ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Prof. Dr. Helmut Bölcskei
Chair for Mathematical Information Science

Solutions to the
Examination on Mathematics of Information
September 2, 2022

Problem 1

(a) We consider the set

$$\{g_k\}_{k\in\mathbb{N}} := \left\{ e_1, \dots, e_{n-1}, \left(\frac{1}{\sqrt{2}}\right) e_n, \left(\frac{1}{\sqrt{2}}\right)^2 e_n, \left(\frac{1}{\sqrt{2}}\right)^3 e_n, \left(\frac{1}{\sqrt{2}}\right)^4 e_n, \dots \right\}$$

and calculate, for $x \in V$,

$$\sum_{k\in\mathbb{N}} |\langle x, g_k\rangle|^2 = \sum_{k=1}^{n-1} |\langle x, e_k\rangle|^2 + \sum_{k=1}^{\infty} \left| \left(\frac{1}{\sqrt{2}}\right)^k \langle x, e_n\rangle \right|^2 \tag{1}$$

$$= \sum_{k=1}^{n-1} |\langle x, e_k\rangle|^2 + |\langle x, e_n\rangle|^2 \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \tag{2}$$

$$= \sum_{k=1}^{n-1} |\langle x, e_k\rangle|^2 + |\langle x, e_n\rangle|^2 \tag{3}$$

$$= \sum_{k=1}^{n} |\langle x, e_k\rangle|^2 = \|x\|^2, \tag{4}$$

where we used that $\{e_k\}_{k=1}^n$ is an orthonormal basis (ONB) for V. We have hence established that $\{g_k\}_{k\in\mathbb{N}}$ is a tight frame with infinitely many elements and frame bound $A = 1$.

(b) We start by establishing that $B \leq \sum_{k\in\mathcal{K}} \|g_k\|^2$. Let $x \in V$ and use the Cauchy-Schwarz inequality to obtain

$$\sum_{k\in\mathcal{K}} |\langle x, g_k\rangle|^2 \leq \sum_{k\in\mathcal{K}} \|x\|^2 \, \|g_k\|^2 = \|x\|^2 \sum_{k\in\mathcal{K}} \|g_k\|^2.$$

Hence, $\sum_{k\in\mathcal{K}} \|g_k\|^2$ is a valid upper frame bound and the tightest possible upper frame bound B thus has to satisfy $B \leq \sum_{k\in\mathcal{K}} \|g_k\|^2$.

Next, we show that

$$nA \leq \sum_{k\in\mathcal{K}} \|g_k\|^2 \leq nB. \tag{5}$$

Since $\{g_k\}_{k\in\mathcal{K}}$ is a frame with (tightest possible) frame bounds A and B, we have, for all $\ell = 1, \dots, n$,

$$A = A\|e_\ell\|^2 \le \sum_{k\in\mathcal{K}} |\langle e_\ell, g_k\rangle|^2 \le B\|e_\ell\|^2 = B.$$

Summing over $\ell \in \{1, \dots, n\}$, we obtain

$$nA \le \sum_{\ell=1}^{n} \sum_{k\in\mathcal{K}} |\langle e_\ell, g_k\rangle|^2 \le nB,$$

which yields (5) upon interchanging the order of summation (Theorem 4 in the Handout) and using the fact that $\{e_\ell\}_{\ell=1}^n$ is an ONB. Finally, rewriting (5), we obtain $A \le \frac{1}{n}\sum_{k\in\mathcal{K}}\|g_k\|^2 \le B$, which completes the proof.

(c) Towards a contradiction, assume that there is a frame $\{g_k\}_{k\in\mathcal{K}}$ satisfying the stated conditions. By the definition of a frame, the tightest possible frame bounds $A, B \in \mathbb{R}$ must satisfy $0 < A \le B < \infty$. Hence, in particular, B must be finite. Further, we have $\frac{1}{n}\sum_{k\in\mathcal{K}}\|g_k\|^2 = \frac{1}{n}\sum_{k\in\mathcal{K}} 1 = \infty$. Together with $B \ge \frac{1}{n}\sum_{k\in\mathcal{K}}\|g_k\|^2$, this establishes the contradiction.

