Efficient recovery of non-periodic multivariate functions from few scattered samples

Felix Bartel*, Kai Lüttgen†, Nicolas Nagel* and Tino Ullrich*
Faculty of Mathematics, Chemnitz University of Technology
09107 Chemnitz, Germany
Email: *{felix.bartel, kai.luettgen, nicolas.nagel, tino.ullrich}@mathematik.tu-chemnitz.de, †luettgen.math@gmail.com

***Abstract*—It has been observed by several authors (see [8], [17], [26]) that well-known periodization strategies like tent or Chebyshev transforms lead to remarkable results for the recovery of multivariate functions from few samples. So far, theoretical guarantees are missing. The goal of this paper is twofold. On the one hand, we give such guarantees and briefly describe the difficulties of the involved proof. On the other hand, we combine these periodization strategies with recent novel constructive methods for the efficient subsampling of finite frames in $\mathbb{C}^m$. As a result we are able to reconstruct non-periodic multivariate functions from very few samples. The used sampling nodes are the result of a two-step procedure. Firstly, a random draw with respect to the Chebyshev measure provides an initial node set. A further sparsification technique selects a significantly smaller subset of these nodes with equal approximation properties. This set of sampling nodes scales linearly in the dimension of the subspace on which we project and works universally for the whole class of functions. The method is based on principles developed by Batson, Spielman, and Srivastava [4] and can be numerically implemented. Samples on these nodes are then used in a (plain) least-squares sampling recovery step on a suitable hyperbolic cross subspace of functions resulting in a near-optimal behavior of the sampling error. Numerical experiments indicate the applicability of our results.**

I. Introduction

We investigate the worst-case approximation of non-periodic functions f on the cube $[-1,1]^d$ via function samples and propose a two-step procedure. Firstly, we have to find a good set of sample points and secondly, we need a method to use this information to get an approximation for functions from the target space.

In the periodic case, it is well-known that choosing sufficiently many uniform random sample nodes in $[-1,1]^d$ together with a least squares recovery using the Fourier basis yields near-optimal reconstructions. The rate of convergence in this case is determined by the function model, in particular the smoothness s (made rigorous in Section II) with polylogarithmic factors determining the precise behaviour in dimension d.

The non-periodic case on the other hand is not so clear. One may try to use a periodization technique to turn f into a periodic function $\tilde{f}$ and apply the above described procedure to $\tilde{f}$. Extending for example f periodically via $\tilde{f}(\mathbf{x}) = f((\mathbf{x} \bmod 2) - \mathbf{1})$, i.e. "copying" f over to other cubes, gives such a periodization but may also introduce discontinuities (as is the case for the function in Figure 1). We thus aim to find a periodization, such that $\tilde{f}$ is at least as smooth as f itself.

In Section II we give such a periodization via the cosine composition operator $\tilde{f} = T_{\cos} f$ acting between suitable Sobolev-type spaces by

$$(T_{\cos} f)(x_1, \dots, x_d) = f(\cos(\pi x_1), \dots, \cos(\pi x_d)). \qquad (1)$$

For us the important consequence of this is that approximation algorithms for periodic functions can be taken over to the non-periodic case, keeping the cosine composition in mind.

If U is a uniform random variable taking values in $[-1,1]$, then $\cos(\pi U)$ is distributed according to the Chebyshev measure $(\pi\sqrt{1-x^2})^{-1}\mathrm{d}x$ on $[-1,1]$. Similarly, since $\tilde{f} = T_{\cos} f$ is even, the d-dimensional Fourier basis (on $[-1,1]^d$) boils down to $\prod_{\ell=1}^d \cos(\pi k_\ell x_\ell)$. Consequently, going over to the non-periodic case, we will approximate f using $\prod_{\ell=1}^d \cos(k_\ell \arccos(x_\ell))$, i.e. tensor product Chebyshev polynomials.

The first part of the previous paragraph suggests that we need to sample points following the Chebyshev distribution

$$\mathrm{d}\varrho_D(\mathbf{x}) = \prod_{\ell=1}^d \left(\pi\sqrt{1-x_\ell^2}\right)^{-1} \mathrm{d}\mathbf{x}. \qquad (2)$$

These random points are near-optimal. To achieve the optimal error behaviour, we will use a subsampling technique to lower the number of used sample points even further while still keeping good approximation properties. This will be explained in Section III.

Lastly, we will describe our method for function recovery in Section IV. Our claims are backed up by numerical experiments with the test function f from (6) (tensored cutouts of a quadratic B-spline), the two-dimensional version being depicted in Figure 1. The resulting L_2-errors as shown in Figure 3 strongly follow a main decay rate of $s = 5/2$, the smoothness of f.

