ON THE RECONSTRUCTION OF FUNCTIONS FROM VALUES AT SUBSAMPLED QUADRATURE POINTS

FELIX BARTEL, LUTZ KÄMMERER, DANIEL POTTS, AND TINO ULLRICH

ABSTRACT. This paper is concerned with function reconstruction from samples. The sampling points used in several approaches are (1) structured points connected with fast algorithms or (2) unstructured points coming from, e.g., an initial random draw to achieve an improved information complexity. We connect both approaches and propose a subsampling of structured points in an offline step. In particular, we start with structured quadrature points (QMC), which provide stable L_2 reconstruction properties. The subsampling procedure consists of a computationally inexpensive random step followed by a deterministic procedure to further reduce the number of points while keeping its information. In these points functions (belonging to a RKHS of bounded functions) will be sampled and reconstructed from whilst achieving state of the art error decay.

Our method is dimension-independent and is applicable as soon as we know some initial quadrature points. We apply our general findings on the d-dimensional torus to subsample rank-1 lattices, where it is known that full rank-1 lattices lose half the optimal order of convergence (expressed in terms of the size of the lattice). In contrast to that, our subsampled version regains the optimal rate since many of the lattice points are not needed. Moreover, we utilize fast and memory efficient Fourier algorithms in order to compute the approximation. Numerical experiments in several dimensions support our findings.

1. INTRODUCTION

The goal of this paper is the reconstruction of multivariate functions $f \colon D \to \mathbb{C}$, $D \subset \mathbb{R}^d$, from sampled function values. There seems to be an increasing interest in this topic with several new and surprising approaches ranging from highly non-constructive to implementable, cf. [24, 4, 16, 25, 28, 29, 45, 46, 36, 20, 47, 33, 37, 13, 1, 19]. Since it is known [4] that the highly optimized and fast algorithms [9, 24, 32, 22, 20] are often not optimal with respect to information complexity, there is a trend to make the recent "complexity optimal" approaches [45, 37, 13, 33] accessible for practitioners, cf. [17, 1]. We continue working in this direction and provide new optimal theoretical results by exploiting a constructive subsampling strategy [1] on the one hand. On the other hand we focus on implementability and fast/memory-efficient algorithms. The latter is tightly connected to the use of structured sampling point sets since they often allow for fast matrix vector multiplication (FFT). In [12] there are several such constructions available mostly from numerical

$$\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\} \qquad \boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J} \qquad \boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'}$$
$$\boldsymbol{W}_{\mathrm{MZ}} \in [0, \infty)^{M \times M} \qquad \boldsymbol{W} \in [0, \infty)^{|\boldsymbol{X}| \times |\boldsymbol{X}|} \qquad \boldsymbol{W}' \in [0, \infty)^{|\boldsymbol{X}'| \times |\boldsymbol{X}'|}$$
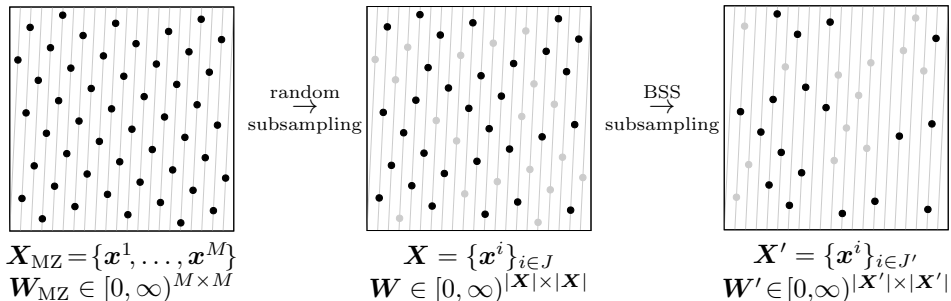
FIGURE 1.1. Two-step subsampling procedure from a rank-1 lattice

integration, e.g. lattice constructions, which allow for fast computation. In addition, if a point set $\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\} \subset D$ and weights $\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \dots, \omega_M)$ exactly integrate functions from a space of the form $V \cdot \bar{V}$, i.e., for $0 < A < \infty$

$$(1.1) \qquad \sum_{i=1}^{M} \omega_i g(\boldsymbol{x}^i) \overline{h(\boldsymbol{x}^i)} = A \int_D g(\boldsymbol{x}) \overline{h(\boldsymbol{x})} \, \mathrm{d}\nu(\boldsymbol{x}) \quad \text{for all} \quad g, h \in V,$$

then, equivalently, they establish an L_2-Marcinkiewicz-Zygmund inequality with constants $A = B$ for all $f \in V$, cf. Theorem 2.3. The concept of these inequalities goes back to J. Marcinkiewicz and A. Zygmund [34] and establishes a connection between the continuous L_2-norm and point evaluations of functions. A set of points $\boldsymbol{X}_{\mathrm{MZ}}$ and weights $\boldsymbol{W}_{\mathrm{MZ}}$ fulfills an L_2-Marcinkiewicz-Zygmund (MZ) inequality with constants $0 < A \le B < \infty$ for a finite-dimensional function space V, if

$$(1.2) \qquad A\|f\|_{L_2}^2 \le \sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2 \le B\|f\|_{L_2}^2 \quad \text{for all} \quad f \in V.$$

Clearly, such points together with appropriate subspaces V are good for sampling recovery in $L_2(\nu)$, cf. Theorem 4.3. For a systematic study of MZ inequalities (also for $p \ne 2$) we refer to the recent series of papers by V.N. Temlyakov and coauthors, see for instance [44, 26].

Note that M in (1.2) might be much larger than $\dim(V)$ which requires an additional "subsampling step" in order to create an optimal least squares approximation method out of an exact quadrature rule. By bridging to the frame terminology we are able to apply recent techniques from [1] in order to reduce the number of points whilst keeping their approximation power. We apply two different techniques, which we use in combination:

- In a first step we use random subsampling in order to obtain an index set $J \subset \{1, \dots, M\}$ such that $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J}$ and new weights $\boldsymbol{W}$ fulfill an L_2-MZ inequality with logarithmic oversampling, i.e., for a fixed constant C we have $|\boldsymbol{X}| \le C \dim(V) \log(\dim(V))$.
- In a second step we use the deterministic BSS-Algorithm from [2, 1] to further reduce the number of points to merely linear oversampling, i.e., for b close to one we compute $J' \subset J$ such that $\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'}$ and weights $\boldsymbol{W}'$ with $|\boldsymbol{X}'| \le \lceil b \dim(V) \rceil$ fulfill an L_2-MZ inequality as well.

This scheme is depicted in Figure 1.1 for initial points from a rank-1 lattice, see Section 5 for details. The corresponding results are found in Theorems 3.1,

3.3, and 3.4. There is freedom in the way of subsampling which allows to control the approximation error for functions from infinite-dimensional function spaces as well. We assume to be given a class of functions $H(K)$ (reproducing kernel Hilbert space with kernel K on D) and an initial set of L_2-MZ points $\boldsymbol{X}_{\mathrm{MZ}}$ for a subspace $V = \operatorname{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}} \subset H(K)$ spanned by $L_2(\nu)$-orthonormal eigenfunctions $\eta_k(\cdot)$ of the operator $W_\nu = \mathrm{Id}^* \circ \mathrm{Id}$ with $\mathrm{Id} \colon H(K) \hookrightarrow L_2(\nu)$ the embedding. We then consider the least squares reconstruction $S_I^{\boldsymbol{X}}$, defined in (2.3), on the possibly smaller subspace $\operatorname{span}\{\eta_k\}_{k \in I}$, $I \subset I_{\mathrm{MZ}}$, where the choice of I is a degree of freedom, and are interested in the L_2-worst-case error

$$(1.3) \qquad \sup_{\|f\|_{H(K)} \leq 1} \|f - S_I^{\boldsymbol{X}} f\|_{L_2(\nu)}^2 .$$

In Theorem 4.6 we show bounds for this error using a random subsample $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J} \subset \boldsymbol{X}_{\mathrm{MZ}}$. For this subsampling we use a special discrete probability distribution given in (4.6) below. The bounds are stated in terms of the eigenvalues λ_k and the approximation power of the initial method $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}}$ expressed by a quantity like (1.3). Moreover, using an additional deterministic subsampling which yields $\boldsymbol{X}' \subset \boldsymbol{X}$ with merely linear oversampling, i.e., $|\boldsymbol{X}'| \leq b|I|$ for b close to one, a special case of the main result in Theorem 4.8 for finite measures ν and kernels K which are constant on the diagonal reads as follows.

Theorem 1.1. *Let $H(K)$ be a separable reproducing kernel Hilbert space with a finite trace (4.2) and bounded (4.1) kernel. Further, let λ_k and η_k denote the eigenvalues and $L_2(\nu)$-orthonormal eigenfunctions of the operator $W_\nu = \mathrm{Id}^* \circ \mathrm{Id}$ with $\mathrm{Id} \colon H(K) \hookrightarrow L_2(\nu)$ the embedding. For $V = \operatorname{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}} \subset H(K)$ a finite-dimensional function space, we assume the points $\boldsymbol{X}_{\mathrm{MZ}} \subset D$ and weights $\boldsymbol{W}_{\mathrm{MZ}} \in [0, \infty)^{M \times M}$ fulfill an L_2-MZ inequality for V with constants A and B. Finally, let $I \subset I_{\mathrm{MZ}} \subset \mathbb{N}$, $|I| \geq 3$, be an index set and $\log(73B/A) + 2 \leq r$ a constant.*

For a target oversampling factor $b > 1 + \frac{1}{|I|}$ we construct a subset of points $\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'} \subset \boldsymbol{X}_{\mathrm{MZ}}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that we have the following bound with probability $1 - 4/|I|$

$$\sup_{\|f\|_{H(K)} \leq 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2}^2 \leq 79\,616\, \frac{B}{A} \frac{(b+1)^2}{(b-1)^3} \log|I| \left(\frac{B}{A} \sup_{k \notin I} \lambda_k + \frac{r}{|I|} \sum_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k \right.$$

$$\left. + \frac{1}{A} \sup_{\|f\|_{H(K)} \leq 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 \right)$$

with weights $\boldsymbol{W}' = \operatorname{diag}(\omega_i/\varrho_i)_{i \in J'}$ and the least squares approximation $S_I^{\boldsymbol{X}'} f = S_I^{\boldsymbol{X}'}(\boldsymbol{W}')f$ defined in (2.3).

To interpret the above result, we assume the index set I contains the indices corresponding to the largest eigenvalues and b fixed. Up to a logarithmic and constant factor we obtain two categories of summands

- The summand $\sup_{k \notin I} \lambda_k$ is the truncation error, i.e., the best possible error from the function space $\operatorname{span}\{\eta_k\}_{k \in I}$. The summand $\frac{1}{|I|} \sum_{k \notin I} \lambda_k$ is of the same order whenever the eigenvalues have polynomial decay smaller in order than -1, i.e., $\lambda_k \lesssim k^{-1-\varepsilon}$. These only depend on the index set I for the least squares approximation.

- The remaining summand is the error from the initial L_2-MZ inequality of the function space $V = \mathrm{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$.

Despite the fact that we restrict ourselves to structured nodes we achieve the same rates as in [37, 1] which are worse by the square root of a logarithm compared to [13]. When going to specific scenarios (see Corollary 1.2 below), these structured nodes come from rank one lattices which can be well-processed in a computer (they are at least rational). The lower bound in [18] indicates that the bound $1/|I| \sum_{k \notin I} \lambda_k$ in [13] is most likely sharp and we are close to optimal. In fact, the term only depending on I_{MZ} can be made arbitrarily small by choosing a good initial L_2-MZ inequality for large enough I_{MZ} without deteriorating the remaining bound or the size of the subsampled points $\boldsymbol{X}'$. Note that often $I_{\mathrm{MZ}} = I$ is a valid choice. We stress again that, by choosing a "good enough" initial L_2-MZ inequality, one achieves the best possible error rate up to a logarithm. Recent progress has shown that the logarithmic factor might be avoided but this utilizes the Kadison-Singer theorem and is not constructive in contrast to our approach, cf. [13].