(d) Fix $\alpha > 0$, let $\gamma := \min\left\{\frac{1}{2}, \frac{\alpha}{2}\right\} \in \left(0, \frac{1}{2}\right)$, and, setting $\mathcal{K} = \{1, \dots, n\}$, consider the set

$$\{g_k\}_{k=1}^n := \left\{\sqrt{\gamma}\, e_1 + \sqrt{1-\gamma}\, e_2,\ \sqrt{\gamma}\, e_1 - \sqrt{1-\gamma}\, e_2,\ e_3, \dots, e_n\right\}.$$

Clearly, $\|g_k\|^2 = \|e_k\|^2 = 1$, for $k = 3, \dots, n$, and

$$\|g_1\|^2 = \|g_2\|^2 = \gamma\|e_1\|^2 + (1-\gamma)\|e_2\|^2 = 1.$$

Next, we observe that every $x \in V$ can be written as $x = \sum_{\ell=1}^n c_\ell\, e_\ell$, with $c_\ell = \langle x, e_\ell\rangle$, for $\ell = 1, \dots, n$, and we calculate

$$\begin{aligned}
\sum_{k=1}^n |\langle x, g_k\rangle|^2 &= \sum_{k=3}^n \left|\left\langle \sum_{\ell=1}^n c_\ell\, e_\ell, e_k\right\rangle\right|^2 + \left|\left\langle \sum_{\ell=1}^n c_\ell\, e_\ell, \sqrt{\gamma}\, e_1 + \sqrt{1-\gamma}\, e_2\right\rangle\right|^2 \\
&\quad + \left|\left\langle \sum_{\ell=1}^n c_\ell\, e_\ell, \sqrt{\gamma}\, e_1 - \sqrt{1-\gamma}\, e_2\right\rangle\right|^2 \\
&= \left(\sum_{k=3}^n c_k^2\right) + \left(c_1\sqrt{\gamma} + c_2\sqrt{1-\gamma}\right)^2 + \left(c_1\sqrt{\gamma} - c_2\sqrt{1-\gamma}\right)^2 \\
&= \left(\sum_{k=3}^n c_k^2\right) + 2c_1^2\gamma + 2c_2^2(1-\gamma), \qquad \forall x \in V. \qquad (6)
\end{aligned}$$

As $\gamma \le (1-\gamma) \le 1$, and $\|x\|^2 = \sum_{k=1}^n c_k^2$, we can upper-bound (6) according to

$$\sum_{k=1}^n |\langle x, g_k\rangle|^2 \le \left(\sum_{k=3}^n c_k^2\right) + 2c_1^2 + 2c_2^2 \le 2\sum_{k=1}^n c_k^2 = 2\|x\|^2, \qquad \forall x \in V.$$

Hence, 2 is a valid upper frame bound for the set $\{g_k\}_{k=1}^n$. Further, using $1 \geq 2\gamma$ and $(1-\gamma) \geq \gamma$, we get

$$\sum_{k=1}^{n} |\langle x, g_k \rangle|^2 \geq 2\gamma \left(\sum_{k=3}^{n} c_k^2 \right) + 2\gamma c_1^2 + 2\gamma c_2^2 = 2\gamma \sum_{k=1}^{n} c_k^2 = 2\gamma \|x\|^2, \qquad \forall x \in V. \quad (7)$$

Therefore, 2γ is a valid lower frame bound for the set $\{g_k\}_{k=1}^n$. As $\gamma > 0$, $\{g_k\}_{k=1}^n$ thus constitutes a frame for V. Finally, we observe that with $x = e_1$, (6) reads

$$\sum_{k=1}^{n} |\langle e_1, g_k \rangle|^2 = 2\gamma = 2\gamma \|e_1\|^2,$$

which shows that the inequality in (7) is tight. The tightest possible frame bound A hence satisfies $A = 2\gamma \leq \alpha$ as desired.

Problem 2

(a) (i) $\mathcal{C}_1 = \left[\frac{1}{3}\,\mathcal{C}_0\right] \bigcup \left[\frac{1}{3}\,\mathcal{C}_0 + \frac{2}{3}\right] = \left[\frac{1}{3}[0,1]\right] \bigcup \left[\frac{1}{3}[0,1] + \frac{2}{3}\right] = \left[0, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, 1\right]$.

Then, $\frac{1}{3}\,\mathcal{C}_1 = \left[0, \frac{1}{9}\right] \bigcup \left[\frac{2}{9}, \frac{1}{3}\right]$, and $\frac{1}{3}\,\mathcal{C}_1 + \frac{2}{3} = \left[\frac{2}{3}, \frac{7}{9}\right] \bigcup \left[\frac{8}{9}, 1\right]$.

Hence, $\mathcal{C}_2 = \left[\frac{1}{3}\,\mathcal{C}_1\right] \bigcup \left[\frac{1}{3}\,\mathcal{C}_1 + \frac{2}{3}\right] = \left[0, \frac{1}{9}\right] \bigcup \left[\frac{2}{9}, \frac{1}{3}\right] \bigcup \left[\frac{2}{3}, \frac{7}{9}\right] \bigcup \left[\frac{8}{9}, 1\right]$.

(ii) We proceed by induction.

