ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Chair for Mathematical Information Science
Prof. Dr. H. Bölcskei
Sternwartstrasse 7
CH-8092 Zürich

Mathematics of Information

Spring semester 2022

Solutions to Problem Set 8

Problem 1 Equivalent Expression for the Restricted Isometry Constants.

It follows directly from the definition that δ_s is the smallest number $\tilde{\delta} \geq 0$ such that

$$\left| \|Ax\|_2^2 - \|x\|_2^2 \right| \leq \tilde{\delta},$$

for all s-sparse $x \in \mathbb{C}^N$ with $\|x\|_2 \leq 1$. Therefore

$$\begin{aligned}
\delta_s &= \sup_{\substack{x \in \mathbb{C}^N \ s\text{-sparse} \\ \|x\|_2 \leq 1}} \left| \|Ax\|_2^2 - \|x\|_2^2 \right| = \sup_{\substack{S \subset [N] \\ \#(S) \leq s}} \sup_{\substack{x \in \mathbb{C}^S \\ \|x\|_2 \leq 1}} \left| \|Ax\|_2^2 - \|x\|_2^2 \right| \\
&= \sup_{\substack{S \subset [N] \\ \#(S) \leq s}} \sup_{\substack{x \in \mathbb{C}^S \\ \|x\|_2 \leq 1}} \left| \left\langle \left(A_S^{\mathrm{H}} A_S - \mathrm{Id} \right) x, x \right\rangle_{\mathbb{C}^S} \right| \\
&= \sup_{\substack{S \subset [N] \\ \#(S) \leq s}} \left\| A_S^{\mathrm{H}} A_S - \mathrm{Id} \right\|_2 ,
\end{aligned}$$

where in the last line we used the fact that $A_S^{\mathrm{H}} A_S - \mathrm{Id}$ is hermitian.

Problem 2 Properties of the Restricted Isometry Constants.

a) For every $\boldsymbol{x} \in \mathbb{C}^N$, which is a 1-sparse vector, there exists $i \in \{1, \dots, N\}$ and $x_i \in \mathbb{C}$ nonzero such that

$$\boldsymbol{x} = (0, \dots, 0, x_i, 0, \dots, 0)^T.$$

Therefore, δ_1 is the smallest $\delta \geq 0$ such that

$$(1 - \delta)|x_i|^2 \leq \|A_i\|_2^2 |x_i|^2 \leq (1 + \delta)|x_i|^2,$$

for all $i \in \{1, \dots, N\}$ and $x_i \in \mathbb{C}$ nonzero, where A_i denotes the i-th column of A. Since, by assumption, the columns of A are normalized, we have $\|A_i\|_2 = 1$, for all $i \in \{1, \dots, N\}$. Therefore, δ_1 is the smallest $\delta \geq 0$ such that

$$(1 - \delta) \leq 1 \leq (1 + \delta),$$

that is, $\delta_1 = 0$.

b) We know from Problem 1, that the resticted isometry constant δ_2 can be equivalently expressed as

$$\delta_2 = \sup_{\substack{i,j\in\{1,\ldots,N\}\\ i\neq j}} \left|\!\left|\!\left| A_{\{i,j\}}^{\mathrm{H}} A_{\{i,j\}} - \mathrm{Id} \right|\!\right|\!\right|_2,$$

where $A_{\{i,j\}}$ is the matrix obtained by retaining only the columns i and j of A. A little computation shows that

$$A_{\{i,j\}}^{\mathrm{H}} A_{\{i,j\}} - \mathrm{Id} = \begin{pmatrix} \|A_i\|_2^2 & \langle A_j, A_i\rangle \\ \langle A_i, A_j\rangle & \|A_j\|_2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Using that the columns of A are normalized yields

$$A_{\{i,j\}}^{\mathrm{H}} A_{\{i,j\}} - \mathrm{Id} = \begin{pmatrix} 0 & \langle A_j, A_i\rangle \\ \langle A_i, A_j\rangle & 0 \end{pmatrix},$$

which is a matrix with eigenvalues $\pm|\langle A_i, A_j\rangle|$. Therefore, the operator norm of this matrix being equal to its largest eigenvalue, we get

$$\delta_2 = \max_{\substack{i,j\in\{1,\ldots,N\}\\ i\neq j}} |\langle A_i, A_j\rangle|,$$

which, by definition, is the coherence of the matrix A.