We have two aims with this paper. Firstly, we want to demonstrate the power and versatility of subsampling by applying it to the Chebyshev setting. And secondly, we want to give a theoretical explanation of why Chebyshev distributed sample points are suitable for the recovery of non-periodic functions. While other methods cap at certain decay rates or only give half the rate due to deterministic components (see Section IV for such methods), our method gives an optimal main decay rate of s.

II. Periodization Operator and Embeddings

In preparation of what is to follow, we fix some notation and give a few definitions. $\mathbb{N}$ denotes the positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The parameter $d \in \mathbb{N}$ is used for the dimension. Bold characters are used for vectors from some d-dimensional space. The real parameter s is always assumed to satisfy $s > 0$. Let $D = [-1,1]^d$. On D we distinguish between two L_2-spaces. Firstly, $L_2(D)$ denotes the usual space of square-integrable functions with respect to the Lebesgue measure. Secondly, $L_2(\varrho_D)$ is the space of square-integrable functions on D with respect to the Chebyshev measure from (2). The inner product on this latter space is denoted by $\langle f, g\rangle_\varrho$.

The torus is $\mathbb{T}^d = (\mathbb{R}/(2\mathbb{Z}))^d$, where we will use D as the fundamental domain. In other words, $\mathbb{T}^d$ is derived from D by identifying opposite sides of D. $L_2(\mathbb{T}^d)$ denotes the usual space of square-integrable periodic functions.

For $m \in \mathbb{N}_0$ the (classical) Sobolev space $H^m_{\text{mix}}(D)$ of dominating mixed smoothness is the space of all $f \in L_2(D)$ such that $\|f\|_{H^m_{\text{mix}}(D)} = \|f\|_{L_2(D)} + \sum_\alpha \|\partial^\alpha f\|_{L_2(D)} < \infty$, where we sum over all $\alpha \in \mathbb{N}_0^d$ with $0 \le \max_{i\in[d]} \alpha_i \le m$. In case of fractional smoothness $s > 0$, $H^s_{\text{mix}}(D)$ is defined via the Besov-type norm

$$\|f\|_{H^s_{\text{mix}}(D)} = \|f\|_{L_2(D)} + \sum_{e\subset[d]} \mathcal{R}^M_e(f) < \infty. \tag{3}$$

Here, $M \ge s$ is a fixed integer and $\mathcal{R}^M_e(f)$ is given by

$$\mathcal{R}^M_e(f)^2 = \sum_{\mathbf{j}\in\mathbb{N}_0^d(e)} 2^{2s|\mathbf{j}|_1} \sup_{|h_\ell|\le 2^{-j_\ell}} \|\Delta^{m,e}_{\mathbf{h}} f\|^2_{L_2(D)},$$

where $\mathbb{N}_0^d(e) = \{\mathbf{j} \in \mathbb{N}_0^d \mid j_\ell = 0 \ \forall \ell \notin e\}$, $\Delta^{m,e}_{\mathbf{h}} = \prod_{\ell\in e} \Delta^m_{h_\ell,\ell}$ and $\Delta^m_{h_\ell,\ell}$ is the mth-order forward difference in the ℓth coordinate direction of step size h_ℓ.

The periodic Sobolev space $H^s_{\text{mix}}(\mathbb{T}^d)$ consists of all $f \in L_2(\mathbb{T}^d)$ for which

$$\|f\|^2_{H^s_{\text{mix}}(\mathbb{T}^d)} = \sum_{\mathbf{k}\in\mathbb{Z}^d} |\hat{f}_{\mathbf{k}}|^2 \prod_{\ell=1}^d \left(1 + |k_\ell|^2\right)^s < \infty,$$

where the Fourier coefficients of a function f are given by

$$\hat{f}_{\mathbf{k}} = 2^{-d} \int_{\mathbb{T}^d} f(\mathbf{x})\, \mathrm{e}^{-\pi \mathrm{i}\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{x}, \quad \mathbf{k} \in \mathbb{Z}^d.$$

Finally, we consider the $L_2(\varrho_D)$-normalized Chebyshev polynomials of the first kind

$$T_k(x) = \sqrt{2}^{\min\{1,k\}} \cos(k \arccos(x)), \quad k \in \mathbb{N}_0. \tag{4}$$

Taking tensor products of these polynomials, we set

$$\eta_{\mathbf{k}}(\mathbf{x}) = \prod_{\ell=1}^d T_{k_\ell}(x_\ell), \quad \mathbf{k} = (k_\ell)_{\ell=1}^d \in \mathbb{N}_0^d. \tag{5}$$

