On sampling discretization in L_2

I. Limonova* and V. Temlyakov†

Abstract

We prove a sampling discretization theorem for the square norm of functions from a finite dimensional subspace satisfying Nikol'skii's inequality with an upper bound on the number of sampling points of the order of the dimension of the subspace.

Keywords and phrases: real and complex sampling discretization, submatrices of orthogonal matrices.

1 Introduction

Let Ω be a nonempty subset of $\mathbb{R}^d$ with the probability measure μ. By L_q, $1 \le q < \infty$, norm we understand

$$\|f\|_q := \|f\|_{L_q(\Omega,\mu)} := \left(\int_\Omega |f|^q d\mu\right)^{1/q}.$$

By discretization of the L_q-norm we understand a replacement of the measure μ by a discrete measure μ_m with support on a set $\xi = \{\xi^j\}_{j=1}^m \subset \Omega$. This means that integration with respect to measure μ is replaced by an appropriate cubature formula. Thus, integration is replaced by evaluation of a function f at a finite set ξ of points. This method of discretization is called *sampling discretization*. Discretization is an important step in making

*Lomonosov Moscow State University, Moscow Center for Fundamental and Applied Mathematics,

†University of South Carolina, Steklov Institute of Mathematics, Lomonosov Moscow State University, and Moscow Center for Fundamental and Applied Mathematics.

a continuous problem computationally feasible. The reader can find a corresponding discussion in a recent survey [3]. The first results in sampling discretization were obtained by Marcinkiewicz and by Marcinkiewicz-Zygmund (see [25]) for discretization of the L_q-norms of the univariate trigonometric polynomials in 1930s. Therefore, sampling discretization results are sometimes referred to as *Marcinkiewicz-type theorems* (see [21], [22], [3]). Recently, substantial progress in sampling discretization has been made in [21], [22], [10], [3], [4], [5], [11].

Let us comment on the values of functions from L_q. We are interested in discretization of the L_q-norms, $1 \le q \le \infty$, of elements of finite dimensional subspaces. By a function $f \in L_q(\Omega, \mu)$ we understand a specific function (not an equivalency class), which is defined almost everywhere with respect to μ on Ω. In other words, for $f \in L_q(\Omega, \mu)$ there exists a set $E(f) \subset \Omega$ such that $\mu(E(f)) = 0$ and $f(x)$ is defined for all $x \in \Omega \setminus E(f)$. We say that a subspace $X_N \subset L_q(\Omega, \mu)$ is an N-dimensional subspace if there are N linearly independent functions $u_i \in X_N$, $i = 1, \dots, N$, such that $X_N = \operatorname{span}(u_1, \dots, u_N)$. In this case, for the subspace X_N there exists a set $E(X_N) \subset \Omega$ such that $\mu(E(X_N)) = 0$ and each $f \in X_N$ is defined for all $x \in \Omega \setminus E(X_N)$. It will be convenient for us to assume that each $f \in X_N$ is defined for all $x \in \Omega$.

In this paper we present results on sampling discretization in the case $q = 2$. We consider two settings: (I) discretization with equal weights and (II) weighted discretization. In Section 2 we prove the main technical results – Lemma 2.2 and its Corollary 2.1 – that are used in the proofs of Theorems 1.1 and 1.2. We also discuss in detail known results related to the main lemma – Lemma 2.2. In Section 3 we prove and discuss Theorems 1.1 and 1.2.

We now proceed to a detailed discussion of our new results and related known results. First, in Subsection 1.1 we formulate the main results of the paper and comment on their novelty and impact. Second, in Subsection 1.2 we present a brief history of discretization with equal weights, which is directly related to our Theorem 1.1. Finally, in Subsection 1.3 we give a historical comment on weighted discretization.

1.1 Main results

I. Equal weights. The following condition is the key to the existence of good discretization with equal weights.

Condition E. We say that an orthonormal system $\{u_i(x)\}_{i=1}^N$ defined on Ω satisfies Condition E with a constant $t>0$ if for all $x\in\Omega$

$$\sum_{i=1}^N |u_i(x)|^2 \le Nt^2.$$

Note that integration of the above inequality over $x\in\Omega$ gives $t\ge 1$.

The following Theorem 1.1 solves (in the sense of order) the problem of discretization with equal weights for N-dimensional subspaces of $L_2(\Omega,\mu)$ satisfying Condition E.

Theorem 1.1. *Let $\Omega\subset\mathbb{R}^d$ be a nonempty set with the probability measure μ. Assume that $\{u_i(x)\}_{i=1}^N$ is a real (or complex) orthonormal system in $L_2(\Omega,\mu)$ satisfying Condition E. Then there is an absolute constant C_1 such that there exists a set $\{\xi^j\}_{j=1}^m\subset\Omega$ of $m\le C_1t^2N$ points with the property: for any $f=\sum_{i=1}^N c_iu_i$ we have*

$$C_2\|f\|_2^2 \le \frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^2 \le C_3t^2\|f\|_2^2,$$

where C_2 and C_3 are absolute positive constants.

It is known that Condition E is equivalent to the fact that the subspace $X_N := \operatorname{span}(u_1,\dots,u_N)$ satisfies the Nikol'skii inequality for the pair $(2,\infty)$ (see Section 3 for a detailed discussion). In Section 3 we reformulate Theorem 1.1 in terms of Nikol'skii's inequality (see Theorem 3.5 and Remark 3.3) and prove it.

II. General weights. Theorem 1.2 solves (in the sense of order) the problem of weighted discretization for arbitrary N-dimensional subspaces of $L_2(\Omega,\mu)$.

Theorem 1.2. *If X_N is an N-dimensional subspace of the complex $L_2(\Omega,\mu)$, then there exist three absolute positive constants C_1', c_0', C_0', a set of $m\le C_1'N$ points $\xi^1,\dots,\xi^m\in\Omega$, and a set of nonnegative weights λ_j, $j=1,\dots,m$, such that*

$$c_0'\|f\|_2^2 \le \sum_{j=1}^m \lambda_j|f(\xi^j)|^2 \le C_0'\|f\|_2^2,\quad \forall f\in X_N.$$

Remark 1.1. *We formulate Theorem 1.2 in terms of the class of nonnegative weights λ_j, $j = 1, \dots, m$. This class of weights is a standard class in numerical integration (cubature formulas) and in discretization. Clearly, the terms with $\lambda_j = 0$ can be dropped and then we come to the class of positive weights, provided at least one of the weights is positive.*

Theorem 1.2 is proved in Section 3. For the reader's convenience it is formulated there again as Theorem 3.3.