c) As $\mathrm{supp}(u) \cap \mathrm{supp}(v) = \varnothing$, we must have $\langle u, v\rangle = 0$. Let us now define the support set $S := \mathrm{supp}(u) \cup \mathrm{supp}(v)$, which has cardinality $\#S = 2s$. We thus have

$$\begin{aligned} |\langle Au, Av\rangle| &= |\langle A_S u, A_S v\rangle| \\ &= \left|\left\langle A_S^{\mathrm{H}} A_S\, u, v\right\rangle - \langle u, v\rangle\right| \\ &= \left|\left\langle \left(A_S^{\mathrm{H}} A_S - \mathrm{Id}\right) u, v\right\rangle\right|. \end{aligned}$$

Now using consecutively Cauchy-Schwarz and the definition of the operator norm gives

$$|\langle Au, Av\rangle| \leq \left\| \left(A_S^{\mathrm{H}} A_S - \mathrm{Id}\right) u \right\|_2 \|v\|_2 \leq \left|\!\left|\!\left| A_S^{\mathrm{H}} A_S - \mathrm{Id} \right|\!\right|\!\right|_2 \|u\|_2 \|v\|_2,$$

which, using the result of Problem 1, yields the desired result

$$|\langle Au, Av\rangle| \leq \delta_{2s} \|u\|_2 \|v\|_2.$$

Problem 3 Restricted Orthogonality (Exam 2020/2021, Problem 1).

a) It follows directly from the definition of $\theta_{s,t}$ that $\theta_{s,t}$ is the smallest number $\tilde{\theta} \geq 0$ such that

$$\left| \frac{|\langle Au, Av\rangle|}{\|u\|_2 \|v\|_2} \right| \leq \tilde{\theta},$$

for all disjointly supported s-sparse and t-sparse vectors $u \in \mathbb{C}^N \backslash \{0\}$ and $v \in \mathbb{C}^N \backslash \{0\}$, respectively. Therefore,

$$\begin{aligned}
\theta_{s,t} &= \max_{\substack{u,v\in\mathbb{C}^N \text{ disjointly } s,t\text{-sparse},\\ \|u\|_2=\|v\|_2=1}} |\langle Au, Av\rangle| \\
&= \max_{\substack{S,T\subset\{1,\dots,N\}, S\cap T=\emptyset,\\ |S|\le s, |T|\le t}} \max_{\substack{u\in\mathbb{C}^{|S|}, v\in\mathbb{C}^{|T|},\\ \|u\|_2=\|v\|_2=1}} |\langle Au, Av\rangle| \\
&= \max_{\substack{S,T\subset\{1,\dots,N\}, S\cap T=\emptyset,\\ |S|\le s, |T|\le t}} \max_{\substack{u\in\mathbb{C}^{|S|}, v\in\mathbb{C}^{|T|},\\ \|u\|_2=\|v\|_2=1}} |\langle A_S u, A_T v\rangle| \\
&= \max_{\substack{S,T\subset\{1,\dots,N\}, S\cap T=\emptyset,\\ |S|\le s, |T|\le t}} \max_{\substack{u\in\mathbb{C}^{|S|}, v\in\mathbb{C}^{|T|},\\ \|u\|_2=\|v\|_2=1}} |\langle (A_T^H A_S) u, v\rangle|.
\end{aligned} \tag{1}$$

Note that for every $S, T \subset \{1, \dots, N\}$ with $S \cap T = \emptyset$, $|S| \le s$, $|T| \le t$ and $u \in \mathbb{C}^{|S|}$, $v \in \mathbb{C}^{|T|}$ with $\|u\|_2 = \|v\|_2 = 1$, the Cauchy–Schwarz inequality yields

$$|\langle (A_T^H A_S) u, v\rangle| \le \|(A_T^H A_S) u\|_2 \|v\|_2 = \|(A_T^H A_S) u\|_2.$$