For $n = 0$, we define the sequence (I_j^0), $j \in \{2^0 = 1\}$, containing one closed interval, namely $I_1^0 = \mathcal{C}_0 = [0,1]$. Properties 1 and 2 hold as $I_1^0 = [0,1] \subseteq [0,1]$ and $|I_1^0| = \frac{1}{3^0} = 1$, and Property 3 does not need to be verified as $n = 0$.

Next, fix $n \in \mathbb{N}_0$, and suppose that there exist 2^n disjoint closed intervals I_j^n, $j \in \{1, \dots, 2^n\}$, such that $\mathcal{C}_n = \bigcup_{j=1}^{2^n} I_j^n$ and Properties 1-3 are satisfied. Then, by definition,

$$\mathcal{C}_{n+1} = \left[\frac{1}{3}\,\mathcal{C}_n\right] \bigcup \left[\frac{1}{3}\,\mathcal{C}_n + \frac{2}{3}\right] = \left[\bigcup_{j=1}^{2^n} \frac{1}{3} I_j^n\right] \bigcup \left[\bigcup_{j=1}^{2^n} \left(\frac{1}{3} I_j^n + \frac{2}{3}\right)\right]. \tag{8}$$

We define (I_j^{n+1}), $j \in \{1, \dots, 2^{n+1}\}$, to be the sequence of intervals satisfying

$$I_j^{n+1} = \frac{1}{3} I_j^n, \text{ and } I_{j+2^n}^{n+1} = \frac{1}{3} I_j^n + \frac{2}{3}, \ \forall j \in \{1, \dots, 2^n\}. \tag{9}$$

By identifying these intervals in (8), one gets

$$\mathcal{C}_{n+1} = \left[\bigcup_{j=1}^{2^n} I_j^{n+1}\right] \bigcup \left[\bigcup_{j=1}^{2^n} I_{j+2^n}^{n+1}\right] = \left[\bigcup_{j=1}^{2^n} I_j^{n+1}\right] \bigcup \left[\bigcup_{j=2^n+1}^{2^{n+1}} I_j^{n+1}\right] = \bigcup_{j=1}^{2^{n+1}} I_j^{n+1}. \tag{10}$$

1. $(I_j^{n+1})_{j \in \{1,\dots,2^{n+1}\}}$ satisfies Property 1, as $I_j^n \subseteq [0,1]$, for $j \in \{1, \dots, 2^n\}$, implies that $I_j^{n+1} = \frac{1}{3} I_j^n \subseteq [0, \frac{1}{3}] \subseteq [0,1]$ and $I_{j+2^n}^{n+1} = \frac{1}{3} I_j^n + \frac{2}{3} \subseteq [\frac{2}{3}, 1] \subseteq [0,1]$, both for $j \in \{1, \dots, 2^n\}$.
2. $(I_j^{n+1})_{j \in \{1,\dots,2^{n+1}\}}$ satisfies Property 2 as $|I_j^n| = \frac{1}{3^n}$ implies that $|I_j^{n+1}| = |\frac{1}{3} I_j^n| = \frac{1}{3^{n+1}}$ and $|I_{j+2^n}^{n+1}| = |\frac{1}{3} I_j^n + \frac{2}{3}| = \frac{1}{3^{n+1}}$, both for $j \in \{1, \dots, 2^n\}$.
3. To show that $(I_j^{n+1})_{j \in \{1,\dots,2^{n+1}\}}$ satisfies Property 3 we argue as follows: First, note that $d(I_j^n, I_{j'}^n) \geq \frac{1}{3^n}$, for all $j, j' \in \{1, \dots, 2^n\}$ with $j \neq j'$, implies both $d(\frac{1}{3} I_j^n, \frac{1}{3} I_{j'}^n) \geq \frac{1}{3^{n+1}}$ and $d(\frac{1}{3} I_j^n + \frac{2}{3}, \frac{1}{3} I_{j'}^n + \frac{2}{3}) \geq \frac{1}{3^{n+1}}$, for all $j, j' \in \{1, \dots, 2^n\}$ with $j \neq j'$. Further, $d(\frac{1}{3} I_j^n, \frac{1}{3} I_{j'}^n + \frac{2}{3}) \overset{(a)}{\geq} \frac{1}{3} \geq \frac{1}{3^{n+1}}$, for $j, j' \in \{1, \dots, 2^n\}$, where (a) follows from the fact that $\frac{1}{3} I_j^n, \frac{1}{3} I_{j'}^n \subseteq \left[0, \frac{1}{3}\right]$.

This concludes the proof.