The novelty of our result is that we start with discrete L_2-MZ points and positive weights which often have a structure that can be exploited towards fast matrix-vector multiplications. These are available for, e.g. the d-dimensional torus $\mathbb{T}^d$ with the trigonometric polynomials with Fourier-like algorithms, the cube with Chebyshev polynomials [40], the two-dimensional sphere $\mathbb{S}^2$ with the spherical harmonics [31], or the rotation group $SO(3)$ with the Wigner-D functions [41]. We make a case study on the d-dimensional torus $\mathbb{T}^d$ with rank-1 lattices, where we apply our theory and conduct numerical experiments. It is known that full rank-1 lattices achieve only half the optimal main rate for Sobolev spaces with dominating mixed smoothness, cf. [4] and Theorem 5.1. With the subsampling approach in this paper, we regain the optimal order of convergence via subsampling:

Corollary 1.2. *Let $H^s_{\mathrm{mix}}(\mathbb{T}^d)$, $s > 1/2$, be the Sobolev space with dominating mixed smoothness (5.3) on the d-torus, $I \subset I_{\mathrm{MZ}} \subset \mathbb{Z}^d$, $|I| \geq 3$, be the hyperbolic cross frequency index sets defined in (5.5), and $\boldsymbol{X}_{\mathrm{MZ}}$ a reconstructing rank-1 lattice for I_{MZ} with M points, cf. (5.4). For $b > 1 + \frac{1}{|I|}$ we construct a subset of points $\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'} \subset \boldsymbol{X}_{\mathrm{MZ}}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that we have the following bound for the plain least squares approximation*

$$(1.4) \qquad \sup_{\|f\|_{H^s_{\mathrm{mix}}} \leq 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2}^2 \leq C_{d,s} \left(\frac{b}{b-1}\right)^3 \log |I| \left(|I|^{-2s}(\log |I|)^{(d-1)2s}\right.$$

$$\left. + |I_{\mathrm{MZ}}|^{-2s+1}(\log |I_{\mathrm{MZ}}|)^{2(d-1)s}\right),$$

with probability $1 - 4/|I|$ and $C_{d,s}$ a constant depending on d and s.

Note, that the size M of the initial reconstructing rank-1 lattice for I_{MZ} has to be roughly quadratic in $|I_{\mathrm{MZ}}|$, cf. Lemma 5.1 (iii). We work with merely linear oversampling ($n = |\boldsymbol{X}'| \leq \lceil b|I| \rceil$) achieving the optimal error decay up to a logarithmic factor. There exists an initial lattice of size M such that the bound in (1.4) above, stated in terms of the number of sample points n and initial lattice size M, is of order less than

$$\log n \left(n^{-2s}(\log n)^{(d-1)2s} + M^{-s+1/2}(\log M)^{(d-1)(s+1)-s}\right)$$

using Theorem 5.1. Whereas full rank-1 lattices achieve only half the main order of convergence, the subsampled rank-1 lattice achieves the optimal main order of

convergence. Moreover, we are still able to utilize the well-tried memory efficient fast Fourier algorithms with the low number of samples.

For the structure of this paper we investigate L_2-MZ points in Section 2 and show their relation to exact quadrature and establish a connection to finite frames. In Section 3 we use this frame characterization to apply different subsampling techniques in order to reduce the number of points and obtain a new L_2-MZ inequality. The infinite-dimensional case is considered in Section 4, where we modify the subsampling such that we control and improve on the approximation error. We finish with an application in Section 5 on the d-dimensional torus $\mathbb{T}^d$ with rank-1 lattices and conduct numerical experiments.

2. Exact quadrature, L_2-MZ inequalities, and least squares

The problem in this section is the recovery of functions $f \colon D \to \mathbb{C}$ defined on some domain $D \subset \mathbb{R}^d$. For that we have given samples $\boldsymbol{f} = (f(\boldsymbol{x}^1), \dots, f(\boldsymbol{x}^M))^\mathsf{T} \in \mathbb{C}^M$ in points $\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\} \subset D$. For now, we assume f to be an element of a finite-dimensional function space $V \subset L_2$ with dimension smaller than M (oversampling). Our method of choice to recover the function from its samples is least squares approximation and, as we will see, L_2-MZ inequalities come into play in a natural way.

At first, for a σ-finite measure $\nu \colon D \to \mathbb{R}$, we define the space of square-integrable functions $L_2 = L_2(D, \nu)$ with

$$\langle f, g \rangle_{L_2} := \int_D f(\boldsymbol{x}) \overline{g(\boldsymbol{x})} \, \mathrm{d}\nu(\boldsymbol{x}) \quad \text{and} \quad \|f\|_{L_2} := \left(\int_D |f(\boldsymbol{x})|^2 \, \mathrm{d}\nu(\boldsymbol{x}) \right)^{1/2}.$$

Having an inner product we calculate an orthonormal basis $\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$ of V and, with that, have an expansion for functions $f \in V$:

$$(2.1) \qquad f(\boldsymbol{x}) = \sum_{k \in I_{\mathrm{MZ}}} a_k \eta_k(\boldsymbol{x}) \quad \text{with} \quad \boldsymbol{a} = (a_k)_{k \in I_{\mathrm{MZ}}} = (\langle f, \eta_k \rangle_{L_2})_{k \in I_{\mathrm{MZ}}}.$$

Now we seek a way to recover the coefficients $\boldsymbol{a}$ from the samples $\boldsymbol{f}$. The following system of linear equations establishes this connection

$$(2.2) \qquad \boldsymbol{L}\boldsymbol{a} = \boldsymbol{f} \quad \text{with} \quad \boldsymbol{L} := \boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}} = \begin{pmatrix} \eta_1(\boldsymbol{x}^1) & \cdots & \eta_{|I_{\mathrm{MZ}}|}(\boldsymbol{x}^1) \\ \vdots & \ddots & \vdots \\ \eta_1(\boldsymbol{x}^M) & \cdots & \eta_{|I_{\mathrm{MZ}}|}(\boldsymbol{x}^M) \end{pmatrix}.$$

Assuming $\boldsymbol{L}$ has full rank, a stable way to solve this overdetermined system of equations is to minimize the residual $\|\boldsymbol{L}\boldsymbol{a} - \boldsymbol{f}\|_2^2$. The solution is obtained by solving the normal equation and the approximation $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f$ takes the form

$$(S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f)(\boldsymbol{x}) := \sum_{k \in I_{\mathrm{MZ}}} a_k \eta_k(\boldsymbol{x}) \quad \text{with} \quad \boldsymbol{a} = (\boldsymbol{L}^* \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{f}.$$

Further, we consider weighted versions of the least squares algorithm. For a diagonal matrix $\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \dots, \omega_M) \in [0, \infty)^{M \times M}$ the weighted least squares approximation $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f = S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}}(\boldsymbol{W}_{\mathrm{MZ}})f$ takes the form

$$(2.3) \qquad (S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f)(\boldsymbol{x}) = \sum_{k \in I_{\mathrm{MZ}}} a_k \eta_k(\boldsymbol{x}) \quad \text{with} \quad \boldsymbol{a} = (\boldsymbol{L}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{f}$$

where $\boldsymbol{a}$ is the minimizer of $\boldsymbol{a} \mapsto \|\boldsymbol{L}\boldsymbol{a} - \boldsymbol{f}\|_{\boldsymbol{W}_{\mathrm{MZ}}}^2 = \|\boldsymbol{W}_{\mathrm{MZ}}^{1/2}(\boldsymbol{L}\boldsymbol{a} - \boldsymbol{f})\|_2^2$. So it is obtained from the unweighted least squares version by a change of measure using the discrete weights $\boldsymbol{W}_{\mathrm{MZ}}$.

If the singular values of $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}$ are non-zero, we immediately obtain that the least squares algorithm $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}}$ reconstructs all functions $f \in V$. The stability of the reconstruction as well as the number of iterations of a conjugate gradient method used to solve the system of equations heavily depends on the condition number of the matrix $(\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}\boldsymbol{L})^{-1}\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}^{1/2}$, cf. [15, Theorem 3.1.1].

Lemma 2.1. [25, Proposition 3.1] *Let* $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L} \in \mathbb{C}^{M \times |I_{\mathrm{MZ}}|}$ *be a matrix with* $|I_{\mathrm{MZ}}| \leq M$ *with non-zero singular values. Then*

$$\sigma_{\max}\left((\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}\boldsymbol{L})^{-1}\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\right) = \frac{1}{\sigma_{\min}(\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L})}$$

and
$$\sigma_{\min}\left((\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}\boldsymbol{L})^{-1}\boldsymbol{L}^*\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\right) = \frac{1}{\sigma_{\max}(\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L})} .$$

This motivates having a look at lower and upper bounds for the singular values of $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}$. The following lemma gives different characterizations for these bounds.

Lemma 2.2. *Let* $\eta_k \colon D \to \mathbb{C}$, $k \in I_{\mathrm{MZ}}$ *be an* L_2-*orthonormal basis of a finite-dimensional function space* V, *let* $\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\} \subset D$ *be points and* $\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \ldots, \omega_M) \in [0, \infty)^{M \times M}$ *weights. Then the following are equivalent:*

(i) *the singular values of the matrix* $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L} \in \mathbb{C}^{M \times |I_{\mathrm{MZ}}|}$, *where* $\boldsymbol{L}$ *is defined in* (2.2), *lie in the interval* $[\sqrt{A}, \sqrt{B}]$, *i.e.,*

$$A\|\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}\boldsymbol{a}\|_2^2 \leq B\|\boldsymbol{a}\|_2^2 \quad \textit{for all} \quad \boldsymbol{a} \in \mathbb{C}^{|I_{\mathrm{MZ}}|} ;$$

(ii) *the rows* $\sqrt{\omega_i}(\eta_k(\boldsymbol{x}^i))_{k \in I_{\mathrm{MZ}}} \in \mathbb{C}^{|I_{\mathrm{MZ}}|}$, $i = 1, \ldots, M$, *of the matrix* $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}$ *form a frame with bounds* A *and* B, *i.e.,*

$$A\|\boldsymbol{a}\|_2^2 \leq \sum_{i=1}^M |\langle \boldsymbol{a}, \sqrt{\omega_i}(\eta_k(\boldsymbol{x}^i))_{k \in I_{\mathrm{MZ}}}\rangle|^2 \leq B\|\boldsymbol{a}\|_2^2 \quad \textit{for all} \quad \boldsymbol{a} \in \mathbb{C}^{|I_{\mathrm{MZ}}|} ;$$

(iii) *the points* $\boldsymbol{X}_{\mathrm{MZ}}$ *and weights* $\boldsymbol{W}_{\mathrm{MZ}}$ *form an* L_2-*MZ inequality for the function space* $V = \mathrm{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$ *with bounds* A *and* B, *i.e.,*

$$(2.4) \qquad A\|f\|_{L_2}^2 \leq \sum_{i=1}^M \omega_i |f(\boldsymbol{x}^i)|^2 \leq B\|f\|_{L_2}^2 \quad \textit{for all} \quad f \in V .$$

Proof. **Step 1.** To show the equivalence of (i) and (ii) we rewrite the matrix-vector product $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}\overline{\boldsymbol{a}}$ using the Euclidian inner products of $\boldsymbol{a}$ and the rows of $\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}$ as follows

$$A\|\boldsymbol{a}\|_2^2 = A\|\overline{\boldsymbol{a}}\|_2^2 \leq \|\boldsymbol{W}_{\mathrm{MZ}}^{1/2}\boldsymbol{L}\overline{\boldsymbol{a}}\|_2^2 = \sum_{i=1}^M |\langle \boldsymbol{a}, \sqrt{\omega_i}(\eta_k(\boldsymbol{x}^i))_{k \in I_{\mathrm{MZ}}}\rangle_2|^2 \leq B\|\overline{\boldsymbol{a}}\|_2^2 = B\|\boldsymbol{a}\|_2^2 .$$

This immediately shows the equivalence of the two stated conditions.