On the other hand, $A_T^H A_S \in \mathbb{C}^{|T|\times|S|}$ and hence $(A_T^H A_S)u \in \mathbb{C}^{|T|}$. Therefore, if $(A_T^H A_S)u \neq 0$, $v = (A_T^H A_S)u / \|(A_T^H A_S)u\|_2$ satisfies $v \in \mathbb{C}^{|T|}$ with $\|v\|_2 = 1$, and we get

$$\begin{aligned}
|\langle (A_T^H A_S) u, v\rangle| &= |\langle (A_T^H A_S) u, (A_T^H A_S) u / \|(A_T^H A_S) u\|_2\rangle| \\
&= \frac{\|(A_T^H A_S) u\|_2^2}{\|(A_T^H A_S) u\|_2} = \|(A_T^H A_S) u\|_2.
\end{aligned}$$

Hence,

$$\max_{\substack{u\in\mathbb{C}^{|S|}, v\in\mathbb{C}^{|T|},\\ \|u\|_2=\|v\|_2=1}} |\langle (A_T^H A_S) u, v\rangle| = \max_{\substack{u\in\mathbb{C}^{|S|},\\ \|u\|_2=1}} \|(A_T^H A_S) u\|_2 = \max_{\substack{u\in\mathbb{C}^{|S|},\\ \|u\|_2\le 1}} \|(A_T^H A_S) u\|_2.$$

Combining this with (1), we obtain

$$\begin{aligned}
\theta_{s,t} &= \max_{\substack{S,T\subset\{1,\dots,N\}, S\cap T=\emptyset,\\ |S|\le s, |T|\le t}} \max_{\substack{u\in\mathbb{C}^{|S|},\\ \|u\|_2\le 1}} \|(A_T^H A_S) u\|_2 \\
&= \max_{\substack{S,T\subset\{1,\dots,N\}, S\cap T=\emptyset,\\ |S|\le s, |T|\le t}} \|A_T^H A_S\|_2.
\end{aligned}$$

b) Let $u, v \in \mathbb{C}^N$ be disjointly supported s-sparse and t-sparse vectors, respectively, let $S := \operatorname{supp}(u) \cup \operatorname{supp}(v)$, and let $u_S, v_S \in \mathbb{C}^{|S|}$ be the restrictions of $u, v \in \mathbb{C}^N$ to S. Since u and v have disjoint supports, we have $\langle u_S, v_S\rangle = 0$ and hence

$$|\langle Au, Av\rangle| = |\langle A_S u_S, A_S v_S\rangle - \langle u_S, v_S\rangle| = |\langle (A_S^H A_S - \mathrm{I}_{|S|}) u_S, v_S\rangle|.$$

Applying the Cauchy–Schwarz inequality and the relation

$$\|u_S\|_2 \|A_S^H A_S - \mathrm{I}_{|S|}\|_2 = \|u_S\|_2 \max_{\substack{x\in\mathbb{C}^{|S|},\\ \|x\|_2\le 1}} \|(A_S^H A_S - \mathrm{I}_{|S|}) x\|_2 \ge \|(A_S^H A_S - \mathrm{I}_{|S|}) u_S\|_2,$$

we get

$$|\langle Au, Av\rangle| \leq \|(A_S^H A_S - \mathrm{I}_{|S|})u_S\|_2 \|v_S\|_2 \leq \|A_S^H A_S - \mathrm{I}_{|S|}\|_2 \|u_S\|_2 \|v_S\|_2.$$

