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
August 28, 2021

Problem 1

(a) (i) For arbitrary $x \in \mathcal{H}$, we compute

$$\sum_{k=1}^{\infty} |\langle x, h_k \rangle|^2 = \sum_{k=1}^{\infty} |\langle x, (-1)^k e_k \rangle|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 = \|x\|^2,$$

where the last equality holds because $\{e_k\}_{k\in\mathbb{N}}$ is an orthonormal basis (ONB) for $\mathcal{H}$ by assumption. This establishes that the set $\{h_k\}_{k\in\mathbb{N}}$ is a tight frame with frame bounds $A = B = 1$.

(ii) For arbitrary $x \in \mathcal{H}$, we compute

$$\sum_{k=1}^{\infty} |\langle x, h_k \rangle|^2 = \sum_{k=1}^{\infty} k \left| \frac{1}{k} \langle x, e_k \rangle \right|^2 = \sum_{k=1}^{\infty} \frac{1}{k} |\langle x, e_k \rangle|^2. \tag{1}$$

Next, we assume towards a contradiction that there exists a lower frame bound $A > 0$, i.e., $A\|x\|^2 \leq \sum_{k=1}^{\infty} |\langle x, h_k \rangle|^2, \forall x \in \mathcal{H}$. Fix an integer N such that $\frac{1}{N} < A$ and evaluate (1) for $x = e_N \in \mathcal{H}$ to get

$$\sum_{k=1}^{\infty} |\langle e_N, h_k \rangle|^2 = \frac{1}{N} |\langle e_N, e_N \rangle|^2 = \frac{1}{N} \|e_N\|_2 < A\|e_N\|^2.$$

This stands in contradiction to the assumption that A is a lower frame bound. As A was arbitrary no lower frame bound can therefore exist and $\{h_k\}_{k\in\mathbb{N}}$ is thus not a frame.

(b) (i) We prove that $\{g_k\}_{k\in\mathbb{N}}$ is complete by showing that the only signal $x \in \mathcal{H}$ that satisfies $\langle x, g_k \rangle = 0$, $\forall k \in \mathbb{N}$, is $x = 0$. Take $x \in \mathcal{H}$ with $0 = \langle x, g_k \rangle = \langle x, e_k + e_{k+1} \rangle$, $\forall k \in \mathbb{N}$. Hence,

$$\langle x, e_k \rangle = -\langle x, e_{k+1} \rangle, \ \forall k \in \mathbb{N},$$

which implies $|\langle x, e_k \rangle| = C, \ \forall k \in \mathbb{N}$, for some $C \geq 0$. Further, owing to $x \in \mathcal{H}$, we have $\|x\| < \infty$ and thus

$$\infty > \|x\|^2 = \sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 = \sum_{k=1}^{\infty} C^2, \tag{2}$$

where we used that $\{e_k\}_{k\in\mathbb{N}}$ is an ONB. The proof is concluded by noting that (2) can hold only if $C = 0$ and thus $x = 0$.

(ii) We prove that $\{g_k\}_{k\in\mathbb{N}}$ is not a frame by showing that no lower frame bound $A > 0$ exists. To this end, we fix $q \in (0, 1)$, consider the signal $x_q = \sum_{\ell=1}^{\infty}(-q)^{\ell-1}e_\ell$ and start by showing that $x_q \in \mathcal{H}$. Since $\{e_k\}_{k\in\mathbb{N}}$ is an ONB by assumption, we can write

$$\|x_q\|^2 = \sum_{k=1}^{\infty} |\langle x_q, e_k\rangle|^2 = \sum_{k=1}^{\infty}(q^2)^{k-1} = \frac{1}{1-q^2} < \infty, \quad \forall q \in (0, 1),$$

which establishes that $x_q \in \mathcal{H}$. Next, we compute

$$\begin{aligned}
\sum_{k=1}^{\infty} |\langle x_q, g_k\rangle|^2 &= \sum_{k=1}^{\infty} \left| \left\langle \sum_{\ell=1}^{\infty}(-q)^{\ell-1}e_\ell, e_k + e_{k+1} \right\rangle \right|^2 \\
&= \sum_{k=1}^{\infty} |(-q)^{k-1} + (-q)^k|^2 \\
&= \sum_{k=1}^{\infty} |(1-q)(-q)^{k-1}|^2 \\
&= (1-q)^2 \sum_{k=1}^{\infty}(q^2)^{k-1} \\
&= (1-q)^2 \frac{1}{1-q^2} = (1-q)^2\|x_q\|^2.
\end{aligned}$$

