left(\frac{1}{M_{\mathcal{A}}} \right)^{\frac{1-\lambda}{4\alpha\lambda}} \\ &\leq (\Phi_{3,\lambda}^{\frac{1-\lambda}{4\alpha\lambda}} + \Phi_{2,\lambda}^{\frac{\lambda}{2}}) \left(\frac{\Phi_{1,\lambda}}{\varphi(N)} \right)^{\frac{1}{2\lambda} \frac{1-\lambda}{4\alpha\lambda}} =: C_{\infty,\epsilon}^{\text{full}} \left(\frac{1}{\varphi(N)} \right)^{\theta_{\infty,\epsilon}^{\text{full}}}, \end{aligned}$$

where $\theta_{\infty,\epsilon}^{\text{full}} := \frac{1-\lambda}{8\alpha\lambda^2} = \frac{1}{2}(\alpha - \frac{1}{2} - \epsilon)(1 - \frac{\epsilon}{\alpha})$.

- FULL LATTICE WITH $p \in [2, \infty]$. We use the interpolation bound $\|g\|_{L_p} \leq \|g\|_{L_2}^{\frac{2}{p}} \|g\|_{L_{\infty}}^{1-\frac{2}{p}}$ to conclude that $C_{p,\epsilon}^{\text{full}} := (C_{2,\epsilon}^{\text{full}})^{\frac{2}{p}} (C_{\infty,\epsilon}^{\text{full}})^{1-\frac{2}{p}}$, and

$$\begin{aligned} \theta_{p,\epsilon}^{\text{full}} &:= \theta_{2,\epsilon}^{\text{full}} \left(\frac{2}{p} \right) + \theta_{\infty,\epsilon}^{\text{full}} \left(1 - \frac{2}{p} \right) = \frac{1}{2}(\alpha - \epsilon) \frac{2}{p} + \frac{1}{2}(\alpha - \frac{1}{2} - \epsilon)(1 - \frac{\epsilon}{\alpha}) \left(1 - \frac{2}{p} \right) \\ &= \frac{1}{2}(\alpha - \frac{1}{2} + \frac{1}{p} - \epsilon(1 - \frac{2}{p}))(1 - \frac{\epsilon}{\alpha}). \end{aligned}$$

- SUBSAMPLED LATTICE WITH $p = 2$. For the set $\mathcal{B}$ we have from (47) and [1, Theorem 4.5] that

$$\frac{\gamma_{\{1\}}}{|\mathcal{B}|^{2\alpha}} \leq \frac{1}{M_B} \leq \Phi_{2,\lambda} \left(\frac{1}{|\mathcal{B}|} \right)^{\frac{1}{\lambda}} \quad \text{and} \quad \frac{6(\log |J| + t)}{|J|} \leq \frac{1}{|\mathcal{B}|} \leq \frac{13(\log |J| + t)}{|J|}. \quad (53)$$

We use $M_B = 1/(2\sqrt{S_N(\mathbf{z})}) \geq 1$, substitute (48) and the lower bound in (47) into (43), and apply the upper bounds in (53) to obtain (51), with

$$\begin{aligned} & \left(\frac{7}{(\gamma_{\{1\}} M_B)^{\frac{1}{2\alpha}}} \left(\frac{\Phi_{3,\lambda}}{M_B} \right)^{\frac{1-\lambda}{2\alpha\lambda}} + \frac{7}{M_B} \right)^{\frac{1}{2}} \leq \left(\frac{7\Phi_{3,\lambda}^{\frac{1-\lambda}{2\alpha\lambda}}}{\gamma_{\{1\}}^{\frac{1}{2\alpha}}} + 7 \right)^{\frac{1}{2}} \left(\frac{1}{M_B} \right)^{\frac{1}{4\alpha\lambda}} \\ & \leq \left(\frac{7\Phi_{3,\lambda}^{\frac{1-\lambda}{2\alpha\lambda}}}{\gamma_{\{1\}}^{\frac{1}{2\alpha}}} + 7 \right)^{\frac{1}{2}} \Phi_{2,\lambda}^{\frac{1}{4\alpha\lambda}} 13^{\frac{1}{4\alpha\lambda^2}} \left(\frac{\log |J| + t}{|J|} \right)^{\frac{1}{4\alpha\lambda^2}} =: C_{2,\epsilon}^{\text{sub}} \left(\frac{\log |J| + t}{|J|} \right)^{\theta_{2,\epsilon}^{\text{sub}}}, \end{aligned}$$

where $\theta_{2,\epsilon}^{\text{sub}} := \frac{1}{4\alpha\lambda^2} = (\alpha - \epsilon)(1 - \frac{\epsilon}{\alpha})$.

- SUBSAMPLED LATTICE WITH $p = \infty$. We use $M_B = 1/(2\sqrt{S_N(\mathbf{z})}) \geq 1$, substitute (48) into (44), and apply again the upper bounds in (53) to obtain (51), with

$$\begin{aligned} & (1 + \sqrt{7}) \left(\frac{\Phi_{3,\lambda}}{M_B} \right)^{\frac{1-\lambda}{4\alpha\lambda}} + \sqrt{12} \sqrt{|\mathcal{B}|} \left(\frac{1}{2M_B} \right)^{\frac{1}{2}} \leq \left((1 + \sqrt{7}) \Phi_{3,\lambda}^{\frac{1-\lambda}{4\alpha\lambda}} \Phi_{2,\lambda}^{\frac{1-\lambda}{4\alpha\lambda^2}} + \sqrt{6} \Phi_{2,\lambda}^{\frac{1}{2\lambda}} \right) \left(\frac{1}{|\mathcal{B}|} \right)^{\frac{1-\lambda}{4\alpha\lambda^2}} \\ & \leq \left((1 + \sqrt{7}) \Phi_{3,\lambda}^{\frac{1-\lambda}{4\alpha\lambda}} \Phi_{2,\lambda}^{\frac{1-\lambda}{4\alpha\lambda^2}} + \sqrt{6} \Phi_{2,\lambda}^{\frac{1}{2\lambda}} \right) 13^{\frac{1-\lambda}{4\alpha\lambda^2}} \left(\frac{\log |J| + t}{|J|} \right)^{\frac{1-\lambda}{4\alpha\lambda^2}} =: C_{\infty,\epsilon}^{\text{sub}} \left(\frac{\log |J| + t}{|J|} \right)^{\theta_{\infty,\epsilon}^{\text{sub}}}, \end{aligned}$$

where $\theta_{\infty,\epsilon}^{\text{sub}} := \frac{1-\lambda}{4\alpha\lambda^2} = (\alpha - \frac{1}{2} - \epsilon)(1 - \frac{\epsilon}{\alpha})$.

