ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Prof. Dr. Helmut Bölcskei
Chair for Mathematical Information Science

Examination on Mathematics of Information
February 8, 2021

Please note:

- Exam duration: 180 minutes
- Maximum number of points: 100
- You are allowed to use any printed or handwritten material (i.e., books, lecture and discussion session notes, summaries), but no computers, tablets, smart phones or other electronic devices.
- Your solutions should be explained in detail and your handwriting needs to be clean and readable.
- Please do not use red or green pens. You may use pencils.
- Please note that the ETHZ "Disziplinarordnung RSETHZ 361.1" applies.

Before you start:

1. The problem statements consist of 7 pages including this page. Please verify that you have received all 7 pages.
2. Please fill in your name, student ID card number and signature below.
3. Please place your student ID card at the front of your desk so we can verify your identity.

During the exam:

4. For your solutions, please use only the empty sheets provided by us. Should you need additional sheets, please let us know.
5. Each problem consists of several subproblems. If you do not provide a solution to a subproblem, you may, whenever applicable, nonetheless assume its conclusion in the ensuing subproblems.

After the exam:

6. Please write your name on every sheet and prepare all sheets in a pile. All sheets, including those containing problem statements, must be handed in.
7. Please clean up your desk and stay seated and silent until you are allowed to leave the room in a staggered manner row by row.
8. Please avoid crowding and leave the building by the most direct route.

Family name: .... First name: ....

Legi-No.: ....

Number of additional sheets handed in: ....

Signature: ....

Problem 1 (25 points)

In this problem, we define the *(s,t)-restricted orthogonality constant* $\theta_{s,t} = \theta_{s,t}(A)$ of a matrix $A \in \mathbb{C}^{m\times N}$ as the smallest $\theta \geq 0$ such that

$$|\langle Au, Av\rangle| \leq \theta \|u\|_2 \|v\|_2$$

for all disjointly supported s-sparse and t-sparse vectors $u \in \mathbb{C}^N$ and $v \in \mathbb{C}^N$, respectively. Here $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{C}^N$ and $m, N \in \mathbb{N}$.

Moreover, for a vector $u \in \mathbb{C}^N$, a matrix $B \in \mathbb{C}^{m\times N}$, and a set $S \subset \{1, \ldots, N\}$, we define $u_S \in \mathbb{C}^{|S|}$ to be the vector obtained from u by keeping only the entries indexed by S, and similarly, we define $B_S \in \mathbb{C}^{m\times |S|}$ to be the matrix obtained from B by keeping only the columns indexed by S.

(a) (12 points) Show that

$$\theta_{s,t} = \max\left\{\|A_T^H A_S\|_2,\ S, T \subset \{1, \ldots, N\},\ S \cap T = \emptyset,\ |S| \leq s,\ |T| \leq t\right\},$$

where $\|\cdot\|_2$ denotes the matrix operator norm with respect to the ℓ^2-norm on $\mathbb{C}^{|S|}$, i.e.,

$$\|A_T^H A_S\|_2 := \max_{\substack{u \in \mathbb{C}^{|S|} \\ \|u\|_2 \leq 1}} \|(A_T^H A_S)u\|_2.$$

(b) (6 points) Recall that the s-restricted isometry constant $\delta_s = \delta_s(A)$ of a matrix $A \in \mathbb{C}^{m\times N}$ is defined as the smallest $\delta \geq 0$ such that

$$(1 - \delta)\|u\|_2^2 \leq \|Au\|_2^2 \leq (1 + \delta)\|u\|_2^2$$

for all s-sparse vectors $u \in \mathbb{C}^N$. Moreover, recall the following result:

Lemma. *The restricted isometry constant δ_s of A can equivalently be expressed as*

$$\delta_s = \max_{S \subset \{1,\ldots,N\}, |S| \leq s} \|A_S^H A_S - \mathrm{I}_{|S|}\|_2,$$

where $\mathrm{I}_{|S|}$ is the $|S| \times |S|$ identity matrix.

Using this lemma, prove the following relation between the restricted isometry constant and the restricted orthogonality constant

$$\theta_{s,t} \leq \delta_{s+t}.$$

(c) (7 points) In this subproblem, we want to show that

$$\theta_{t,r} \leq \sqrt{\frac{t}{s}}\,\theta_{s,r},$$

where $r, s, t \geq 1$ are such that $t \geq s$. To this end, let $u \in \mathbb{C}^N$ be t-sparse, $v \in \mathbb{C}^N$ r-sparse and u and v are disjointly supported. Furthermore, let $T = \{j_1, j_2, \ldots, j_t\}$ denote the support set of u and consider the t subsets $S_1, S_2, \ldots, S_t \subset T$ of cardinality s defined by

$$S_i = \{j_i, j_{i+1}, \ldots, j_{i+s-1}\}, \quad \text{for all} \quad i \in \{1, 2, \ldots, t\},$$

where the indices are understood to be modulo t.