The equality $\sum_{k=1}^{\infty} |\langle x_q, g_k\rangle|^2 = (1-q)^2\|x_q\|^2$ then establishes that there can be no $A > 0$ such that $A\|x\|^2 \le \sum_{k=1}^{\infty} |\langle x, g_k\rangle|^2$, $\forall x \in \mathcal{H}$, as we can always find a $q \in (0, 1)$ so that $(1-q)^2 < A$.

(c) (i) For all $g, f \in L^2(\mathbb{R})$, we have

$$\begin{aligned}
\langle \mathbb{W}_{m,n}^{(T,F)} g, f\rangle &= \int_{-\infty}^{\infty} e^{2\pi inFt} g(t-mT)\overline{f(t)}\, dt && (3)\\
&= \int_{-\infty}^{\infty} e^{2\pi inF(t'+mT)} g(t')\overline{f(t'+mT)}\, dt' && (4)\\
&= \int_{-\infty}^{\infty} g(t')\overline{e^{-2\pi inmTF} e^{2\pi i(-n)Ft'} f(t'-(-m)T)}\, dt' && (5)\\
&= \int_{-\infty}^{\infty} g(t')\overline{e^{-2\pi inmTF} (\mathbb{W}_{-m,-n}^{(T,F)} f)(t')}\, dt' && (6)\\
&= \langle g, e^{-2\pi inmTF}\, \mathbb{W}_{-m,-n}^{(T,F)} f\rangle && (7)\\
&= \langle g, \left(\mathbb{W}_{m,n}^{(T,F)}\right)^* f\rangle, && (8)
\end{aligned}$$

which establishes that $\left(\mathbb{W}_{m,n}^{(T,F)}\right)^* = e^{-2\pi inmTF}\, \mathbb{W}_{-m,-n}^{(T,F)}$.

(ii) For every $g \in L^2(\mathbb{R})$, we have

$$\left(\mathbb{W}_{m,n}^{(T,F)}\right)^* \mathbb{W}_{m,n}^{(T,F)} g = \left(\mathbb{W}_{m,n}^{(T,F)}\right)^* \left(e^{2\pi i n F \bullet} g(\bullet - mT)\right) \tag{9}$$

$$= e^{-2\pi i n m T F}\, \mathbb{W}_{-m,-n}^{(T,F)} \left(e^{2\pi i n F \bullet} g(\bullet - mT)\right) \tag{10}$$

$$= e^{-2\pi i n m T F} e^{-2\pi i n F \bullet} e^{2\pi i n F(\bullet + mT)} g(\bullet + mT - mT) \tag{11}$$

$$= g, \tag{12}$$

and

$$\mathbb{W}_{m,n}^{(T,F)} \left(\mathbb{W}_{m,n}^{(T,F)}\right)^* g = \mathbb{W}_{m,n}^{(T,F)} \left(e^{-2\pi i n m T F}\, \mathbb{W}_{-m,-n}^{(T,F)} g\right) \tag{13}$$

$$= e^{-2\pi i n m T F}\, \mathbb{W}_{m,n}^{(T,F)} \left(e^{-2\pi i n F \bullet} g(\bullet + mT)\right) \tag{14}$$

$$= e^{-2\pi i n m T F} \left(e^{2\pi i n F \bullet} e^{-2\pi i n F(\bullet - mT)} g(\bullet - mT + mT)\right) \tag{15}$$

$$= g. \tag{16}$$

Problem 2

(a) (i) We solve the linear system $D\widehat{x} = 0$ to obtain

$$\mathcal{N}(D) = \operatorname{span}\left(\begin{pmatrix} \frac{7}{5} \\ -\frac{3}{5} \\ -\frac{3}{5} \\ 1 \end{pmatrix}\right).$$

(ii) In (i) we saw that $\mathcal{N}(D) \neq \{0\}$. Therefore, the columns of D are linearly dependent which implies $\operatorname{spark}(D) \leq 4$. As $\|x\|_0 = 3$ the condition $\|x\|_0 < \frac{\operatorname{spark}(D)}{2}$ hence does not hold.