- SUBSAMPLED LATTICE WITH $p \in [2, \infty]$. Again we use interpolation to conclude that $C_{p,\epsilon}^{\text{sub}} := (C_{2,\epsilon}^{\text{sub}})^{\frac{2}{p}} (C_{\infty,\epsilon}^{\text{sub}})^{1-\frac{2}{p}}$, and

$$\begin{aligned} \theta_{p,\epsilon}^{\text{sub}} &:= \theta_{2,\epsilon}^{\text{sub}} \left(\frac{2}{p} \right) + \theta_{\infty,\epsilon}^{\text{sub}} \left(1 - \frac{2}{p} \right) = (\alpha - \epsilon) \left(1 - \frac{\epsilon}{\alpha} \right) \frac{2}{p} + \left(\alpha - \frac{1}{2} - \epsilon \right) \left(1 - \frac{\epsilon}{\alpha} \right) \left(1 - \frac{2}{p} \right) \\ &= \left(\alpha - \frac{1}{2} + \frac{1}{p} - \epsilon \right) \left(1 - \frac{\epsilon}{\alpha} \right). \end{aligned}$$

From the definition of J and the upper and lower bounds in (47) and (45), we have the following connection between $|J|$ and N :

$$\begin{aligned} & 12 |\mathcal{B}| (\log |\mathcal{B}| + t) \leq |J| \leq 13 |\mathcal{B}| (\log |\mathcal{B}| + t), \\ & \left(\frac{\varphi(N)}{\Phi_{1,\lambda}} \right)^{\frac{1}{4\alpha\lambda}} \left(\frac{\gamma_{\{1\}}}{2} \right)^{\frac{1}{2\alpha}} \leq \left(\frac{\gamma_{\{1\}}}{2\sqrt{S_N(\mathbf{z})}} \right)^{\frac{1}{2\alpha}} \leq |\mathcal{B}| \leq \left(\frac{\Phi_{2,\lambda}}{2\sqrt{S_N(\mathbf{z})}} \right)^{\lambda} \leq \frac{\Phi_{2,\lambda}^{\lambda} N^{\alpha\lambda}}{(8\zeta(2\alpha)\gamma_{\{1\}})^{\frac{\lambda}{2}}}. \end{aligned}$$

This yields

$$(\varphi(N))^{\frac{1}{4\alpha\lambda}} \log \varphi(N) \lesssim |J| \lesssim N^{\alpha\lambda} \log N, \quad \left(\frac{|J|}{\log |J|} \right)^{\frac{1}{\alpha\lambda}} \lesssim N \quad \text{and} \quad \varphi(N) \lesssim \left(\frac{|J|}{\log |J|} \right)^{4\alpha\lambda}.$$

If $N = q^m$ is a prime power, then $\varphi(N) = q^m(1 - \frac{1}{q}) \geq \frac{N}{2}$, so $\frac{1}{N} \leq \frac{1}{\varphi(N)} \leq \frac{2}{N}$, which simplifies the connection between $|J|$ and N .

The constants $\Phi_{1,\lambda}$, $\Phi_{2,\lambda}$, $\Phi_{3,\lambda}$ all depend on s , α , γ and $\lambda = \frac{1}{2(\alpha-\epsilon)}$ with arbitrarily small $\epsilon > 0$. Consequently, the constants $C_{2,\epsilon}^{\text{full}}$, $C_{\infty,\epsilon}^{\text{sub}}$, $C_{2,\epsilon}^{\text{sub}}$, $C_{\infty,\epsilon}^{\text{sub}}$ all depend on s , α , γ and ϵ . For $p \in (2, \infty)$, the constants $C_{p,\epsilon}^{\text{full}}$, $C_{p,\epsilon}^{\text{sub}}$ all depend on s , α , γ , ϵ , and depend trivially on p . To ensure that all constants are bounded independently of s , we use $\max(1, |\mathbf{u}|) \leq (e^{1/\epsilon})^{|\mathbf{u}|}$ and then assume that the weights γ satisfy (52). This completes the proof.

The estimate $(\log |J| + t) \leq c_\delta |J|^\delta$ holds for $c_\delta > 0$ and arbitrarily small $\delta > 0$ and $|J|$ sufficiently large. This can be used to simplify (51) into

$$e^{\text{app}}(A_{\text{ker}}^{X_J}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq e^{\text{app}}(S_B^{X_J}, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \leq c_\delta C_{p,\epsilon}^{\text{sub}} \left(\frac{1}{|J|} \right)^{\theta_{p,\epsilon}^{\text{sub}}(1-\delta)}. \quad (54)$$

5. Error analysis

The error for our method (1)–(3) with the full lattice $\mathbf{X}$ is bounded as follows:

$$\begin{aligned} |I(f) - Q_\psi^X(f)| &= \left| \int_{[0,1]^s} (F(\mathbf{y}) - (A_{\text{ker}}^X F)(\mathbf{y})) \mu(\mathbf{y}) d\mathbf{y} \right| \\ &\leq \inf_{p \in [2, \infty]} \left\{ \|F - A_{\text{ker}}^X F\|_{L_p([0,1]^s)} \|\mu\|_{L_{p'}([0,1]^s)} \right\} \\ &\leq \inf_{p \in [2, \infty]} \left\{ e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \|F\|_{\text{Kor},s,\alpha,\gamma} \|\mu\|_{L_{p'}([0,1]^s)} \right\} \\ &\leq \inf_{p \in [2, \infty]} \left\{ e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \|\mu\|_{L_{p'}([0,1]^s)} \|f\|_{\text{Sob},s,\alpha,\beta,\omega} \right\}, \end{aligned} \quad (55)$$

where we first use Hölder's inequality with the conjugate pair $1/p + 1/p' = 1$, and then apply the worst case L_p approximation error bound assuming that $F \in H_{s,\alpha,\gamma}^{\text{Kor}}$, and finally we connect the Korobov norm of the transformed F to the Sobolev norm of the original integrand f with rescaled weights β and a weight function ω , see Lemma 3. An analogous bound holds when the full lattice $\mathbf{X}$ is replaced by a subsampled lattice $\mathbf{X}_J$.

With $1/p + 1/p' = 1$, we have $\|\mu\|_{L_{p'}([0,1]^s)} = \xi_p^s$, where

$$\xi_p := \frac{\pi}{2} \left(\int_0^1 |\sin(2\pi y)|^{p'} dy \right)^{1/p'} = \frac{\pi}{2} \left(\frac{\Gamma(\frac{p'+1}{2})}{\sqrt{\pi} \Gamma(\frac{p'+2}{2})} \right)^{1/p'} = \begin{cases} \frac{\pi}{\sqrt{8}} \leq 1.110721 & \text{if } p = 2, \\ 1 & \text{if } p = \infty, \end{cases} \quad (56)$$

and $\Gamma(\cdot)$ is the Gamma function. For example, we have also $\xi_{2.5} \approx 1.07999$, $\xi_3 \approx 1.062643$, $\xi_{10} \approx 1.01554$, $\xi_{20} \approx 1.0075$.

5.1. Main theorem

We can now substitute the L_p approximation error bounds from Theorem 7 into (55) to arrive at our main theorem.

Theorem 8. *Let the Korobov space $H_{s,\alpha,\gamma}^{\text{Kor}}$ and Sobolev space $H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ with integer $\alpha \geq 1$ be related by their weights as defined in (21). For $N \geq 7$ sufficiently large, obtain $\mathbf{z} \in \mathbb{Z}^s$ by the component-by-component construction in [34, Algorithm 3.2] so that $S_N(\mathbf{z}) = S_{s,\alpha,\gamma,N}(\mathbf{z}) \leq 1/4$ satisfies (45). Define the frequency index set $\mathcal{B}$ as in (49). Then, for all $f \in H_{s,\alpha,\beta,\omega}^{\text{Sob}}$ and arbitrarily small $\epsilon > 0$, we have the following results:*

(a) *The method (1)–(3) with the full lattice $\mathbf{X} = \{ \frac{k}{N} \mathbf{z} \bmod 1 \}_{k=0}^{N-1}$ satisfies*

$$\begin{aligned} |I(f) - Q_\psi^X(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_{p,\epsilon}^{\text{full}} \left(\frac{1}{\varphi(N)} \right)^{\theta_{p,\epsilon}^{\text{full}}} \xi_p^s \|f\|_{\text{Sob},s,\alpha,\beta,\omega} \right\}, \\ \theta_{p,\epsilon}^{\text{full}} &\rightarrow \frac{\alpha}{2} - \frac{1}{4} + \frac{1}{2p} \quad \text{as } \epsilon \rightarrow 0. \end{aligned} \quad (57)$$

