Sampling discretization of integral norms. Lecture 1

Vladimir Temlyakov

Chemnitz, September, 2019
Two settings

We are interested in discretizing the $L_q$, $1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:
Two settings

We are interested in discretizing the $L_q, 1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an $N$-dimensional subspace $X_N$. We call such results the Marcinkiewicz-type discretization theorems.
We are interested in discretizing the $L_q$, $1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an $N$-dimensional subspace $X_N$. We call such results the Marcinkiewicz-type discretization theorems.
- Functions belong to a given function class.
We are interested in discretizing the $L_q$, $1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an $N$-dimensional subspace $X_N$. We call such results the Marcinkiewicz-type discretization theorems.

- Functions belong to a given function class. There are different settings and different ingredients, which play important role in this problem.
We are interested in discretizing the $L_q$, $1 \leq q \leq \infty$, norm of functions defined on a compact subset of $\mathbb{R}^d$. We distinguish right a way two settings:

- Functions belong to an $N$-dimensional subspace $X_N$. We call such results the Marcinkiewicz-type discretization theorems.

- Functions belong to a given function class. There are different settings and different ingredients, which play important role in this problem.

We begin with the Marcinkiewicz-type discretization.
Let $\Omega$ be a compact subset of $\mathbb{R}^d$ with the probability measure $\mu$. We say that a linear subspace $X_N$ of the $L_q(\Omega)$, $1 \leq q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters $m$ and $q$ if there exist a set $\{\xi^\nu \in \Omega, \nu = 1, \ldots, m\}$ and two positive constants $C_j(d, q)$, $j = 1, 2$, such that for any $f \in X_N$ we have

$$C_1(d, q)\|f\|_q^q \leq \frac{1}{m} \sum_{\nu=1}^{m} |f(\xi^\nu)|^q \leq C_2(d, q)\|f\|_q^q. \quad (1)$$
Marcinkiewicz problem

Let $\Omega$ be a compact subset of $\mathbb{R}^d$ with the probability measure $\mu$. We say that a linear subspace $X_N$ of the $L_q(\Omega)$, $1 \leq q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters $m$ and $q$ if there exist a set $\{\xi^\nu \in \Omega, \nu = 1, \ldots, m\}$ and two positive constants $C_j(d, q)$, $j = 1, 2$, such that for any $f \in X_N$ we have

$$C_1(d, q)\|f\|_q^q \leq \frac{1}{m} \sum_{\nu=1}^{m} |f(\xi^\nu)|^q \leq C_2(d, q)\|f\|_q^q. \quad (1)$$

In the case $q = \infty$ we define $L_\infty$ as the space of continuous on $\Omega$ functions and ask for

$$C_1(d)\|f\|_\infty \leq \max_{1 \leq \nu \leq m} |f(\xi^\nu)| \leq \|f\|_\infty. \quad (2)$$
Let $\Omega$ be a compact subset of $\mathbb{R}^d$ with the probability measure $\mu$. We say that a linear subspace $X_N$ of the $L_q(\Omega)$, $1 \leq q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters $m$ and $q$ if there exist a set $\{\xi^\nu \in \Omega, \nu = 1, \ldots, m\}$ and two positive constants $C_j(d, q)$, $j = 1, 2$, such that for any $f \in X_N$ we have

$$C_1(d, q)\|f\|_q^q \leq \frac{1}{m} \sum_{\nu=1}^{m} |f(\xi^\nu)|^q \leq C_2(d, q)\|f\|_q^q. \quad (1)$$

In the case $q = \infty$ we define $L_\infty$ as the space of continuous on $\Omega$ functions and ask for

$$C_1(d)\|f\|_\infty \leq \max_{1 \leq \nu \leq m} |f(\xi^\nu)| \leq \|f\|_\infty. \quad (2)$$

We will also use a brief way to express the above property: the $\mathcal{M}(m, q)$ theorem holds for a subspace $X_N$ or $X_N \in \mathcal{M}(m, q)$. 
Discretization for the trigonometric polynomials