(iii) From the proof of Theorem 3.2 in the lecture notes we know that $\operatorname{spark}(D) \geq 1 + 1/\mu(D)$. The condition

$$\|x\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(D)}\right)$$

together with $\operatorname{spark}(D) \leq 4$ would hence require $\|x\|_0 < 2$. This is not satisfied as we have $\|x\|_0 = 3$.

(iv) We have $\mathcal{X} = x + \mathcal{N}(D)$, where $x = (1\ 1\ 1\ 0)^T$ is the particular solution from the problem statement. Hence,

$$\mathcal{X} = \left\{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{7}{5} \\ -\frac{3}{5} \\ -\frac{3}{5} \\ 1 \end{pmatrix} : \lambda \in \mathbb{R}\right\}. \tag{17}$$

(v) (P0) identifies the vector

$$\underset{\widehat{x} \in \mathcal{X}}{\arg\min} \|\widehat{x}\|_0,$$

where $\mathcal{X}$ denotes the solution set characterized in (17). We notice, with $\lambda = \frac{5}{3}$ in (17), that the vector

$$x' := \begin{pmatrix} 1 + \frac{7}{3} \\ 0 \\ 0 \\ \frac{5}{3} \end{pmatrix}$$

is contained in the solution set, i.e., $x' \in \mathcal{X}$. Since $\|x'\|_0 = 2 < \|x\|_0 = 3$, it follows that the solution to (P0) is not equal to x. Hence, x is not recovered through (P0).

(vi) (P1) identifies the vector that minimizes

$$\min_{\widehat{x}\in\mathcal{X}} \|\widehat{x}\|_1 = \min_{\lambda\in\mathbb{R}} \left\| \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{7}{5} \\ -\frac{3}{5} \\ -\frac{3}{5} \\ 1 \end{pmatrix} \right\|_1 \tag{18}$$

$$= \min_{\lambda\in\mathbb{R}} \left|1 + \frac{7}{5}\lambda\right| + 2\left|1 - \frac{3}{5}\lambda\right| + |\lambda| \tag{19}$$

$$=: \min_{\lambda\in\mathbb{R}} f(\lambda). \tag{20}$$

Noting that $f(\lambda)$ has its unique minimum at $\lambda = 0$, it follows that (P1) recovers x uniquely.

In order to formally establish that $f(\lambda)$ is minimized at $\lambda = 0$, we observe that $f(0) = 3$ and for every $\delta \in (0, \frac{5}{7})$, we have

$$f(0+\delta) = 1 + \frac{7}{5}\delta + 2 - \frac{6}{5}\delta + \delta = 3 + \frac{6}{5}\delta > 3$$

and

$$f(0-\delta) = 1 - \frac{7}{5}\delta + 2 + \frac{6}{5}\delta + \delta = 3 + \frac{4}{5}\delta > 3.$$

Hence, $\lambda = 0$ is a local minimum and because f is the sum of strictly convex functions also a global minimum.

(b) (i) We first establish the inequality provided in the Hint.

$$|\langle u, v\rangle| = \left|\sum_{k=1}^{N} u_k \overline{v_k}\right| \le \sum_{k=1}^{N} |u_k||v_k| \le \max_{\ell} |v_\ell| \sum_{k=1}^{N} |u_k| = \|v\|_\infty \|u\|_1.$$

Next, we show that under (C1) for all $z \neq x$ with $Dz = y = Dx$, we have $\|z\|_1 > \|x\|_1$ as this implies that the minimization problem (P1) has the unique solution x as desired. To this end, we define $v := z - x$ and note that $Dv = D(z-x) = Dz - Dx = 0$, i.e., $v \in \mathcal{N}(D)$. We now bound

$$\|z\|_1 = \|v + x\|_1 = \|v_S + x_S\|_1 + \|v_{S^c}\|_1 \tag{21}$$

$$= \|v_S + x_S\|_1 \|\mathrm{sgn}(x_S)\|_\infty + \|v_{S^c}\|_1 \tag{22}$$

$$\ge |\langle x_S + v_S, \mathrm{sgn}(x_S)\rangle| + \|v_{S^c}\|_1 \tag{23}$$

$$> |\langle x_S + v_S, \mathrm{sgn}(x_S)\rangle| + |\langle v_S, \mathrm{sgn}(x_S)\rangle| \tag{24}$$