Based on the lemma in the problem statement and using $\|u_S\|_2 = \|u\|_2$, $\|v_S\|_2 = \|v\|_2$, this allows us to conclude that

$$|\langle Au, Av\rangle| \leq \delta_{s+t}\|u\|_2\|v\|_2,$$

which, in turn, proves

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

c) i) Note that each $j \in T$ belongs to exactly s sets S_i, so 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) Using c)i) and the triangle inequality, we get

$$\begin{aligned} |\langle Au, Av\rangle| &\leq \frac{1}{s}\sum_{i=1}^{t} |\langle Au_{S_i}, Av\rangle| \leq \frac{1}{s}\sum_{i=1}^{t} \theta_{s,r}\|u_{S_i}\|_2\|v\|_2 \\ &= \theta_{s,r}\,\frac{1}{s}\left(\sum_{i=1}^{t}\|u_{S_i}\|_2\right)\|v\|_2, \end{aligned} \tag{2}$$

where in the second inequality we used that u_{S_i} and v are disjointly supported s-sparse and r-sparse vectors, respectively. Moreover, note that the Cauchy-Schwarz inequality yields

$$\left(\sum_{i=1}^{t}\|u_{S_i}\|_2\right)^2 \leq \left(\sum_{i=1}^{t}\|u_{S_i}\|_2^2\right)\left(\sum_{i=1}^{t}1\right) = t\left(\sum_{i=1}^{t}\|u_{S_i}\|_2^2\right).$$

This together with (2) and subproblem (c)(i) yields

$$\begin{aligned} |\langle Au, Av\rangle| &\leq \theta_{s,r}\,\frac{\sqrt{t}}{s}\left(\sum_{i=1}^{t}\|u_{S_i}\|_2^2\right)^{1/2}\|v\|_2 \\ &= \theta_{s,r}\sqrt{\frac{t}{s}}\,\|u\|_2\|v\|_2. \end{aligned}$$

Problem 4 Johnson-Lindenstrauss with sub-Gaussian Matrices.

Let $\boldsymbol{a}_j^T$ be the j-th row of $\boldsymbol{A}$, and set $X_j := \frac{\sqrt{k}}{\|\boldsymbol{u}\|}\boldsymbol{a}_j^T\boldsymbol{u}$. Note that

$$\boldsymbol{a}_j^T\boldsymbol{u} = \sum_{i=1}^{n}\boldsymbol{a}_{i,j}\boldsymbol{u}_i$$

is a linear combination of independent sub-Gaussian random variables. We first prove that it implies that X_j is a sub-Gaussian random variable with parameter 1. Using the definition of a sub-Gaussian random variable, together with the independence of the $\{\boldsymbol{a}_{i,j}\}_{1\leq i\leq n}$, we have

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda X_j}\right] &= \mathbb{E}\left[e^{\lambda \sum_{i=1}^{n} \frac{\sqrt{k}}{\|\boldsymbol{u}\|}\boldsymbol{a}_{i,j}\boldsymbol{u}_i}\right] \\
&= \prod_{i=1}^{n} \mathbb{E}\left[e^{\frac{\lambda\sqrt{k}\boldsymbol{u}_i}{\|\boldsymbol{u}\|}\boldsymbol{a}_{i,j}}\right] \\
&\leq \prod_{i=1}^{n} e^{\frac{\lambda^2 k \boldsymbol{u}_i^2}{2\|\boldsymbol{u}\|^2}\frac{1}{k}} \\
&= e^{\frac{\lambda^2}{2}},
\end{aligned}$$

for all $\lambda \in \mathbb{R}$. Therefore, X_j is indeed a sub-Gaussian random variable with parameter 1. Now, taking $\lambda \in (0, 1/2)$, and applying Markov's inequality, we obtain

$$\begin{aligned}
\mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \geq (1+\epsilon)\|\boldsymbol{u}\|^2\right] &= \mathbb{P}\left[\sum_{i=1}^{k} X_j^2 \geq (1+\epsilon)k\right] \\
&= \mathbb{P}\left[e^{\lambda\sum_{i=1}^{k} X_j^2} \geq e^{\lambda(1+\epsilon)k}\right] \\
&\leq e^{-\lambda(1+\epsilon)k}\mathbb{E}\left[e^{\lambda\sum_{i=1}^{k} X_j^2}\right] \\
&= e^{-\lambda(1+\epsilon)k}\left(\mathbb{E}\left[e^{\lambda X_1^2}\right]\right)^k.
\end{aligned}$$