We briefly present well known results related to the Marcinkiewicz-type discretization theorems for the trigonometric polynomials.

\[ \Pi(N) := \prod_{j=1}^{d} [ -N_j, N_j ] \times \cdots \times [ -N_d, N_d ], \quad N_j \in \mathbb{N} \text{ or } N_j = 0, \quad j = 1, \ldots, d. \]

Denote \[ P'(N) := \{ n = (n_1, \ldots, n_d) : n_j - \text{are natural numbers}, \quad 0 \leq n_j \leq 4N_j - 1, \quad j = 1, \ldots, d \} \]
and set \[ x(n) := (\pi n_1 2^{N_1}, \ldots, \pi n_d 2^{N_d}), \quad n \in P'(N). \]
We briefly present well known results related to the Marcinkiewicz-type discretization theorems for the trigonometric polynomials. We begin with the case

\[ \Pi(\mathbf{N}) := [-N_1, N_1] \times \cdots \times [-N_d, N_d], \quad N_j \in \mathbb{N} \text{ or } N_j = 0, \]

\[ j = 1, \ldots, d, \quad \mathbf{N} = (N_1, \ldots, N_d). \]
We briefly present well known results related to the Marcinkiewicz-type discretization theorems for the trigonometric polynomials. We begin with the case 

\[ \Pi(\mathbf{N}) := [-N_1, N_1] \times \cdots \times [-N_d, N_d], \quad N_j \in \mathbb{N} \text{ or } N_j = 0, \quad j = 1, \ldots, d, \quad \mathbf{N} = (N_1, \ldots, N_d). \]

Denote 

\[ P'(\mathbf{N}) := \{ \mathbf{n} = (n_1, \ldots, n_d), \quad n_j - \text{ are natural numbers,} \]

\[ 0 \leq n_j \leq 4N_j - 1, \quad j = 1, \ldots, d \} \]
We briefly present well known results related to the Marcinkiewicz-type discretization theorems for the trigonometric polynomials. We begin with the case
\[ \Pi(N) := [-N_1, N_1] \times \cdots \times [-N_d, N_d], \ N_j \in \mathbb{N} \text{ or } N_j = 0, \]
\[ j = 1, \ldots, d, \ N = (N_1, \ldots, N_d). \]
Denote
\[ P'(N) := \left\{ n = (n_1, \ldots, n_d), \quad n_j - \text{ are natural numbers,} \right\} \]
and set
\[ x(n) := \left( \frac{\pi n_1}{2N_1}, \ldots, \frac{\pi n_d}{2N_d} \right), \quad n \in P'(N). \]
Marcinkiewicz-type theorem for $\mathcal{T}(\Pi(\mathbb{N}))$

In the case $N_j = 0$ we assume $x_j(n) = 0$. Denote $\overline{N} := \max(N, 1)$ and $\nu(N) := \prod_{j=1}^{d} \overline{N}_j$. Then the following Marcinkiewicz-type discretization theorem is known for all $1 \leq q \leq \infty$: for any $f \in \mathcal{T}(\Pi(\mathbb{N}))$

$$C_1(d, q)\|t\|_q^q \leq \nu(4N)^{-1} \sum_{n \in \mathcal{P}'} |f(x(n))|^q \leq C_2(d, q)\|t\|_q^q,$$  \hspace{1cm} (5)
Marcinkiewicz-type theorem for $\mathcal{T}(\Pi(\mathbb{N}))$

In the case $N_j = 0$ we assume $x_j(n) = 0$. Denote $\overline{N} := \max(N, 1)$ and $v(N) := \prod_{j=1}^{d} \overline{N}_j$. Then the following Marcinkiewicz-type discretization theorem is known for all $1 \leq q \leq \infty$: for any $f \in \mathcal{T}(\Pi(\mathbb{N}))$

$$C_1(d, q)\|t\|_q^q \leq v(4N)^{-1} \sum_{n \in P'(N)} |f(x(n))|^q \leq C_2(d, q)\|t\|_q^q,$$