(b) (i) Fix $n \in \mathbb{N}_0$. Consider the sequence $(I_j^n)_{j \in \{1,\dots,2^n\}}$, defined in subproblem a)(ii), along with the sequence $(X_j^n)_{j \in \{1,\dots,2^n\}}$ of corresponding central points

$$X_j^n = \frac{\min I_j^n + \max I_j^n}{2}, \quad \text{for } j \in \{1, \dots, 2^n\}. \tag{11}$$

Let $j \in \{1, \ldots, 2^n\}$. By Property 2, $|X_j^n - x| \leq \frac{1}{2 \cdot 3^n} = \varepsilon_n$, for all $x \in I_j^n$. Since $\mathcal{C}_n = \bigcup_{j=1}^{2^n} I_j^n$, there exists a $j \in \{1, \ldots, 2^n\}$ such that $|X_j^n - x| \leq \varepsilon_n$, for all $x \in \mathcal{C}_n$. Therefore, $(X_j^n)_{j \in \{1,\ldots,2^n\}}$ constitutes an ε_n-covering of $\mathcal{C}_n$, of cardinality $\#(X_j^n)_{j \in \{1,\ldots,2^n\}} = 2^n$.

(ii) Fix $n \in \mathbb{N}_0$. Consider the sequence $(X_j^n)_{j \in \{1,\ldots,2^n\}}$ defined in the solution of subproblem b)(i). By Properties 2 and 3, $|X_j^n - X_{j'}^n| \geq 2 \cdot \frac{1}{3^n} > \frac{1}{3^n} = \varepsilon_n$, for all $j, j' \in \{1, \ldots, 2^n\}$ with $j \neq j'$. Hence, $(X_j^n)_{j \in \{1,\ldots,2^n\}}$ constitutes an ε_n-packing of $\mathcal{C}_n$, of cardinality $\#(X_j^n)_{j \in \{1,\ldots,2^n\}} = 2^n$.

(iii) Fix $n \in \mathbb{N}_0$. One has

$$2^n \overset{(a)}{\leq} M\left(\frac{1}{3^n}, \mathcal{C}_n, |\cdot|\right) = M\left(2 \cdot \frac{1}{2 \cdot 3^n}, \mathcal{C}_n, |\cdot|\right) \overset{(b)}{\leq} N\left(\frac{1}{2 \cdot 3^n}, \mathcal{C}_n, |\cdot|\right) \overset{(c)}{\leq} 2^n, \tag{12}$$

where (a) follows from subproblem b)(ii) and the fact that the $\frac{1}{3^n}$-packing number of $\mathcal{C}_n$ is the cardinality of the largest $\frac{1}{3^n}$-packing of $\mathcal{C}_n$, (b) is by using Theorem 1 in the Handout, and (c) follows from subproblem b)(i) and the fact that the $\frac{1}{2 \cdot 3^n}$-covering number of $\mathcal{C}_n$ is the cardinality of the smallest $\frac{1}{2 \cdot 3^n}$-covering of $\mathcal{C}_n$. From (12) we then have

$$N\left(\frac{1}{2 \cdot 3^n}, \mathcal{C}_n, |\cdot|\right) = 2^n. \tag{13}$$

(c) (i) Fix $n \in \mathbb{N}_0$ and $\varepsilon \in \left[\frac{1}{2 \cdot 3^{n+1}}, \frac{1}{2 \cdot 3^n}\right]$. Note that $\mathcal{C}_\infty = \bigcap_{\ell \geq 0} \mathcal{C}_\ell \subseteq \mathcal{C}_{n+1}$. From Theorem 2 in the Handout, we know that $\varepsilon \mapsto N(\varepsilon, \mathcal{C}_\infty, |\cdot|)$ is non-increasing. Hence, for every $\varepsilon \in \left[\frac{1}{2 \cdot 3^{n+1}}, \frac{1}{2 \cdot 3^n}\right]$, it holds that

$$N(\varepsilon, \mathcal{C}_\infty, |\cdot|) \overset{(a)}{\leq} N\left(\frac{1}{2 \cdot 3^{n+1}}, \mathcal{C}_\infty, |\cdot|\right) \overset{(b)}{\leq} N\left(\frac{1}{2 \cdot 3^{n+1}}, \mathcal{C}_{n+1}, |\cdot|\right) \overset{(c)}{=} 2^{n+1}, \tag{14}$$

where (a) follows from Theorem 2 in the Handout, (b) is a consequence of $\mathcal{C}_\infty \subseteq \mathcal{C}_{n+1}$, and (c) follows from (13).