With these definitions in mind we move on to the topic of periodization. Of the many ways to extend functions in a periodic fashion we choose to study the operator $T_{\cos}$ as defined in (1). On the one hand, thinking purely in terms of smoothness, the analysis of $T_{\cos}$ comes down to the chain rule and its generalizations. On the other hand, $T_{\cos}$ is not bounded as an operator from $L_2(D)$ to $L_2(\mathbb{T}^d)$, i.e. $T_{\cos}$ does not mesh well with integrability. Since Sobolev spaces measure a combination of both smoothness and integrability, the study of the action of $T_{\cos}$ on these spaces is somewhat delicate. Nonetheless, employing tools such as maximal functions, the Fefferman-Stein inequality and the Riesz-Thorin interpolation theorem we are able to prove the following theorem in [21].

Theorem 2.1: If $s > 1/2$, the linear operator $T_{\cos} : H^s_{\text{mix}}(D) \to H^s_{\text{mix}}(\mathbb{T}^d)$ is bounded.

There is another interesting property of the periodization operator $T_{\cos}$. A simple calculation reveals that $(T_{\cos} f)^\wedge_{\mathbf{k}} = \langle f, \eta_{\mathbf{k}}\rangle_\varrho \prod_{\ell=1}^d \sqrt{2}^{-\min\{1,k_\ell\}}$. In light of the above theorem we conclude:

Corollary 2.2: If $s > 1/2$ and $f \in H^s_{\text{mix}}(D)$, we have

$$\sum_{\mathbf{k}\in\mathbb{N}_0^d} |\langle f, \eta_{\mathbf{k}}\rangle_\varrho|^2 \prod_{\ell=1}^d \left(1 + |k_\ell|^2\right)^s \lesssim \|f\|^2_{H^s_{\text{mix}}(D)} < \infty.$$

Below, in Section IV, we consider an explicit example function constructed from tensor products of B-splines (see also Figure 1).

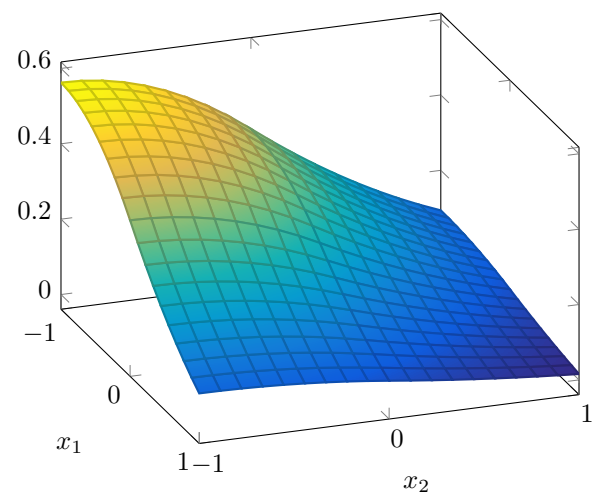

Fig. 1: Surface plot of the test function from (6) for $d = 2$.

Remark 2.3: For the purposes of this short note we have formulated the results of this section in the Hilbert space setting. However, with the same techniques we are able to prove analogous results for the more general Besov spaces $S^r_{p,q}B(D)$ of dominating mixed smoothness as well. The definition of these spaces is based on generalizations of the norm in equation (3).

Within the context of Besov spaces the operator $T_{\cos}$ can be interpreted not just as a periodization, but also as a generalized change of variables. In this regard our work is closely related to [6], [13], [27] and [28].

Finally, in analogy to the classical function spaces on the torus based on the Fourier system it makes sense to define function spaces based on Chebyshev coefficients. A theory of such spaces in an isotropic setting is developed in the works [5], [12], [14], [15], [18], [19], [31], [33] among others. Our results give novel insight into the relationship between

the classical Sobolev spaces on D and these Chebyshev–type spaces in the d-dimensional setting.

III. Subsampling

To efficiently recover functions on D we need to find suitable sample points $\mathbf{x}^i \in D, i \in [n]$. This class of points will be used for the whole function space and thus needs to be calculated only once. To do so, we start by sampling $\mathcal{O}(m \log m)$ random points according to the probability measure ϱ_D from (2) and use a subsampling technique based on Weaver's KS_2-conjecture (proven in [22] and further developed in [29], see [24] for a complete account) to reduce the budget to $\mathcal{O}(m)$. We start with the following theorem from [25], [29].