(5)

which implies the following relation

$$\mathcal{T}(\Pi(\mathbb{N})) \in \mathcal{M}(v(4N), q), \quad 1 \leq q \leq \infty.$$ 

Note that $v(4N) \leq C(d) \dim \mathcal{T}(\Pi(\mathbb{N})).$
In the case $N_j = 0$ we assume $x_j(n) = 0$. Denote $\overline{N} := \max(N, 1)$ and $v(N) := \prod_{j=1}^{d} \overline{N}_j$. Then the following Marcinkiewicz-type discretization theorem is known for all $1 \leq q \leq \infty$: for any $f \in \mathcal{T}(\Pi(N))$

$$C_1(d, q) \|t\|_q^q \leq v(4N)^{-1} \sum_{n \in P'(N)} |f(x(n))|^q \leq C_2(d, q) \|t\|_q^q, \quad (5)$$

which implies the following relation

$$\mathcal{T}(\Pi(N)) \in \mathcal{M}(v(4N), q), \quad 1 \leq q \leq \infty.$$ 

Note that $v(4N) \leq C(d) \dim \mathcal{T}(\Pi(N))$. It is clear from the above construction that the set $\{x(n) : n \in P'(N)\}$ depends substantially on $N$. The main goal of this paper is to construct for a given $q$ and $M$ a set, which satisfies an analog of (5) for all $N$ with $v(N) \leq M$. 

Vladimir Temlyakov
Sampling discretization of integral norms. Lecture 1
Let $Q$ be a finite subset of $\mathbb{Z}^d$. We denote

$$T(Q) := \{ f : f = \sum_{k \in Q} c_k e^{i(k,x)} \}.$$
Let $Q$ be a finite subset of $\mathbb{Z}^d$. We denote

$$T(Q) := \{ f : f = \sum_{k \in Q} c_k e^{i(k,x)} \}.$$ 

The following result was obtained by VT, 2017.

**Theorem (1; VT, 2017)**

There are three positive absolute constants $C_1$, $C_2$, and $C_3$ with the following properties: For any $d \in \mathbb{N}$ and any $Q \subset \mathbb{Z}^d$ there exists a set of $m \leq C_1 |Q|$ points $\xi_j \in \mathbb{T}^d$, $j = 1, \ldots, m$ such that for any $f \in T(Q)$ we have

$$C_2 \| f \|_2^2 \leq \frac{1}{m} \sum_{j=1}^{m} |f(\xi_j)|^2 \leq C_3 \| f \|_2^2.$$
The above theorem is based on the following lemma from S. Nitzan, A. Olevskii, and A. Ulanovskii, 2016.

**Lemma (NOU, 2016)**

Let a system of vectors $v_1, \ldots, v_M$ from $\mathbb{C}^N$ have the following properties: for all $w \in \mathbb{C}^N$ we have $\sum_{j=1}^{M} |\langle w, v_j \rangle|^2 = \|w\|^2_2$ and $\|v_j\|^2_2 = N/M$, $j = 1, \ldots, M$. Then there is a subset $J \subset \{1, 2, \ldots, M\}$ such that for all $w \in \mathbb{C}^N$

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

where $c_0$ and $C_0$ are some absolute positive constants.
The above Lemma was derived from the following theorem from A. Marcus, D.A. Spielman, and N. Srivastava, 2015, which solved the **Kadison-Singer problem**.

**Theorem (MSS, 2015)**

Let a system of vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_M \) from \( \mathbb{C}^N \) have the following properties: for all \( \mathbf{w} \in \mathbb{C}^N \) we have \( \sum_{j=1}^{M} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 = \| \mathbf{w} \|^2_2 \) and \( \| \mathbf{v}_j \|^2_2 \leq \epsilon \).