Novelty and impact. Theorems 1.1 and 1.2 solve (in the sense of order) two sampling discretization problems at a reasonable level of generality. In the case of the Marcinkiewicz-type discretization with equal weights we only impose Condition E, which is a standard condition in this case (see Subsection 1.2 for details). Theorem 1.2 provides a discretization result for any subspace of L_2, which is important for applications. A preprint version of this paper has been published in [13]. It has made an immediate impact on research of the optimal sampling recovery. Theorem 1.2 was used in [23] for proving an important inequality for the optimal sampling recovery. For the reader's convenience we formulate it here. Recall the setting of the optimal recovery. For a fixed m and a set of points $\xi := \{\xi^j\}_{j=1}^m \subset \Omega$, let Φ_ξ be a linear operator from $\mathbb{C}^m$ into $L_p(\Omega, \mu)$. Denote for a class $\mathbf{F}$ (usually, centrally symmetric and compact subset of $L_p(\Omega, \mu)$)

$$\varrho_m(\mathbf{F}, L_p) := \inf_{\text{linear } \Phi_\xi;\, \xi} \sup_{f \in \mathbf{F}} \|f - \Phi_\xi(f(\xi^1), \dots, f(\xi^m))\|_p.$$

The following statement was proved in [23]. There exist two positive absolute constants b and B such that for any compact subset Ω of $\mathbb{R}^d$, any probability measure μ on it, and any compact subset $\mathbf{F}$ of $\mathcal{C}(\Omega)$ we have

$$\varrho_{bn}(\mathbf{F}, L_2(\Omega, \mu)) \le B d_n(\mathbf{F}, L_\infty). \tag{1.1}$$

Here, $d_n(\mathbf{F}, L_\infty)$ is the Kolmogorov width of $\mathbf{F}$ in the uniform norm. In turn, inequality (1.1) was used in [24] to prove new bounds for the optimal sampling recovery of functions with small mixed smoothness.

The proof of Theorem 1.1 is based on Lemma 2.2. A version of this lemma (see Remark 2.2 below) was used in [16] for a breakthrough result on the sampling recovery.

Note that our proofs of Theorems 1.1 and 1.2 are not technically involved because they are based on deep known results.

1.2 Historical comments on discretization with equal weights

We begin with the formulation of Rudelson's result from [19]. In the paper [19] it is formulated in terms of submatrices of an orthogonal matrix. We reformulate it in our setting. Note that Theorem 1.3 can be derived from the original result of Rudelson in the same way as we derive Theorem 3.1 from Lemma 2.2 (see Section 3 below).

Theorem 1.3 ([19]). *Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu_M(x^j) = 1/M$, $j = 1, \dots, M$. Assume that a real orthonormal system $\{u_i(x)\}_{i=1}^N$ satisfies Condition E on Ω_M. Then for every $\epsilon > 0$ there exists a set $J \subset \{1, \dots, M\}$ of indices with cardinality*

$$m := |J| \le C\frac{t^2}{\epsilon^2} N \log \frac{Nt^2}{\epsilon^2} \tag{1.2}$$

such that for any $f = \sum_{i=1}^N c_i u_i$ we have

$$(1-\epsilon)^2 \|f\|_2^2 \le \frac{1}{m} \sum_{j \in J} f(x^j)^2 \le (1+\epsilon)^2 \|f\|_2^2.$$

In [22] it was demonstrated how the Bernstein-type concentration inequalities for random matrices can be used to prove an analog of Theorem 1.3 for a general Ω. The proof in [22] is based on a different idea than the Rudelson's proof. Here is the corresponding result.

Theorem 1.4 ([22, Theorem 6.6]). *Let $\{u_i(x)\}_{i=1}^N$ be a real orthonormal in $L_2(\Omega, \mu)$ system satisfying Condition E. Then for every $\epsilon > 0$ there exists a set $\{\xi^j\}_{j=1}^m \subset \Omega$ with*

$$m \le C\frac{t^2}{\epsilon^2} N \log N$$

such that for any $f = \sum_{i=1}^N c_i u_i$ we have

$$(1-\epsilon) \|f\|_2^2 \le \frac{1}{m} \sum_{j=1}^m f(\xi^j)^2 \le (1+\epsilon) \|f\|_2^2.$$

We note that Theorem 1.4 is more general and slightly stronger than Theorem 1.3. Theorem 1.4 provides the Marcinkiewicz-type discretization theorem for a general domain Ω instead of a discrete set Ω_M. Also, in Theorem 1.4 we have an extra factor $\log N$ instead of $\log \frac{Nt^2}{\epsilon^2}$ in (1.2). A typical necessary condition for the Marcinkiewicz-type discretization theorem to hold for an N-dimensional subspace X_N is $m \geq N$. For instance, such a necessary condition holds when X_N is an N-dimensional subspace of continuous functions with $\Omega = [0,1]$ and the Lebesgue measure on it. Both Theorem 1.3 and Theorem 1.4 provide sufficient conditions on m (the upper bound) for existence of a good set of cardinality m for sampling discretization. These sufficient conditions are close to the necessary condition, which is $m \geq N$, but still have an extra $\log N$ factor in the bound for m. The main goal of this paper is to prove a sufficient condition on m without an extra $\log N$ factor in the upper bound, which guarantees the Marcinkiewicz-type discretization theorem in L_2. This is done in Theorem 1.1. The first result in that direction was obtained under a condition stronger than Condition E.