Theorem 3.1: Let $c, A, B > 0$ and $\mathbf{u}^1, \dots, \mathbf{u}^M \in \mathbb{C}^m$ with $\|\mathbf{u}^i\|_2^2 \le c\frac{m}{M}$ for all $i = 1, \dots, M$ and

$$A\|\mathbf{w}\|_2^2 \le \sum_{i=1}^{M} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 \le B\|\mathbf{w}\|_2^2$$

for all $\mathbf{w} \in \mathbb{C}^m$. Then there is a $J \subset [M]$ of size $\#J \le bm$ with

$$A' \cdot \frac{m}{M}\|\mathbf{w}\|_2^2 \le \sum_{j \in J} |\langle \mathbf{w}, \mathbf{u}^j \rangle|^2 \le B' \cdot \frac{m}{M}\|\mathbf{w}\|_2^2$$

for all $\mathbf{w} \in \mathbb{C}^m$, where $b, A', B' > 0$ only depend on c, A, B (and can be given explicitly).

In Theorem 3.1 the constants A and B are called the frame bounds and the system $\mathbf{u}^1, \dots, \mathbf{u}^M$ is called a frame. The strategy here is to use $\mathbf{u}^i = [\eta_{\mathbf{k}}(\mathbf{x}^i)]_{\mathbf{k} \in \Lambda}$ as the frame, where Λ is a set of frequencies. The indices $J \subset [M]$ as chosen in the above theorem will give sample points $\{\mathbf{x}^j\}_{j \in J}$ suitable for approximation.

Theorem 3.1 only shows the existence of such a good set of sampling points. It also leaves no control on the oversampling factor b (e.g. if $c = 1$, $A = 1/2$ and $B = 3/2$, Theorem 3.1 gives $b = 3284$, $A' = 11.65$ and $B' = 4926$). Constructive subsampling methods with arbitrary oversampling factors $b > 1 + \varepsilon$ were studied in [2].

Theorem 3.2: Let $\mathbf{u}^1, ..., \mathbf{u}^M \in \mathbb{C}^m$, choose $b > 1 + \frac{1}{m}$ and assume $M \ge bm$. There is a polynomial time algorithm to construct a $J \subset [M]$ with $\#J \le \lceil bm \rceil$ and

$$\frac{1}{M}\sum_{i=1}^{M} |\langle \mathbf{w}, \mathbf{u}^i \rangle|^2 \le 89\frac{(b+1)^2}{(b-1)^3} \cdot \frac{1}{m}\sum_{j \in J} |\langle \mathbf{w}, \mathbf{u}^j \rangle|^2$$

for all $\mathbf{w} \in \mathbb{C}^m$.

Theorem 3.2 enables us to choose an oversampling factor b close to 1, but it also has the disadvantage of only giving a lower bound. However, this suffices for our purposes. We will use Algorithms 1 and 3 from [2] for our numerical experiments.

IV. Numerical Experiments

As an example, we follow [1] and consider the function

$$f(\mathbf{x}) = \prod_{\ell=1}^{d} B_2(x_\ell) \tag{6}$$

with

$$B_2(x) = \begin{cases} -\frac{x^2}{4} - \frac{x}{2} + \frac{1}{2} & , -1 \le x \le 0 \\ \frac{x^2}{8} - \frac{x}{2} + \frac{1}{2} & , 0 < x \le 1 \end{cases}.$$

B_2 is a cutout of a piecewise quadratic B-spline. Since spaces of dominating mixed smoothness have the cross norm property we observe that $f \in B^{5/2}_{2,\infty}(D)$ or, on the scale of Sobolev spaces, $f \in H^s_{\text{mix}}(D)$ if and only if $s < 5/2$ (see (7) below). This fact is reflected in the observed decay rates in Figure 3. In terms of the $L_2(\varrho_D)$-normalized Chebyshev polynomials of the first kind T_k from (4) we have

$$\begin{aligned} B_2(x) = \frac{15}{32}T_0(x) - \frac{\pi - 1}{2\pi\sqrt{2}}T_1(x) - \frac{1}{32\sqrt{2}}T_2(x) \\ - \frac{3}{2\pi\sqrt{2}}\sum_{k=3}^{\infty}\frac{\sin(k\pi/2)}{k(k^2-4)}T_k(x). \end{aligned} \tag{7}$$