Then there exists a partition of \( \{1, \ldots, M\} \) into two sets \( S_1 \) and \( S_2 \), such that for each \( i = 1, 2 \) we have for all \( \mathbf{w} \in \mathbb{C}^N \)

\[
\sum_{j \in S_i} |\langle \mathbf{w}, \mathbf{v}_j \rangle|^2 \leq \frac{(1 + \sqrt{2\epsilon})^2}{2} \| \mathbf{w} \|^2_2.
\]
Let $\Pi(\mathbf{N}) := [-N_1, N_1] \times \cdots \times [-N_d, N_d]$, $N_j \in \mathbb{N}$ or $N_j = 0$, $j = 1, \ldots, d$, $\mathbf{N} = (N_1, \ldots, N_d)$. The following result is obtained by VT, 2017.
Discretization for the trigonometric polynomials in $L_1$

Let $\Pi(N) := [-N_1, N_1] \times \cdots \times [-N_d, N_d]$, $N_j \in \mathbb{N}$ or $N_j = 0$, $j = 1, \ldots, d$, $N = (N_1, \ldots, N_d)$. The following result is obtained by VT, 2017.

**Theorem (2; VT, 2017)**

Let $d \in \mathbb{N}$. For any $n \in \mathbb{N}$ and any $Q \subset \Pi(N)$ with $N = (2^n, \ldots, 2^n)$ there exists a set of $m \leq C_1(d)|Q|n^{7/2}$ points $\xi_j \in \mathbb{T}^d$, $j = 1, \ldots, m$ such that for any $f \in \mathcal{T}(Q)$ we have

$$C_2(d)\|f\|_1 \leq \frac{1}{m} \sum_{j=1}^{m} |f(\xi_j)| \leq C_3(d)\|f\|_1.$$
We discussed in Kashin and Temlyakov, 2018, the following setting of the discretization problem of the uniform norm. Let $S_m := \{\xi_j\}_{j=1}^m \subset \mathbb{T}^d$ be a finite set of points. Clearly,

$$\|f\|_{L^\infty(S_m)} := \max_{1 \leq j \leq m} |f(\xi_j)| \leq \|f\|_{\infty}.$$ 

We are interested in estimating the following quantities

$$D(Q, m) := D(Q, m, d) := \inf_{S_m} \sup_{f \in \mathcal{T}(Q)} \frac{\|f\|_{\infty}}{\|f\|_{L^\infty(S_m)}},$$

$$D(N, m) := D(N, m, d) := \sup_{Q, |Q|=N} D(Q, m, d).$$
Certainly, one should assume that $m \geq N$. Then the characteristic $D(Q, m)$ guarantees that there exists a set of $m$ points $S_m$ such that for any $f \in \mathcal{T}(Q)$ we have

$$\|f\|_\infty \leq D(Q, m) \|f\|_{L_\infty(S_m)}.$$ 

In the case $d = 1$ and $Q = [-n, n]$ classical Marcinkiewicz theorem gives for $m \geq 4n$ that $D([-n, n], 4n) \leq C$. Similar relation holds for $D([-n_1, n_1] \times \cdots \times [-n_d, n_d], (4n_1) \times \cdots \times (4n_d))$. 
It was proved in Kashin and Temlyakov, 2018, that for a pair \( N, m \), such that \( m \asymp N \) we have \( D(N, m) \asymp N^{1/2} \). We formulate this result as a theorem.

**Theorem (KT, 2018)**

*For any constant \( c \geq 1 \) there exists a positive constant \( C \) such that for any pair of parameters \( N, m \), with \( m \leq cN \) we have*

\[
D(N, m) \geq CN^{1/2}.
\]

*Also, there are two positive absolute constants \( c_1 \) and \( C_1 \) with the following property: For any \( d \in \mathbb{N} \) we have for \( m \geq c_1 N \)*

\[
D(N, m, d) \leq C_1 N^{1/2}.
\]
Recall that the set of hyperbolic cross polynomials is defined as

\[ \mathcal{T}(N) := \mathcal{T}(N, d) := \left\{ f : f = \sum_{k \in \Gamma(N)} c_k e^{i(k,x)} \right\}, \]

where \( \Gamma(N) \) is the hyperbolic cross

\[ \Gamma(N) := \Gamma(N, d) := \left\{ k \in \mathbb{Z}^d : \prod_{j=1}^d \max\{|k_j|, 1\} \leq N \right\}. \]