Theorem 1.5 ([21, Theorem 4.7]). *Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu_M(x^j) = 1/M$, $j = 1,\dots,M$. Assume that $\{u_i(x)\}_{i=1}^N$ is an orthonormal on Ω_M system (real or complex). Assume in addition that this system has the following property: for all $j = 1,\dots,M$ we have*

$$\sum_{i=1}^N |u_i(x^j)|^2 = N. \tag{1.3}$$

Then there is an absolute constant C_1 such that there exists a subset $J \subset \{1,2,\dots,M\}$ with the property: $m := |J| \leq C_1 N$ and for any $f = \sum_{i=1}^N c_i u_i$ we have

$$C_2\|f\|_2^2 \leq \frac{1}{m}\sum_{j\in J} |f(x^j)|^2 \leq C_3\|f\|_2^2,$$

where C_2 and C_3 are absolute positive constants.

In Theorem 3.1 we strengthen Theorem 1.5 by showing that the same is true under the weaker Condition E instead of (1.3).

1.3 Historical comments on weighted discretization

In the case of weighted discretization, namely, when instead of $\frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^2$ we use the weighted sum $\sum_{j=1}^m \lambda_j |f(\xi^j)|^2$, the problem of discretization is solved in the sense of order in the case of real subspaces X_N. It is pointed out in [22] that the paper by J. Batson, D.A. Spielman, and N. Srivastava [1] basically solves the discretization problem with weights. We present an explicit formulation of this important result in our notation.

Theorem 1.6 ([1, Theorem 3.1]). *Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu_M(x^j) = 1/M$, $j = 1, \dots, M$, and let X_N be an N-dimensional subspace of real functions defined on Ω_M. Then for any number $b > 1$ there exists a set of weights $\lambda_j \geq 0$ such that $|\{j : \lambda_j \neq 0\}| \leq \lceil bN \rceil$ so that for any $f \in X_N$ we have*

$$\|f\|_2^2 \leq \sum_{j=1}^M \lambda_j f(x^j)^2 \leq \frac{b+1+2\sqrt{b}}{b+1-2\sqrt{b}} \|f\|_2^2.$$

As observed in [3, Theorem 2.13], this last theorem with a general probability space (Ω, μ) in place of the discrete space (Ω_M, μ_M) remains true (with other constant in the right hand side) if $X_N \subset L_4(\Omega, \mu)$. It was proved in [5] that the additional assumption $X_N \subset L_4(\Omega, \mu)$ can be dropped as well.

Theorem 1.7 ([5, Theorem 6.3]). *If X_N is an N-dimensional subspace of the real $L_2(\Omega, \mu)$, then for any $b \in (1, 2]$, there exist a set of $m \leq \lceil bN \rceil$ points $\xi^1, \dots, \xi^m \in \Omega$ and a set of nonnegative weights λ_j, $j = 1, \dots, m$, such that*

$$\|f\|_2^2 \leq \sum_{j=1}^m \lambda_j f(\xi^j)^2 \leq \frac{C}{(b-1)^2} \|f\|_2^2, \quad \forall f \in X_N,$$

where $C > 1$ is an absolute constant.

In this paper we obtain analogs of Theorems 1.6 and 1.7 in the case of complex subspaces X_N (see Theorems 3.2, 3.3, and Remark 3.2 in Section 3). Moreover, we provide two different proofs of Theorem 1.2 – one is based on results from [1] and the other is based on results from [15].

We note that there are related results on the Banach–Mazur distance between two finite dimensional spaces of the same dimension (see, for instance, [2], [20], [7]).

2 Main lemma

Results of this section are based on the following result by A. Marcus, D.A. Spielman and N. Srivastava.

Theorem 2.1 ([15, Corollary 1.5 with $r = 2$]). *Let a system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ from $\mathbb{C}^N$ have the following properties: for all $\mathbf{w} \in \mathbb{C}^N$*

$$\sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 = \|\mathbf{w}\|_2^2 \tag{2.1}$$

and for some $\epsilon > 0$

$$\|\mathbf{v}_j\|_2^2 \leq \epsilon, \qquad j = 1, \dots, M.$$

Then there is a partition of $\{1, 2, \dots, M\}$ into two sets S_1 and S_2 such that for all $\mathbf{w} \in \mathbb{C}^N$ and for each $i = 1, 2$

$$\sum_{j \in S_i} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \leq \frac{(1 + \sqrt{2\epsilon})^2}{2} \|\mathbf{w}\|_2^2.$$

The following Lemma 2.1 was derived from Theorem 2.1 in [17] (also see [18, Lemma 10.22, p.105]).

Lemma 2.1 ([17, Lemma 2]). *Let a system of vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ from $\mathbb{C}^N$ satisfy* (2.1) *for all $\mathbf{w} \in \mathbb{C}^N$ and*

$$\|\mathbf{v}_j\|_2^2 = N/M, \qquad j = 1, \dots, M.$$

Then there is a subset $J \subset \{1, 2, \dots, M\}$ such that for all $\mathbf{w} \in \mathbb{C}^N$

$$c_0 \|\mathbf{w}\|_2^2 \leq \frac{M}{N} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \leq C_0 \|\mathbf{w}\|_2^2,$$

where c_0 and C_0 are some absolute positive constants.

Lemma 2.1 does not control the cardinality of the set J, which we need for applications in discretization. The following simple remark was made in [21].

Remark 2.1 ([21]). *For the cardinality of the subset J from Lemma 2.1 we have*

$$c_0 N \leq |J| \leq C_0 N.$$

The following lemma is the main lemma for the proof of Theorem 1.1.

Lemma 2.2 (Main lemma)**.** *Let a system of vectors* $\mathbf{v}_1, \dots, \mathbf{v}_M$ *from* $\mathbb{C}^N$ *satisfy* (2.1) *for all* $\mathbf{w} \in \mathbb{C}^N$ *and*

$$\|\mathbf{v}_j\|_2^2 \le \theta N/M, \qquad \theta \le M/N, \qquad j = 1, \dots, M. \tag{2.2}$$

Then there is a subset $J \subset \{1, 2, \dots, M\}$ *such that for all* $\mathbf{w} \in \mathbb{C}^N$

$$c_0\theta\|\mathbf{w}\|_2^2 \le \frac{M}{N} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le C_0\theta\|\mathbf{w}\|_2^2, \quad |J| \le C_1\theta N, \tag{2.3}$$

where c_0*,* C_0*, and* C_1 *are some absolute positive constants.*

Remark 2.2. *The proof of Lemma 2.2 gives a slightly stronger result than Lemma 2.2 – the tight frame condition (2.1) can be replaced by a frame condition*

$$A\|\mathbf{w}\|_2^2 \le \sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le B\|\mathbf{w}\|_2^2, \quad 0 < A \le B < \infty.$$

This stronger version of Lemma 2.2 was used in the followup paper [16] for sampling recovery.