$$\ge |\langle x_S, \mathrm{sgn}(x_S)\rangle| - |\langle v_S, \mathrm{sgn}(x_S)\rangle| + |\langle v_S, \mathrm{sgn}(x_S)\rangle| \tag{25}$$

$$= \left|\sum_{k\in S} x_k \frac{\overline{x_k}}{|x_k|}\right| = \|x_S\|_1 = \|x\|_1, \tag{26}$$

where in (22) we used $\|\text{sgn}(x_S)\|_\infty = 1$, for $x_S \neq 0$, in (23) we applied the Hint, in (24) we used (C1), and in (25) we employed the reverse triangle inequality.

(ii) First note that for all $v \in \mathcal{N}(D)\backslash\{0\}$, we have

$$0 = Dv = D_S v_S + D_{S^c} v_{S^c} \tag{27}$$

and hence $D_S v_S = -D_{S^c} v_{S^c}$. Further, we realize that $(D^H h)_S = (D_S)^H h,\ \forall h \in \mathbb{C}^m$. Next, we assume that (C2) holds and show that this implies (C1). Let $h \in \mathbb{C}^m$ be such that

$$(D^H h)_S = \text{sgn}(x_S) \ \text{ and } \ \|(D^H h)_{S^c}\|_\infty < 1. \tag{28}$$

Such an $h \in \mathbb{C}^m$ exists by assumption (C2). With $(D^H h)_S = (D_S)^H h$ this implies (C1) as follows,

$$\begin{aligned}
\left|\sum_{j\in S} v_j \overline{(\text{sgn}(x))_j}\right| &= \langle v_S, \text{sgn}(x_S)\rangle = \langle v_S, (D^H h)_S\rangle && (29)\\
&= |\langle v_S, (D_S)^H h\rangle| = |\langle D_S v_S, h\rangle| && (30)\\
&= |\langle -D_{S^c} v_{S^c}, h\rangle| = |\langle v_{S^c}, (D_{S^c})^H h\rangle| && (31)\\
&= |\langle v_{S^c}, (D^H h)_{S^c}\rangle| \leq \|v_{S^c}\|_1 \|(D^H h)_{S^c}\|_\infty && (32)\\
&< \|v_{S^c}\|_1. && (33)
\end{aligned}$$

To see that the final inequality is, indeed, strict, we note that $v_{S^c} \neq 0$ as otherwise $v \neq 0$ would imply $v_S \neq 0$ and (27) would imply $D_S v_S = 0$. This would, however, stand in contradiction to $\mathcal{N}(D_S) = \{0\}$. Therefore, we have $\|v_{S^c}\|_1 \neq 0$. Together with $\|(D^H h)_{S^c}\|_\infty < 1$, which is by (28), this guarantees strict inequality.

Problem 3

(a) First, recall the definition of the operator norm as $\|P_{\mathcal{W}}\|_{2\to 2} := \sup_{\|g\|_2=1} \|P_{\mathcal{W}}g\|_2$, which implies that, for all $h \in L^2(\mathbb{R})$ with $h \neq 0$,

$$\|P_{\mathcal{W}}h\|_2 = \left\| P_{\mathcal{W}} \frac{h}{\|h\|_2} \right\|_2 \|h\|_2 \leq \sup_{\|g\|_2=1} \|P_{\mathcal{W}}\, g\|_2 \|h\|_2 = \|P_{\mathcal{W}}\|_{2\to 2} \|h\|_2. \tag{34}$$

We follow the Hint and observe that for all $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$,

$$\|P_{\mathcal{W}}g\|_2^2 \overset{\text{Plancherel}}{=} \left\| \widehat{P_{\mathcal{W}}g} \right\|_2^2 = \int_{\mathcal{W}} |\hat{g}(w)|^2 \, dw \leq \int_{-\infty}^{\infty} |\hat{g}(w)|^2 \, dw \overset{\text{Plancherel}}{=} \|g\|_2^2,$$

and that if moreover g has its Fourier transform $\hat{g}$ supported on $\mathcal{W}$, then $P_{\mathcal{W}}g = g$. We have therefore proven that