Throughout this section, we define

\[ \alpha_d := \sum_{j=1}^d \frac{1}{j} \quad \text{and} \quad \beta_d := d - \alpha_d. \]

We use the following notation here. For \( x \in \mathbb{T}^d \) and \( j \in \{1, \ldots, d\} \) we denote \( x^j := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d) \).
The following result was obtained by Dai, Prymak, Temlyakov, and Tikhonov, 2018.

**Theorem (DPTT, 2018)**

For each $d \in \mathbb{N}$ and each $N \in \mathbb{N}$ there exists a set $W(N, d)$ of at most $C_d N^\alpha d (\log N)^\beta d$ points in $[0, 2\pi)^d$ such that for all $f \in T(N)$, 

$$
\|f\|_\infty \leq C(d) \max_{w \in W(N, d)} |f(w)|.
$$
Some historical remarks

It is well known that

\[ T(\Pi(N)) \in \mathcal{M}(C(d)N^d, \infty), \]

\[ \Pi(N) := \{ k \in \mathbb{Z}^d : |k_j| \leq N, j = 1, \ldots, d \}. \]

In particular, this implies that

\[ T(N) \in \mathcal{M}(C(d)N^d, \infty). \]

Theorem DPTT shows that we can improve the above relation to

\[ T(N) \in \mathcal{M}(C(d)N^{\alpha_d}(\log N)^{\beta_d}, \infty). \]

Note that \( \alpha_d \asymp \ln d \).
A trivial lower bound for $m$ in the inclusion $T(N) \in \mathcal{M}(m, \infty)$ is $m \geq \dim(T(N)) \asymp N(\log N)^{d-1}$. The following nontrivial lower bound was obtained in Kashin and Temlyakov, 1998.

**Theorem (KT, 1998)**

Let a set $W \subset \mathbb{T}^2$ have a property:

$$\forall t \in T(N) \quad \|t\|_{\infty} \leq b(\log N)^{\alpha} \max_{w \in W} |t(w)|$$

with some $0 \leq \alpha < 1/2$. Then

$$|W| \geq C_1 N \log Ne^{C_2 b^{-2}(\log N)^{1-2\alpha}}.$$ 

In particular, Theorem KT with $\alpha = 0$ implies that a necessary condition on $m$ for inclusion $T(N) \in \mathcal{M}(m, \infty)$ is $m \geq \dim(T(N))N^c$ with positive absolute constant $c$. 
An operator $T_N$ with the following properties was constructed in Temlyakov, 1993. The operator $T_N$ has the form

$$T_N(f) = \sum_{j=1}^{m} f(x^j)\psi_j(x), \quad m \leq c(d)N(\log N)^{d-1}, \quad \psi_j \in \mathcal{T}(N2^d)$$

and

$$T_N(f) = f, \quad f \in \mathcal{T}(N), \quad (3)$$

$$\|T_N\|_{L_\infty \rightarrow L_\infty} \asymp (\log N)^{d-1}. \quad (4)$$
Points \( \{x^j\} \) are from the Smolyak net. Properties (3) and (4) imply that all \( f \in \mathcal{T}(N) \) satisfy the discretization inequality

\[
\|f\|_\infty \leq C(d)(\log N)^{d-1} \max_{1 \leq j \leq m} |f(x^j)|.
\]
Some remarks for the case $q = 2$