To derive a complex case analog of Theorem 1.6 (i.e., Theorem 3.2) from results in [15] we need to prove the following corollary.

Corollary 2.1. *Let a system of vectors* $\mathbf{v}_1, \dots, \mathbf{v}_M$ *from* $\mathbb{C}^N$ *satisfy* (2.1) *for all* $\mathbf{w} \in \mathbb{C}^N$*. Then there exists a set of weights* $\lambda_j \ge 0$*,* $j = 1, \dots, M$*, such that* $|\{j : \lambda_j \ne 0\}| \le 2C_1 N$ *and for all* $\mathbf{w} \in \mathbb{C}^N$ *we have*

$$c_0\|\mathbf{w}\|_2^2 \le \sum_{j=1}^{M} \lambda_j |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le C_0\|\mathbf{w}\|_2^2.$$

where c_0*,* C_0*, and* C_1 *are absolute positive constants from Lemma 2.2.*

Proof. Without loss of generality we assume that $\|\mathbf{v}_1\|_2 = \min_{j=1,\dots,M} \|\mathbf{v}_j\|_2$. Let $n_1, \dots, n_M$ be natural numbers such that for every j, $1 \le j \le M$,

$$\|\mathbf{v}_1\|_2^2 \le \frac{\|\mathbf{v}_j\|_2^2}{n_j} < 2\|\mathbf{v}_1\|_2^2. \tag{2.4}$$

Denote

$$M' = \sum_{j=1}^{M} n_j. \tag{2.5}$$

We build a system V of vectors $\mathbf{v}'_1, \ldots, \mathbf{v}'_{M'}$ from $\mathbb{C}^N$ in the following way: for every j, $1 \le j \le M$, we include in V n_j copies of the vector $\mathbf{v}_j/\sqrt{n_j}$. Let us check that V satisfies (2.1) and (2.2) with $\theta = 2$. By construction and by our assumption that the system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_M$ satisfies (2.1), we have

$$\sum_{j=1}^{M'} |\langle \mathbf{w}, \mathbf{v}'_j \rangle|^2 = \sum_{j=1}^{M} n_j |\langle \mathbf{w}, \mathbf{v}_j/\sqrt{n_j} \rangle|^2 = \|\mathbf{w}\|_2^2. \tag{2.6}$$

By construction of the system V we obtain from (2.4) and (2.5) that

$$\|\mathbf{v}_1\|_2^2 M' \le \sum_{j=1}^{M} n_j \frac{\|\mathbf{v}_j\|_2^2}{n_j} = \sum_{j=1}^{M'} \|\mathbf{v}'_j\|_2^2. \tag{2.7}$$

Let e_i, $i = 1, \ldots, N$, be the canonical basis of $\mathbb{C}^N$. Then from (2.6) we obtain

$$\sum_{j=1}^{M'} \|\mathbf{v}'_j\|_2^2 = \sum_{j=1}^{M'} \sum_{i=1}^{N} |\langle e_i, \mathbf{v}'_j \rangle|^2 = \sum_{i=1}^{N} \sum_{j=1}^{M'} |\langle e_i, \mathbf{v}'_j \rangle|^2 = \sum_{i=1}^{N} \|e_i\|^2 = N. \tag{2.8}$$

Thus, from (2.7) and (2.8) we have $\|\mathbf{v}_1\|_2^2 \le N/M'$. By construction for each $j = 1, \ldots, M'$, there is a number $k(j) \in \{1, \ldots, M\}$ such that $\mathbf{v}'_j = \mathbf{v}_{k(j)}/\sqrt{n_{k(j)}}$. Therefore, by (2.4) we get

$$\|\mathbf{v}'_j\|_2^2 = \frac{\|\mathbf{v}_{k(j)}\|_2^2}{n_{k(j)}} < 2\|\mathbf{v}_1\|_2^2 \le 2\frac{N}{M'}, \qquad j = 1, \ldots, M'.$$

The above inequality implies that the system V satisfies condition (2.2) and equality (2.6) implies condition (2.1). We apply Lemma 2.2 to the system V and obtain a subset $J \subset \{1, \ldots, M'\}$ with $|J| \le 2C_1 N$ such that for all $\mathbf{w} \in \mathbb{C}^N$

$$c_0 \|\mathbf{w}\|_2^2 \le \frac{M'}{2N} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{v}'_j \rangle|^2 \le C_0 \|\mathbf{w}\|_2^2.$$

It is clear that

$$\frac{M'}{2N} \sum_{j \in J} |\langle \mathbf{w}, \mathbf{v}'_j \rangle|^2 = \sum_{j=1}^{M} \lambda_j |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2$$

for some nonnegative $\lambda_j, j = 1, \ldots, M$, so that $|\{j : \lambda_j \ne 0\}| \le 2C_1 N$. $\square$

Note that condition (2.1) implies that $M \geq N$. Lemma 2.2 in some sense improves the celebrated result of M. Rudelson [19] where a result similar to Lemma 2.2 was proved with $|J| \leq C_1(t)N \log N$ and with bounds depending on ϵ (see Theorem 1.3 in Introduction). Proof of Lemma 2.2 uses the iteration method suggested by A. Lunin [14]. We also refer the reader to the papers [8], [9], [12] for a discussion of recent outstanding progress in the area of submatrices of orthogonal matrices.

Proof of Lemma 2.2. We use the following known results (for Proposition 2.1 see Corollary B from [17], Corollary 10.19 from [18], p.104, or [6], and for Lemma 2.3 see Lemma 1 in [17] or Lemma 10.20 in [18], p.104).