(ii) Fix $n \in \mathbb{N}_0$ and $\varepsilon \in \left[\frac{1}{2 \cdot 3^{n+1}}, \frac{1}{2 \cdot 3^n}\right]$. We have

$$2^n = 3^{n \log_3(2)} \overset{(a)}{\leq} \left(\frac{1}{\varepsilon}\right)^{\log_3(2)}, \tag{15}$$

where (a) follows from $\varepsilon \leq \frac{1}{2 \cdot 3^n} \leq \frac{1}{3^n}$. Then,

$$N(\varepsilon, \mathcal{C}_\infty, |\cdot|) \leq 2^{n+1} = 2 \cdot 2^n \leq 2 \cdot \left(\frac{1}{\varepsilon}\right)^{\log_3(2)}. \tag{16}$$

It follows that

$$\log_2\left(N(\varepsilon, \mathcal{C}_\infty, |\cdot|)\right) \leq 1 + \log_3(2)\log_2(\varepsilon^{-1}). \tag{17}$$

The term on the right hand side of (17) does not depend on n, hence (17) holds for all $\varepsilon \in \bigcup_{n \geq 0} \left[\frac{1}{2 \cdot 3^{n+1}}, \frac{1}{2 \cdot 3^n}\right] = (0, 1/2]$. Finally, since $\log_3(2)\log_2(\varepsilon^{-1}) =$

$\log_3(\varepsilon^{-1})$, we have

$$\log_2\left(N(\varepsilon, \mathcal{C}_\infty, |\cdot|)\right) \leq 1 + \log_3(\varepsilon^{-1}), \quad \text{for all } \varepsilon \in (0, 1/2]. \tag{18}$$

Problem 3

(a) The function f is continuous on $[0,1]$ and twice differentiable on $(0,1)$. Its derivative on $(0,1)$ is given by

$$f'(y) = \delta_s - \delta_t + \theta_{s,t}\frac{1-2y}{\sqrt{y(1-y)}},$$

and the second derivative on $(0,1)$ is

$$f''(y) = -\theta_{s,t}\left\{\frac{(1-2y)^2}{2(y(1-y))^{3/2}} + \frac{2}{\sqrt{y(1-y)}}\right\} \leq 0.$$

Therefore f is concave on $[0,1]$. This implies that there exists a $y^* \in [0,1]$ such that f is nondecreasing on $[0,y^*]$ and nonincreasing on $[y^*,1]$.

(b) Take $u \in \mathbb{C}^N$ to be the vector obtained from x by retaining the s components with largest absolute value and setting the other ones to zero, and $v \in \mathbb{C}^N$ the vector obtained from x by retaining the t remaining components. Then, u and v are s-sparse and t-sparse vectors of disjoint support, and they trivially satisfy $x = u+v$. Moreover, since u and v have disjoint support, it holds that

$$1 = \|x\|_2^2 = \|u\|_2^2 + \|v\|_2^2. \tag{19}$$

We need to prove that $\|u\|_2^2 \in [s/(s+t), 1]$. The inequality $\|u\|_2^2 \leq 1$ is a direct consequence of (19). By construction,

$$\frac{1}{s}\|u\|_2^2 \geq \frac{1}{t}\|v\|_2^2 \overset{(19)}{=} \frac{1}{t}\left(1 - \|u\|_2^2\right).$$

Rearranging terms yields the desired result

$$\|u\|_2^2 \geq \frac{s}{s+t}. \tag{20}$$

(c) The result is established through the following direct calculation

$$\begin{aligned}
\|Ax\|_2^2 &= \|A(u+v)\|_2^2 \\
&= \langle A(u+v), A(u+v)\rangle \\
&= \langle Au, Au\rangle + \langle Av, Av\rangle + \langle Au, Av\rangle + \langle Av, Au\rangle \\
&= \langle Au, Au\rangle + \langle Av, Av\rangle + \langle Au, Av\rangle + \overline{\langle Au, Av\rangle} \\
&= \|Au\|_2^2 + \|Av\|_2^2 + 2\,\mathrm{Re}\langle Au, Av\rangle.
\end{aligned}$$

(d) The result is proven through the following direct calculation

$$\begin{aligned}
\left|\|Ax\|_2^2 - \|x\|_2^2\right| &\overset{(*)}{=} \left|\|Au\|_2^2 + \|Av\|_2^2 + 2\,\mathrm{Re}\langle Au, Av\rangle - \|u\|_2^2 - \|v\|_2^2\right| \\
&\leq \left|\|Au\|_2^2 - \|u\|_2^2\right| + \left|\|Av\|_2^2 - \|v\|_2^2\right| + 2\left|\langle Au, Av\rangle\right| \\
&\leq \delta_s\|u\|_2^2 + \delta_t\|v\|_2^2 + 2\theta_{s,t}\|u\|_2\|v\|_2,
\end{aligned}$$

where in $(*)$ we used the result of subproblem (c) and (19), and the last inequality is by the definition of the restricted isometry constant and the definition of the restricted orthogonality constant.