Step 2. Now we show the equivalence of (ii) and (iii). Using the series expansion (2.1) of f, we have

$$\sum_{i=1}^{M} |\langle \overline{\boldsymbol{a}}, \sqrt{\omega_i}(\eta_k(\boldsymbol{x}^i))_{k \in I_{\mathrm{MZ}}} \rangle_2|^2 = \sum_{i=1}^{M} \omega_i \Big| \sum_{k \in I_{\mathrm{MZ}}} a_k \eta_k(\boldsymbol{x}^i) \Big|^2 = \sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2$$

and further using Parseval's identity, we obtain

$$\|\overline{\boldsymbol{a}}\|_2^2 = \|\boldsymbol{a}\|_2^2 = \Big\| \sum_{k \in I_{\mathrm{MZ}}} a_k \eta_k(\boldsymbol{x}^i) \Big\|_{L_2}^2 = \|f\|_{L_2}^2 \,.$$

Plugging these two formulas into the frame condition, we obtain a reformulation in terms of functions, which is an L_2-MZ inequality for the function space V with bounds A and B. ∎

The L_2-MZ characterization supports the availability of points with well-behaved least squares matrices. With Lemma 2.2, using a set of points $\boldsymbol{X}_{\mathrm{MZ}}$ coming from an L_2-MZ inequality with well-behaved constants, we automatically have a good point set to use in least squares approximation, and vice versa. Point sets fulfilling L_2-MZ inequalities are widely available and well-studied, cf. [35, 27, 14, 7, 8, 44]. Furthermore, the next theorem shows an equivalence to an exact integration condition when the constants in the L_2-MZ inequality coincide. This widens the applicability even further, cf. [9, 48, 23].

Theorem 2.3. *The points $\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\} \subset D$ and weights $\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \ldots, \omega_M) \in [0, \infty)^{M \times M}$ obey an L_2-MZ inequality (2.4) on the function space $V = \mathrm{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$ with constants $A = B$ if and only if we have exact quadrature on $V \cdot \overline{V} := \{f = \overline{g} \cdot h : g, h \in V\}$, i.e., we have*

$$\sum_{i=1}^{M} \omega_i g(\boldsymbol{x}^i)\overline{h(\boldsymbol{x}^i)} = A \int_D g(\boldsymbol{x})\overline{h(\boldsymbol{x})}\,\mathrm{d}\nu(\boldsymbol{x}) \quad \text{for all} \quad g, h \in V \,.$$

Proof. Starting with the L_2-MZ inequality, we have

$$\sum_{i=1}^{M} \omega_i |f(\boldsymbol{x}^i)|^2 = A \int_D |f(\boldsymbol{x})|^2 \,\mathrm{d}\nu(\boldsymbol{x}) \quad \text{for all} \quad f \in V \,.$$

By the parallelogram law the corresponding inner products also coincide, which is on direction of the assertion. The reverse is achieved by using $g = h$. ∎

3. Subsampling of L_2-MZ inequalities

In this section we utilize recent random and deterministic frame subsampling techniques from [1] to select subsets of L_2-MZ point sets keeping their good approximation properties following the illustration in Figure 1.1. The connection from L_2-MZ points to frames was already done in Lemma 2.2. The best possible start are tight frames, i.e., with frame bounds $A = B$, which is equivalent to having exact quadrature as we have seen in Theorem 2.3.

The next theorem covers random subsampling.

Theorem 3.1. *Let $0 < C \leq 1$, $\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$ be an L_2-orthonormal basis of the finite-dimensional function space V, let $\boldsymbol{X} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\} \subset D$ be points and*

$\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \ldots, \omega_M) \in [0, \infty)^{M \times M}$ *weights fulfilling an L_2-MZ inequality for V with constants A and B. Further, let $t > 0$, $I \subset I_{\mathrm{MZ}}$, and $n \in \mathbb{N}$ be such that*

$$(3.1) \qquad n \geq \frac{12B}{AC} |I|(\log |I| + t).$$

We draw a set $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J}$, $|J| = n$, of points i.i.d. from $\boldsymbol{X}_{\mathrm{MZ}}$ with respect to the discrete probability weights ϱ_i (with duplicates), which fulfill

$$(3.2) \qquad \varrho_i \geq \frac{C\omega_i \sum_{k \in I} |\eta_k(\boldsymbol{x}^i)|^2}{\sum_{j=1}^M \omega_j \sum_{k \in I} |\eta_k(\boldsymbol{x}^j)|^2} \quad for \quad i = 1, \ldots, M.$$

Then, with probability larger than $1 - 2\exp(-t)$, there holds the subsampled L_2-MZ inequality

$$(3.3) \qquad \frac{1}{2} A\|f\|_{L_2}^2 \leq \sum_{i \in J} \frac{\omega_i}{n\varrho_i} |f(\boldsymbol{x}^i)|^2 \leq \frac{3}{2} B\|f\|_{L_2}^2 \quad for \ all \quad f \in \mathrm{span}\{\eta_k\}_{k \in I}.$$

Proof. We will use [1, Theorem A.3] (which is a reformulation of [49, Theorem 1.1]) with $\boldsymbol{A}_i = \frac{\omega_i}{n\varrho_i} \boldsymbol{y}^i \otimes \boldsymbol{y}^i$, $\boldsymbol{y}^i = (\eta_k(\boldsymbol{x}^i))^{k \in I}$ and $t = 1/2$ (this is not the same t as in the present assertion). By the property of ϱ_i and the given L_2-MZ inequality, we have

$$\lambda_{\max}(\boldsymbol{A}_i) = \frac{\omega_i}{n\varrho_i} \|\boldsymbol{y}^i\|_2^2 \leq \frac{1}{nC} \sum_{j=1}^M \omega_j \sum_{k \in I} |\eta_k(\boldsymbol{x}^j)|^2 \leq \frac{B|I|}{nC}.$$

Further,

$$\sum_{i=1}^n \mathbb{E}(\boldsymbol{A}_i) = \sum_{i=1}^n \sum_{j=1}^M \varrho_j \frac{\omega_j}{n\varrho_j} \boldsymbol{y}^j \otimes \boldsymbol{y}^j = \boldsymbol{L}_{\mathrm{MZ}}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\mathrm{MZ}},$$

where $\boldsymbol{L}_{\mathrm{MZ}}$ is the system matrix with the points $\boldsymbol{X}_{\mathrm{MZ}}$ and $\boldsymbol{W}_{\mathrm{MZ}} = \mathrm{diag}(\omega_1, \ldots, \omega_M)$. With Lemma 2.2, the eigenvalues of the matrix $\boldsymbol{L}_{\mathrm{MZ}}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\mathrm{MZ}}$ are bounded by $A \leq \lambda_{\min}(\boldsymbol{L}_{\mathrm{MZ}}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\mathrm{MZ}}) \leq \lambda_{\max}(\boldsymbol{L}_{\mathrm{MZ}}^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\mathrm{MZ}}) \leq B$. Note, $A/2 \leq \lambda_{\min}(\sum_{i=1}^n \boldsymbol{A}_i)$ and $\lambda_{\max}(\sum_{i=1}^n \boldsymbol{A}_i) \leq 3B/2$ are the target inequalities.

By applying [1, Theorem A.3] and using the assumption on n, both inequalities of the assertion hold with probability

$$1 - |I| \exp\left(-\frac{An}{12CB|I|} \right) \geq 1 - \exp(-t).$$

Union bound gives the overall probability and, thus, the assertion. ∎

Thus, we found a subset $\boldsymbol{X}$ of $\boldsymbol{X}_{\mathrm{MZ}}$ with logarithmic oversampling (independent of the original size M) which fulfills an L_2-MZ inequality with similar constants. The probability density stated in (3.2) is also used in Christoffel-weighted least squares, cf. [38].

Remark 3.2. *We could apply the orthogonalization trick from [1, Lemma 4.3] to eliminate the factor B/A in the assumption on the number of points n in (3.1). But this demands for setting up the matrix $\boldsymbol{L}$ from (2.2) whereas in the above formulation, we only need the evaluation of the Christoffel function $N(I) = \sum_{k \in I} |\eta_k(\boldsymbol{x})|^2$.*

Next, we subsample $\boldsymbol{X}$ further to obtain merely linear oversampling.

Theorem 3.3. *Let the assumptions from Theorem 3.1 hold and $\boldsymbol{X}$ be such that (3.3) holds. Further, let $b > \kappa^2$ with*

$$\kappa = \frac{3B}{2A} + \frac{1}{2} + \sqrt{\left(\frac{3B}{2A} + \frac{1}{2}\right)^2 - 1}\,.$$

Then BSS-*subsampling (cf.* [1, Algorithm 1]*) the points $\boldsymbol{X}$, we obtain points $\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i \in J'} \subset \boldsymbol{X}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ and non-negative weights s_i, $i \in J'$ such that there holds the subsampled L_2-MZ inequality*

$$\frac{1}{2}A\|f\|_{L_2}^2 \leq \sum_{i \in J'} \frac{\omega_i s_i}{n \varrho_i} |f(\boldsymbol{x}^i)|^2 \leq \frac{3}{2} \frac{(\sqrt{b}+1)^2}{(\sqrt{b}-1)(\sqrt{b}-\kappa)} B\|f\|_{L_2}^2 \quad \forall f \in \mathrm{span}\{\eta_k\}_{k \in I}\,.$$

Proof. The result is an immediate consequence of applying [1, Theorem 3.1] to the randomly subsampled points of Theorem 3.1. ■

The previous result is based on [2] where tight frames were subsampled, which was used [33] for subsampling random points. This was then extended in [1] for non-tight frames as well, which we use here. The BSS-algorithm gives no control over the weights s_i, however. One alternative would be to use Weaver-subsampling to lose the weights, but this is highly nonconstructive and increases the involved constants tremendously, cf. [37]. A clever extension of the BSS-algorithm makes it possible to lose the weights, regain the constructiveness, choose a smaller oversampling factor b, and save the left-hand side of the L_2-MZ inequality, which is the important one as it allows for the reconstruction from the function evaluations $f(\boldsymbol{x}^i)$.

Theorem 3.4. *Let the assumptions from Theorem 3.1 hold and $\boldsymbol{X}$ be such that (3.3) holds. Further, let $I \subset I_{\mathrm{MZ}}$ and $b > 1 + \frac{1}{|I|}$. Then* PlainBSS-*subsampling (cf.* [1, Algorithm 3]*) the points $\boldsymbol{X}$, we obtain points $\boldsymbol{X}' = (\boldsymbol{x}^i)_{i \in J'} \subset \boldsymbol{X}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that there holds the subsampled left L_2-MZ inequality*

$$\frac{(b-1)^3}{178\,(b+1)^2}A\|f\|_{L_2}^2 \leq \frac{1}{|I|} \sum_{i \in J'} \frac{\omega_i}{\varrho_i} |f(\boldsymbol{x}^i)|^2 \quad \text{for all} \quad f \in V\,.$$

Proof. The result is an immediate consequence of applying [1, Corollary 4.5] to the randomly subsampled points of Theorem 3.1. ■

4. Sampling recovery in reproducing kernel Hilbert spaces

In this section we tackle the same problem as in Section 2, namely function recovery from samples, but now for infinite-dimensional function spaces. We consider functions from separable reproducing kernel Hilbert spaces $H(K) \subset \mathbb{C}^D$ with Hermitian positive definite kernels $K(\cdot,\cdot)$ on $D \times D$, where $D \subset \mathbb{R}^d$ is some domain. Fundamental is the reproducing property

$$f(\boldsymbol{x}) = \langle f, K(\cdot,\boldsymbol{x})\rangle_{H(K)}$$

for all $\boldsymbol{x} \in D$, which implies that point evaluations are continuous functionals on $H(K)$. Introductory information about these spaces can be found in [3, Chapter 4] or [6, Chapter 1]. In the framework of this paper, we assume the kernels to be bounded, i.e.,

$$(4.1) \qquad \|K\|_{\ell_\infty(D)} := \sqrt{\sup_{\boldsymbol{x} \in D} K(\boldsymbol{x},\boldsymbol{x})} < \infty\,.$$

With $\|\cdot\|_{H(K)}^2 = \langle\cdot,\cdot\rangle_{H_K}$, this condition implies $\|f\|_{\ell_\infty(D)} \leq \|K\|_{\ell_\infty(D)} \cdot \|f\|_{H(K)}$, which, in other words, is the continuous embedding of $H(K)$ into the space of essentially bounded functions $\ell_\infty(D)$. Furthermore, we assume the finite trace property

$$(4.2) \qquad \|K\|_2^2 := \int_D K(\boldsymbol{x},\boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}) < \infty\,,$$

which, for finite measures, follows from (4.1). This integrability condition ensures the compactness of the embedding operator $\mathrm{Id}\colon H(K) \hookrightarrow L_2$ which is Hilbert-Schmidt. Now we consider the compact and self-adjoint linear operator $W_\nu = \mathrm{Id}^* \circ \mathrm{Id}\colon H(K) \to H(K)$. Applying the spectral theorem, we obtain the representation

$$W_\nu f = \sum_{k=1}^\infty \lambda_k \langle f, e_k\rangle_{H(K)} e_k\,,$$

where $(\lambda_k)_{k=1}^\infty$ is the non-increasing rearrangement of the eigenvalues of W_ν and $(e_k)_{k=1}^\infty \subset H(K)$ the system of eigenfunctions of W_ν which form an orthonormal system in $H(K)$. Using the reproducing property, we obtain a representation of the kernel in terms of the eigenfunctions

$$(4.3) \qquad K(\boldsymbol{x},\boldsymbol{y}) = \sum_{k\in\mathbb{N}} \langle K(\boldsymbol{x},\cdot), e_k\rangle_{H(K)} e_k(\boldsymbol{y}) = \sum_{k\in\mathbb{N}} \overline{e_k(\boldsymbol{x})} e_k(\boldsymbol{y})\,.$$

Further, we have

$$\langle e_k, e_l\rangle_{L_2} = \langle \mathrm{Id}(e_k), \mathrm{Id}(e_l)\rangle_{L_2} = \langle W_\nu(e_k), e_l\rangle_{H(K)} = \lambda_k \langle e_k, e_l\rangle_{H(K)} = \lambda_k \delta_{k,l}\,.$$

Therefore, with $\eta_k(\boldsymbol{x}) := \sigma_k^{-1} e_k(\boldsymbol{x})$ where $\sigma_k = \sqrt{\lambda_k}$ are the singular numbers of Id, we obtain an orthonomal system in L_2.