We approximate functions over D via the following procedure. Given a hyperbolic cross

$$\Lambda = \Lambda_{d,R} = \{\mathbf{k} \in \mathbb{N}_0^d : \prod_{\ell=1}^{d} \max\{1, k_\ell\} \le R\}$$

with $d, R \in \mathbb{N}$, set $m := \#\Lambda$ (for estimates on the size of $\Lambda_{d,R}$ see [7] and [20], but note that the hyperbolic cross is defined slightly different there). We now draw $M := \lceil 4m \log m \rceil$ random points according to the probability measure ϱ_D (see Figure 2 (left plot) and note the higher concentration of points at the boundary). Letting this set be denoted by $X = \{\mathbf{x}^i : i \in [M]\}$, we define the vectors $\mathbf{u}^i = \left[\eta_{\mathbf{k}}(\mathbf{x}^i)\right]_{\mathbf{k} \in \Lambda}$ with $\eta_{\mathbf{k}}(\mathbf{x}^i)$ as in (5). Using an oversampling factor of $b = 1.1$ the algorithm from Theorem 3.2 gives a set $J = \{j_1, \dots, j_n\} \subset [M]$ of size $\#J = n \le \lceil bm \rceil$, such that the $\mathcal{O}(m)$ points $\{\mathbf{x}^j\}_{j \in J}$ approximate functions on $[-1, 1]^d$ just as well (asymptotically) as the $\mathcal{O}(m \log m)$ random points X.

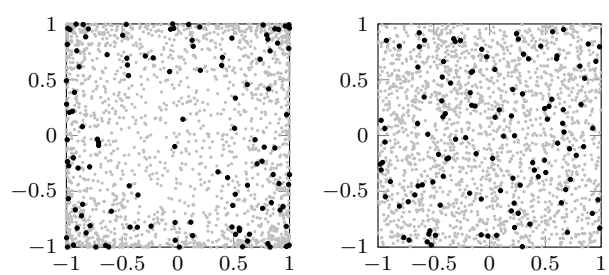

Fig. 2: $M = 2000$ random Chebyshev [uniform] sample points (light gray, $A \approx 0.766$ [0.375], $B \approx 1.225$ [0.628]), subsampled to $n = 117$ [116] points (black, $A \approx 0.026$ [0.01], $B \approx 1.872$ [1.068]), using $m = 107$ frequencies (hyperbolic cross, oversampling factor $b = 1.1$)

Figure 2 (left plot) shows the subsampling step for $d = 2$ and $R = 20$, so that $m = 107$ and $M = 2000$. The algorithm

finds a subset of $n = 117 \leq \lceil 1.1 \cdot 107 \rceil$ points that are suitable for function approximation. The quality of a point set in this case is measured by the smallest and largest singular values (i.e. the condition) of the matrix $(\#X)^{-1/2} \left[\eta_{\mathbf{k}}(\mathbf{x})\right]_{\mathbf{x}\in X, \mathbf{k}\in\Lambda}$. For the $M = 2000$ random points the smallest and largest singular values respectively are $A \approx 0.766$ and $B \approx 1.225$. For the $n = 117$ subsampled points these come out as $A \approx 0.026$ and $B \approx 1.872$.

Calculating coefficients $c_{\mathbf{k}}, \mathbf{k} \in \Lambda$ via the least squares system

$$\begin{bmatrix} \eta_{\mathbf{k}_1}(\mathbf{x}^{j_1}) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^{j_1}) \\ \vdots & & \vdots \\ \eta_{\mathbf{k}_1}(\mathbf{x}^{j_n}) & \cdots & \eta_{\mathbf{k}_m}(\mathbf{x}^{j_n}) \end{bmatrix} \begin{bmatrix} c_{\mathbf{k}_1} \\ \vdots \\ c_{\mathbf{k}_m} \end{bmatrix} \approx \begin{bmatrix} f(\mathbf{x}^{j_1}) \\ \vdots \\ f(\mathbf{x}^{j_n}) \end{bmatrix}$$

we approximate $f \approx \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}}\eta_{\mathbf{k}}$. Figure 3 shows the sampling error $\|f - \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}}\eta_{\mathbf{k}}\|_{L_2(\varrho_D)}$ against the number of samples n (after the subsampling step). Note that this procedure is inherently probabilistic, explaining the fluctuations in the plot. We compare this to the theoretical bound $n^{-2.5}(\log n)^{2.5(d-1)+0.5}$ from [25], Section 7 (building on concentration inequalities for matrices from [23]) and improving over [16], see also [32]. The authors in [10] give the final sharp, however non-constructive, version based on Theorem 3.1, which solved an outstanding open problem from [11].