We describe the properties of the subspace $X_N$ in terms of a system $\mathcal{U}_N := \{u_i\}_{i=1}^N$ of functions such that $X_N = \text{span}\{u_i, i = 1, \ldots, N\}$. In the case $X_N \subset L_2$ we assume that the system is orthonormal on $\Omega$ with respect to measure $\mu$. In the case of real functions we associate with $x \in \Omega$ the matrix $G(x) := [u_i(x)u_j(x)]_{i,j=1}^N$. Clearly, $G(x)$ is a symmetric positive semi-definite matrix of rank 1. It is easy to see that for a set of points $\xi^k \in \Omega$, $k = 1, \ldots, m$, and $f = \sum_{i=1}^N b_i u_i$ we have

$$\sum_{k=1}^m \lambda_k f(\xi^k)^2 - \int_{\Omega} f(x)^2 \, d\mu = b^T \left( \sum_{k=1}^m \lambda_k G(\xi^k) - I \right) b,$$

where $b = (b_1, \ldots, b_N)^T$ is the column vector and $I$ is the identity matrix.
Therefore, the $\mathcal{M}^w(m, 2)$ problem is closely connected with a problem of approximation (representation) of the identity matrix $I$ by an $m$-term approximant with respect to the system $\{G(x)\}_{x \in \Omega}$. It is easy to understand that under our assumptions on the system $\mathcal{U}_N$ there exist a set of knots $\{\xi^k\}_{k=1}^m$ and a set of weights $\{\lambda_k\}_{k=1}^m$, with $m \leq N^2$ such that

$$I = \sum_{k=1}^{m} \lambda_k G(\xi^k)$$

and, therefore, we have for any $X_N \subset L_2$ that

$$X_N \in \mathcal{M}^w(N^2, 2, 0).$$
We begin with formulation of the Rudelson result from 1999. Let $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu(x^j) = 1/M$, $j = 1, \ldots, M$. Assume that $\{u_i(x)\}_{i=1}^N$ is a real orthonormal on $\Omega_M$ system satisfying the following condition: 

**Condition E.** For all $j$

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

with some $t \geq 1$. 
Then for every $\epsilon > 0$ there exists a set $J \subset \{1, \ldots, M\}$ of indices with cardinality

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

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

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

Theorem (VT, 2017)

Let \( \{u_i\}_{i=1}^N \) be an orthonormal system, satisfying condition \( \mathbf{E} \). Then for every \( \epsilon > 0 \) there exists a set \( \{\xi_j\}_{j=1}^m \subset \Omega \) with

\[
m \leq C \frac{t^2}{\epsilon^2} N \log N
\]

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

\[
(1 - \epsilon)\|f\|_2^2 \leq \frac{1}{m} \sum_{j=1}^m f(\xi_j)^2 \leq (1 + \epsilon)\|f\|_2^2.
\]
We now comment on a recent breakthrough result by J. Batson, D.A. Spielman, and N. Srivastava, 2012. We formulate their result in our notations. Let as above $\Omega_M = \{x^j\}_{j=1}^M$ be a discrete set with the probability measure $\mu(x^j) = 1/M$, $j = 1, \ldots, M$. Assume that $\{u_i(x)\}_{i=1}^N$ is a real orthonormal on $\Omega_M$ system. Then for any number $d > 1$ there exist a set of weights $w_j \geq 0$ such that $|\{j : w_j \neq 0\}| \leq dN$ so that for any $f \in \text{span}\{u_1, \ldots, u_N\}$ we have

$$\|f\|_2^2 \leq \sum_{j=1}^M w_j f(x^j)^2 \leq \frac{d + 1 + 2\sqrt{d}}{d + 1 - 2\sqrt{d}} \|f\|_2^2.$$
The proof of this result is based on a delicate study of the $m$-term approximation of the identity matrix $I$ with respect to the system $\mathcal{D} := \{ G(x) \}_{x \in \Omega}$, $G(x) := [u_i(x)u_j(x)]_{i,j=1}^N$ in the spectral norm. The authors control the change of the maximal and minimal eigenvalues of a matrix, when they add a rank one matrix of the form $wG(x)$. Their proof provides an algorithm for construction of the weights $\{w_j\}$. In particular, this implies that

$$X_N(\Omega_M) \in M^w(m, 2, \epsilon) \quad \text{provided} \quad m \geq CN\epsilon^{-2}$$