(i) (2 points) Show that

$$u = \frac{1}{s}\sum_{i=1}^{t} u_{S_i} \quad \text{and} \quad \|u\|_2^2 = \frac{1}{s}\sum_{i=1}^{t} \|u_{S_i}\|_2^2.$$

(ii) (5 points) Use (c)(i) to establish that

$$|\langle Au, Av\rangle| \leq \sqrt{\frac{t}{s}}\,\theta_{s,r}\,\|u\|_2\|v\|_2.$$

Problem 2 (25 points)

In this problem, for a finite set A, we denote by $\mathrm{card}(A) \in \mathbb{N}_0$ the cardinality of A. One way to define *compressibility* of a vector $x \in \mathbb{C}^N$ is to say that x is *compressible* if the number

$$\mathrm{card}\big(\{j \in \{1, \ldots, N\} \colon |x_j| \geq t\}\big)$$

of its significant components is small. This leads to the idea of quantifying compressibility through the following quasinorm on $\mathbb{C}^N$

$$\|x\|_{2,\infty} = \inf\left\{ M \geq 0 \colon \mathrm{card}\big(\{j \in \{1, \ldots, N\} \colon |x_j| \geq t\}\big) \leq \frac{M^2}{t^2}, \text{ for all } t > 0 \right\}.$$

In this problem, we will verify that $\|x\|_{2,\infty}$, indeed, constitutes a quasinorm, i.e., it satisfies the norm axioms, except for the triangle inequality which is replaced by

$$\|x + y\|_{2,\infty} \leq K(\|x\|_{2,\infty} + \|y\|_{2,\infty}) \tag{1}$$

for some $K > 0$. We will also compare the properties of $\|x\|_{2,\infty}$ with those of the usual ℓ_2-norm.

(a) (7 points) Show that for every $x \in \mathbb{C}^N$ and $\lambda \in \mathbb{C}$,

  (i) (4 points) $\|x\|_{2,\infty} = 0 \implies x = 0$,

  (ii) (3 points) $\|\lambda x\|_{2,\infty} = |\lambda| \|x\|_{2,\infty}$.

(b) (8 points) Next, we show that the quasinorm $\|x\|_{2,\infty}$ does not satisfy the triangle inequality. To this end, let us fix $x = (1, 2^{-1/2})$ and $y = (2^{-1/2}, 1)$.

  (i) (4 points) Calculate $\|x\|_{2,\infty}$ and $\|y\|_{2,\infty}$.

  (ii) (4 points) Calculate $\|x + y\|_{2,\infty}$ and use (b)(i) to conclude that in fact

$$\|x + y\|_{2,\infty} > \|x\|_{2,\infty} + \|y\|_{2,\infty}.$$

  Hint: Use (a)(ii) to simplify your calculations.

(c) (8 points) Next, we establish (1) by proving a more general result. To this end, let us fix $x^1 = (x^1_1, \ldots, x^1_N), x^2 = (x^2_1, \ldots, x^2_N), \ldots, x^k = (x^k_1, \ldots, x^k_N) \in \mathbb{C}^N$ and $t > 0$.

  (i) (2 points) Show that

$$\big\{j \in \{1, \ldots, N\} \colon |x^1_j + \cdots + x^k_j| \geq t\big\} \subset \bigcup_{i \in \{1, \ldots, k\}} \big\{j \in \{1, \ldots, N\} \colon |x^i_j| \geq t/k\big\}.$$

(ii) (3 points) Use (c)(i) to prove that

$$\|x^1 + \cdots + x^k\|_{2,\infty} \le k\big(\|x^1\|_{2,\infty}^2 + \cdots + \|x^k\|_{2,\infty}^2\big)^{1/2},$$

where $\|x^i\|_{2,\infty}^2$ stands for $(\|x^i\|_{2,\infty})^2$.

(iii) (3 points) Now employ (c)(ii) to show that

$$\|x^1 + \cdots + x^k\|_{2,\infty} \le k\big(\|x^1\|_{2,\infty} + \cdots + \|x^k\|_{2,\infty}\big).$$

(d) (2 points) Next, we compare the quasinorm $\|x\|_{2,\infty}$ with the usual ℓ_2-norm. To this end, let us fix $x = (x_1, \ldots, x_N) \in \mathbb{C}^N$ and assume that

$$\|x\|_{2,\infty} = \max_{k\in\{1,\ldots,N\}} k^{1/2} x_k^*, \tag{2}$$

where $x^* \in \mathbb{R}_+^N$ denotes the nonincreasing rearrangement of $(|x_1|, \ldots, |x_N|) \in \mathbb{R}_+^N$. Use (2) to prove that

$$\|x\|_{2,\infty} \le \|x\|_2.$$

Problem 3 (25 points)