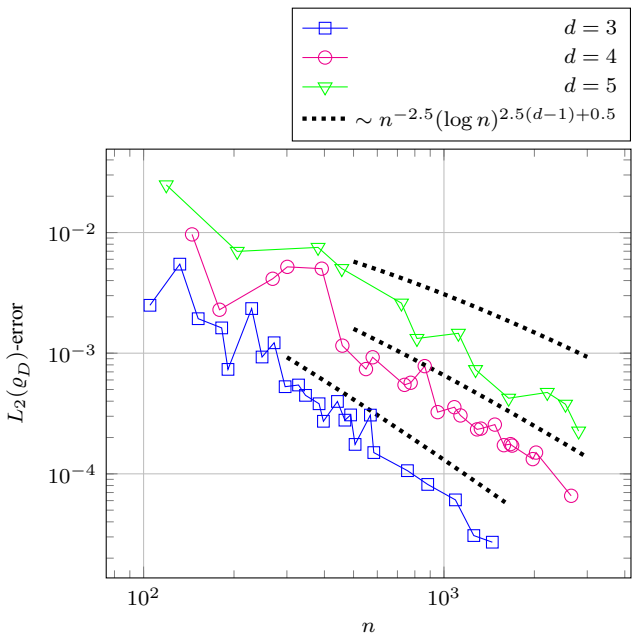

Fig. 3: Sampling errors in the $L_2(\varrho_D)$-norm using the Chebyshev basis for $f \in H^{5/2-\varepsilon}_{\text{mix}}(D)$ from (6).

This gives an improved rate compared to the approximation with the half period cosine basis $V_k(x) = \sqrt{2}^{-\delta_{0,k}} \cos\left(\pi k \frac{x+1}{2}\right)$ ($L_2(D)$-normalized) over $[-1, 1]$, where we have

$$B_2(x) = \frac{23}{24\sqrt{2}} V_0(x) + \sum_{k=1}^{\infty} \left(\frac{6\sin(\pi k/2)}{\pi^3 k^3} - \frac{(-1)^k}{\pi^2 k^2} \right) V_k(x).$$

Compared to the above, we only need to take uniform samples over $[-1, 1]^d$ and change the basis functions to $\psi_{\mathbf{k}}(\mathbf{x}^i) = \prod_{\ell=1}^d V_{k_\ell}(x^i_\ell)$. Figure 2 (right plot) demonstrates the subsampling step and Figure 4 shows the sampling error (with respect to $L_2(D)$). While other, smaller order factors (e.g. polylogarithmic) are not known, we still observe a main decay rate of $3/2$ (compared to $5/2$ in the Chebyshev setting).

Note that the half period cosine is the natural basis of $H^1_{\text{mix}}(D)$ and the linear decay is guaranteed. In [1], the basis of $H^2_{\text{mix}}(D)$ was considered with guaranteed quadratic decay. In the numerical experiment there the rate $5/2$ was obtained similar to the Chebyshev setting.

We also want to emphasize that the error in Figure 3 is measured in the stronger $L_2(\varrho_D)$-norm compared to the $L_2(D)$-norm used in Figure 4. This is due to

$$\int_D |g(\mathbf{x})|^2 \mathrm{d}\varrho_D(\mathbf{x}) \geq \pi^{-d} \int_D |g(\mathbf{x})|^2 \mathrm{d}\mathbf{x}$$

for functions g, for which the integrals are defined (as is the case for continuous functions on D), while no comparable bound can be made in the other direction.

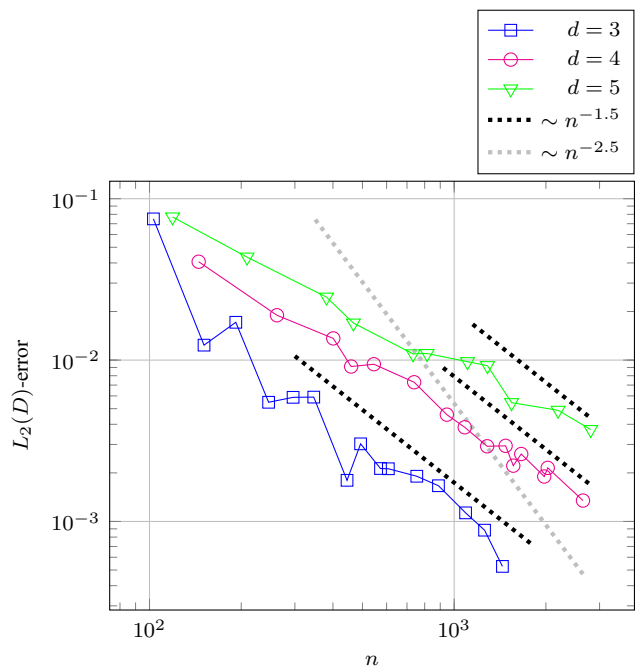

Fig. 4: Sampling errors in the $L_2(D)$-norm using the half period cosine basis for $f \in H^{5/2-\varepsilon}_{\text{mix}}(D)$ from (6).