Remark 4.1. *Let* $V = \mathrm{span}\{\eta_k\}_{k\in I}$, η_k L_2*-orthonormal, and*

$$(4.4) \qquad P_I f(\boldsymbol{x}) = \sum_{k\in I} \langle f, \eta_k\rangle_{L_2} \eta_k(\boldsymbol{x})$$

the L_2*-projection onto* V*. Then*

$$\sup_{\|f\|_{H(K)}\leq 1} \|f - P_I f\|_{\ell_\infty(D)}^2 = \sup_{\boldsymbol{x}\in D} \sup_{\|f\|_{H(K)}\leq 1} \left|\sum_{k\notin I} \langle f, e_k\rangle_{H(K)} e_k(\boldsymbol{x})\right|^2$$

$$= \sup_{\boldsymbol{x}\in D} \sup_{\|\boldsymbol{a}\|_2^2\leq 1} \left|\sum_{k\notin I} a_k e_k(\boldsymbol{x})\right|^2\,.$$

Interpreting this as a functional on $\ell_2(\mathbb{N}\setminus I)$ *applied to* $\boldsymbol{a} = (a_k)_{k\notin I}$*, we obtain that the above equals the norm of the Riesz representer* $(e_k(\boldsymbol{x}))_{k\notin I}$*. Thus,*

$$\sup_{\|f\|_{H(K)}\leq 1} \|f - P_I f\|_{\ell_\infty(D)}^2 = \sup_{\boldsymbol{x}\in D} \sum_{k\notin I} |e_k(\boldsymbol{x})|^2\,.$$

4.1. **Sampling recovery with L_2-MZ points.** A reconstruction model using finitely many samples cannot reconstruct every function in an infinite-dimensional function space. Thus, we make an error. In particular, for least squares we reconstruct functions $f \in V = \mathrm{span}\{\eta_k\}_{k\in I_{\mathrm{MZ}}}$ but none in $H(K)\setminus V$. In this section we quantify the worst-case error of the least-squares method using samples in L_2-MZ points. Similar results can be found in [16].

We consider the following set of assumptions on the initial points $\boldsymbol{X}_{\mathrm{MZ}}$ and weights $\boldsymbol{W}_{\mathrm{MZ}}$ for the theorems of this section.

Assumption 4.2. *Let $H(K)$ be a separable reproducing kernel Hilbert space with a finite trace (4.2) and bounded (4.1) kernel. Further, let λ_k and η_k denote the eigenvalues and $L_2(\nu)$-orthonormal eigenfunctions of the operator $W_\nu = \mathrm{Id}^* \circ \mathrm{Id}$ with $\mathrm{Id}\colon H(K) \hookrightarrow L_2(\nu)$ the embedding.*

For $V = \mathrm{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}} \subset H(K)$ a finite-dimensional function space, we assume the points $\boldsymbol{X}_{\mathrm{MZ}} \subset D$ and weights $\boldsymbol{W}_{\mathrm{MZ}} \in [0,\infty)^{M \times M}$ fulfill an L_2-MZ inequality for V with constants A and B.

Theorem 4.3. *Let Assumptions 4.2 hold. Then the error of the reconstruction operator $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} = S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}}(\boldsymbol{W}_{\mathrm{MZ}})$ is zero for functions in $V = \mathrm{span}\{\eta_k\}_{k \in I_{\mathrm{MZ}}}$ and otherwise bounded as follows:*

$$\sup_{\|f\|_{H(K)} \leq 1} \left\| f - S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f \right\|_{L_2}^2 \leq \sup_{k \notin I_{\mathrm{MZ}}} \lambda_k + \frac{\sum_{i=1}^M \omega_i}{A} \sup_{\|f\|_{H(K)} \leq 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 .$$

Proof. The reconstructing property for functions in V is immediate. For functions $f \in H(K)$ we use orthogonality of the projection (4.4) to obtain

$$(4.5) \qquad \|f - S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f\|_{L_2}^2 = \|f - P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 + \|P_{I_{\mathrm{MZ}}} f - S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f\|_{L_2}^2$$

of which we now estimate each summand individually.

Step 1. We bound the first summand of (4.5) by

$$\|f - P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 = \left\| \sum_{k \notin I_{\mathrm{MZ}}} \langle f, e_k \rangle_{H(K)} e_k \right\|_{L_2}^2 = \sum_{k \notin I_{\mathrm{MZ}}} \lambda_k |\langle f, e_k \rangle_{H(K)}|^2$$
$$\leq \|f\|_{H(K)}^2 \sup_{k \notin I_{\mathrm{MZ}}} \lambda_k .$$

Step 2. For the second summand of (4.5) we use the reconstructing property of $S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}}$ for functions in V

$$\|P_{I_{\mathrm{MZ}}} f - S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f\|_{L_2}^2 = \|S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} (P_{I_{\mathrm{MZ}}} f - f)\|_{L_2}^2$$
$$\leq \left\| ((\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^{-1} (\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^* \boldsymbol{W}_{\mathrm{MZ}}^{1/2} \right\|_{2 \to 2}^2 \left\| \boldsymbol{W}_{\mathrm{MZ}}^{1/2} \left((P_{I_{\mathrm{MZ}}} f - f)(\boldsymbol{x}^i) \right)_{i=1}^M \right\|_2^2 .$$

By Lemmata 2.1 and 2.2 we have for the first factor

$$\left\| ((\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^* \boldsymbol{W}_{\mathrm{MZ}} \boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^{-1} (\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}})^* \boldsymbol{W}_{\mathrm{MZ}}^{1/2} \right\|_{2 \to 2}^2 \leq \frac{1}{A} .$$

Finally, we obtain

$$\sup_{\|f\|_{H(K)} \leq 1} \|P_{I_{\mathrm{MZ}}} f - S_{I_{\mathrm{MZ}}}^{\boldsymbol{X}_{\mathrm{MZ}}} f\|_{L_2}^2 \leq \frac{1}{A} \sup_{\|f\|_{H(K)} \leq 1} \sum_{i=1}^M \omega_i |(P_{I_{\mathrm{MZ}}} f - f)(\boldsymbol{x}^i)|^2$$
$$\leq \frac{\sum_{i=1}^M \omega_i}{A} \sup_{\|f\|_{H(K)} \leq 1} \|P_{I_{\mathrm{MZ}}} f - f\|_{\ell_\infty(D)}^2 . \qquad \blacksquare$$

The above are rather basic bounds which hold in a general setting. In special cases like the exponential functions, Chebyshev polynomials, and the half-period cosine functions this topic is examined in detail in [32] for rank-1 lattices. In particular for function spaces with mixed smoothness there exist almost tight error bounds in [4] which improve on Theorem 4.3.

4.2. **Sampling recovery with subsampled L_2-MZ points.** The next step is to subsample the L_2-MZ points like in Section 3 whilst paying attention to the approximation error. With that we obtain better error bounds and will relate the error using the subsampled points to the error of the initial point set. In contrast to Section 2, we subsample two frames and one Bessel sequence simultaneously from the points $\boldsymbol{X}_{\mathrm{MZ}}$. In order to control the subsampling we use a change of measure via a convex combination of the respective probability densities. Each point $\boldsymbol{x}^i$ from $\boldsymbol{X}_{\mathrm{MZ}}$ will be drawn with the following probability density:

$$(4.6) \qquad \varrho_i = \frac{\omega_i \sum_{k\in I} |\eta_k(\boldsymbol{x}^i)|^2}{3 \sum_{j=1}^{M} \omega_j \sum_{k\in I} |\eta_k(\boldsymbol{x}^j)|^2} + \frac{\omega_i \sum_{k\in I_{\mathrm{MZ}}\setminus I} |e_k(\boldsymbol{x}^i)|^2}{3 \sum_{j=1}^{M} \omega_j \sum_{k\in I_{\mathrm{MZ}}\setminus I} |e_k(\boldsymbol{x}^j)|^2} + \frac{\omega_i}{3} \, .$$

Since $\varrho_i \geq 0$ and $\sum_{i=1}^{M} \varrho_i = 1$, these are proper density weights. This is based on an idea first used in [28].

For the later analysis we need some preparations, starting with a concentration inequality for random infinite matrices from [36].

Proposition 4.4 ([36, Proposition 3.8]). *Let $\boldsymbol{u}^1, \ldots, \boldsymbol{u}^n$ be i.i.d. random sequences from $\ell_2(\mathbb{N})$. Let further $n \geq 3$, $t > 0$, $M > 0$ such that $\|\boldsymbol{u}^i\|_2 \leq M$ almost surely and $\mathbb{E}(\boldsymbol{u}^i \otimes \boldsymbol{u}^i) = \boldsymbol{\Lambda}$ for all $i = 1, \ldots, n$. Then*

$$\Big\| \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{u}^i \otimes \boldsymbol{u}^i \Big\|_{2\to2} \leq \frac{21(\log(n)+t)}{n} M^2 + 2\|\boldsymbol{\Lambda}\|_{2\to2}$$

with probability exceeding $1 - 2^{3/4} \exp(-t)$.

Lemma 4.5. *Let Assumptions 4.2 hold and let $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i\in J}$, $|J| = n$, be points drawn i.i.d. from $\boldsymbol{X}_{\mathrm{MZ}}$ with respect to the discrete density weights ϱ_i, defined in (4.6). Further, for $I \subset I_{\mathrm{MZ}}$, we define the matrices*

$$\boldsymbol{W} = \mathrm{diag}(\omega_i/\varrho_i)_{i\in J} \quad and \quad \boldsymbol{\Phi}_{\boldsymbol{X}}^{I'} = (e_k(\boldsymbol{x}^i))_{i\in J, k\in I_{\mathrm{MZ}}\setminus I} \, .$$

Then we have with probability exceeding $1 - 2^{3/4} \exp(-t)$

$$\frac{1}{n} \Big\| \boldsymbol{W}^{1/2} \boldsymbol{\Phi}_{\boldsymbol{X}}^{I_{\mathrm{MZ}}\setminus I} \Big\|_{2\to2}^2 \leq \frac{63(\log(n)+t)}{n} B \sum_{k\in I_{\mathrm{MZ}}\setminus I} \lambda_k + 2B \sup_{k\notin I} \lambda_k \, .$$

Proof. Since points with probability $\varrho_i = 0$ will not be drawn, we assume $\varrho_i > 0$. Let

$$\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}\setminus I} = \Big(\eta_k(\boldsymbol{x}^i) \Big)_{\boldsymbol{x}^i \in \boldsymbol{X}_{\mathrm{MZ}}, k\in I_{\mathrm{MZ}}\setminus I} = \boldsymbol{\Phi}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}}\setminus I} \mathrm{diag}(\sigma_k^{-1})_{k\in I_{\mathrm{MZ}}\setminus I}.$$

We want to apply Proposition 4.4. To this end we define $\boldsymbol{u}^i = \sqrt{\omega_i/\varrho_i}(e_k(\boldsymbol{x}^i))_{k\in I_{\mathrm{MZ}}\setminus I}$. Note,

$$\sum_{i\in J} \boldsymbol{u}^i \otimes \boldsymbol{u}^i = \Big(\boldsymbol{\Phi}_{\boldsymbol{X}}^{I_{\mathrm{MZ}}\setminus I} \Big)^* \boldsymbol{W} \boldsymbol{\Phi}_{\boldsymbol{X}}^{I_{\mathrm{MZ}}\setminus I} \, .$$

Now we estimate $\|\boldsymbol{u}^i\|_2^2$ and $\|\mathbb{E}\boldsymbol{u}^i \otimes \boldsymbol{u}^i\|_{2\to2}$.