(b) *For $t \geq 4$, let $J \subseteq \{0, \dots, N-1\}$ be a multiset of uniformly i.i.d. drawn integers with $|J| := \lceil 12 |\mathcal{B}| (\log |\mathcal{B}| + t) \rceil$. Then the method (1)–(3) with the subsampled lattice $\mathbf{X}_J = \{ \frac{k}{N} \mathbf{z} \bmod 1 \}_{k \in J}$ satisfies, with probability exceeding $1 - 3 \exp(-t)$,*

$$\begin{aligned} |I(f) - Q_\psi^{X_J}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_{p,\epsilon}^{\text{sub}} \left(\frac{\log |J| + t}{|J|} \right)^{\theta_{p,\epsilon}^{\text{sub}}} \xi_p^s \|f\|_{\text{Sob},s,\alpha,\beta,\omega} \right\}, \\ \theta_{p,\epsilon}^{\text{sub}} &\rightarrow \alpha - \frac{1}{2} + \frac{1}{p} \quad \text{as } \epsilon \rightarrow 0. \end{aligned} \quad (58)$$

All the constants $C_{p,\epsilon}^{\text{full}}$ and $C_{p,\epsilon}^{\text{sub}}$ are bounded independently of s if the weights γ satisfy (52).

Hence, for the case of the subsampled lattice, at the two extremes we have close to the rate $|J|^{-(\alpha-1/2)}$ for $p = \infty$ and $|J|^{-\alpha} (1.110721)^s$ for $p = 2$. Our result holds for the infimum over all values of $p \in [2, \infty]$, with a trade-off between the polynomial rate in $|J|$ and the exponential dependence on s . We stress that our method does not depend on the value of p and we do not need to specify a value of p for the error bound to hold. We will automatically get the smallest error bound over all possible values of p .

5.2. Doubling the rate

In [46] a doubling of the rate effect was proven when the function has *twice* the smoothness of the Hilbert space corresponding to the kernel, with no need to change the points or the method. Independently, this effect was investigated in [30, 53] with the focus on classical Sobolev spaces, where also intermediate spaces (i.e., smoothness increased by a factor between 1 and 2) were investigated and the sufficient condition for the increased rate effect was proven to be also necessary. Specializing to the weighted Korobov space which is embedded in $L_2([0, 1]^s)$, the crucial property is, for $\sigma \in [1, 2]$, $F \in H_{s,\sigma\alpha,\gamma}^{\text{Kor}}$, and $G \in H_{s,(2-\sigma)\alpha,\gamma^{2-\sigma}}^{\text{Kor}}$,

$$\begin{aligned} |\langle F, G \rangle_{\text{Kor},s,\alpha,\gamma}| &= \sum_{h \in \mathbb{Z}^s} (r_{s,\alpha,\gamma}(h))^{\sigma/2} |\hat{F}_h| (r_{s,\alpha,\gamma}(h))^{(2-\sigma)/2} |\hat{G}_h| \\ &= \sum_{h \in \mathbb{Z}^s} (r_{s,\sigma\alpha,\gamma^\sigma}(h))^{1/2} |\hat{F}_h| (r_{s,(2-\sigma)\alpha,\gamma^{2-\sigma}}(h))^{1/2} |\hat{G}_h| \\ &\leq \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma} \|G\|_{\text{Kor},s,(2-\sigma)\alpha,\gamma^{2-\sigma}}. \end{aligned} \quad (59)$$

In the following theorem we state the doubling of the rate effect and extend it on the one hand to intermediate spaces $\sigma \in [1, 2]$ using techniques from [53], and on the other hand to general L_p approximation for $p \in [1, \infty]$. Note that changing weights in a weighted Korobov space does not change the membership of the space; it only changes the size of the unit ball.

The advantage compared with replacing the kernel of smoothness α by the kernel of smoothness $\sigma\alpha$ is that a lower smoothness method is often more numerically stable and that lower order kernels are easier to compute.

Theorem 9. Let A_{ker}^X be defined by the kernel of the weighted Korobov space $H_{s,\alpha,\gamma}^{\text{Kor}}$ with smoothness α and weights γ . For all $p \in [1, \infty]$, $\sigma \in [1, 2]$, and $F \in H_{s,\sigma\alpha,\gamma^\sigma}^{\text{Kor}}$ we have

$$\|F - A_{\text{ker}}^X F\|_{L_p([0,1]^s)} \leq e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \left(e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_2) \right)^{\sigma-1} \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma}. \quad (60)$$

Equivalently, for all $p \in [1, \infty]$ and $\sigma \in [1, 2]$ we have

$$e^{\text{app}}(A_{\text{ker}}^X, H_{s,\sigma\alpha,\gamma^\sigma}^{\text{Kor}}, L_p) \leq e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \left(e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_2) \right)^{\sigma-1}.$$

Analogous results hold if the full lattice X is replaced by a subsampled lattice X_J .

PROOF. Throughout the proof let $F \in H_{s,\sigma\alpha,\gamma^\sigma}^{\text{Kor}}$ and $G \in H_{s,(2-\sigma)\alpha,\gamma^{2-\sigma}}^{\text{Kor}}$. Using Hölder's inequality with parameters $1/(2-\sigma)$ and $1/(\sigma-1)$ we have

$$\begin{aligned} \|G\|_{\text{Kor},s,(2-\sigma)\alpha,\gamma^{2-\sigma}} &= \left(\sum_{h \in \mathbb{Z}^s} (r_{s,\alpha,\gamma}(h))^{2-\sigma} |\hat{G}_h|^{4-2\sigma} |\hat{G}_h|^{2\sigma-2} \right)^{1/2} \\ &\leq \left(\left(\sum_{h \in \mathbb{Z}^s} r_{s,\alpha,\gamma}(h) |\hat{G}_h|^2 \right)^{2-\sigma} \left(\sum_{h \in \mathbb{Z}^s} |\hat{G}_h|^2 \right)^{\sigma-1} \right)^{1/2} \\ &= \|G\|_{\text{Kor},s,\alpha,\gamma}^{2-\sigma} \|G\|_{L_2([0,1]^s)}^{\sigma-1}. \end{aligned} \quad (61)$$

Now we use the fact that A_{ker}^X is an orthogonal projection from $H_{s,\alpha,\gamma}^{\text{Kor}}$ onto $\text{span}\{K_{s,\alpha,\gamma}^{\text{Kor}}(t_k, \cdot) : k = 0, \dots, N-1\}$, which follows from the reproducing property of the kernel and the interpolation property. Using also (59) and (61), we then obtain

$$\begin{aligned} \|F - A_{\text{ker}}^X F\|_{\text{Kor},s,\alpha,\gamma}^2 &= \langle F, F - A_{\text{ker}}^X F \rangle_{\text{Kor},s,\alpha,\gamma} \\ &\leq \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma} \|F - A_{\text{ker}}^X F\|_{\text{Kor},s,(2-\sigma)\alpha,\gamma^{2-\sigma}} \\ &\leq \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma} \|F - A_{\text{ker}}^X F\|_{\text{Kor},s,\alpha,\gamma}^{2-\sigma} \|F - A_{\text{ker}}^X F\|_{L_2([0,1]^s)}^{\sigma-1}. \end{aligned}$$

Canceling the common factor on both sides gives

$$\|F - A_{\text{ker}}^X F\|_{\text{Kor},s,\alpha,\gamma} \leq \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma}^{1/\sigma} \|F - A_{\text{ker}}^X F\|_{L_2([0,1]^s)}^{(\sigma-1)/\sigma}.$$

For all $p \in [1, \infty]$ this yields

$$\begin{aligned} \|F - A_{\text{ker}}^X F\|_{L_p([0,1]^s)} &= \|(F - A_{\text{ker}}^X F) - A_{\text{ker}}^X (F - A_{\text{ker}}^X F)\|_{L_p([0,1]^s)} \\ &\leq e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \|F - A_{\text{ker}}^X F\|_{\text{Kor},s,\alpha,\gamma} \\ &\leq e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma}^{1/\sigma} \|F - A_{\text{ker}}^X F\|_{L_2([0,1]^s)}^{(\sigma-1)/\sigma}. \end{aligned} \quad (62)$$

Taking $p = 2$ in (62) and canceling a common factor gives $\|F - A_{\text{ker}}^X F\|_{L_2([0,1]^s)}^{1/\sigma} \leq e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_2) \|F\|_{\text{Kor},s,\sigma\alpha,\gamma^\sigma}^{1/\sigma}$. Substituting this back into (62) for general p gives (60).