$$1 \leq \|P_{\mathcal{W}}\|_{2\to 2} \leq 1,$$

and hence

$$\|P_{\mathcal{W}}\|_{2\to 2} = 1. \tag{35}$$

Next, we note that

$$\begin{aligned}
\|f - P_{\mathcal{W}}P_{\mathcal{T}}f\|_2 &\overset{\text{triang. ineq.}}{\leq} \|f - P_{\mathcal{W}}f\|_2 + \|P_{\mathcal{W}}f - P_{\mathcal{W}}P_{\mathcal{T}}f\|_2 \\
&\overset{(34)}{\leq} \|f - P_{\mathcal{W}}f\|_2 + \|P_{\mathcal{W}}\|_{2\to 2}\|f - P_{\mathcal{T}}f\|_2 \\
&\overset{(35)}{=} \|f - P_{\mathcal{W}}f\|_2 + \|f - P_{\mathcal{T}}f\|_2 \leq \varepsilon_{\mathcal{W}} + \varepsilon_{\mathcal{T}},
\end{aligned} \tag{36}$$

where the last inequality holds as f is $\varepsilon_{\mathcal{T}}$-concentrated to $\mathcal{T}$ and simultaneously $\varepsilon_{\mathcal{W}}$-concentrated to $\mathcal{W}$.

(b) As it has been assumed in the problem statement that $\|f\|_2 = 1$, applying the reverse triangle inequality yields the desired result according to

$$\begin{aligned}
\|P_{\mathcal{W}}P_{\mathcal{T}}\|_{2\to 2} &\overset{(34)}{\geq} \|P_{\mathcal{W}}P_{\mathcal{T}}f\|_2 = \|f - (f - P_{\mathcal{W}}P_{\mathcal{T}}f)\|_2 \\
&\overset{\text{RTI}}{\geq} \|f\|_2 - \|f - P_{\mathcal{W}}P_{\mathcal{T}}f\|_2 \overset{(36)}{\geq} 1 - \varepsilon_{\mathcal{T}} - \varepsilon_{\mathcal{W}}.
\end{aligned}$$

(c) Plugging in the definition of $P_{\mathcal{W}}P_{\mathcal{T}}g$, we obtain

$$\begin{aligned}
(P_{\mathcal{W}}P_{\mathcal{T}}g)(s) &= \int_{\mathcal{W}} e^{2\pi i w s} \left(\widehat{\mathbb{1}_{\mathcal{T}}\, g} \right)(w) \, dw \\
&= \int_{\mathcal{W}} \int_{-\infty}^{\infty} e^{2\pi i w(s-t)} \, \mathbb{1}_{\mathcal{T}}(t) \, g(t) \, dt \, dw && (37) \\
&\overset{(*)}{=} \int_{-\infty}^{\infty} \left\{ \int_{\mathcal{W}} e^{2\pi i w(s-t)} \, dw \, \mathbb{1}_{\mathcal{T}}(t) \right\} g(t) \, dt \\
&= \int_{-\infty}^{\infty} q(s,t) g(t) \, dt, && (38)
\end{aligned}$$

where $(*)$ follows from Fubini's theorem, and we set

$$q(s,t) = \int_{\mathcal{W}} e^{2\pi i w(s-t)} dw \, \mathbb{1}_{\mathcal{T}}(t).$$

The condition for the application of Fubini's theorem, namely absolute integrability in (37), is satisfied as $\mathcal{T}$ and $\mathcal{W}$ are bounded sets. Now, fixing $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\|g\|_2 = 1$ and using (38), we obtain

$$\begin{aligned}
\|P_{\mathcal{W}} P_{\mathcal{T}} g\|_2^2 &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} q(s,t) g(t) \, dt \right|^2 ds \\
&\overset{\text{C.S.}}{\leq} \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} |q(s,t)|^2 \, dt \underbrace{\int_{-\infty}^{\infty} |g(u)|^2 \, du}_{=1} \right\} ds \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |q(s,t)|^2 \, dt \, ds, \qquad (39)
\end{aligned}$$

where C.S. stands for 'Cauchy-Schwarz inequality'. As the right hand side of (39) does not depend on g, we can conclude, by taking the supremum over all g satisfying $\|g\|_2 = 1$, that, as desired,

$$\|P_{\mathcal{W}} P_{\mathcal{T}}\|_{2\to 2}^2 = \sup_{\|g\|_2 = 1} \|P_{\mathcal{W}} P_{\mathcal{T}} g\|_2^2 \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |q(s,t)|^2 dt \, ds.$$