with large enough $C$. 
Let $X$ be a Banach space and let $B_X$ denote the unit ball of $X$ with the center at $0$. Denote by $B_X(y, r)$ a ball with center $y$ and radius $r$: $\{x \in X : \|x - y\| \leq r\}$. For a compact set $A$ and a positive number $\varepsilon$ we define the covering number $N_\varepsilon(A, X)$ as follows

$$N_\varepsilon(A, X) := \min\{n : \exists y_1, \ldots, y_n, y_j \in A : A \subseteq \bigcup_{j=1}^{n} B_X(y_j, \varepsilon)\}.$$
Let $X$ be a Banach space and let $B_X$ denote the unit ball of $X$ with the center at $0$. Denote by $B_X(y, r)$ a ball with center $y$ and radius $r$: $\{x \in X : \|x - y\| \leq r\}$. For a compact set $A$ and a positive number $\varepsilon$ we define the covering number $N_{\varepsilon}(A, X)$ as follows

$$N_{\varepsilon}(A, X) := \min\{n : \exists y^1, \ldots, y^n, y^j \in A : A \subseteq \bigcup_{j=1}^n B_X(y^j, \varepsilon)\}.$$ 

It is convenient to consider along with the entropy $H_{\varepsilon}(A, X) := \log_2 N_{\varepsilon}(A, X)$ the entropy numbers $\varepsilon_k(A, X)$:

$$\varepsilon_k(A, X) := \inf\{\varepsilon : \exists y^1, \ldots, y^{2^k} \in A : A \subseteq \bigcup_{j=1}^{2^k} B_X(y^j, \varepsilon)\}.$$ 

In our definition of $N_{\varepsilon}(A, X)$ and $\varepsilon_k(A, X)$ we require $y^j \in A$. In a standard definition of $N_{\varepsilon}(A, X)$ and $\varepsilon_k(A, X)$ this restriction is not imposed. However, it is well known that these characteristics may differ at most by a factor 2.
Theorem (4; VT2017)

Suppose that a real $N$-dimensional subspace $X_N$ satisfies the following condition on the entropy numbers of the unit ball $X_N^1 := \{ f \in X_N : \|f\|_1 \leq 1 \}$ with $B \geq 1$

$$\varepsilon_k(X_N^1, L_\infty) \leq B \begin{cases} \frac{N}{k}, & k \leq N, \\ 2^{-k/N}, & k \geq N. \end{cases}$$

Then there exists a set of $m \leq C_1 NB(\log_2(2N \log_2(8B)))^2$ points $\xi^j \in \Omega, j = 1, \ldots, m$, with large enough absolute constant $C_1$, such that for any $f \in X_N$ we have

$$\frac{1}{2} \|f\|_1 \leq \frac{1}{m} \sum_{j=1}^{m} |f(\xi^j)| \leq \frac{3}{2} \|f\|_1.$$
Concentration measure lemma

The following lemma is from J. Bourgain, J. Lindenstrauss and V. Milman, 1989.

**Lemma (BLM, 1989)**

Let \( \{g_j\}_{j=1}^m \) be independent random variables with \( \mathbb{E}g_j = 0 \), \( j = 1, \ldots, m \), which satisfy

\[
\|g_j\|_1 \leq 2, \quad \|g_j\|_\infty \leq M, \quad j = 1, \ldots, m.
\]

Then for any \( \eta \in (0, 1) \) we have the following bound on the probability

\[
P\left\{ \left| \sum_{j=1}^m g_j \right| \geq m\eta \right\} < 2 \exp\left(-\frac{m\eta^2}{8M}\right).
\]
We now consider measurable functions \( f(x) \), \( x \in \Omega \). For \( 1 \leq q < \infty \) define

\[
L^q_z(f) := \frac{1}{m} \sum_{j=1}^m |f(x^j)|^q - \|f\|_q^q, \quad z := (x^1, \ldots, x^m).
\]