We bound $\|\boldsymbol{u}^i\|_2$ by using the L_2-MZ inequality and $\|e_k\|_{L_2}^2 = \lambda_k$

$$\|\boldsymbol{u}^i\|_2^2 \leq 3 \sum_{k \in I_{\mathrm{MZ}} \setminus I} \sum_{j=1}^M \omega_j |e_k(\boldsymbol{x}^j)|^2 \leq 3B \sum_{k \in I_{\mathrm{MZ}} \setminus I} \|e_k\|_{L_2}^2 = 3B \sum_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k \, .$$

We further have by the compatibility of the spectral norm and Lemma 2.2

$$
\begin{aligned}
\left\| \mathbb{E} \boldsymbol{u}^i \otimes \boldsymbol{u}^i \right\|_{2 \to 2} &= \left\| \boldsymbol{W}_{\mathrm{MZ}}^{1/2} \boldsymbol{\Phi}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}} \setminus I} \right\|_{2 \to 2}^2 \\
&\leq \left\| \boldsymbol{W}_{\mathrm{MZ}}^{1/2} \boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}^{I_{\mathrm{MZ}} \setminus I} \right\|_{2 \to 2}^2 \left\| \mathrm{diag}(\sigma_k)_{I_{\mathrm{MZ}} \setminus I} \right\|_{2 \to 2}^2 \\
&\leq B \sup_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k \, ,
\end{aligned}
$$

which gives the assertion after applying Proposition 4.4. ∎

We now formulate a central result of the paper, which bounds the worst-case reconstruction error for the least squares method where the points are drawn randomly from a discrete set of points fulfilling an L_2-MZ inequality, cf. middle of Figure 1.1.

Theorem 4.6. *Let Assumption 4.2 hold and $I \subset I_{\mathrm{MZ}} \subset \mathbb{N}$, $|I| \geq 3$, be an index set. For $n \in \mathbb{N}$, $t > 0$, and $r \geq 1$ such that*

$$n := \left\lceil \frac{36B}{A} |I|(\log |I| + t) \right\rceil \leq |I|^r,$$

let $\boldsymbol{X} = (\boldsymbol{x}^i)_{i \in J}$, $|J| = n$, be points drawn i.i.d. from $\boldsymbol{X}_{\mathrm{MZ}}$ with respect to the discrete density weights ϱ_i, defined in (4.6). Then we have with probability exceeding $1 - 4\exp(-t)$

$$
\sup_{\|f\|_{H(K)} \leq 1} \|f - S_I^{\boldsymbol{X}} f\|_{L_2}^2 \leq \frac{9B}{A} \sup_{k \notin I} \lambda_k + \frac{7r}{|I|} \sum_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k
$$

$$
+ \frac{12}{A} \sup_{\|f\|_{H(K)} \leq 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 \, ,
$$

where $\boldsymbol{W} = \mathrm{diag}(\omega_i / \varrho_i)_{i \in J}$ and $S_I^{\boldsymbol{X}} = S_I^{\boldsymbol{X}}(\boldsymbol{W})$.

Proof. Without loss of generality let $\varrho_i > 0$. We begin by defining two events. The first one bounds the spectral norm of the least squares matrix

$$S_1 = \left\{ \boldsymbol{X} \subset \boldsymbol{X}_{\mathrm{MZ}} : \left\| \left(\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L} \right)^{-1} \boldsymbol{L}^* \boldsymbol{W}^{1/2} \right\|_{2 \to 2}^2 \leq \frac{2}{An} \right\}.$$

Using Lemma 2.1, we have, that this is equivalent to a lower bound of the singular values of $\boldsymbol{W}^{1/2} \boldsymbol{L}$. Via Lemma 2.2 we use the L_2-MZ characterization and the assumption on n in order to apply Theorem 3.1 with $C = 1/3$ and obtain that the above holds with probability $1 - 2\exp(-t)$.

For $\boldsymbol{Y} \in \{\boldsymbol{X}_{\mathrm{MZ}}, \boldsymbol{X}\}$ we define the matrix

$$\boldsymbol{\Phi}_{\boldsymbol{Y}}^{I_{\mathrm{MZ}} \setminus I} = \left(e_k(\boldsymbol{x}^i) \right)_{\boldsymbol{x}^i \in \boldsymbol{Y}, k \in I_{\mathrm{MZ}} \setminus I}.$$

Event S_2 is the inequalities of Lemma 4.5 which holds with probability at least $1 - 2^{3/4} \exp(-t)$. Together we obtain the desired probability

$$\mathbb{P}(S_1 \cap S_2) \geq 1 - \mathbb{P}S_1^{\complement} - \mathbb{P}S_2^{\complement} \geq 1 - 4\exp(-t) \, .$$

It remains to show the bound based on the events S_1 and S_2. On that account we decompose the recovery error using triangle inequality

$$\|f - S_I^{\boldsymbol{X}} f\|_{L_2}^2 \le \|f - P_I f\|_{L_2}^2 + \|P_I f - S_I^{\boldsymbol{X}} f\|_{L_2}^2$$

$$(4.7) \qquad \le \|f - P_I f\|_{L_2}^2 + 2\|P_I f - S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 + 2\|S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f - S_I^{\boldsymbol{X}} f\|_{L_2}^2 \,.$$

In the following we estimate each of the three summands individually.

Step 1. We estimate the first summand of (4.7) in the same manner as in the first step of the proof of Theorem 4.3:

$$\|f - P_I f\|_{L_2}^2 \le \|f\|_{H(K)}^2 \sup_{k \notin \mathcal{I}} \lambda_k \,.$$

Step 2. Using the invariance of $S_I^{\boldsymbol{X}}$ on $P_I f$, we estimate the second summand of (4.7) by

$$\left\| P_I f - S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f \right\|_{L_2}^2 = \left\| S_I^{\boldsymbol{X}} (P_{I_{\mathrm{MZ}}} f - P_I f) \right\|_{L_2}^2$$

$$= \left\| (\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{W} ((P_{I_{\mathrm{MZ}}} f - P_I f)(\boldsymbol{x}^i))_{i \in J} \right\|_2^2$$

$$\le \left\| (\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{W}^{1/2} \right\|_{2 \to 2}^2 \sum_{i \in J} \frac{\omega_i}{\varrho_i} |(P_{I_{\mathrm{MZ}}} f - P_I f)(\boldsymbol{x}^i)|^2 \,.$$

We estimate the first factor by the means of event S_1 and write the latter in terms of its coefficients in order to apply event S_2

$$\left\| S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f - S_I^{\boldsymbol{X}} f \right\|_{L_2}^2 \le \frac{2}{An} \left\| \boldsymbol{W}^{1/2} \boldsymbol{\Phi}_{I_{\mathrm{MZ}} \setminus I}^{\boldsymbol{X}} \left(\langle f, e_k \rangle_{H(K)} \right)_{k \in I_{\mathrm{MZ}} \setminus I} \right\|_2^2$$

$$\le \frac{2}{An} \left\| \boldsymbol{W}^{1/2} \boldsymbol{\Phi}_{I_{\mathrm{MZ}} \setminus I}^{\boldsymbol{X}} \right\|_{2 \to 2}^2 \sum_{k \in I_{\mathrm{MZ}} \setminus I} |\langle f, e_k \rangle_{H(K)}|^2$$

$$\le \left(\frac{126 B (\log(n) + t)}{An} \sum_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k + \frac{4B}{A} \sup_{k \notin I} \lambda_k \right) \|f\|_{H(K)}^2 \,.$$

Step 3. We start estimating the third summand of (4.7) analogously to Step 2

$$\|P_I f - S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 \le \frac{2}{An} \sum_{i \in J} \frac{\omega_i}{\varrho_i} |(f - P_{I_{\mathrm{MZ}}} f)(\boldsymbol{x}^i)|^2 \,.$$

Now we use the third part of ϱ_i, to obtain

$$\|P_I f - S_I^{\boldsymbol{X}} P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 \le \frac{6}{An} \sum_{i \in J} |(f - P_{I_{\mathrm{MZ}}} f)(\boldsymbol{x}^i)|^2$$

$$\le \frac{6}{A} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 \,.$$

Overall, applying the estimates $1. - 3.$ in (4.7) we obtain

$$\sup_{\|f\|_{H(K)} \le 1} \|f - S_I^{\boldsymbol{X}} f\|_{L_2}^2 \le \sup_{k \notin I} \lambda_k + \frac{252 B (\log(n) + t)}{An} \sum_{k \in I_{\mathrm{MZ}} \setminus I} \lambda_k + \frac{8B}{A} \sup_{k \notin I} \lambda_k$$

$$+ \frac{12}{A} \sup_{\|f\|_{H(K)} \le 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 \,.$$

By the assumption on $|I|$ we have

$$\frac{252B(\log(n)+t)}{An} \leq \frac{7(\log(n)+t)}{|I|(\log(|I|)+t)} \leq \frac{7r(\log(|I|)+t)}{|I|(\log(|I|)+t)} = \frac{7r}{|I|} \leq 3\,,$$

and obtain the assertion. ∎

Remark 4.7. *(i) Given a bounded orthonormal system, i.e., $\|\eta_k\|_{\ell_\infty(D)} = C < \infty$ for all k we can alter events S_2 and S_3 such that we won't need to sample with respect to the density weights ϱ_i and sample with respect to the weights ω_i coming from the L_2-MZ inequality directly.*

(ii) Whenever the Christoffel-function $\sum_{k\in I} |\eta_k(\boldsymbol{x})|^2$ is constant, independent of the underlying index set I, the three summands in density weights ϱ_i in (4.6) coincide and the number of random points may be divided by three whilst achieving the same error bound.

(iii) Up to the quotient of the constants for the L_2-MZ inequality, the first two summands of Theorem 4.6 are the same as in [25] where points were drawn with respect to a continuous probability measure. The difference is in the latter summand, which only depends on the initial point set. By choosing a suitable initial point set, i.e., a point set satisfying a L_2-MZ inequality for large enough I_{MZ}, we can make this as small as needed. In particular smaller than the first two terms, which, therefore, determine the error decay behaviour.

Next, we use the unweighted frame subsampling from [1] to prove the main result which lowers the number of points to be linear in $|I|$, cf. right of Figure 1.1.

Theorem 4.8. *Let Assumption 4.2 hold and $I \subset I_{\mathrm{MZ}} \subset \mathbb{N}$, $|I| \geq 3$ be an index set. For $n \in \mathbb{N}$, $t > 0$, and $r \geq 1$ such that*

$$n := \left\lceil \frac{36B}{A}|I|(\log|I|+t) \right\rceil \leq |I|^r,$$

let $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i\in J}$, $|J| = n$ be points drawn i.i.d. from $\boldsymbol{X}_{\mathrm{MZ}}$ with respect to the discrete density weights ϱ_i, defined in (4.6). For $b > 1 + \frac{1}{|I|}$, $\texttt{PlainBSS}$ subsampling (cf. [1, Algorithm 3]) the points $\boldsymbol{X}$, we obtain points $\boldsymbol{X}' = \{\boldsymbol{x}^i\}_{i\in J'} \subset \boldsymbol{X}$ with $|\boldsymbol{X}'| \leq \lceil b|I| \rceil$ such that we have with probability exceeding $1 - 4\exp(-t)$

$$\sup_{\|f\|_{H(K)}\leq 1} \|f - S_I^{\boldsymbol{X}'} f\|_{L_2}^2$$

$$\leq 39\,808\, \frac{B}{A} \frac{(b+1)^2}{(b-1)^3} (\log|I|+t) \left(\frac{B}{A} \sup_{k\notin I} \lambda_k + \frac{r}{|I|} \sum_{k\in I_{\mathrm{MZ}}\setminus I} \lambda_k \right.$$

$$\left. + \frac{1}{A} \sup_{\|f\|_{H(K)}\leq 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty(D)}^2 \right)$$

with weights $\boldsymbol{W}' = \mathrm{diag}(\omega_i/\varrho_i)_{i\in J'}$ and the least squares approximation $S_I^{\boldsymbol{X}'} f = S_I^{\boldsymbol{X}'}(\boldsymbol{W}')f$ defined in (2.3).