(d) We observe that

$$q(s+t,t) = \int_{-\infty}^{\infty} e^{2\pi i w s} \mathbb{1}_{\mathcal{W}}(w) \, dw \cdot \mathbb{1}_{\mathcal{T}}(t) = \mathcal{F}^{-1}\{\mathbb{1}_{\mathcal{W}}\}(s) \cdot \mathbb{1}_{\mathcal{T}}(t),$$

where $\mathcal{F}^{-1}\{\mathbb{1}_{\mathcal{W}}\}(s)$ is the inverse Fourier transform of the indicator function $\mathbb{1}_{\mathcal{W}}$ evaluated at s. This yields

$$\begin{aligned}
\int_{-\infty}^{\infty} |q(s,t)|^2 \, ds &= \int_{-\infty}^{\infty} |q(s+t,t)|^2 \, ds \\
&= \int_{-\infty}^{\infty} \left| \mathcal{F}^{-1}\{\mathbb{1}_{\mathcal{W}}\}(s) \right|^2 ds \cdot \mathbb{1}_{\mathcal{T}}(t) \\
&\overset{\text{Pl.}}{=} \int_{-\infty}^{\infty} \left| \mathcal{F}\mathcal{F}^{-1}\{\mathbb{1}_{\mathcal{W}}\}(w) \right|^2 dw \cdot \mathbb{1}_{\mathcal{T}}(t) \\
&= \int_{-\infty}^{\infty} \mathbb{1}_{\mathcal{W}}(w) \, dw \cdot \mathbb{1}_{\mathcal{T}}(t), \qquad (40)
\end{aligned}$$

where we used the Plancherel identity, abbreviated as 'Pl.'. Upon integration over t, (40) results in

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |q(s,t)|^2 ds \, dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{1}_{\mathcal{W}}(w) \mathbb{1}_{\mathcal{T}}(t) \, dw \, dt = |\mathcal{W}||\mathcal{T}|. \qquad (41)$$

As $\mathcal{T}$ and $\mathcal{W}$ are bounded sets by assumption, the right hand side of (41) is finite and we can hence apply Fubini's theorem to conclude, as desired, that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|q(s,t)|^2 dt\,ds = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|q(s,t)|^2 ds\,dt = |\mathcal{W}||\mathcal{T}|.$$

(e) We combine the results established in the previous subproblems according to

$$|\mathcal{W}||\mathcal{T}| \overset{(d)}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|q(s,t)|^2 dt\,ds \overset{(c)}{\geq} \|P_{\mathcal{W}}P_{\mathcal{T}}\|_{2\to 2}^2 \overset{(b)}{\geq} \left(1-(\varepsilon_{\mathcal{T}}+\varepsilon_{\mathcal{W}})\right)^2.$$

Problem 4

In this solution, to avoid confusion, we write x_i for the i-th vector in the set $\{x_1, \ldots, x_N\}$ and $x_{(i)}$ for the i-th component of the vector x.

(a) (i) The first inequality is obtained by taking the square root in the following inequality

$$\|x\|_2^2 = \sum_{i=1}^{n} \left|x_{(i)}\right|^2 \leq \sum_{i=1}^{n} \left|x_{(i)}\right|^2 + \sum_{i,j,i\neq j} \left|x_{(i)}\right| \left|x_{(j)}\right| = \left(\sum_{i=1}^{n} \left|x_{(i)}\right|\right)^2 = \|x\|_1^2, \quad (42)$$

and the second one follows by application of the Cauchy-Schwarz inequality according to

$$\|x\|_1 = \langle x, \mathrm{sgn}(x)\rangle \overset{\text{C.S.}}{\leq} \|x\|_2 \|\mathrm{sgn}(x)\|_2 \leq \sqrt{n}\|x\|_2, \quad (43)$$

with $\mathrm{sgn}(x) \in \mathbb{R}^n$ defined as

$$\mathrm{sgn}(x)_{(i)} := \begin{cases} -1, & \text{if } x_{(i)} < 0, \\ +1, & \text{if } x_{(i)} > 0, \\ 0, & \text{if } x_{(i)} = 0. \end{cases}$$