We can now extend Theorem 8 to the case where f is smoother compared to the reproducing kernel used in the method. For the sake of simplicity we only state the result with double the smoothness, i.e., $\sigma = 2$.

Corollary 10. Consider the Korobov space $H_{s,\alpha,\gamma}^{\text{Kor}}$ and Sobolev space $H_{s,2\alpha,\eta^2,\omega}^{\text{Sob}}$ with integer $\alpha \geq 1$ and weights related by

$$\gamma_u^2 := c_{2\alpha}^{|\mathbf{u}|} \eta_u^2, \quad \text{with} \quad c_{2\alpha} = \sum_{\lambda=1}^{2\alpha} S(2\alpha, \lambda)^2.$$

Let A_{ker}^X be defined by the kernel of $H_{s,\alpha,\gamma}^{\text{Kor}}$. Suppose that $f \in H_{s,2\alpha,\eta^2,\omega}^{\text{Sob}}$. Under the same remaining setup as Theorem 8, the error bound (57) for the full lattice is replaced by

$$\begin{aligned} |I(f) - Q_{\psi}^X(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_{p,\epsilon}^{\text{full}} C_{2,\epsilon}^{\text{full}} \left(\frac{1}{\varphi(N)} \right)^{\theta_{p,\epsilon}^{\text{full}} + \theta_{2,\epsilon}^{\text{full}}} \xi_p^{\epsilon_s} \right\} \|f\|_{\text{Sob},s,2\alpha,\eta^2,\omega}, \\ \theta_{p,\epsilon}^{\text{full}} + \theta_{2,\epsilon}^{\text{full}} &\rightarrow \alpha - \frac{1}{4} + \frac{1}{2p} \quad \text{as} \quad \epsilon \rightarrow 0, \end{aligned} \quad (63)$$

while the error bound (58) for the subsampled lattice is replaced by

$$\begin{aligned} |I(f) - Q_{\psi}^{X_J}(f)| &\leq \inf_{p \in [2, \infty]} \left\{ C_{p,\epsilon}^{\text{sub}} C_{2,\epsilon}^{\text{sub}} \left(\frac{\log |J| + t}{|J|} \right)^{\theta_{p,\epsilon}^{\text{sub}} + \theta_{2,\epsilon}^{\text{sub}}} \xi_p^{\epsilon_s} \right\} \|f\|_{\text{Sob},s,2\alpha,\eta^2,\omega}, \\ \theta_{p,\epsilon}^{\text{sub}} + \theta_{2,\epsilon}^{\text{sub}} &\rightarrow 2\alpha - \frac{1}{2} + \frac{1}{p} \quad \text{as} \quad \epsilon \rightarrow 0. \end{aligned} \quad (64)$$

PROOF. Applying (60) with $\sigma = 2$ after the first inequality in (55) gives

$$|I(f) - Q_{\psi}^X(f)| \leq \inf_{p \in [2, \infty]} \left\{ e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_p) e^{\text{app}}(A_{\text{ker}}^X, H_{s,\alpha,\gamma}^{\text{Kor}}, L_2) \xi_p^{\epsilon_s} \right\} \|F\|_{\text{Kor},s,2\alpha,\gamma^2}.$$

Using Lemma 3 with α and γ_u replaced by 2α and γ_u^2 , respectively, we obtain $\|F\|_{\text{Kor},s,2\alpha,\gamma^2} \leq \|f\|_{\text{Sob},s,2\alpha,\eta^2,\omega}$, where the sequence $\eta = (\eta_u)_{u \in \mathbb{N}}$ is defined by $\gamma_u^2 := c_{2\alpha}^{|\mathbf{u}|} \eta_u^2$. The result now follows by applying the L_p approximation error bounds from Theorem 7.

s	$I(f)$ with $q = 2.5$	$I(f)$ with $q = 6$
2	1.1032908827	1.0986484606
4	1.1040739938	1.0986487483
64	1.1041644920	1.0986487490

Table 4

Computed reference value for integral of reciprocal function (66).

6. Numerical experiments

In this section we present results for the weighted quadrature rule for both the full and subsampled lattice for a range of values $\alpha \in \{1, 2, 3\}$ in dimensions up to $s = 64$. The values of N tested are prime numbers (roughly doubling at each step) in the range

$$N \in \{131, 257, 521, 1031, 2053, 4099, 8209, 16411, 32771, 65537, 131101, 262147, 524309\}$$

For the full lattice, the weights are computed using the FFT via (26) as described in Section 3.3 and for the subsampled lattice the system (28) is solved using the conjugate gradient iterative method with FFT (Method 3) as described in Section 3.4. The subsampling set J is chosen randomly (with replacement) with size $|J| = \lceil \sqrt{N} \log N \rceil$, as in [1]. The generating vector is constructed using the component-by-component construction for approximation [18] using POD weights given by

$$\gamma_{\mathbf{u}} = \Gamma_{|\mathbf{u}|} \prod_{j \in \mathbf{u}} \gamma_j \quad \text{with} \quad \gamma_j = \frac{1}{j^q}, \quad \Gamma_{\ell} = \ell!, \quad (65)$$

where $q > 1$ is a parameter matched to the decay of the variables in each test integrand below.

All computations were performed on the computational cluster Katana [42] at UNSW Sydney.

6.1. Example 1: reciprocal function

First, we test our method for the reciprocal function, for $q \geq 2$ given by

$$f(\mathbf{x}) = \frac{1}{1 + \sum_{j=1}^s (x_j - \frac{1}{2})/j^q}, \quad (66)$$

which has been studied previously in, e.g., [17, 1]. We study the convergence of the new quadrature rule applied to f for two different decays $q \in \{2.5, 6\}$ and dimension $s \in \{2, 4, 64\}$.

The error is computed by comparing the weighted quadrature rules (23) and (27) to a reference integral, which was computed to accuracy $< 10^{-10}$ using a combination of explicitly integrating the first two dominant dimensions then applying a randomly-shifted rank-1 lattice rule with 2^{23} points and 64 random shifts (2^{29} points in total). The computed reference value of the integral is given in Table 4. Since the first two dimensions are integrated analytically, the value of the integral is exact for $s = 2$.