Let \( \mu \) be a probabilistic measure on \( \Omega \). Denote \( \mu^m := \mu \times \cdots \times \mu \) the probabilistic measure on \( \Omega^m := \Omega \times \cdots \times \Omega \). We need the following inequality, which is a corollary of the above Lemma.
Proposition (VT, 2017)

Let \( f_j \in L_1(\Omega) \) be such that

\[ \|f_j\|_1 \leq 1/2, \quad j = 1, 2; \quad \|f_1 - f_2\|_\infty \leq \delta. \]

Then

\[ \mu^m \{ z : |L_z^1(f_1) - L_z^1(f_2)| \geq \eta \} < 2 \exp \left( -\frac{m\eta^2}{16\delta} \right). \] (5)
We consider the case $X$ is $C(\Omega)$ the space of functions continuous on a compact subset $\Omega$ of $\mathbb{R}^d$ with the norm
\[ \|f\|_\infty := \sup_{x \in \Omega} |f(x)|. \]

We use the abbreviated notations
\[ \varepsilon_n(W) := \varepsilon_n(W, C). \]

In our case
\[ W := W(Q) := \{t \in T(Q) : \|t\|_1 = 1/2\}. \]  
(6)
The entropy bound

Theorem (6; VT, 2017)

For any $Q \subset \Pi(N)$ with $N = (2^n, \ldots, 2^n)$ we have

$$\varepsilon_k(T(Q)_1, L_\infty) \leq 2\varepsilon_k := 2C_4(d) \left\{ \begin{array}{ll} n^{3/2}(|Q|/k), & k \leq 2|Q|, \\ n^{3/2}2^{-k/(2|Q|)}, & k \geq 2|Q|. \end{array} \right.$$
The entropy bound

**Theorem (6; VT, 2017)**

For any $Q \subset \Pi(N)$ with $N = (2^n, \ldots, 2^n)$ we have

$$
\varepsilon_k(T(Q)_1, L_{\infty}) \leq 2\varepsilon_k := 2C_4(d) \left\{ \begin{array}{ll}
n^{3/2}(|Q|/k), & k \leq 2|Q|, \\
n^{3/2}2^{-k/(2|Q|)}, & k \geq 2|Q|.
\end{array} \right.
$$

Specify $\eta = 1/4$. Denote $\delta_j := \varepsilon_{2j}$, $j = 0, 1, \ldots$, and consider minimal $\delta_j$-nets $\mathcal{N}_j \subset W$ of $W$ in $C(\mathbb{T}^d)$. We use the notation $\mathcal{N}_j := |\mathcal{N}_j|$. Let $J$ be the minimal $j$ satisfying $\delta_j \leq 1/16$. For $j = 1, \ldots, J$ we define a mapping $A_j$ that associates with a function $f \in W$ a function $A_j(f) \in \mathcal{N}_j$ closest to $f$ in the $C$ norm.
Building a chain

Then, clearly,

$$\|f - A_j(f)\|_c \leq \delta_j.$$ 

We use the mappings $A_j, j = 1, \ldots, J$ to associate with a function $f \in W$ a sequence (a chain) of functions $f_j, f_{j-1}, \ldots, f_1$ in the following way

$$f_j := A_j(f), \quad f_j := A_j(f_{j+1}), \quad j = J - 1, \ldots, 1.$$
Reduction to simple events

Set

\[ \eta_j := \frac{1}{16nd}, \quad j = 1, \ldots, J. \]

Rewriting

\[ L_z^1(f_j) = L_z^1(f_j) - L_z^1(f_{j-1}) + \cdots + L_z^1(f_2) - L_z^1(f_1) + L_z^1(f_1) \]

we conclude that if \(|L_z^1(f)| \geq 1/4\) then at least one of the following events occurs:

\[ |L_z^1(f_j) - L_z^1(f_{j-1})| \geq \eta_j \quad \text{for some} \quad j \in (1, J) \quad \text{or} \quad |L_z^1(f_1)| \geq \eta_1. \]

In the rest of the proof we use the Proposition to estimate accurately the probability of the above events.