(ii) Let $\{y_1, \ldots, y_N\} \subset \mathbb{R}^n$ be an ε-covering of K with respect to the $\|\cdot\|_1$-norm. For every $y \in K$, there hence exists an index i, $1 \leq i \leq N$, such that $\|y - y_i\|_1 \leq \varepsilon$ and therefore

$$\|y - y_i\|_2 \overset{(42)}{\leq} \|y - y_i\|_1 \leq \varepsilon.$$

We have hence established that every ε-covering of K with respect to the $\|\cdot\|_1$-norm is also an ε-covering of K with respect to the $\|\cdot\|_2$-norm, which in turn implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|_2) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|_1).$$

Likewise, let $\{z_1, \ldots, z_N\} \subset \mathbb{R}^n$ be an $(\varepsilon/\sqrt{n})$-covering of K with respect to the $\|\cdot\|_2$-norm. For every $z \in K$, there hence exists an index i, $1 \leq i \leq N$, such that $\|z - z_i\|_2 \leq \varepsilon/\sqrt{n}$ and therefore

$$\|z - z_i\|_1 \overset{(43)}{\leq} \sqrt{n}\|z - z_i\|_2 \leq \sqrt{n}\frac{\varepsilon}{\sqrt{n}} = \varepsilon.$$

We have hence established that every $(\varepsilon/\sqrt{n})$-covering of K with respect to the $\|\cdot\|_2$-norm is also an ε-covering of K with respect to the $\|\cdot\|_1$-norm, which in turn implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|_1) \leq \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2).$$

(iii) First note that from $\mathcal{N}(\varepsilon; K, \|\cdot\|_2) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|_1)$, one has for all $\varepsilon > 0$,

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_2)}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_1)}{\log_2(1/\varepsilon)}.$$

Taking the limit $\varepsilon \to 0^+$ on both sides yields

$$\dim_{\|\cdot\|_2}(K) \leq \dim_{\|\cdot\|_1}(K). \tag{44}$$

Likewise, it follows from $\mathcal{N}(\varepsilon; K, \|\cdot\|_1) \leq \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)$, that for all $\varepsilon > 0$,

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_1)}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)}{\log_2(1/\varepsilon)} = \frac{\log_2 \mathcal{N}(\varepsilon/\sqrt{n}; K, \|\cdot\|_2)}{\log_2(\sqrt{n}/\varepsilon) - \log_2(\sqrt{n})}.$$

Taking the limit $\varepsilon \to 0^+$ on both sides yields

$$\dim_{\|\cdot\|_1}(K) \leq \dim_{\|\cdot\|_2}(K). \tag{45}$$

Combining (44) and (45), we get the desired result

$$\dim_{\|\cdot\|_1}(K) = \dim_{\|\cdot\|_2}(K).$$

(iv) We proceed as above but for general norms $\|\cdot\|$ and $\|\cdot\|'$ on $\mathbb{R}^n$. From the norm equivalence in finite dimensions, it follows that there exists a constant $C \geq 1$ such that

$$C^{-1}\|x\| \leq \|x\|' \leq C\|x\|,$$

for all $x \in \mathbb{R}^n$. Let $\{y_1, \dots, y_N\} \subset \mathbb{R}^n$ be a $(C^{-1}\varepsilon)$-covering of K with respect to the $\|\cdot\|$-norm. For every $y \in K$, there hence exists an index i, $1 \leq i \leq N$, such that $\|y - y_i\| \leq C^{-1}\varepsilon$ and therefore

$$\|y - y_i\|' \leq C\|y - y_i\| \leq C\, C^{-1}\varepsilon = \varepsilon.$$

We have hence established that every $(C^{-1}\varepsilon)$-covering of K with respect to the $\|\cdot\|$-norm is an ε-covering with respect to the $\|\cdot\|'$-norm, which implies

$$\mathcal{N}(\varepsilon; K, \|\cdot\|') \leq \mathcal{N}(C^{-1}\varepsilon; K, \|\cdot\|). \tag{46}$$