Proposition 2.1 ([17, Corollary B]). *Let* $\mathbf{v}_1, \dots, \mathbf{v}_M \in \mathbb{C}^N$ *and* $\delta > 0$ *be such that* $\|\mathbf{v}_j\|_2^2 \leq \delta$ *for all* $j = 1, \dots, M$. *If*

$$\alpha\|\mathbf{w}\|_2^2 \leq \sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \leq \beta\|\mathbf{w}\|_2^2, \qquad \forall \mathbf{w} \in \mathbb{C}^N,$$

with some numbers $\beta \geq \alpha > \delta$, *then there exists a partition of* $\{1, \dots, M\}$ *into* S_1 *and* S_2 *such that for each* $i = 1, 2$:

$$\frac{1 - 5\sqrt{\delta/\alpha}}{2}\alpha\|\mathbf{w}\|_2^2 \leq \sum_{j \in S_i} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \leq \frac{1 + 5\sqrt{\delta/\alpha}}{2}\beta\|\mathbf{w}\|_2^2, \quad \forall \mathbf{w} \in \mathbb{C}^N.$$

Lemma 2.3 ([17, Lemma 1]). *Let* $0 < \delta < 1/100$, *and let* $\alpha_j, \beta_j, j = 0, 1, \dots,$ *be defined inductively*

$$\alpha_0 = \beta_0 = 1, \quad \alpha_{j+1} := \alpha_j \frac{1 - 5\sqrt{\delta/\alpha_j}}{2}, \quad \beta_{j+1} := \beta_j \frac{1 + 5\sqrt{\delta/\alpha_j}}{2}.$$

Then there exist a positive absolute constant C *and a number* $L \in \mathbb{N}$ *such that*

$$\alpha_j \geq 100\delta, \quad j \leq L, \quad 25\delta \leq \alpha_{L+1} < 100\delta, \quad \beta_{L+1} < C\alpha_{L+1}.$$

If $\delta := \theta N/M \geq 1/100$, then (2.3) holds with $J = \{1, 2, \dots, M\}$ and $C_1 = 1/\delta \leq 100$, $c_0 = 1$, $C_0 = 100$. Assume $\delta < 1/100$. Let α_j, β_j be as defined in Lemma 2.3; then the vectors $\mathbf{v}_1, \dots, \mathbf{v}_M$ satisfy the assumptions of Proposition 2.1 with $\alpha = \beta = 1$. We apply Proposition 2.1 and choose a subset of the obtained partition with a smaller cardinality. We obtain a set $J_1 \subset \{1, 2, \dots, M\}$ with $|J_1| \le M/2$ such that for all $\mathbf{w} \in \mathbb{C}^N$

$$\alpha_1 \|\mathbf{w}\|_2^2 \le \sum_{i \in J_1} |\langle \mathbf{w}, \mathbf{v}_i \rangle|^2 \le \beta_1 \|\mathbf{w}\|_2^2.$$

Since $\alpha_1 > 25\delta$ we can apply Proposition 2.1 again and obtain $J_2 \subset J_1$ with $|J_2| \le M/2^2$, for which we have two-sided inequalities with $\alpha_2 > 0$ and β_2. Let L be the number from Lemma 2.3. We iteratively apply Proposition 2.1 (choosing at each step the subset S_i with the smallest cardinality) and find $J_1 \supset J_2 \supset \cdots \supset J_{L+1}$ with the property

$$\frac{1 - 5\sqrt{\delta/\alpha_L}}{2} \alpha_L \|\mathbf{w}\|_2^2 \le \sum_{j \in J_{L+1}} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \le \frac{1 + 5\sqrt{\delta/\alpha_L}}{2} \beta_L \|\mathbf{w}\|_2^2, \quad \forall \mathbf{w} \in \mathbb{C}^N.$$

By Lemma 2.3 we obtain

$$\frac{1 - 5\sqrt{\delta/\alpha_L}}{2} \alpha_L = \alpha_{L+1} \ge 25\delta,$$

$$\frac{1 + 5\sqrt{\delta/\alpha_L}}{2} \beta_L = \beta_{L+1} \le C\alpha_{L+1} < 100C\delta.$$

Thus, for $J := J_{L+1}$ we have

$$25\theta \frac{N}{M} \|\mathbf{w}\|_2^2 \le \sum_{i \in J} |\langle \mathbf{w}, \mathbf{v}_i \rangle|^2 \le 100C\theta \frac{N}{M} \|\mathbf{w}\|_2^2.$$

Note that $2^{-L-1} \le \beta_{L+1} < 100C\delta$, therefore $|J_{L+1}| \le M/2^{L+1} \le 100CM\delta = 100C\theta N$ as required. $\square$

3 Application to discretization

The following corollary of Lemma 2.2 is a generalization of Theorem 4.7 from [21] (see Theorem 1.5 in Introduction). In [21] instead of condition (3.1) a stronger assumption (1.3) was imposed.

Theorem 3.1. *Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu_M(x^j) = 1/M$, $j = 1, \dots, M$. Assume that $\{u_i(x)\}_{i=1}^N$ is an orthonormal on Ω_M system (real or complex). Assume in addition that this system has the following property: for all $j = 1, \dots, M$ and for some $t > 0$ we have*

$$\sum_{i=1}^N |u_i(x^j)|^2 \leq Nt^2. \tag{3.1}$$

Then there is an absolute constant C_1 such that there exists a subset $J \subset \{1, 2, \dots, M\}$ with the property: $m := |J| \leq C_1 t^2 N$ and for any $f = \sum_{i=1}^N c_i u_i$ we have

$$C_2 \|f\|_2^2 \leq \frac{1}{m} \sum_{j \in J} |f(x^j)|^2 \leq C_3 t^2 \|f\|_2^2, \tag{3.2}$$

where C_2 and C_3 are absolute positive constants.