Using that X_1 is a sub-Gaussian random variable with parameter 1, it satisfies $\mathbb{E}\left[e^{\lambda X_1^2}\right] \leq \frac{1}{\sqrt{1-2\lambda}}$, and therefore, it holds that

$$\mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \geq (1+\epsilon)\|\boldsymbol{u}\|^2\right] \leq \left(\frac{e^{-2\lambda(1+\epsilon)}}{1-2\lambda}\right)^{k/2}.$$

Taking the particular value $\lambda = \frac{\epsilon}{2(1+\epsilon)}$, we obtain

$$\mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \geq (1+\epsilon)\|\boldsymbol{u}\|^2\right] \leq \left((1+\epsilon)e^{-\epsilon}\right)^{k/2} < e^{-k\frac{\epsilon^2-\epsilon^3}{4}},$$

where the last inequality comes from the relation

$$1+\epsilon < e^{\epsilon-(\epsilon^2-\epsilon^3)/2}. \tag{3}$$

Repeating the arguments above, we can also prove that

$$\mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \leq (1-\epsilon)\|\boldsymbol{u}\|^2\right] < e^{-k\frac{\epsilon^2-\epsilon^3}{4}}.$$

Finally, applying the union bound yields the desired result

$$\begin{aligned}
\mathbb{P}\left[\left|\|\boldsymbol{A}\boldsymbol{u}\|^2 - \|\boldsymbol{u}\|^2\right| \geq \epsilon\|\boldsymbol{u}\|^2\right] &= \mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \leq (1-\epsilon)\|\boldsymbol{u}\|^2\right] + \mathbb{P}\left[\|\boldsymbol{A}\boldsymbol{u}\|^2 \geq (1+\epsilon)\|\boldsymbol{u}\|^2\right] \\
&< 2e^{-(\epsilon^2-\epsilon^3)k/4}.
\end{aligned}$$

Remark: We conclude this problem by providing an explicit proof of the inequality (3). We need to prove that, for all $0 < \epsilon < 1$, it holds that

$$1 + \epsilon < e^{\epsilon - (\epsilon^2 - \epsilon^3)/2}.$$

Taking logarithm on both sides, we equivalently wish to prove

$$\log(1 + \epsilon) < \epsilon - (\epsilon^2 - \epsilon^3)/2.$$

A Taylor expansion yields

$$\begin{aligned}
\epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{2} - \log(1+\epsilon) &= \epsilon - \frac{\epsilon^2}{2} + \frac{\epsilon^3}{2} + \sum_{n=1}^{\infty} \frac{(-\epsilon)^n}{n} \\
&= \frac{\epsilon^3}{2} - \frac{\epsilon^3}{3} + \sum_{n=4}^{\infty} \frac{(-\epsilon)^n}{n} \\
&= \frac{\epsilon^3}{2} - \frac{\epsilon^3}{3} + \sum_{n=2}^{\infty} \frac{\epsilon^{2n}}{2n} - \frac{\epsilon^{2n+1}}{2n+1} \\
&= \frac{\epsilon^3}{6} + \sum_{n=2}^{\infty} \frac{2n(1-\varepsilon)+1}{2n(2n+1)} \epsilon^{2n} > 0.
\end{aligned}$$

Where the last inequality is obtained by observing that we have a sum of positive terms when $0 < \epsilon < 1$. This is the desired result.

Problem 5 Random Matrices (Exam 2020, Problem 1).

a) i) We have

$$\sum_{j=1}^{N} \tilde{\mu}_j^2 = \sum_{j=1}^{N} \max_{k \in [N]} |\langle g_k, h_j \rangle|^2 \geq \sum_{j=1}^{N} |\langle g_1, h_j \rangle|^2 = \|g_1\|_2^2 = 1,$$

where we used the fact that $\{h_j\}_{j=1}^N$ is an orthonormal basis for $\mathbb{C}^N$.