Proof. For $g \in V$ we have

$$\|S_I^{\boldsymbol{X}'} g\|_{L_2}^2 \leq \left\| ((\boldsymbol{L}')^*\boldsymbol{W}'\boldsymbol{L}')^{-1}(\boldsymbol{L}')^*(\boldsymbol{W}')^{1/2} \right\|_{2\to2}^2 \sum_{i\in J'} \frac{\omega_i}{\varrho_i}|g(\boldsymbol{x}^i)|^2\,.$$

By Theorem 3.4 and Lemma 2.1 we have

$$
\begin{aligned}
\|S_I^{\boldsymbol{X}'} g\|_{L_2}^2 &\leq \frac{89(b+1)^2}{(b-1)^3} \frac{n}{|I|} \frac{2}{An} \sum_{i \in J'} \frac{\omega_i}{\varrho_i} |g(\boldsymbol{x}^i)|^2 \\
&\leq \frac{3234(b+1)^2 B}{(b-1)^3 A} (\log(|I|) + t) \Big(\frac{2}{An} \sum_{i \in J'} \frac{\omega_i}{\varrho_i} |g(\boldsymbol{x}^i)|^2 \Big),
\end{aligned}
$$

where $|I| \geq 3$ was used in

$$
n = \Big\lceil \frac{36B}{A} |I| (\log |I| + t) \Big\rceil \leq \frac{109B}{3A} |I| (\log |I| + t).
$$

Using this estimate in step 2 and 3 of the proof of Theorem 4.6 we obtain the assertion. ∎

We payed the logarithmic factor in the bound to work with linearly many points whilst achieving the same error bound as in [37]. Recent progress has shown that the logarithmic factor can be avoided but this utilizes the Kadison-Singer theorem and is not constructive, cf. [13].

5. Application: Sampling recovery on $\mathbb{T}^d$ with rank-1 lattices

In this section we apply our general theory from Section 4.2 to the d-dimensional torus $\mathbb{T}^d \simeq (\mathbb{R}/\mathbb{Z})^d$. First we deal with function spaces on $\mathbb{T}^d$. To begin with, we consider a weight function $w(\mathbf{k})\colon \mathbb{Z}^d \to (0,\infty)$ for $\mathbf{k} \in \mathbb{Z}^d$ and define the kernel

$$
(5.1) \qquad K_w(\mathbf{x}, \mathbf{y}) := \sum_{\mathbf{k} \in \mathbb{Z}^d} \frac{\exp(2\pi i \langle \mathbf{k}, (\mathbf{y} - \mathbf{x}) \rangle)}{w(\boldsymbol{k})^2} \quad \text{for} \quad \mathbf{x}, \mathbf{y} \in \mathbb{T}^d,
$$

which is a reproducing kernel for $H^w := H(K_w)$. The kernel defined in (5.1) is associated to the inner product

$$
(5.2) \qquad \langle f, g \rangle_{H^w} := \sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}_{\boldsymbol{k}} \overline{\hat{g}_{\boldsymbol{k}}} w(\boldsymbol{k})^2
$$

and the corresponding norm. These definitions make sense if $\sum_{\boldsymbol{k} \in \mathbb{Z}} w(\boldsymbol{k})^{-2} < \infty$.

An interesting case are Sobolev spaces with mixed smoothness. For $s \in \mathbb{N}$ we define the space $H_{\text{mix}}^s = H_{\text{mix}}^s(\mathbb{T}^d)$ as the Hilbert space with the inner product

$$
(5.3) \qquad \langle f, g \rangle_{H_{\text{mix}}^s} := \sum_{\boldsymbol{j} \in \{0,s\}^d} \langle D^{(\boldsymbol{j})} f, D^{(\boldsymbol{j})} g \rangle_{L_2}, \quad \text{with} \quad D^{(\boldsymbol{j})} = \frac{\partial^{\|j\|_1}}{\partial x_1^{j_1} \cdots \partial x_d^{j_d}}.
$$

In the univariate case the kernel equals

$$
K_s^1(x, y) := \sum_{k \in \mathbb{Z}} \frac{\exp(2\pi i k(y - x))}{w_s(k)^2} \quad \text{with} \quad w_s(k) = \Big(1 + (2\pi |k|)^{2s} \Big)^{1/2}
$$

for $k \in \mathbb{Z}$ and $x, y \in \mathbb{T}$. This directly leads to the d-dimensional kernel

$$
K_s^d(\boldsymbol{x}, \boldsymbol{y}) := K_s^1(x_1, y_1) \otimes \cdots \otimes K_s^1(x_d, y_d) \quad \text{for} \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{T}^d.
$$

This construction of the kernel extends to $s \in \mathbb{R}$, $s > 1/2$ naturally. These function spaces are suitable for high-dimensional approximation. By collecting the most important frequencies with respect to the singular values we end up with a special structure for the frequency index set, namely a hyperbolic cross. Detailed information on approximation in these spaces can be found in [43, 10].

Next, we deal with L_2-MZ inequalities on the torus. In particular we are interested in those coming from exact quadrature rules, cf. Lemma 2.3. The canonical example would be equispaced points

$$\boldsymbol{X}_{\mathrm{MZ}} = \Big\{ \frac{1}{\sqrt[d]{M}} \boldsymbol{i} : \boldsymbol{i} \in \{1, \ldots, \sqrt[d]{M}\}^d \Big\}$$

(assuming $\sqrt[d]{M}$ is integer) which fulfill an L_2-MZ inequality with equal weights $\omega_i = 1/M$ and $A = B$ for the frequency index set $I_{\mathrm{MZ}} = \{-\sqrt[d]{M}/2, \ldots, \sqrt[d]{M}/2 - 1\}^d$. However, these points are unfeasible in high-dimensional approximation if we do not subsample. Another candidate for quadrature rules suitable for Sobolev spaces with mixed smoothness are rank-1 lattices

$$\boldsymbol{X}_{\mathrm{MZ}} := \Big\{ \frac{1}{M} (i\boldsymbol{z} \bmod M\mathbb{1}) \in \mathbb{T}^d : i = 0, \ldots, M - 1 \Big\}$$

which were studied in [42, 21, 23, 39, 11]. We say a rank-1 lattices $\boldsymbol{X}_{\mathrm{MZ}} = \{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\}$ has a reconstructing property for a frequency index set I_{MZ} when

$$(5.4) \qquad \frac{1}{M} \sum_{i=1}^{M} \exp(2\pi\mathrm{i}\langle \boldsymbol{k} - \boldsymbol{l}, \boldsymbol{x}^i \rangle) = \delta_{\boldsymbol{l},\boldsymbol{k}} \,,$$

for all $\boldsymbol{k}, \boldsymbol{l} \in I_{\mathrm{MZ}}$, which is equivalent to an L_2-MZ inequality (2.4) for $A = B$ with equal weights ω_i. There exist algorithms which, given a frequency index set I_{MZ} and M, compute a generating vector $\boldsymbol{z}$ such that the rank-1 lattice $\boldsymbol{X}_{\mathrm{MZ}}$ has this reconstructing property, cf. [21, 32]. In particular, the probabilistic approach recently presented in [22] seems to be particularly suitable for downstream subsampling. On the one hand the component-by-component (CBC) rank-1 lattice construction is extremely efficient with computational costs that are linear in $|I_{\mathrm{MZ}}|$ up to a few logarithmic factors. On the other hand, the sizes M of the resulting rank-1 lattices might be slightly (but only up to a factor of two with very high probability) larger than those resulting from the deterministic approaches. In summary, the CBC construction costs are not a major concern, and random subsampling the rank-1 lattice reduces the number of samples actually used to a number that does not depend on the slightly too large M.

Next, we consider asymptotic approximation results for rank-1 lattices in Sobolev spaces with dominating mixed smoothness from [4]. For simplicity we omit constants in the following. For two sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty} \subset \mathbb{R}$ we write $a_n \lesssim b_n$ if there exists a constant $c > 0$ such that $a_n \leq cb_n$ for all n. We will write $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$.

Theorem 5.1 ([30, 5, 4] Approximation with rank-1 lattices in H_{mix}^s). *Let $I_{\mathrm{MZ}} \subset \mathbb{N}$ be a hyperbolic cross frequency index set, i.e.,*

$$(5.5) \qquad I_{\mathrm{MZ}} = \Big\{ \boldsymbol{k} \in \mathbb{Z}^d : \prod_{j=1}^{d} \max\Big\{1, \frac{|k_j|}{\gamma}\Big\} \leq R \Big\}$$

for some $\gamma > 0$, $R \in (1, \infty)$, and $\boldsymbol{X}_{\mathrm{MZ}}$ a corresponding reconstructing rank-1 lattice with M points, cf. (5.4).

(i) The size of a hyperbolic cross is asymptotically $|I_{\mathrm{MZ}}| \asymp R(\log R)^{d-1}$.

(ii) The best possible error of functions in H_{mix}^s behaves as follows

$$\sup_{\|f\|_{H_{\mathrm{mix}}^s} \leq 1} \|f - P_{I_{\mathrm{MZ}}} f\|_{L_2}^2 \asymp |I_{\mathrm{MZ}}|^{-2s} (\log |I_{\mathrm{MZ}}|)^{2(d-1)s}$$

and

$$\sup_{\|f\|_{H^s_{\mathrm{mix}}} \le 1} \|f - P_{I_{\mathrm{MZ}}} f\|^2_{\ell_\infty(\mathbb{T}^d)} \asymp |I_{\mathrm{MZ}}|^{-2s+1/2} (\log |I_{\mathrm{MZ}}|)^{2(d-1)s} .$$

(iii) *The size of a reconstructing rank-1 lattice for I_{MZ} is bounded by*

$$|I_{\mathrm{MZ}}|^2 (\log |I_{\mathrm{MZ}}|)^{-2(d-1)} \lesssim M \lesssim |I_{\mathrm{MZ}}|^2 (\log |I_{\mathrm{MZ}}|)^{-d} ,$$

where the lower inequality holds for all reconstructing rank-1 lattices and there exists a rank-1 lattice fulfilling the upper one.

(iv) *The error of the least squares operator using rank-1 lattices is bounded as follows:*

$$M^{-s} \lesssim \sup_{\|f\|_{H^s_{\mathrm{mix}}} \le 1} \|f - S^{\boldsymbol{X}_{\mathrm{MZ}}}_{I_{\mathrm{MZ}}} f\|^2_{L_2} \lesssim \begin{cases} |I_{\mathrm{MZ}}|^{-2s} (\log |I_{\mathrm{MZ}}|)^{(d-1)(2s+1)} \\ M^{-s} (\log M)^{(d-2)s+d-1} , \end{cases}$$

where the lower inequality holds for all reconstructing rank-1 lattices and there exists a rank-1 lattice fulfilling the upper one.

Proof. There are different definitions for the hyperbolic cross and the Sobolev spaces with mixed smoothness. Up to constants the considered quantities coincide, cf. [30]. The assertion (i) is given in [4, Lemma 2]. To show (ii), we use $\sup_{\|f\|_{H^s_{\mathrm{mix}}} \le 1} \|f - P_{I_{\mathrm{MZ}}} f\|^2_{L_2} = R^{-s}$ and (i). The second part of (ii) is given in [5, Theorem 6.11 (iii)].

Assertion (iii) follows from Lemmata 2 and 3 in [4]. To show (iv), we use [4, Theorem 2] with $\alpha = s$, $\beta = \gamma = 0$ and Lemmata 2 and 3 from [4] again.