Proof. Define the column vectors

$$\mathbf{v}_j := M^{-1/2}(u_1(x^j), \dots, u_N(x^j))^T, \qquad j = 1, \dots, M.$$

Then our assumption (3.1) implies that the system $\mathbf{v}_1, \dots, \mathbf{v}_M$ satisfies (2.2) with $\theta = t^2$. For any $\mathbf{w} = (w_1, \dots, w_N)^T \in \mathbb{C}^N$ we have

$$\sum_{j=1}^M |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 = \frac{1}{M} \sum_{j=1}^M \sum_{i,k=1}^N w_i \bar{w}_k \bar{u}_i(x^j) u_k(x^j) = \sum_{i=1}^N |w_i|^2$$

by the orthonormality assumption. This implies that the system $\mathbf{v}_1, \dots, \mathbf{v}_M$ satisfies (2.1).

Note that the necessary condition for (3.2) to hold is $m \geq N$. Applying Lemma 2.2 we complete the proof of Theorem 3.1. $\square$

The following Theorem 3.2, which is a complex analog of Theorem 1.6, can be derived from Corollary 2.1 in the same way as we have derived Theorem 3.1 from Lemma 2.2 above.

Theorem 3.2. *Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu_M(x^j) = 1/M$, $j = 1, \dots, M$. Assume that $\{u_i(x)\}_{i=1}^N$ is an orthonormal on Ω_M system (real or complex). Then there is an absolute constant*

C_1 such that there exists a set of weights $\lambda_j \geq 0$, $j = 1, \dots, M$, with the property: $m := |\{j : \lambda_j \neq 0\}| \leq C_1 N$ and for any $f = \sum_{i=1}^{N} c_i u_i$ we have

$$c_0 \|f\|_2^2 \leq \sum_{j=1}^{M} \lambda_j |f(x^j)|^2 \leq C_0 \|f\|_2^2,$$

where c_0 and C_0 are from Lemma 2.2.

Further, using Theorem 3.2 and repeating the argument in the proof of Theorem 6.3 from [5] (with natural modifications from the real case to the complex case), which was used to derive Theorem 1.7 from Theorems 1.6 and 1.4, we obtain the complex analog of Theorem 1.7 – Theorem 1.2, which we formulate below as Theorem 3.3 for the reader's convenience. Note that the complex version of Theorem 1.4 can be proved in the same way as Theorem 1.4 was proved in [22].

Theorem 3.3. *If X_N is an N-dimensional subspace of the complex $L_2(\Omega, \mu)$, then there exist three absolute positive constants C_1', c_0', C_0', a set of $m \leq C_1' N$ points $\xi^1, \dots, \xi^m \in \Omega$, and a set of nonnegative weights λ_j, $j = 1, \dots, m$, such that*

$$c_0' \|f\|_2^2 \leq \sum_{j=1}^{m} \lambda_j |f(\xi^j)|^2 \leq C_0' \|f\|_2^2, \quad \forall f \in X_N.$$

Remark 3.1. *A combination of the proof of Theorem 6.3 from [5] with Theorem 3.5 (in the proof of Theorem 6.3 from [5] we use Theorem 3.5 instead of Theorem 1.2 (Theorem 1.4 above)) gives Theorem 3.3 with $C_1' = C_1'$, $c_0' = C_2'$, and $C_0' = C_3'$, where C_i', $i = 1, 2, 3$, are from Theorem 3.5. It is another way to prove Theorem 3.3.*

It is important to emphasize that in the proofs of Theorems 3.2 and 3.3, which are complex companions of Theorems 1.6 and 1.7, we did not use Theorems 1.6 and 1.7. Thus, our arguments give other proofs of analogs of Theorems 1.6 and 1.7. Note that constants in Theorems 3.2 and 3.3 are not as good as constants in Theorems 1.6 and 1.7.

A comment on connection between real and complex weighted discretization. We show here that good discretization of the $L_2(\Omega, \mu)$-norm of functions from real subspaces of dimension $2N$ implies good discretization of the $L_2(\Omega, \mu)$-norm of functions from complex subspaces of dimension N.

Definition 3.1. *Let X_N be a subspace of $L_2(\Omega, \mu)$. For $m \in \mathbb{N}$ and positive constants $C_1 \le C_2$ we write $X_N \in \mathcal{M}^w(m, 2, C_1, C_2)$ if there exist a set of points $\xi^1, \dots, \xi^m \in \Omega$ and a set of weights λ_ν, $\nu = 1, \dots, m$, such that for any $f \in X_N$ we have*

$$C_1 \|f\|_2^2 \le \sum_{\nu=1}^{m} \lambda_\nu |f(\xi^\nu)|^2 \le C_2 \|f\|_2^2. \tag{3.3}$$

Proposition 3.1. *Let $X_N = \operatorname{span}(w_1, \dots, w_N)$ be a subspace of complex $L_2(\Omega, \mu)$. Suppose that $w_j = u_j + iv_j$, where u_j, v_j are real functions, $j = 1, \dots, N$. Denote $Y_S := \operatorname{span}(u_1, \dots, u_N, v_1, \dots, v_N)$, $S := \dim Y_S \le 2N$, a real subspace of $L_2(\Omega, \mu)$. Then*

$$Y_S \in \mathcal{M}^w(m, 2, C_1, C_2) \quad \textit{implies} \quad X_N \in \mathcal{M}^w(m, 2, C_1, C_2).$$

Moreover, for discretization of X_N we can use the same points and weights as for discretization of Y_S.

Proof. Take an $f \in X_N$ and write

$$f = f_R + if_I, \quad f_R, f_I \in Y_S.$$

Assume that a set of points $\xi^1, \dots, \xi^m \in \Omega$ and a set of weights λ_ν, $\nu = 1, \dots, m$, are such that for any $g \in Y_S$ we have

$$C_1 \|g\|_2^2 \le \sum_{\nu=1}^{m} \lambda_\nu |g(\xi^\nu)|^2 \le C_2 \|g\|_2^2. \tag{3.4}$$

Then on one hand

$$\sum_{\nu=1}^{m} \lambda_\nu |f(\xi^\nu)|^2 = \sum_{\nu=1}^{m} \lambda_\nu (|f_R(\xi^\nu)|^2 + |f_I(\xi^\nu)|^2) \le C_2(\|f_R\|_2^2 + \|f_I\|_2^2) = C_2 \|f\|_2^2.$$