(e) From subproblem (d) we have

$$\begin{aligned}\left|\|Ax\|_2^2 - \|x\|_2^2\right| &\leq \delta_s\|u\|_2^2 + \delta_t\|v\|_2^2 + 2\theta_{s,t}\|u\|_2\|v\|_2 \\ &\overset{(19)}{=} \delta_s\|u\|_2^2 + \delta_t\big(1 - \|u\|_2^2\big) + 2\theta_{s,t}\|u\|_2\sqrt{1-\|u\|_2^2} \\ &= f\big(\|u\|_2^2\big).\end{aligned}$$

Now, as f is nonincreasing on $[y^*, 1]$, and since we have

$$y^* \leq \frac{s}{s+t} \overset{(20)}{\leq} \|u\|_2^2 \leq 1,$$

it holds that

$$f\big(\|u\|_2^2\big) \leq f\left(\frac{s}{s+t}\right),$$

and therefore

$$\left|\|Ax\|_2^2 - \|x\|_2^2\right| \leq f\left(\frac{s}{s+t}\right)$$

as desired.

(f) We have established that

$$\left|\|Ax\|_2^2 - \|x\|_2^2\right| \leq f\left(\frac{s}{s+t}\right) = f\left(\frac{s}{s+t}\right)\|x\|_2^2.$$

A straightforward derivation shows that

$$f\left(\frac{s}{s+t}\right) = \frac{1}{s+t}\left(s\delta_s + t\delta_t + 2\sqrt{st}\,\theta_{s,t}\right).$$

By definition of the restricted isometry constant δ_{s+t}, one must then have

$$\delta_{s+t} \leq f\left(\frac{s}{s+t}\right) = \frac{1}{s+t}\left(s\delta_s + t\delta_t + 2\sqrt{st}\,\theta_{s,t}\right).$$

This concludes the proof.

Problem 4

As suggested by the hint, we first treat the cases $n = 1$ and $n = 2$.

For $n = 1$, $\mathbb{R}^n$ is the real line and the closed ball $B(x_0, r)$ is an interval centered at x_0 and of length $2r$. We prove that, in this case, the VC dimension of $\mathcal{B}$ equals 2, a result that we have already seen in the exercise session. To see this, consider the points $x_1 = 0$ and $x_2 = 1$ and note that they are shattered by $\mathcal{B}$ since choosing the values (see Fig. 1 Left)

$$\begin{cases} x_0 = -1, r = 1/2, \\ x_0 = 0, r = 1/2, \\ x_0 = 1, r = 1/2, \\ x_0 = 0, r = 1, \end{cases} \quad \text{yields the respective labeling} \quad \begin{cases} (0,0), \\ (1,0), \\ (0,1), \\ (1,1). \end{cases}$$

On the other hand, given three distinct points $\{x_1, x_2, x_3\}$, satisfying $x_1 < x_2 < x_3$, the labeling $(1, 0, 1)$ cannot be generated. Indeed, the closed ball $B(x_0, r)$ by virtue of being a convex set, must contain x_2 if it is to contain x_1 and x_3 (see Fig. 1 Right). This proves that the VC dimension of $\mathcal{B}$ equals 2.

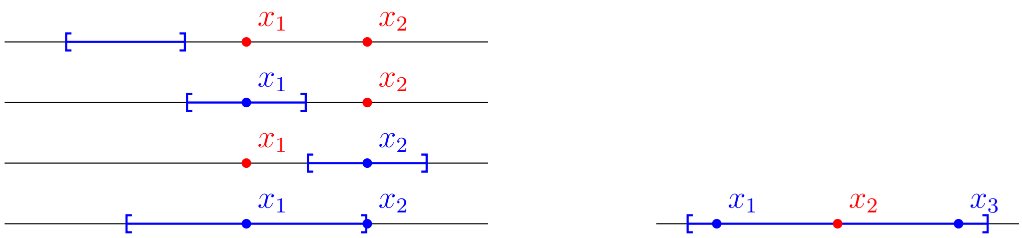


Fig. 1: Left: All 4 possible labelings. Right: Every interval containing x_1 and x_3 must contain x_2 as well.