Note, that the Fourier coefficients of the approximation in [4] (formula 2.3) are computed by applying the adjoint of the Fourier matrix. Because of the reconstructing property, this coincides with the least squares approximation presented here, i.e., with $\boldsymbol{W} = \mathrm{diag}(1/M, \ldots, 1/M)$ we have

$$S^{\boldsymbol{X}_{\mathrm{MZ}}}_{I_{\mathrm{MZ}}} f = \sum_{\boldsymbol{k} \in I_{\mathrm{MZ}}} a_{\boldsymbol{k}} \exp(2\pi \mathrm{i}\langle \boldsymbol{k}, \boldsymbol{x}\rangle) \quad \text{with} \quad \boldsymbol{a} = (\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L})^{-1} \boldsymbol{L}^* \boldsymbol{W} \boldsymbol{f} = \frac{1}{M} \boldsymbol{L}^* \boldsymbol{f} . \quad \blacksquare$$

Similar bounds were obtained for the kernel method operating on rank-1 lattice points, cf. [20]. For random points one has the rate M^{-2s} and additional logarithmic terms, however. So, using full rank-1 lattices we lose half the rate of convergence in the main order. But we recover the order of convergence by subsampling as stated in Corollary 1.2.

Proof of Corollary 1.2. Since we are dealing with a reconstructing lattice, we have a tight frame $A = B$. We need to apply Theorem 4.8 and Remark 4.7 and use the following two inequalities

$$\sum_{\boldsymbol{k} \notin I} \lambda_{\boldsymbol{k}} = \sum_{\boldsymbol{k} \notin I} \prod_{j=1}^d w_s(k_j)^{-2} \lesssim |I|^{-2s+1} (\log |I|)^{2(d-1)s}$$

from [10, (2.3.2)]. Using Hölder's inequality and the above inequality again, we obtain

$$
\begin{aligned}
\|f - P_{I_{\mathrm{MZ}}} f\|_{\ell_\infty}^2 &\leq \Big| \sum_{k \notin I_{\mathrm{MZ}}} \lambda_{\boldsymbol{k}}^{-1} \lambda_{\boldsymbol{k}} |\hat{f}_{\boldsymbol{k}}| \Big|^2 \\
&\leq \sum_{k \notin I_{\mathrm{MZ}}} \lambda_{\boldsymbol{k}}^{-2} \sum_{k \notin I_{\mathrm{MZ}}} \lambda_{\boldsymbol{k}}^2 |\hat{f}_{\boldsymbol{k}}|^2 \\
&\lesssim |I_{\mathrm{MZ}}|^{-2s+1} (\log |I_{\mathrm{MZ}}|)^{2(d-1)s} \|f\|_{H_{\mathrm{mix}}^s}^2 . \qquad \blacksquare
\end{aligned}
$$

Remark 5.2 (Optimality for rank-1 lattices). *With the initial hyperbolic cross I_{MZ} in Corollary 1.2 slightly bigger than I we recover the optimal error bound $n^{-s}(\log n)^{(d-1)s}$, cf. [10, (2.3.2)], up to a logarithmic factor. Thus, the phrase "This main rate (without logarithmic factors) is half the optimal main rate [...] and turns out to be the best possible among all algorithms taking samples on lattices" from [4] has to be stated more precisely to "...samples on* full *lattices". We conjecture that the non-constructive approach in [13] (based on Kadison-Singer and Weaver subsampling [37]) may lead to a bound without additional logarithm.*

5.1. **Numerical experiments.** Because of the one-dimensional structure of rank-1 lattices, the matrix-vector product with the system matrix $\boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}}$ can be carried out using a one-dimensional fast Fourier transform in $\mathcal{O}(M \log M)$. We may use this algorithm for the subsampled points $\boldsymbol{X} = \{\boldsymbol{x}^i\}_{i \in J}$ and system matrix $\boldsymbol{L}$ too:

$$
(5.6) \qquad \boldsymbol{L} = \boldsymbol{P} \boldsymbol{L}_{\boldsymbol{X}_{\mathrm{MZ}}} \quad \text{where} \quad \boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ & \vdots & & & & \vdots & \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ \vdots \\ n \end{matrix} .
$$
$$
 j_1 \quad\;\; j_2 \qquad\qquad j_n
$$

Here, we use hyperbolic crosses for the frequency index sets with size $M \lesssim |I_{\mathrm{MZ}}|^2 (\log |I_{\mathrm{MZ}}|)^{-d}$, cf. Theorem 5.1 (iii). Using the Fast Fourier Transform this yields the following complexity for the matrix-vector product with the full rank-1 lattice and, subsequently, the subsampled one:

$$
\mathcal{O}(M \log M) = \mathcal{O}(|I_{\mathrm{MZ}}|^2 (\log |I_{\mathrm{MZ}}|)^{1-d})
$$

with the same memory usage. In contrast, if we naively set up the matrix with the $n \sim |I_{\mathrm{MZ}}| \log |I_{\mathrm{MZ}}|$ random points we have a complexity of $\mathcal{O}(|I_{\mathrm{MZ}}|^2 \log |I_{\mathrm{MZ}}|)$ for the number of arithmetic operations and memory usage of the matrix-vector product. If we further use BSS subsampling this reduces to $\mathcal{O}(|I_{\mathrm{MZ}}|^2)$, which is still slower than using the Fast Fourier Transform. In general, whenever we start with a rank-1 lattice with fewer than quadratic points, we gain computation speed.

Experiment 1. For the numerical experiments, we follow [4, 25] and use the five-dimensional kink function

$$
f(\boldsymbol{x}) = f((x_1, \ldots, x_d)^{\mathsf{T}}) = \prod_{j=1}^5 \Big(\frac{5^{3/4} 15}{4\sqrt{3}} \max \Big\{ \frac{1}{5} - \Big(x_j - \frac{1}{2} \Big)^2, 0 \Big\} \Big)
$$

of which we know the exact Fourier coefficients. We remark that $f \in H_{\mathrm{mix}}^{3/2-\varepsilon}(\mathbb{T}^5)$ for $\varepsilon > 0$ and $\|f\|_{L_2} = 1$.
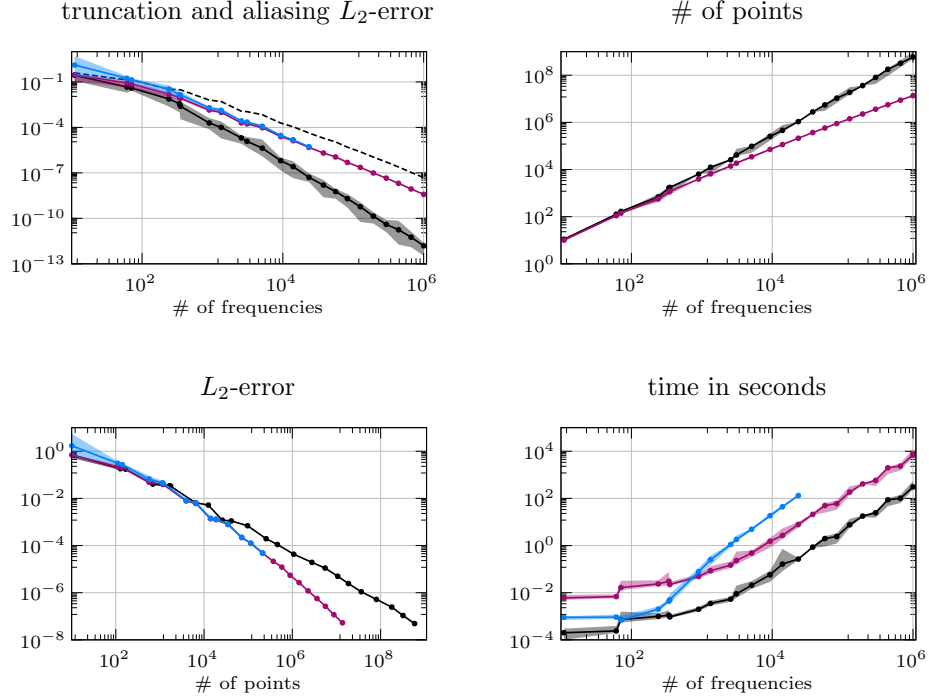
FIGURE 5.1. Five-dimensional experiment 1 on the torus for different point sets and algorithms. Black: full lattice, magenta: randomly subsampled lattice, azure: continuously random points, and dashed the truncation error.

For frequency sets we use hyperbolic crosses with different radii R, defined in (5.5), where we choose the shape parameter $\gamma = 1/2$. Using the probabilistic algorithm from [22] we computed a reconstructing rank-1 lattice $\boldsymbol{X}_{\mathrm{MZ}}$ for I. We used three different techniques for approximation:

- We then solved the least squares system (2.3) for the full rank-1 lattice. Using Lemma 2.3 for $\boldsymbol{L}^* \boldsymbol{W} \boldsymbol{L} = \boldsymbol{I}$, we omit the inverse matrix. The special structure of the points was used to perform the matrix-vector product using a one-dimensional fast Fourier transform in $\mathcal{O}(M \log M)$.
- We randomly subsampled the full rank-1 lattices according to Theorem 4.6 (all points have equal probability $\varrho_i = 1/M$) such that we have $n = \lceil |I| \log |I| \rceil$ points , cf. Figure 1.1. We solved the least squares system (2.3) iteratively using the same one-dimensional fast Fourier transform as described in (5.6) and set an iteration limit of 10.
- For comparison, we drew random samples with respect to the Lebesgue measure as suggested in [25] with the same number of points $n = \lceil |I| \log |I| \rceil$. We set up the system matrix and solved the least squares system (2.3) with the n random points iteratively. We limit the iterations to at most 10 as before.

We computed the L_2-error decomposed into truncation error

$$\|f - P_I f\|_{L_2}^2 = \|f\|_{L_2}^2 - \sum_{k \in I} |\hat{f}_k|^2$$

and the aliasing error

$$\|P_I f - S_I^X f\|_{L_2}^2 = \sum_{k \in I} |\hat{f}_k - \hat{g}_k|^2$$

with $\hat{g}_k$ the Fourier coefficients of the approximation. Further we measured the elapsed time for the computations and stopped the computations when more then 100 gigabytes of memory were used. We repeated this experiment ten times and the minimal, maximal and average results can be seen in Figure 5.1.

- In the upper left figure, we see that for all proposed methods the aliasing error is below the truncation error (dashed line). We emphasize at this point that we have chosen $I = I_{\mathrm{MZ}}$, which is of course reasonable since the bottleneck is the truncation error, which cannot be prevented. Differences in the sets I and I_{MZ} might only have a positive effect on the aliasing error, which seems to be superfluous in our setting.
- In the upper right figure, we see that the full rank-1 lattice has more points and is worth subsampling. In particular for 10^6 frequencies we have an oversampling factor of 489 for the full lattice whereas the oversampling for the random choices is $\log(10^6) \le 14$.
- In the lower left we see the slower error decay of the full rank-1 lattice and the faster error decay of the continuously random points with respect to the number of points. The error of the subsampled rank-1 lattice is the minimum of these two. In particular, for increasing number of points we obtain the better error decay for random points as well as for the randomly subsampled rank-1 lattice.
- In the lower right we see the computation time of the respective methods. The full rank-1 lattice and the subsampled rank-1 lattice differ by a factor of ten. This is due to the same underlying algorithm and the iteration count of ten for the subsampled version. Longer than both took the continuous random points as there the full matrix has to be used in contrast to using the highly tuned Fast Fourier Transform. As we showed at the beginning of this section this is slower in regards to the complexity ($\mathcal{O}(|I|^2 (\log |I|)^{1-d})$ versus $\mathcal{O}(|I| \log |I|)$). For $23\,483$ frequencies the computations took 258 seconds for the full matrix, whereas merely 4 seconds for the subsampled rank-1 lattice. Furthermore storing the matrix needs $23\,483 \cdot 212\,432 \cdot 16$ bytes $\approx$ 80 gigabytes, where 16 bytes is the size of a complex floating point number with double precision (which is why the experiment for the continuously random points stopped early). In contrast to that the rank-1 lattice uses around $1\,000\,000$ points and therefore $1\,000\,000 \cdot 16$ bytes $= 16$ megabytes of memory. Even the largest considered rank-1 lattice uses merely 8 gigabytes.