Fix a real number $0 < \varepsilon < 1$ and an integer $d \geq 1$, and define $n := \lfloor d\varepsilon/2 \rfloor$ to be the largest integer smaller than $d\varepsilon/2$. The present problem proves the following upper bound

$$\frac{\log M(\varepsilon; \mathbb{H}^d, d_H)}{d} \leq D\big((n/d)\|(1/2)\big) + \frac{\log(d+1)}{d} \tag{3}$$

on the ε-packing number $M(\varepsilon; \mathbb{H}^d, d_H)$ of the binary hypercube $\mathbb{H}^d = \{0,1\}^d$ equipped with the normalized Hamming distance $d_H(\theta, \theta') = \frac{1}{d}\sum_{j=1}^d \mathbb{1}[\theta_j \neq \theta'_j]$ and where we define $D\big(\alpha\|(1/2)\big) := \alpha \log\left(\frac{\alpha}{1/2}\right) + (1-\alpha)\log\left(\frac{1-\alpha}{1/2}\right)$, for all $0 < \alpha < 1/2$.

(a) (11 points) Let $\{\theta^1, \ldots, \theta^M\}$ be an ε-packing of $\mathbb{H}^d$, with $M = M(\varepsilon; \mathbb{H}^d, d_H)$. We define the sets $\mathbb{H}_i$, for all $i = 1, 2, \ldots, M$, as

$$\mathbb{H}_i := \{\theta \in \mathbb{H}^d \mid d_H(\theta, \theta^i) \leq \varepsilon/2\}.$$

(i) (3 points) Prove that the sets $\{\mathbb{H}_i\}_{i=1}^M$ are disjoint.

(ii) (3 points) Prove that, for all $1 \leq i \leq M$, the set $\mathbb{H}_i$ has cardinality

$$|\mathbb{H}_i| = \sum_{k=0}^{n} \binom{d}{k}.$$

(iii) (5 points) Using the results of (a)(i) and (a)(ii), prove that

$$\log M(\varepsilon; \mathbb{H}^d, d_H) \leq d \log 2 - \log \binom{d}{n}. \tag{4}$$

Hint: First establish the upper bound $M(\varepsilon; \mathbb{H}^d, d_H) \sum_{k=0}^d \binom{d}{n} \leq |\mathbb{H}^d|$.

(b) (10 points) Let Y be a binomial random variable with parameters $(d, n/d)$, i.e.,

$$\mathbb{P}[Y = \ell] = \binom{d}{\ell}(n/d)^\ell(1 - n/d)^{d-\ell}, \quad \text{for } 0 \leq \ell \leq d.$$

(i) (4 points) Prove that $\mathbb{P}[Y = \ell] \leq \mathbb{P}[Y = n]$, for $0 \leq \ell \leq d$.
Hint: Study the ratio $\frac{\mathbb{P}[Y=\ell]}{\mathbb{P}[Y=\ell-1]}$ *for* $\ell \leq n$ *and for* $\ell > n$.

(ii) (6 points) Prove that

$$\log \binom{d}{n} \geq d\phi(n/d) - \log(d+1), \tag{5}$$

where $\phi(t) = -t\log t - (1-t)\log(1-t)$.
Hint: Use subproblem (b)(i) to find an upper bound on $\sum_{\ell=0}^d \mathbb{P}[Y = \ell]$.

(c) (4 points) Combining inequalities (4) and (5), prove the upper bound (3).

Problem 4 (25 points)

(a) (8 points) Compute the VC dimension of the class

$$\mathcal{H}_1 := \left\{ h_{(a,b)} \colon \mathbb{R} \to \{0,1\} \mid a \le b \right\}$$

of closed intervals of $\mathbb{R}$, with

$$h_{(a,b)}(x) = \begin{cases} 1, & \text{if } a \le x \le b, \\ 0, & \text{otherwise.} \end{cases}$$

(b) (17 points) Let

$$\mathcal{H}_2 := \left\{ h_{(a_1,a_2,b_1,b_2)} \colon \mathbb{R}^2 \to \{0,1\} \mid a_i \le b_i, i = 1,2 \right\}$$

be the class of axis-aligned rectangles with

$$h_{(a_1,a_2,b_1,b_2)}(x_1,x_2) = \begin{cases} 1, & \text{if } a_i \le x_i \le b_i,\ i = 1,2, \\ 0, & \text{otherwise.} \end{cases}$$

(i) (5 points) Provide a set of 4 points shattered by $\mathcal{H}_2$.

(ii) (3 points) Show that there is no $h_{(a_1,a_2,b_1,b_2)} \in \mathcal{H}_2$ such that

$$\begin{cases} h(0,0) = 0, \\ h(-1,0) = h(0,-1) = h(1,0) = h(0,1) = 1. \end{cases}$$

(iii) (7 points) Show that no set of 5 distinct points can be shattered by $\mathcal{H}_2$.

<u>Hint</u>: Taking subproblem (b)(ii) as an example, show that there is at least one point which cannot be labelled 0 if all the other points are labelled 1.

<u>Remark</u>: A formal and detailed proof is not required. Describing the main steps in a clear fashion is enough to get full credit.

(iv) (2 points) Deduce the VC dimension of the class $\mathcal{H}_2$.