ii) Let $\boldsymbol{t}$ be a random variable taking values in $[N] := \{1, 2, \ldots, N\}$ and distributed according to $\widetilde{\mathcal{P}}$. We then have

$$\kappa^{-1} = \min_{j \in [N]} p_j \, \tilde{\mu}_j^{-2} \leq p_{\boldsymbol{t}} \, \tilde{\mu}_{\boldsymbol{t}}^{-2}.$$

Taking the expectation of both sides with respect to $\boldsymbol{t}$ yields $\kappa^{-1} \leq \mathbb{E}_{\boldsymbol{t} \sim \widetilde{\mathcal{P}}} \left[p_{\boldsymbol{t}} \, \tilde{\mu}_{\boldsymbol{t}}^{-2} \right]$.

iii) We have

$$\begin{aligned}
\kappa^{-1} \leq \mathbb{E}_{\boldsymbol{t} \sim \widetilde{\mathcal{P}}} \left[p_{\boldsymbol{t}} \, \tilde{\mu}_{\boldsymbol{t}}^{-2} \right] &= \sum_{j=1}^{N} \mathbb{P}_{\boldsymbol{t} \sim \widetilde{\mathcal{P}}} (\boldsymbol{t} = j) \cdot p_j \tilde{\mu}_j^{-2} \\
&= \sum_{j=1}^{N} \frac{\tilde{\mu}_j^2}{\tilde{\mu}_1^2 + \cdots + \tilde{\mu}_N^2} p_j \tilde{\mu}_j^{-2} = \frac{1}{\tilde{\mu}_1^2 + \cdots + \tilde{\mu}_N^2} \underbrace{\sum_{j=1}^{N} p_j}_{=1} = \frac{1}{\tilde{\mu}_1^2 + \cdots + \tilde{\mu}_N^2},
\end{aligned}$$

and hence $\kappa \geq \sum_{j=1}^N \tilde{\mu}_j^2$. Moreover, equality holds if and only if

$$\min_{j\in[N]} p_j \, \tilde{\mu}_j^{-2} = \mathbb{E}_{\boldsymbol{t}\sim\widetilde{\mathcal{P}}} \left[p_{\boldsymbol{t}} \tilde{\mu}_{\boldsymbol{t}}^{-2} \right].$$

This is the case if and only if the random variable $p_{\boldsymbol{t}} \, \tilde{\mu}_{\boldsymbol{t}}^{-2}$ is constant, i.e., there exists a $c \in \mathbb{R}$ such that $p_j \, \tilde{\mu}_j^{-2} = c$, for $j \in [N]$. If this is the case, then, using that $\mathcal{P}$ is a probability distribution, we have

$$c \sum_{k=1}^N \tilde{\mu}_k^2 = \sum_{k=1}^N p_k = 1,$$

and thus $p_j = c\tilde{\mu}_j^2 = \left(\sum_{k=1}^N \tilde{\mu}_k^2\right)^{-1} \tilde{\mu}_j^2$, for all $j \in [N]$, i.e., $\mathcal{P} = \widetilde{\mathcal{P}}$.

b) i) We have

$$\begin{aligned}
\mathbb{E}\left[|\langle x, Y_\ell\rangle|^2\right] &= \mathbb{E}\left[\langle x, Y_\ell\rangle \overline{\langle x, Y_\ell\rangle}\right] = \mathbb{E}\left[\sum_{k,k'=1}^N A_{\ell k} x_k \overline{A_{\ell k'} x_{k'}}\right] \\
&= \sum_{k,k'=1}^N x_k \overline{x_{k'}} \, m^{-1} \mathbb{E}_{\boldsymbol{t}\sim\mathcal{P}} \left[p_{\boldsymbol{t}}^{-1} \langle g_k, h_{\boldsymbol{t}}\rangle \overline{\langle g_{k'}, h_{\boldsymbol{t}}\rangle} \right] \\
&= \sum_{k,k'=1}^N x_k \overline{x_{k'}} \, m^{-1} \sum_{j=1}^N p_j \cdot p_j^{-1} \langle g_k, h_j\rangle \overline{\langle g_{k'}, h_j\rangle} \\
&= \sum_{k,k'=1}^N x_k \overline{x_{k'}} \, m^{-1} \left\langle \sum_{j=1}^N \langle g_k, h_j\rangle h_j, g_{k'} \right\rangle \\
&= \sum_{k,k'=1}^N x_k \overline{x_{k'}} \, m^{-1} \underbrace{\langle g_k, g_{k'}\rangle}_{=\delta_{kk'}} = m^{-1} \|x\|_2^2 = m^{-1},
\end{aligned}$$