In Figure 3 we plot the error for the reciprocal function with the faster decay $q = 6$ (more smooth). Subfigures (a), (b), (c) give results in different dimensions for $\alpha = 1, 2, 3$, respectively, and subfigure (d) compares the full and subsampled lattice for different α in dimension 64. The two dashed lines indicate two rates we expect from our theory as summarised in Table 1: green is the best possible rate with dimension dependence ($p = 2$) and doubling the rate effect, whereas magenta is the slower dimension independent rate ($p = \infty$) without the doubling the rate effect. Note that the number of function evaluations on the horizontal scale is not the same across the figures. For this faster decay (i.e., more smooth) case, we observe the best possible rate $N^{-\alpha}$ for the full lattice for all values of α , with even faster convergence for $\alpha = 1$ or in lower dimensions for $\alpha = 2, 3$. Comparing the full lattice with the subsampled lattice for dimension $s = 64$ in Figure 3(d), we observe that the full and subsampled lattices exhibit similar behaviour, with both the full and subsampled lattice converging at a rate of roughly α .

In Figure 4 we plot the results for the slower decay $q = 2.5$ (less smooth), where again subfigures (a), (b), (c) give results for $\alpha = 1, 2, 3$, respectively, and (d) compares the full and subsampled lattice. For this more difficult function,

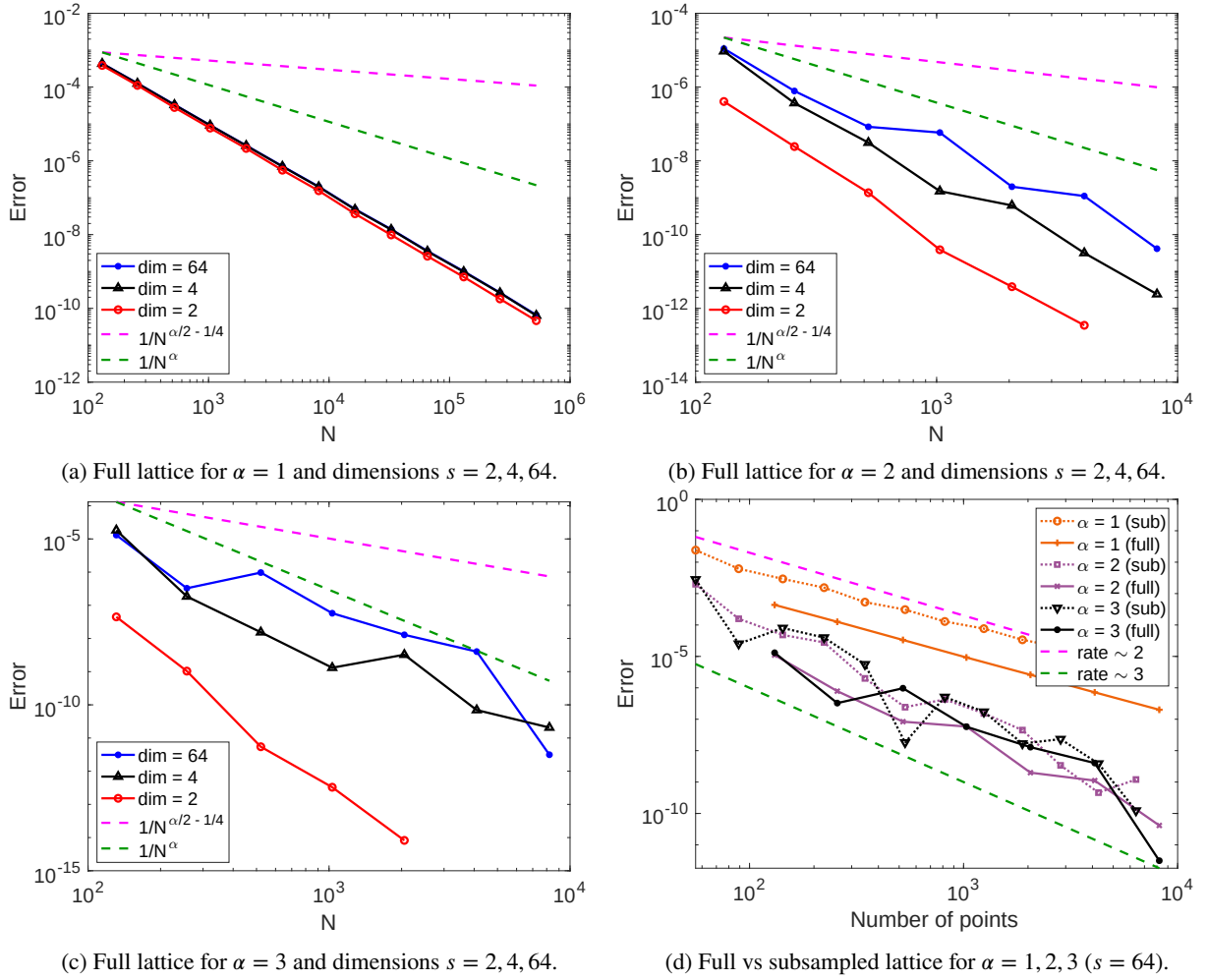


Figure 3: Convergence of error for reciprocal function (66) for decay $q = 6$.

we only observe the best possible rate of $N^{-\alpha}$ for $\alpha = 1$ or for lower dimensions when $\alpha = 2, 3$. For $\alpha = 2$ in dimension 64, we only begin to observe close to the best rate of 2 for higher N , whereas for $\alpha = 3$ we only observe the slower rate of $\frac{\alpha}{2} - \frac{1}{4} = 1.25$, i.e., we do not observe the doubling the rate effect. Similar to the $q = 6$ case, the full and subsampled rules exhibit comparable behavior when using a comparable number of points overall.

6.2. Example 2: product function

For a given $\alpha \geq 1$ and $q \geq 1$, we also consider the product function

$$f(\mathbf{x}) = \prod_{j=1}^s \frac{1 + x_j^{\alpha+1/2}/j^q}{1 + 1/[(\alpha + 3/2)j^q]} \quad (67)$$

for which the integral is equal to 1 in all dimensions. Note that for a given α , we have that $f \in H_{s,\alpha,\beta,\omega}^{\text{Sob}}$, but f does not belong to $H_{s,\alpha+1,\beta,\omega}^{\text{Sob}}$, and so we do not expect to see the full doubling the rate effect.

For this product function we again tested two decays $q \in \{2.5, 6\}$ and $\alpha \in \{1, 2, 3\}$. The results for $q = 6$ are given in Figure 5. We observe similar behavior as for the reciprocal function, however, since the product function is not in $H_{s,\alpha+1,\beta,\omega}^{\text{Sob}}$, for higher α and decay ($\alpha = 3, q = 6$ in Figure 5(c)) we do not observe the full doubling the rate effect,