We now turn to the case $n = 2$. Here, $\mathbb{R}^n$ is the Euclidean plane and $B(x_0, r)$ is a closed disk of radius r centered at x_0. We consider the following three points $x_1 = (0, 0)$, $x_2 = (1, 0)$, and $x_3 = (0, 1)$ in the plane. These points are shattered by $\mathcal{B}$ as choosing the values (see Fig. 2)

$$\begin{cases} x_0 = (-1, 0), r = 1/2, \\ x_0 = (0, 0), r = 1/2, \\ x_0 = (1, 0), r = 1/2, \\ x_0 = (0, 1), r = 1/2, \\ x_0 = (1, 0), r = 1, \\ x_0 = (0, 1), r = 1, \\ x_0 = (1, 1), r = 1, \\ x_0 = (1, 1), r = \sqrt{2}, \end{cases} \quad \text{yields the respective labeling} \quad \begin{cases} (0,0,0), \\ (1,0,0), \\ (0,1,0), \\ (0,0,1), \\ (1,1,0), \\ (1,0,1), \\ (0,1,1), \\ (1,1,1). \end{cases}$$

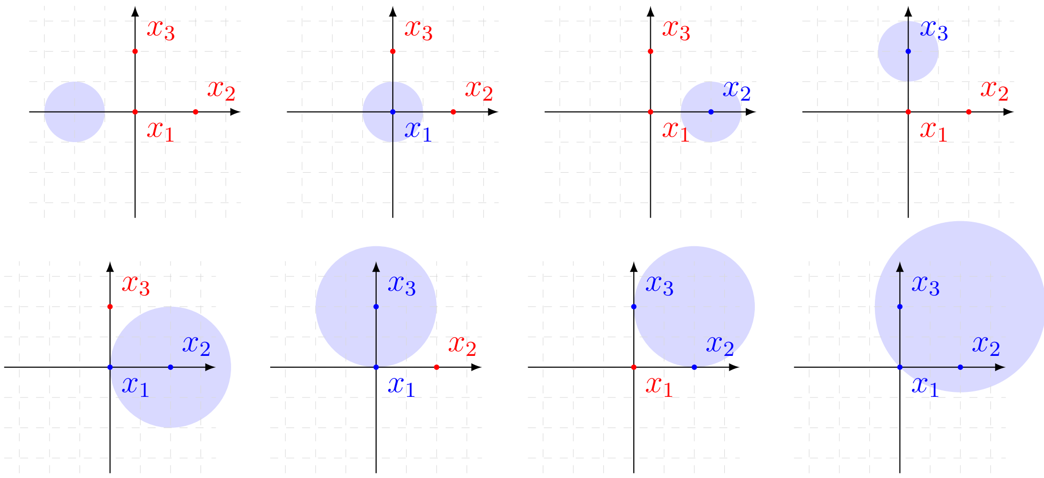


Fig. 2: All 8 possible labelings.

On the other hand, take any four distinct points x_1, x_2, x_3, and x_4 in the plane. First assume that there is one point in the set $\{x_1, x_2, x_3, x_4\}$ which is contained in the convex hull of the other three points, w.l.o.g. let x_4 be this point (see Fig. 3 Left). Then, the labeling $(1, 1, 1, 0)$ cannot be generated. Indeed, the closed ball $B(x_0, r)$, by virtue of being a convex set, must contain x_4 if it is to contain x_1, x_2, and x_3. If none of the points belongs to the convex hull of the other three, then x_1, x_2, x_3, and x_4 are the vertices of a convex polygon (see Fig. 3 Right). W.l.o.g. we assume that the diagonals of this polygon are (x_1, x_3) and (x_2, x_4) with $\|x_4 - x_2\|_2 \leq \|x_3 - x_1\|_2$. Then, the polygon is contained in a closed ball containing x_1 and x_3, and it is thus not possible to generate the labeling $(1, 0, 1, 0)$. This proves that the VC dimension of $\mathcal{B}$ equals 3.

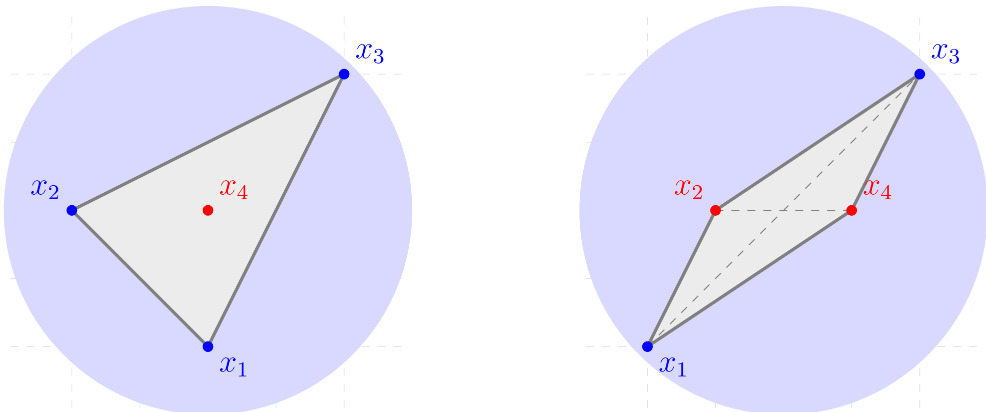


Fig. 3: Left: x_4 belongs to the convex hull (in gray) of x_1, x_2, and x_3 and therefore to any ball containing $\{x_1, x_2, x_3\}$. Right: The vertices x_2 and x_4 of the shorter diagonal are contained in all balls containing x_1 and x_3.