Experiment 2. We repeat Experiment 1 but additionally `BSS` subsample the randomly subsampled rank-1 lattice further to an oversampling factor of $b = 2$ (the `BSS` algorithm is available at `https://github.com/felixbartel/BSSsubsampling.jl`). We need to stop earlier as the `BSS` algorithm is (so far) not suitable for arbitrarily large matrices. The results can be seen in Figure 5.2.

truncation and aliasing L_2-error   # of points
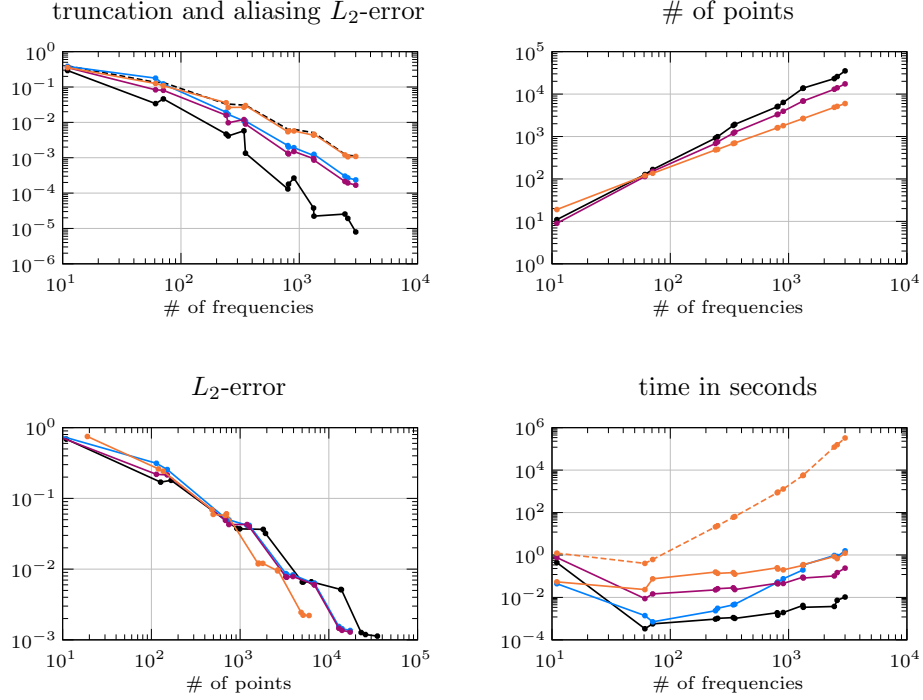
L_2-error   time in seconds

FIGURE 5.2. Five-dimensional experiment 2 on the torus for different point sets and algorithms. Black: full lattice, magenta: randomly subsampled lattice, orange: random and BSS subsampled lattice (dashed: time of the BSS precomputation step) azure: continuously random points, and dashed the truncation error

As the random subsampling step was already evaluated in the experiment above, we focus on the BSS subsampling. We obtain an even smaller number of points $(|\boldsymbol{X}'| = 2|I|)$ while still having the aliasing error slightly smaller than the truncation error. This results in a faster error decay with respect to the number of points, cf. lower left of Figure 5.2. We also see that the BSS algorithm takes way more time, cf. lower right of Figure 5.2. But this is a precomputation step and only has to be done once with the actual iterative solver for the solution not suffering from this.

References

1. F. Bartel, M. Schäfer, and T. Ullrich, *Constructive subsampling of finite frames with applications in optimal function recovery*, Appl. Comput. Harmon. Anal. **to appear** (2023).

2. J. D. Batson, D. A. Spielman, and N. Srivastava, *Twice-Ramanujan sparsifiers*, STOC'09—Proceedings of the 2009 ACM International Symposium on Theory of Computing, ACM, New York, 2009, pp. 255–262. MR 2780071

3. A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*, Kluwer Academic Publishers, Boston, MA, 2004. MR 2239907

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

5. Glenn Byrenheid, Dinh Dũng, Winfried Sickel, and Tino Ullrich, *Sampling on energy-norm based sparse grids for the optimal recovery of sobolev type functions in h^γ*, J. Approx. Theory **207** (2016), 207–231.

6. A. Christmann and I. Steinwart, *Support vector machines*, Springer, 2008.

7. A. Cohen, M. A. Davenport, and D. Leviatan, *On the stability and accuracy of least squares approximations*, Found. Comput. Math. **13** (2013), no. 5, 819–834. MR 3105946

8. A. Cohen and G. Migliorati, *Optimal weighted least-squares methods*, SMAI J. Comput. Math. **3** (2017), 181–203. MR 3716755

9. R. Cools and D. Nuyens, *An overview of fast component-by-component constructions of lattice rules and lattice sequences*, PAMM **7** (2007), 1022609–1022610.

10. D. Dũng, V. N. Temlyakov, and T. Ullrich, *Hyperbolic cross approximation*, Advanced Courses in Mathematics. CRM Barcelona, Birkhäuser/Springer, Cham, 2018. MR 3887571

11. J. Dick, P. Kritzer, and F. Pillichshammer, *Lattice rules—numerical integration, approximation, and discrepancy*, Springer Series in Computational Mathematics, vol. 58, Springer, Cham, [2022] ©2022, With an appendix by Adrian Ebert. MR 4472208

12. J. Dick, F. Y. Kuo, and I. H. Sloan, *High-dimensional integration: The quasi-Monte Carlo way*, Acta Numer. **22** (2013), 133–288.

13. M. Dolbeault, D. Krieg, and M. Ullrich, *A sharp upper bound for sampling numbers in L_2*, arXiv:2204.12621 (2022).

14. F. Filbir and H. N. Mhaskar, *Marcinkiewicz-Zygmund measures on manifolds*, J. Complexity **27** (2011), 568–598.

15. A. Greenbaum, *Iterative methods for solving linear systems*, SIAM, Philadelphia, PA, USA, 1997.

16. K. Gröchenig, *Sampling, Marcinkiewicz-Zygmund inequalities, approximation, and quadrature rules*, J. Approx. Theory **257** (2020), 105455, 20. MR 4124840

17. C. Haberstich, A. Nouy, and G. Perrin, *Boosted optimal weighted least-squares*, Math. Comp. **91** (2022), no. 335, 1281–1315. MR 4405496

18. Aicke Hinrichs, David Krieg, Erich Novak, and Jan Jan Vybíral, *Lower bounds for integration and recovery in L_2*, J. Complexity **72** (2022), Paper No. 101662, 15. MR 4438175

19. T. Jahn, T. Ullrich, and F. Voigtlaender, *Sampling numbers of smoothness classes via ℓ_1-minimization*, arXiv:2212.00445 (2022).

20. 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** (2021), no. 1, 33–77.

21. L. Kämmerer, *High Dimensional Fast Fourier Transform Based on Rank-1 Lattice Sampling*, Dissertation, Universitätsverlag Chemnitz, 2014.

22. L. Kämmerer, *A fast probabilistic component-by-component construction of exactly integrating rank-1 lattices and applications*, arXiv:2012.14263 (2020).

23. L. Kämmerer, D. Potts, and T. Volkmer, *Approximation of multivariate periodic functions by trigonometric polynomials based on rank-1 lattice sampling*, J. Complexity **31** (2015), 543–576.

24. L. Kämmerer, D. Potts, and T. Volkmer, *High-dimensional sparse FFT based on sampling along multiple rank-1 lattices*, Appl. Comput. Harmon. Anal. **51** (2021), 225–257. MR 4185099

25. L. Kämmerer, T. Ullrich, and T. Volkmer, *Worst case recovery guarantees for least squares approximation using random samples*, Constr. Approx. **54** (2021), 295–352.

26. B. Kashin, E. Kosov, I. Limonova, and V. N. Temlyakov, *Sampling discretization and related problems*, J. Complexity **71** (2022), Paper No. 101653, 55. MR 4420519

27. J. Keiner, S. Kunis, and D. Potts, *Efficient reconstruction of functions on the sphere from scattered data*, J. Fourier Anal. Appl. **13** (2007), no. 4, 435–458.

28. D. Krieg and M. Ullrich, *Function values are enough for L_2-approximation*, Found. Comput. Math. **21** (2021), no. 4, 1141–1151. MR 4298242

29. D. Krieg and M. Ullrich, *Function values are enough for L_2-approximation: Part II*, J. Complexity **66** (2021), Paper No. 101569, 14. MR 4292366

30. T. Kühn, W. Sickel, and T. Ullrich, *Approximation of mixed order Sobolev functions on the d-torus: asymptotics, preasymptotics, and d-dependence*, Constr. Approx. **42** (2015), no. 3, 353–398. MR 3416161

31. S. Kunis and D. Potts, *Fast spherical Fourier algorithms*, J. Comput. Appl. Math. **161** (2003), no. 1, 75–98. MR 2018576

32. F.Y. Kuo, G. Migliorati, F. Nobile, and D. Nuyens, *Function integration, reconstruction and approximation using rank-1 lattices*, Math. Comp. **90** (2021), no. 330, 1861–1897. MR 4273118

33. I. Limonova and V. N. Temlyakov, *On sampling discretization in L_2*, J. Math. Anal. Appl. **515** (2022), no. 2, 126457.

34. J. Marcinkiewicz and A. Zygmund, *Sur les fonctions indépendantes*, Fundamenta Mathematicae **29** (1937), no. 1, 60–90 (fre).

35. H. N. Mhaskar, F. J. Narcowich, and J. D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature*, Math. Comput. **70** (2001), 1113–1130, Corrigendum on the positivity of the quadrature weights in 71:453–454, 2002.

36. M. Moeller and T. Ullrich, *L_2-norm sampling discretization and recovery of functions from RKHS with finite trace*, Sampl. Theory Signal Process. Data Anal. **19** (2021), no. 2, Paper No. 13, 31. MR 4354442

37. N. Nagel, M. Schäfer, and T. Ullrich, *A new upper bound for sampling numbers*, Found. Comput. Math. **22** (2022), no. 2, 445–468. MR 4407748

38. A. Narayan, J. D. Jakeman, and T. Zhou, *A Christoffel function weighted least squares algorithm for collocation approximations*, Math. Comp. **86** (2017), no. 306, 1913–1947. MR 3626543

39. G. Plonka, D. Potts, G. Steidl, and M. Tasche, *Numerical Fourier Analysis*, Applied and Numerical Harmonic Analysis, Birkhäuser, 2018.

40. D. Potts, *Fast algorithms for discrete polynomial transforms on arbitrary grids*, vol. 366, 2003, Special issue on structured matrices: analysis, algorithms and applications (Cortona, 2000), pp. 353–370. MR 1987729

41. D. Potts, J. Prestin, and A. Vollrath, *A fast algorithm for nonequispaced Fourier transforms on the rotation group*, Numer. Algorithms **52** (2009), no. 3, 355–384. MR 2563947

42. I. H. Sloan and S. Joe, *Lattice methods for multiple integration*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1994. MR 1442955

43. V. N. Temlyakov, *Approximation of periodic functions*, Computational Mathematics and Analysis Series, Nova Science Publishers Inc., Commack, NY, 1993. MR 1373654 (96j:41001)

44. V. N. Temlyakov, *The Marcinkiewicz-type discretization theorems*, Constr. Approx. **48** (2018), no. 2, 337–369. MR 3848042

45. V. N. Temlyakov, *On optimal recovery in L_2*, J. Complexity **65** (2021), Paper No. 101545, 11. MR 4264807

46. V. N. Temlyakov and T. Ullrich, *Bounds on Kolmogorov widths and sampling recovery for classes with small mixed smoothness*, J. Complexity **67** (2021), Paper No. 101575, 19. MR 4311528

47. V. N. Temlyakov and T. Ullrich, *Approximation of functions with small mixed smoothness in the uniform norm*, J. Approx. Theory **277** (2022), Paper No. 105718, 23. MR 4395317

48. L. N. Trefethen, *Approximation theory and approximation practice*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.

49. J. A. Tropp, *User-friendly tail bounds for sums of random matrices*, Found. Comput. Math. **12** (2012), no. 4, 389–434. MR 2946459

FACULTY OF MATHEMATICS, CHEMNITZ UNIVERSITY OF TECHNOLOGY, D-09107 CHEMNITZ
Email address: `felix.bartel@mathematik.tu-chemnitz.de`

FACULTY OF MATHEMATICS, CHEMNITZ UNIVERSITY OF TECHNOLOGY, D-09107 CHEMNITZ
Email address: `lutz.kaemmerer@mathematik.tu-chemnitz.de`

FACULTY OF MATHEMATICS, CHEMNITZ UNIVERSITY OF TECHNOLOGY, D-09107 CHEMNITZ
Email address: `daniel.potts@mathematik.tu-chemnitz.de`

FACULTY OF MATHEMATICS, CHEMNITZ UNIVERSITY OF TECHNOLOGY, D-09107 CHEMNITZ
Email address: `tino.ullrich@mathematik.tu-chemnitz.de`