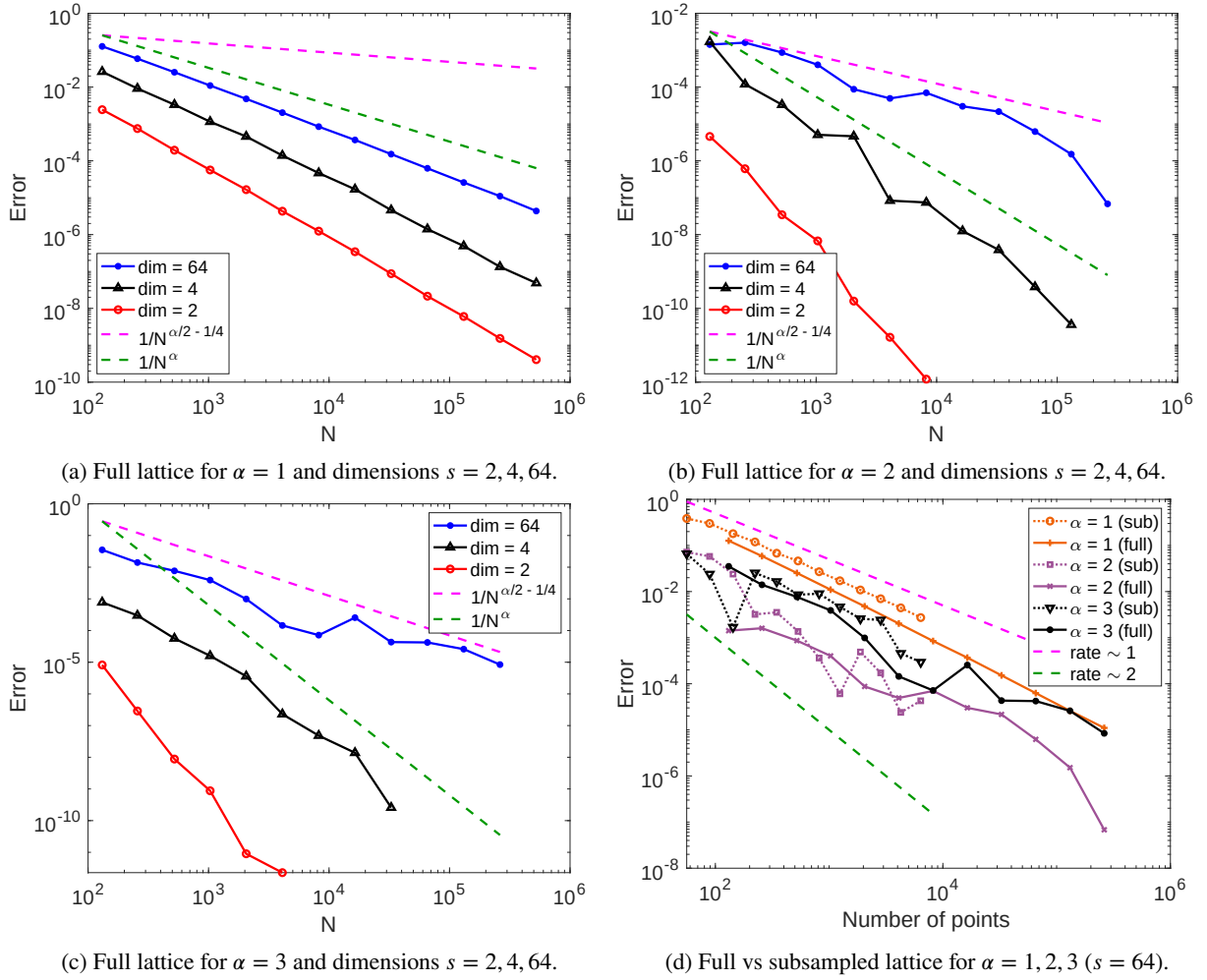


Figure 4: Convergence of error for reciprocal function (66) for decay $q = 2.5$.

but instead see an intermediate rate as predicted by Theorem 9. The results for $q = 2.5$ are similar to the reciprocal function and so have not been included.

7. Other related strategies

7.1. A common periodization method

Many papers (see for example [31] and [44]) have used a coordinate transformation to achieve a somewhat smooth periodic integrand, as in (for the univariate case)

$$\int_0^1 f(x) dx = \int_0^1 f(\phi(y)) \phi'(y) dy.$$

Here ϕ is a smooth increasing function mapping $[0, 1]$ onto $[0, 1]$, with one or more vanishing derivatives at both 0 and 1. Examples are $\phi(t) := 3t^2 - 2t^3$ due to Korobov [31] and $\phi(t) := t - \sin(2\pi t)/(2\pi)$ due to Sidi [44]. Because of the vanishing derivatives at the end points the resulting integrand $f(\phi(y)) \phi'(y)$ can be considered as a single period of a somewhat smooth periodic function on $\mathbb{R}$, whose integral can then be approximated by the trapezoidal rule, which from the Euler–Maclaurin expansion can be highly accurate for integration over a full period of a periodic function.

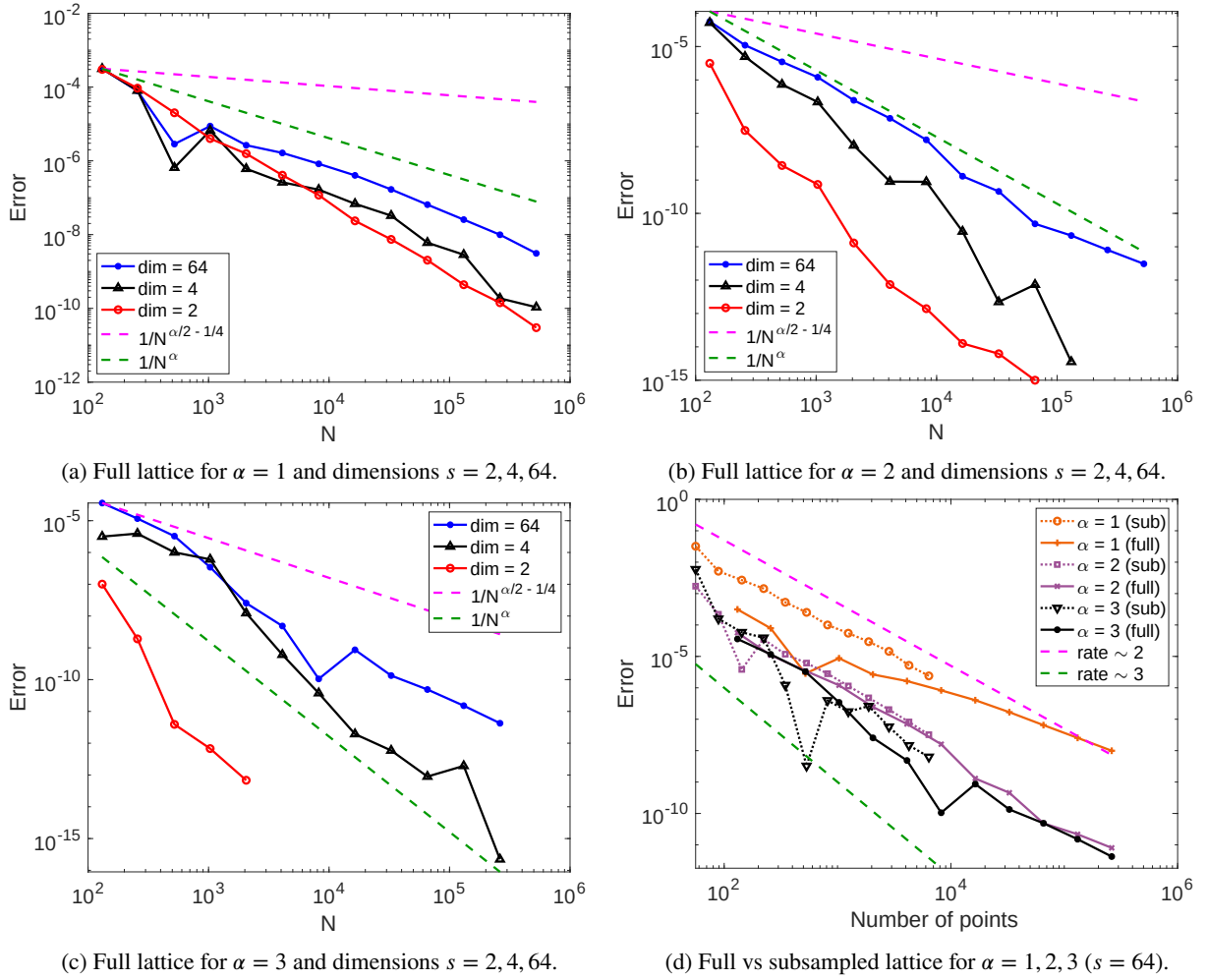


Figure 5: Convergence of error for product function (67) for decay $q = 6$.