On the other hand

$$\sum_{\nu=1}^{m} \lambda_\nu |f(\xi^\nu)|^2 = \sum_{\nu=1}^{m} \lambda_\nu (|f_R(\xi^\nu)|^2 + |f_I(\xi^\nu)|^2) \ge C_1(\|f_R\|_2^2 + \|f_I\|_2^2) = C_1 \|f\|_2^2.$$

The above inequalities prove Proposition 3.1. $\square$

Remark 3.2. *Proposition 3.1 and Theorem 1.7 imply that Theorem 1.2 holds with $m \le \lceil 2bN \rceil$, $b \in (1,2]$, and $c_0' = 1$, $C_0' = C(b-1)^{-2}$, where C is an absolute constant from Theorem 1.7.*

A remark on sampling recovery. We mentioned in the Introduction that inequality (1.1) was proved in [23] with the help of Theorem 1.2. We now point out that if, in the proof of (1.1) (from [23]), we replace Theorem 1.2 with either Theorem 1.7 (real case) or Remark 3.2 (complex case), then we obtain the following version of (1.1).

Theorem 3.4. *For any $b \in (1,2]$ there exists a positive constant $B = B(b)$ such that for any compact subset Ω of $\mathbb{R}^d$, any probability measure μ on it, and any compact subset $\mathbf{F}$ of $\mathcal{C}(\Omega)$ we have in the real case*

$$\varrho_{\lceil b(n+1) \rceil}(\mathbf{F}, L_2(\Omega,\mu)) \le B d_n(\mathbf{F}, L_\infty)$$

and in the complex case

$$\varrho_{\lceil b(2n+1) \rceil}(\mathbf{F}, L_2(\Omega,\mu)) \le B d_n(\mathbf{F}, L_\infty).$$

Let Ω be a nonempty compact set in $\mathbb{R}^d$ and let X_N be an N-dimensional subspace of real (or complex) space of continuous functions $\mathcal{C}(\Omega)$. Let μ be a probability measure on Ω and let $\{u_i(x)\}_{i=1}^N$ be an orthonormal basis for X_N.

Nikol'skii inequality. We say that X_N satisfies the Nikol'skii inequality for the pair $(2,\infty)$ if there exists a constant $t > 0$ such that

$$\|f\|_\infty \le t N^{\frac{1}{2}} \|f\|_2, \quad \forall f \in X_N. \tag{3.5}$$

We point out that condition (3.5) with $X_N = \operatorname{span}(u_1,\dots,u_N)$, where $\{u_j\}_{j=1}^N$ is an orthonormal system, is equivalent to Condition E. This can be seen from the following simple well-known result, which is a corollary of the Cauchy inequality.

Proposition 3.2. *Let X_N be an N-dimensional subspace of $\mathcal{C}(\Omega)$. Then for any orthonormal basis $\{u_i\}_{i=1}^N$ of $X_N \subset L_2(\Omega,\mu)$ we have that for $x \in \Omega$*

$$\sup_{f \in X_N; f \ne 0} |f(x)|/\|f\|_2 = \left(\sum_{i=1}^N |u_i(x)|^2\right)^{1/2}.$$

The following simple result can be found in [3]. Note that only the real case is discussed in [3]. However, the same argument works for the complex case as well.

Proposition 3.3 ([3, Proposition 2.1]). *Let $Y_N := \operatorname{span}(u_1(x), \dots, u_N(x))$ with $\{u_i(x)\}_{i=1}^N$ being a real (or complex) orthonormal on Ω with respect to a probability measure μ basis for Y_N. Assume that $\|u_i\|_4 := \|u_i\|_{L_4(\Omega,\mu)} < \infty$ for all $i = 1, \dots, N$. Then for any $\delta > 0$ there exists a set $\Omega_M = \{x^j\}_{j=1}^M \subset \Omega$ such that for any $f \in Y_N$*

$$\big|\|f\|^2_{L_2(\Omega,\mu)} - \|f\|^2_{L_2(\Omega_M,\mu_M)}\big| \le \delta \|f\|^2_{L_2(\Omega,\mu)},$$

where

$$\|f\|^2_{L_2(\Omega_M,\mu_M)} := \frac{1}{M}\sum_{j=1}^M |f(x^j)|^2.$$

The following generalization of Theorem 3.1, which is equivalent to Theorem 1.1, is the main result of the paper.

Theorem 3.5. *Let $\Omega \subset \mathbb{R}^d$ be a nonempty compact set with the probability measure μ. Assume that $X_N \subset \mathcal{C}(\Omega)$ satisfies the Nikol'skii inequality (3.5). Then there is an absolute constant C_1' such that there exists a set $\{\xi^j\}_{j=1}^m \subset \Omega$ of $m \le C_1' t^2 N$ points with the property: for any $f \in X_N$ we have*

$$C_2'\|f\|_2^2 \le \frac{1}{m}\sum_{j=1}^m |f(\xi^j)|^2 \le C_3' t^2 \|f\|_2^2, \tag{3.6}$$

where C_2' and C_3' are absolute positive constants.

Proof. For a given $\delta \in (0,1)$, taking into account Proposition 3.3, we find a set $\Omega_M = \{x^j\}_{j=1}^M$ such that for any $f \in X_N$

$$\big|\|f\|^2_{L_2(\Omega,\mu)} - \|f\|^2_{L_2(\Omega_M,\mu_M)}\big| \le \delta \|f\|^2_{L_2(\Omega,\mu)}. \tag{3.7}$$

Specify $\delta = 1/2$. Then, clearly, subspace X_N restricted to Ω_M (denote it by Y_l) satisfies the Nikol'skii inequality (3.5) with t replaced by $2t$. Let $u_1, \dots, u_l$, $l \le N$, be an orthonormal basis of Y_l. By Proposition 3.2 inequality (3.5) is equivalent to (3.1). Now applying Theorem 3.1 to Y_l we find a subset

$J \subset \{1, 2, \dots, M\}$ with the property: $m := |J| \le C_1(2t)^2N$ and for any $f \in X_N$ we have

$$C_2\|f\|^2_{L_2(\Omega_M,\mu_M)} \le \frac{1}{m}\sum_{j\in J}|f(x^j)|^2 \le C_3t^2\|f\|^2_{L_2(\Omega_M,\mu_M)},$$

where C_2 and C_3 are absolute positive constants from Theorem 3.1. From here and (3.7) with $\delta = 1/2$ we obtain (3.6). $\square$

Remark 3.3. *In Theorem 3.5 we assume that $X_N \subset \mathcal{C}(\Omega)$. It is done for convenience. The statement of Theorem 3.5 holds if instead of continuity assumption we require that X_N is a subspace of the space $\mathcal{B}(\Omega, \mu)$ of functions, which are bounded and measurable with respect to μ on Ω, where Ω is a nonempty subset of $\mathbb{R}^d$.*

Acknowledgements. The authors are grateful to Boris Kashin for very useful comments, in particular, for bringing the paper [17] to their attention.