Note that our procedure of finding the appropriate sampling nodes is probabilistic in nature. However, once the node set is fixed it works for the whole class of functions (unit ball in a certain Sobolev space) unlike the Monte-Carlo methods like in [9], where the node set might change for each function to approximate. One may ask for deterministic constructions of point sets to approximate non-periodic functions over D. In [30] and [34] so called rank-1 Chebyshev lattices (see Figure 5) were considered. In the same way as above, one uses function evaluations at those lattice points to get the coefficients for

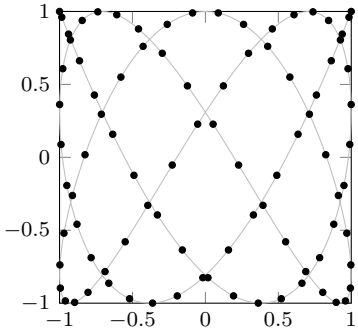

Fig. 5: Rank-1 Chebyshev lattice based on a Fibonacci lattice

tensored Chebyshev basis functions $\eta_{\mathbf{k}}$ at frequencies given by a hyperbolic cross. Example 3.24 from [34] also uses a cutout of a quadratic B-spline. While the theory there only gives a decay rate of $n^{-(1-\varepsilon)}$ for all $\varepsilon > 0$ (note the difference in notation there), the plots suggest an actual rate of $n^{-(1.25-\varepsilon)}$, which can be verified by our analysis. These constructive methods therefore only give half the rate compared to our method. Still, it should be noted that on lattices one can use FFT-techniques to accelerate the computation which are not possible for our unstructured (random) sample points. Also see [3] for further methods using sparse grids.

V. Conclusion

We give analytic and numerical evidence that for the recovery of non-periodic functions belonging to classical spaces over $[-1,1]^d$ sampling points following a Chebyshev distribution together with Chebyshev polynomials (see Figure 3) achieve state of the art error decay. Section II gives the theoretical framework for this where the periodization operator $T_{\cos}$ plays a crucial role in the analysis. We conclude that transforming non-periodic functions to periodic ones using the cosine composition $T_{\cos}$ helps in finding efficient approximation algorithms.

Acknowledgement

The authors would like to thank Winfried Sickel and Erich Novak for suggesting further references. Felix Bartel would like to thank the Deutscher Akademischer Austauschdienst (DAAD) for funding his research scholarship. Nicolas Nagel is funded by the ESF ReSIDA-H2 project. Kai Lüttgen is supported by a *Saxon Scholarship* granted by the Free State of Saxony.

References

[1] Felix Bartel, *Error guarantees for least squares approximation with noisy samples in domain adaptation*, accepted in SMAI J. Comput. Math. (2023).

[2] Felix Bartel, Martin Schäfer, and Tino Ullrich, *Constructive subsampling of finite frames with applications in optimal function recovery*, Applied and Computational Harmonic Analysis **to appear** (2023).

[3] Volker Barthelmann, Erich Novak, and Klaus Ritter, *High dimensional polynomial interpolation on sparse grids*, Advances in Computational Mathematics **12** (2000), 273–288.

[4] Joshua D. Batson, Daniel A. Spielman, and Nikhil 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

[5] D Berthold, W Hoppe, and B Silbermann, *A fast algorithm for solving the generalized airfoil equation*, Journal of computational and applied mathematics **43** (1992), no. 1-2, 185–219.

[6] Gérard Bourdaud, *Changes of variable in Besov spaces II*, Forum Math. **12** (2000), no. 5, 545–563.

[7] Alexey Chernov and Dinh Dũng, *New explicit-in-dimension estimates for the cardinality of high-dimensional hyperbolic crosses and approximation of functions having mixed smoothness*, Journal of Complexity **32** (2016), no. 1, 92–121.

[8] Ronald Cools, Frances Y. Kuo, Dirk Nuyens, and Gowri Suryanarayana, *Tent-transformed lattice rules for integration and approximation of multivariate non-periodic functions*, Journal of Complexity **36** (2016), 166–181.

[9] Matthieu Dolbeault and Albert Cohen, *Optimal pointwise sampling for L_2-approximation*, Journal of Complexity **68** (2022), 101602.

[10] Matthieu Dolbeault, David Krieg, and Mario Ullrich, *A sharp upper bound for sampling numbers in L_2*, Applied and Computational Harmonic Analysis **63** (2023), 113–134.