Since periodicity is enforced by requiring $\phi'(0) = 0 = \phi'(1)$, the end result is a rule of the form

$$\int_0^1 f(x) dx \approx \frac{1}{N} \sum_{k=1}^{N-1} f(\phi(t_k)) \phi'(t_k), \quad \text{with } t_k = \frac{k}{N} \text{ for } k = 1, \dots, N-1.$$

Note that that this commonly used strategy (even with its lattice generalization to the multivariate case) is completely different from the method introduced in this paper. Firstly, our function ψ in (4) is *not* monotone. To the contrary, our function ψ traverses the interval $[0, 1]$ twice with respect to x (see Figure 1), first increasing from 0 to 1, then smoothly reversing. Secondly, our transformed integrand $f(\psi(y)) \mu(y)$ makes no pretence to be smooth — rather, the weight function $\mu(y) = |\psi'(y)|/2$ has a simple discontinuity in its first derivative.

Multivariate applications of Korobov's periodization technique using lattice points have been considered since the paper [31], but mostly not giving an error analysis in the setting of weighted mixed derivative Hilbert spaces, as in the current paper. An exception is [36], whose main conclusion was that this periodization strategy may fail in high dimensions, in that the error can grow exponentially with s because the Sobolev space norm of the periodized integrand $f(\phi(y)) \phi'(y)$ can be exponentially large compared to the Sobolev space norm of the original function f . Finally, the paper [3] explores a number of periodizing transformations in the multivariate-plus-lattice context, including a favoured transcendental transformation.

7.2. The method of product integration

The method of the present paper can be seen as a multivariate variant of the so-called method of *product integration*, see [54, 16], which was designed to handle univariate integrals of the form $\int_0^1 f(x) \nu(x) dx$, where f is smooth and ν is (weakly) singular. The essence of the method is to approximate f to high accuracy by polynomial interpolation at suitable points $t_1, t_2, \dots, t_N \in [0, 1]$, and then to evaluate the integral of a polynomial basis set times ν numerically to high accuracy. The approximate integral is expressed finally in the form

$$\int_0^1 f(x) \nu(x) dx \approx \sum_{k=1}^N w_k f(t_k),$$

where the weights w_k incorporate the influence of the weakly singular function ν . In the present method the polynomial basis is replaced by a set of translated kernel functions, and the weakly singular function ν is replaced by the weight function μ .

In some situations (for example interpolation at the Clenshaw–Curtis points or the zeros of Jacobi polynomials) the computed product-integration weights w_k are known to have a very desirable stability property if ν is non-negative, namely that of “asymptotic positivity” [47], meaning that

$$\lim_{N \rightarrow \infty} \sum_{k=1}^N w_k = \lim_{N \rightarrow \infty} \sum_{k=1}^N |w_k|.$$

This property implies that for N sufficiently large the contribution of negative weights to the sum of absolute values becomes vanishingly small. The asymptotic positivity property is not known to hold for the present multivariate method.

7.3. Integrating the tent-transformed integrand

The method of integrating after application of the so-called *tent transform*, given in the univariate case by $x = \phi(y) := \max\{2y, 1 - 2y\}$, see e.g., [24, 20, 49, 9], can be seen as an extreme example of the periodization method in Subsection 7.1, except that here ϕ is not monotone, instead (in the univariate case) traversing $[0, 1]$ twice. Because this $\phi(y)$ does not have any vanishing derivatives at the end points and is not smooth, the tent transform method does not permit convergence of arbitrary order. However, it does have the advantage that the resulting density is trivially equal to 1, and also, as shown in [24], that it can achieve the convergence order appropriate to smoothness $\alpha = 2$.

We note that this method can also be seen as an extreme example of the product-integration method from Section 7.2, but with a periodizing function that, unlike ψ given by (4), is very far from being smooth.

7.4. Chebyshev lattices

We have already mentioned in Subsection 3.2 that our transformed lattice points are related to Chebyshev lattices [10]. These are typically defined for an even number of points N . Also, the transformation $x = \cos(\pi y)$ in [10, Definition 1] differs from ψ in (4) only to the extent of mapping $[-1, 1]$ to $[-1, 1]$ instead of $[0, 1]$ to $[0, 1]$, as here. An alternative interpretation as described in [33] is that the points are first tent-transformed and then cosine-transformed, which allows also for odd N . For both odd and even N , there is duplication of points due to the tent transform. There will be exactly $\lfloor N/2 + 1 \rfloor$ distinct points if $\gcd(z_j, N) = 1$ for at least one component z_j of the lattice generating vector $\mathbf{z}$, see [33, Lemma 12 and Subsection 4.3].

7.5. Higher-order digital nets

A key competitor to the method presented in this paper is the method of *high-order digital nets* as initiated by Dick [11]. These are methods of a completely different kind. They are quasi-Monte Carlo methods, that is, they are equal weight rules, with $w_k = 1/N$. For the integration problem they can achieve the full order of convergence for functions in H_α^{mix} [21] and can be computed by a component-by-component construction [19]. But order of convergence is not the only consideration. Also of importance is the implied constant, including its rate of growth with respect to dimensionality and order α . There are also reported difficulties in the implementation of higher-order digital nets, in that high-precision arithmetic may be required. A detailed comparison is beyond the scope of this paper, but should be carried out in the future.

Acknowledgments

FB, ADG, FYK and IHS acknowledge financial support from the Australian Research Council Discovery Project (DP240100769). MG acknowledges support from the Hausdorff Center for Mathematics (HCM) in Bonn, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/2 – 390685813. This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney.