We finally turn to the proof for the case of general n and first show that one can find $n+1$ points in $\mathbb{R}^n$ that are shattered by $\mathcal{B}$. To this end, we consider the following family of points $\{x_i\}_{i=1}^{n+1}$ in $\mathbb{R}^n$,

$$x_1 = (0, \ldots, 0) \quad \text{and} \quad x_i = (0, \ldots, 0, \underbrace{1}_{(i-1)\text{-th position}}, 0, \ldots, 0), \quad \text{for } i = 2, \ldots, n+1.$$

We denote the corresponding labelings as $(y_1, \dots, y_{n+1})$, where $y_i \in \{0, 1\}$, for $i = 1, \dots, n+1$. The labeling $y_i = 0$, for all $i = 1, \dots, n+1$, is realized by the ball centered at

$$x_0 = (-1, 0, \dots, 0) \quad \text{with radius} \quad r = 1/2.$$

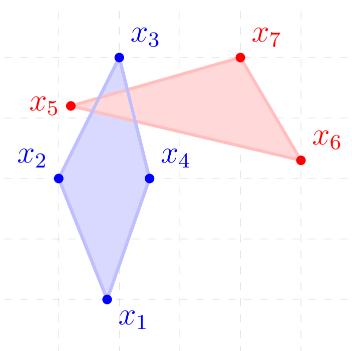
The labeling $y_1 = 1$ and $y_i = 0$, for all $i = 2, \dots, n+1$, is realized by the ball centered at

$$x_0 = (0, 0, \dots, 0) \quad \text{with radius} \quad r = 1/2.$$

For all other labelings, there are $k \geq 1$ points different from x_1 labeled as 1, that is, the set $\{i \mid 2 \leq i \leq n+1 \text{ and } y_i = 1\}$ has cardinality $k \geq 1$. The corresponding labelings are realized by the ball centered at

$$x_0 = (y_2, \dots, y_{n+1}) \quad \text{with radius} \quad r = \begin{cases} \sqrt{k}, & \text{if } y_1 = 1, \\ \sqrt{k-1}, & \text{if } y_1 = 0 \text{ and } k \geq 2, \\ 1/2, & \text{if } y_1 = 0 \text{ and } k = 1. \end{cases}$$

(Note that this construction generalizes those above for $n = 1$ and $n = 2$.) Therefore, the VC dimension of $\mathcal{B}$ is at least equal to $n+1$. We now show that no set of $n+2$ points can be shattered by $\mathcal{B}$. To this end, we argue by way of contradiction and choose S to be a set of $n+2$ points shattered by $\mathcal{B}$. By Radon's theorem (see the Handout), S can be partitioned into two disjoint sets T_1 and T_2 such that their convex hulls intersect. However, since S is shattered by $\mathcal{B}$ by assumption, there exists a ball B_1 containing T_1 (and therefore its convex hull) but no point of T_2, and there exists a ball B_2 containing T_2 (and therefore its convex hull) but no point of T_1. This implies that there exists a hyperplane separating T_1 and T_2 (if B_1 and B_2 are disjoint, this is obvious, otherwise, the intersection of B_1 and B_2 does not contain any point of S and we can take the hyperplane spanned by the common chord), which stands in contradiction to the fact that the convex hulls of T_1 and T_2 intersect (see Fig. 4). Therefore, the VC dimension of $\mathcal{B}$ is at most equal to $n+1$. This proves that the VC dimension of $\mathcal{B}$ is exactly $n+1$.



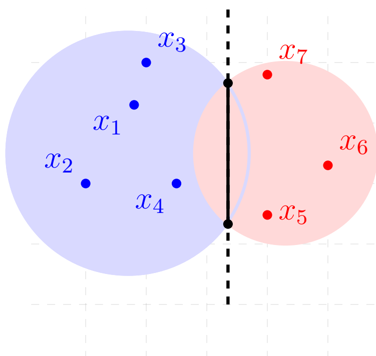


Fig. 4: T_1 (blue) and T_2 (red) cannot have their convex hulls intersecting (Left), and at the same time be separated by a hyperplane (Right). The black line in the right picture represents the common chord of the blue and red balls, and the dashed line represents the hyperplane spanned by the common chord.