and so $\mathbb{E}[X_\ell] = \mathbb{E}[|\langle x, Y_\ell\rangle|^2] - \frac{1}{m} = 0$. Next, for $\ell \in [m]$, applying the Cauchy-Schwarz inequality and $|\langle g_j, h_{\boldsymbol{t}_\ell}\rangle| \leq \tilde{\mu}_{\boldsymbol{t}_\ell}$ yields

$$\begin{aligned}
|\langle x, Y_\ell\rangle|^2 &\leq \Big(\sum_{j\in\operatorname{supp}(x)} |A_{\ell j}|^2 \Big) \|x\|_2^2 = \sum_{j\in\operatorname{supp}(x)} m^{-1} p_{\boldsymbol{t}_\ell}^{-1} |\langle g_j, h_{\boldsymbol{t}_\ell}\rangle|^2 \\
&\leq \sum_{j\in\operatorname{supp}(x)} m^{-1} p_{\boldsymbol{t}_\ell}^{-1} \tilde{\mu}_{\boldsymbol{t}_\ell}^2 = s m^{-1} p_{\boldsymbol{t}_\ell}^{-1} \tilde{\mu}_{\boldsymbol{t}_\ell}^2 \leq s m^{-1} \kappa,
\end{aligned}$$

and so

$$|X_\ell| \leq \max\left\{ |\langle x, Y_\ell\rangle|^2, \tfrac{1}{m} \right\} \leq m^{-1} \max\{s\kappa, 1\} = s m^{-1} \kappa,$$

where the second inequality follows from $\kappa \geq \sum_{j=1}^N \tilde{\mu}_j^2 \geq 1$. Finally, we estimate

$$\begin{aligned}
\mathbb{E}\left[|X_\ell|^2\right] &= \mathbb{E}\left[|\langle x, Y_\ell\rangle|^4\right] - \tfrac{2}{m} \mathbb{E}\left[|\langle x, Y_\ell\rangle|^2\right] + \tfrac{1}{m^2} \\
&\leq s m^{-1} \kappa \, \mathbb{E}\left[|\langle x, Y_\ell\rangle|^2\right] - \tfrac{2}{m} \mathbb{E}\left[|\langle x, Y_\ell\rangle|^2\right] + \tfrac{1}{m^2} \\
&= s m^{-1} \kappa \tfrac{1}{m} - \tfrac{2}{m}\tfrac{1}{m} + \tfrac{1}{m^2} < s m^{-2} \kappa,
\end{aligned}$$

for all $\ell \in [m]$, as desired.

ii) We have

$$\|Ax\|_2^2 - 1 = \sum_{\ell=1}^{m} |\langle x, Y_\ell \rangle|^2 \; - 1 = \sum_{\ell=1}^{m} X_\ell,$$

and thus obtain using Bernstein's inequality

$$\begin{aligned}\mathbb{P}\left(\left|\|Ax\|_2^2 - 1\right| \geq t\right) = \mathbb{P}\left(\left|\sum_{\ell=1}^{m} X_\ell\right| \geq t\right) &\leq 2\exp\left(-\frac{t^2/2}{m \cdot sm^{-2}\kappa + sm^{-1}\kappa\frac{t}{3}}\right) \\ &\leq 2\exp\left(-\frac{t^2/2}{sm^{-1}\kappa + \frac{1}{3}sm^{-1}\kappa}\right) = 2\exp\left(-\frac{3t^2 m}{8s\kappa}\right),\end{aligned}$$

for $t \in (0, 1)$, as desired.