References

- [1] F. Bartel, A. D. Gilbert, F. Y. Kuo, and I. H. Sloan. Minimal subsampled rank-1 lattices for multivariate approximation with optimal convergence rate. *arXiv*, 2506.07729, 2025.
- [2] F. Bartel, M. Schäfer, and T. Ullrich. Constructive subsampling of finite frames with applications in optimal function recovery. *Applied and Computational Harmonic Analysis*, 65:209–248, 2023.
- [3] M. Beckers and A. Haegemans. Transformation of integrands for lattice rules. In T. O. Espelid and A. Genz, editors, *Numerical Integration: Recent Developments, Software and Applications*, pages 329–340. Springer Netherlands, Dordrecht, 1992.
- [4] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004.
- [5] G. Byrenheid, L. Kämmerer, T. Ullrich, and T. Volkmer. Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness. *Numerische Mathematik*, 136(4):993–1034, 2017.
- [6] A. Chkifa, M. Dolbeault, D. Krieg, and M. Ullrich. Constructive discretization and approximation in reproducing kernel Hilbert spaces. *arXiv*, 2602.18719, 2026.
- [7] R. Cools, F. Y. Kuo, D. Nuyens, and I. H. Sloan. Lattice algorithms for multivariate approximation in periodic spaces with general weight parameters. In *75 years of mathematics of computation*, volume 754 of *Contemp. Math.*, pages 93–113. Amer. Math. Soc., Providence, RI, 2020.
- [8] R. Cools, F. Y. Kuo, D. Nuyens, and I. H. Sloan. Fast component-by-component construction of lattice algorithms for multivariate approximation with POD and SPOD weights. *Math. Comp.*, 90(328):787–812, 2021.
- [9] R. Cools, F. Y. Kuo, D. Nuyens, and G. Suryanarayana. Tent-transformed lattice rules for integration and approximation of multivariate non-periodic functions. *J. Complexity*, 36:166–181, 2016.
- [10] R. Cools and K. Poppe. Chebyshev lattices, a unifying framework for cubature with Chebyshev weight function. *BIT Numer. Math.*, (2):275–288, 2011.
- [11] J. Dick. Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.*, 46(3):1519–1553, 2008.
- [12] J. Dick, P. Kritzer, and F. Pillichshammer. *Lattice rules—numerical integration, approximation, and discrepancy*, volume 58 of *Springer Series in Computational Mathematics*. Springer, Cham, 2022.
- [13] J. Dick, F. Y. Kuo, and I. H. Sloan. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer.*, 22:133–288, 2013.
- [14] J. Dick and F. Pillichshammer. *Digital nets and sequences*. Cambridge University Press, Cambridge, 2010.
- [15] M. Dolbeault, D. Krieg, and M. Ullrich. A sharp upper bound for sampling numbers in L_2 . *Applied and Computational Harmonic Analysis*, 63:113–134, 2023.
- [16] D. Elliott and D. F. Paget. Product-integration rules and their convergence. *Nordisk Tidskr. Informationsbehandling (BIT)*, 16(1):32–40, 1976.
- [17] R. N. Gantner and C. Schwab. Computational higher order quasi-Monte Carlo integration. In R. Cools and D. Nuyens, editors, *Monte Carlo and Quasi-Monte Carlo Methods*, pages 271–288, Cham, 2016. Springer International Publishing.
- [18] A. D. Gilbert and I. H. Sloan. Novel component-by-component lattice algorithms for multivariate approximation. *In preparation*, 2026.
- [19] T. Goda. Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces. *J. Computational and Applied Mathematics*, 285:279–294, 2015.
- [20] T. Goda, K. Suzuki, and T. Yoshiki. The b -adic tent transformation for quasi-Monte Carlo integration using digital nets. *J. Approx. Theory*, 194:62–86, 2015.
- [21] T. Goda, K. Suzuki, and T. Yoshiki. Optimal order quasi-Monte Carlo integration in weighted sobolev spaces of arbitrary smoothness. *IMA J. Numer. Anal.*, 37(1):505–518, 2017.
- [22] M. Griebel, C. Rieger, and P. Zaspel. Kernel-based stochastic collocation for the random two-phase Navier-Stokes equations. *Int. J. Uncertain. Quantif.*, 9(5):471–492, 2019.
- [23] M. Griebel and U. Seidler. On quadratures with optimal weights for spaces with unbounded mixed derivatives.
- [24] F. J. Hickernell. Obtaining $O(N^{-2+\epsilon})$ convergence for lattice quadrature rules. In Fang, Kai-Tai and Niederreiter, Harald and Hickernell, Fred J., editor, *Monte Carlo and Quasi-Monte Carlo Methods 2000*, pages 274–289, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [25] V. Kaarnioja, Y. Kazashi, F. Y. Kuo, F. Nobile, and I. H. Sloan. Fast approximation by periodic kernel-based lattice-point interpolation with application in uncertainty quantification. *Numer. Math.*, 150(1):33–77, 2022.
- [26] V. Kaarnioja, F. Y. Kuo, and I. H. Sloan. Lattice-based kernel approximation and serendipitous weights for parametric PDEs in very high dimensions. In A. Hinrichs, P. Kritzer, and F. Pillichshammer, editors, *Monte Carlo and Quasi-Monte Carlo Methods*, pages 81–103, Cham, 2024. Springer International Publishing.
- [27] V. Kaarnioja and C. Schillings. Quasi-Monte Carlo for Bayesian design of experiment problems governed by parametric PDEs. *Numer. Math.*, 2026.
- [28] L. Kämmerer. *High Dimensional Fast Fourier Transform Based on Rank-1 Lattice Sampling*. Dissertation, Technische Universität Chemnitz, Chemnitz, Germany, 2015.

- [29] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on rank-1 lattice sampling. *J. Complexity*, 31(4):543–576, 2015.
- [30] T. Karvonen, G. Santin, and T. Wenzel. General superconvergence for kernel-based approximation. *arXiv*, 2505.11435, 2025.
- [31] N. M. Korobov. On number-theoretic methods in approximate analysis. In *Probl. Numer. Math. Comp. Techn. (Russian)*, pages 36–44. Gosudarstv. Naučno-Tehn. Izdat. Mašinstroït. Lit., Moscow, 1963.
- [32] D. Krieg, K. Pozharska, M. Ullrich, and T. Ullrich. Sampling recovery in L_2 and other norms. *Mathematics of Computation*, 2025.
- [33] F. Y. Kuo, G. Migliorati, F. Nobile, and D. Nuyens. Function integration, reconstruction and approximation using rank-1 lattices. *Math. Comp.*, 90(330):1861–1897, 2021.
- [34] F. Y. Kuo, W. Mo, and D. Nuyens. Constructing embedded lattice-based algorithms for multivariate function approximation with a composite number of points. *Constr. Approx.*, 61(1):81–113, 2025.
- [35] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski. Lattice rules for multivariate approximation in the worst case setting. In *Monte Carlo and quasi-Monte Carlo methods 2004*, pages 289–330. Springer, Berlin, 2006.
- [36] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski. Periodization strategy may fail in high dimensions. *Numer. Algorithms*, 46(4):369–391, 2007.
- [37] N. Nagel, M. Schäfer, and T. Ullrich. A new upper bound for sampling numbers. *Found. Comput. Math.*, 22(2):445–468, 2021.
- [38] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [39] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*, volume 6 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2008.
- [40] J. Oettershagen. *Construction of Optimal Cubature Algorithms with Applications to Econometrics and Uncertainty Quantification*. Dissertation, Institut für Numerische Simulation, Universität Bonn, 2017.
- [41] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, Cham, 2023.
- [42] PVC (Research Infrastructure), UNSW Sydney. Katana. DOI:10.26190/669XA286.
- [43] T. H. Savits. Some statistical applications of Faa di Bruno. *J. Multivariate Anal.*, 97(10):2131–2140, 2006.
- [44] A. Sidi. A new variable transformation for numerical integration. In *Numerical integration, IV (Oberwolfach, 1992)*, volume 112 of *Internat. Ser. Numer. Math.*, pages 359–373. Birkhäuser, Basel, 1993.
- [45] I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1994.
- [46] I. H. Sloan and V. Kaarnioja. Doubling the rate: Improved error bounds for orthogonal projection with application to interpolation. *BIT*, 65(1):10, 2025.
- [47] I. H. Sloan and W. E. Smith. Properties of interpolatory product integration rules. *SIAM J. Numer. Anal.*, 19(2):427–442, 1982.
- [48] I. H. Sloan and H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity*, 14(1):1–33, 1998.
- [49] G. Suryanarayana, D. Nuyens, and R. Cools. Reconstruction and collocation of a class of non-periodic functions by sampling along tent-transformed rank-1 lattices. *J. Fourier Anal. Appl.*, 22:187–214, 2016.
- [50] M. Ullrich and T. Ullrich. The role of Frolov’s cubature formula for functions with bounded mixed derivative. *SIAM J. Numer. Anal.*, 54(2):969–993, 2016.
- [51] G. Wahba. *Spline models for observational data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
- [52] H. Wendland. *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.
- [53] T. Wenzel. Sharp inverse statements for kernel approximation: Superconvergence and saturation. *arXiv*, 2601.01808, 2026.
- [54] A. Young. Approximate product-integration. *Proc. Roy. Soc. London Ser. A*, 224:552–561, 1954.
- [55] X. Zeng, P. Kritzer, and F. J. Hickernell. Spline methods using integration lattices and digital nets. *Constr. Approx.*, 30(3):529–555, 2009.
- [56] X. Zeng, K.-T. Leung, and F. J. Hickernell. Error analysis of splines for periodic problems using lattice designs. In H. Niederreiter and D. Talay, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 501–514, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.