The work was supported by the Russian Science Foundation (project no. 21-11-00131) at Lomonosov Moscow State University (Sections 1,2) and by the Russian Federation Government Grant №14.W03.31.0031 (Section 3)

References

[1] J. Batson, D.A. Spielman, and N. Srivastava, Twice-Ramanujan Sparsifiers, *SIAM Review*, **56** (2014), 315–334.

[2] J. Bourgain, J. Lindenstrauss and V. Milman, Approximation of zonoids by zonotopes, *Acta Math.*, **162** (1989), 73–141.

[3] F. Dai, A. Prymak, V.N. Temlyakov, and S.U. Tikhonov, Integral norm discretization and related problems, *Russian Math. Surveys* **74:4** (2019), 579–630. Translation from *Uspekhi Mat. Nauk* **74:4(448)** (2019), 3–58; arXiv:1807.01353v1.

[4] F. Dai, A. Prymak, A. Shadrin, V. Temlyakov, S. Tikhonov, Sampling Discretization of Integral Norms, *Constr. Approx.* (2021); https://doi.org/10.1007/s00365-021-09539-0; arXiv:2001.09320v1 [math.CA] 25 Jan 2020.

[5] F. Dai, A. Prymak, A. Shadrin, V. Temlyakov, and S. Tikhonov, Entropy numbers and Marcinkiewicz-type discretization theorem, arXiv:2001.10636v1 [math.CA] 28 Jan 2020.

[6] N.J. Harvey and N. Olver, Pipage rounding, pessimistic estimators and matrix concentration, *Proc. of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms* (2014), 926–945.

[7] W.B. Johnson and G. Schechtman, Finite dimensional subspaces of L_p, Handbook of the geometry of Banach spaces, Vol. 1 (2001), 837–870, North-Holland, Amsterdam.

[8] B.S. Kashin, Lunin's method for selecting large submatrices with small norm, *Sb. Math.*, **206:7** (2015), 980–987.

[9] B.S. Kashin, I.V. Limonova, Decomposing a Matrix into two Submatrices with Extremally Small $(2,1)$–Norm, Math. Notes, **106**:1 (2019), 63–70. Translation from *Mat. Zametki*, **106:1** (2019), 53–61.

[10] B.S. Kashin and V.N. Temlyakov, Observations on discretization of trigonometric polynomials with given spectrum, *Russian Math. Surveys*, **73:6** (2018), 1128–1130. Translation from *Uspekhi Mat. Nauk* **73:6** (2018) 197–198.

[11] E. Kosov, Marcinkiewicz-type discretization of L^p-norms under the Nikolskii-type inequality assumption, arXiv:2005.01674v1 [math.FA] 4 May 2020.

[12] I.V. Limonova, Decomposing a Matrix into Two Submatrices with Extremely Small Operator Norm, *Math. Notes*, **108:1** (2020), 137—141. Translation from *Mat. Zametki*, **108:1** (2020), 153–157; arXiv:2003.00979v1.

[13] I. Limonova and V. Temlyakov, On sampling discretization in L_2, arXiv:2009.10789v1 [math.FA] 22 Sep 2020.

[14] A. A. Lunin, Operator norms of submatrices, *Math. Notes*, **45:3** (1989), 248–252. Translation from *Mat. Zametki*, **45:3** (1989) 94–100.

[15] A. Marcus, D.A. Spielman, and N. Srivastava, Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem, *Annals of Math.*, **182:1** (2015), 327–350.

[16] N. Nagel, M. Schäfer, T. Ullrich, A new upper bound for sampling numbers, arXiv:2010.00327v1 [math.NA] 30 Sep 2020.

[17] S. Nitzan, A. Olevskii, and A. Ulanovskii, Exponential frames on unbounded sets, *Proc. Amer. Math. Soc.*, **144:1** (2016), 109–118.

[18] A.M. Olevskii and A. Ulanovskii, Functions with Disconnected Spectrum: Sampling, Interpolation, Translates, Amer. Math. Soc., University Lecture Series, **65**, Providence, Rhode Island, 2016.

[19] M. Rudelson, Almost orthogonal submatrices of an orthogonal matrix, *Israel J. Math.*, **111** (1999), 143–155.

[20] G. Schechtman, More on embedding subspaces of L_p in ℓ_r^n, *Compositio Math.*, **61:2** (1987), 159–169.

[21] V.N. Temlyakov, The Marcinkiewicz-type discretization theorems for the hyperbolic cross polynomials, *Jaen Journal on Approximation*, **9** (2017), No. 1, 37–63; arXiv: 1702.01617v2.

[22] V.N. Temlyakov, The Marcinkiewicz-Type Discretization Theorems, *Constructive Approximation*, **48** (2018), 337–369.

[23] V.N. Temlyakov, On optimal recovery in L_2, J. Complexity (2020), Available online 19 December 2020, 101545; doi: https://doi.org/10.1016/j.jco.2020.101545; arXiv:2010.03103v1 [math.NA] 7 Oct 2020.

[24] V.N. Temlyakov and T. Ullrich, Bounds on Kolmogorov widths and sampling recovery for classes with small mixed smoothness, arXiv:2012.09925v1 [math.NA] 17 Dec 2020.

[25] A. Zygmund, Trigonometric Series, Cambridge University Press, 1959.