[11] Dinh Dung, Vladimir N. Temlyakov, and Tino Ullrich, *Hyperbolic Cross Approximation*, Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser/Springer (2018).

[12] Hans-Jürgen Glaeske and Thomas Runst, *The discrete Jacobi transform of generalized functions*, Mathematische Nachrichten **132** (1987), no. 1, 239–251.

[13] Masahiro Ikeda, Isao Ishikawa, and Koichi Taniguchi, *Boundedness of composition operators on higher order besov spaces in one dimension*, Mathematische Annalen (2023), 1–24.

[14] Kamen Ivanov, Pencho Petrushev, and Yuan Xu, *Decomposition of spaces of distributions induced by tensor product bases*, Journal of Functional Analysis **263** (2012), no. 5, 1147–1197.

[15] Peter Junghanns, Giuseppe Mastroianni, and Incoronata Notarangelo, *Weighted polynomial approximation and numerical methods for integral equations*, Springer Nature, 2021.

[16] Lutz Kämmerer, Tino Ullrich, and Toni Volkmer, *Worst case recovery guarantees for least squares approximation using random samples*, Constr. Approx. **54** (2021), 295–352.

[17] Frances Y. Kuo, Giovanni Migliorati, Fabio Nobile, and Dirk Nuyens, *Function integration, reconstruction and approximation using rank-1 lattices*, Math. Comp. **90** (2021), 1861–1897.

[18] George Kyriazis, Pencho Petrushev, and Yuan Xu, *Decomposition of weighted Triebel–Lizorkin and Besov spaces on the ball*, Proceedings of the London Mathematical Society **97** (2008), no. 2, 477–513.

[19] ———, *Jacobi decomposition of weighted Triebel–Lizorkin and Besov spaces*, Stud. Math. **186** (2008), 2.

[20] Thomas Kühn, Winfried Sickel, and Tino Ullrich, *Approximation of mixed order Sobolev functions on the d-torus – asymptotics, preasymptotics and d-dependence*, Constructive Approximation **42** (2015), 353–398.

[21] Kai Lüttgen and Tino Ullrich, *Properties of a cosine composition operator on product-type function spaces*, work in progress.

[22] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem*, Annals of Mathematics **182** (2015), 327—-350.

[23] Moritz Moeller and Tino Ullrich, *L_2-norm sampling discretization and recovery of functions from RKHS with finite trace*, Sampl. Theory Signal Process. Data Anal. **13** (2021), Article No. 17, 31 pp.

[24] Nicolas Nagel, *On the Kadison-Singer problem and Weaver's conjecture with implications for Fourier systems over unbounded sets*, Bachelor's thesis, Faculty of Mathematics, TU Chemnitz, 2020.

[25] Nicolas Nagel, Martin Schäfer, and Tino Ullrich, *A new upper bound for sampling numbers*, Foundations of Computational Mathematics **22** (2022), 445–468.

[26] Robert Nasdala and Daniel Potts, *Efficient multivariate approximation on the cube*, Numerische Mathematik **147** (2021), 393—-429.

[27] Van Kien Nguyen, *Function spaces of dominating mixed smoothness, Weyl and Bernstein numbers*, Ph.D. thesis, Dissertation, Jena, Friedrich-Schiller-Universität Jena, 2017.

[28] Van Kien Nguyen, Mario Ullrich, and Tino Ullrich, *Change of variable in spaces of mixed smoothness and numerical integration of multivariate*

functions on the unit cube, Constructive Approximation **46** (2017), 69–108.

[29] Shahaf Nitzan, Alexander Olevskii, and Alexander Ulanovskii, *Exponential frames on unbounded sets*, Proceedings of the American Mathematical Society **144** (2015), 109—118.

[30] Daniel Potts and Toni Volkmer, *Fast and exact reconstruction of arbitrary multivariate algebraic polynomials in Chebyshev form*, 2015 International Conference on Sampling Theory and Applications (SampTA), 2015, pp. 392–396.

[31] Thomas Runst and Winfried Sickel, *On strong summability of Jacobi-Fourier-expansions and smoothness properties of functions*, Mathematische Nachrichten **99** (1980), no. 1, 77–85.

[32] V. Temlyakov, *On optimal recovery in L_2*, Journal of Complexity **65** (2021), 101545.

[33] Hans Triebel, *Theory of function spaces*, Modern Birkhäuser Classics (2010).

[34] Toni Volkmer, *Multivariate approximation and high-dimensional sparse FFT based on rank-1 lattice sampling*, Dissertation, Universitätsverlag Chemnitz, 2017.