A similar argument with the roles of the $\|\cdot\|$-norm and the $\|\cdot\|'$-norm reversed yields

$$\mathcal{N}(C\varepsilon; K, \|\cdot\|) \leq \mathcal{N}(\varepsilon; K, \|\cdot\|'). \tag{47}$$

Combining (46) and (47) allows us to conclude that

$$\frac{\log_2 \mathcal{N}(C\varepsilon; K, \|\cdot\|)}{\log_2(1/(C\varepsilon)) + \log_2(C)} \leq \frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|')}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(C^{-1}\varepsilon; K, \|\cdot\|)}{\log_2(1/(C^{-1}\varepsilon)) - \log_2(C)}.$$

Taking the limit $\varepsilon \to 0^+$ yields the desired result

$$\dim_{\|\cdot\|}(K) = \dim_{\|\cdot\|'}(K).$$

(b) (i) We first prove the result for $R = 1$. Following the Hint, we use the relation between metric entropy and the volume ratio provided in the Handout and applied to the unit ball $\mathcal{B} = \mathcal{B}' = B_{\|\cdot\|}(0,1)$ to obtain

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|) \leq \left(\frac{2}{\varepsilon} + 1\right)^n, \tag{48}$$

where we used

$$vol\left(\frac{2}{\varepsilon}\mathcal{B} + \mathcal{B}'\right) = vol\left(\left(\frac{2}{\varepsilon} + 1\right)\mathcal{B}\right) = \left(\frac{2}{\varepsilon} + 1\right)^n vol\left(\mathcal{B}\right).$$

The bounds in (48) now yield $\mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|) \asymp \varepsilon^{-n}$, which in turn implies

$$\dim(B_{\|\cdot\|}(0,1)) = \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon; \mathcal{B}, \|\cdot\|)}{\log_2(1/\varepsilon)} = n.$$

For general $R > 0$, we observe that, by scaling, according to the Hint, we have

$$\mathcal{N}(\varepsilon; B_{\|\cdot\|}(0,1), \|\cdot\|) = \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|), \tag{49}$$

which yields

$$\begin{aligned}
\dim(B_{\|\cdot\|}(0,R)) &= \lim_{\varepsilon' \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon'; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/\varepsilon')} \\
&= \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/(R\varepsilon))} \\
&= \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(R\varepsilon; B_{\|\cdot\|}(0,R), \|\cdot\|)}{\log_2(1/\varepsilon)} \\
&\overset{(49)}{=} \lim_{\varepsilon \to 0^+} \frac{\log_2 \mathcal{N}(\varepsilon; B_{\|\cdot\|}(0,1), \|\cdot\|)}{\log_2(1/\varepsilon)} \\
&= \dim(B_{\|\cdot\|}(0,1)) = n,
\end{aligned}$$

where we took $\varepsilon' = R\varepsilon$.

(ii) Take $R > 0$ large enough such that $K \subset B_\infty(0,R)$, where $B_\infty(0,R)$ is the ball with respect to the infinity norm $\|\cdot\|_\infty$ centered at the origin and of radius R. Such an R exists as K is compact. This inclusion now implies a bound on the covering number according to $\mathcal{N}(\varepsilon; K, \|\cdot\|_\infty) \leq \mathcal{N}(\varepsilon; B_\infty(0,R), \|\cdot\|_\infty)$, for all $\varepsilon > 0$, and consequently also on the following ratio

$$\frac{\log_2 \mathcal{N}(\varepsilon; K, \|\cdot\|_\infty)}{\log_2(1/\varepsilon)} \leq \frac{\log_2 \mathcal{N}(\varepsilon; B_\infty(0,R), \|\cdot\|_\infty)}{\log_2(1/\varepsilon)}.$$

Taking the limit as $\varepsilon \to 0^+$ yields the bound $\dim_{\|\cdot\|_\infty}(K) \leq \dim_{\|\cdot\|_\infty}(B_\infty(0,R))$. The desired bound according to

$$\dim(K) \leq \dim(B_\infty(0,R)) \overset{(b)(i)}{=} n,$$

is now a consequence of the result in (a)(iv).

(iii) Consider $K = \{x\}$, for $x \in \mathbb{R}^n$. For every $\varepsilon > 0$, we have $\mathcal{N}(\varepsilon; K, \|\cdot\|_\infty) = 1$, which yields $\dim(K) = 0 < n$.