Chapter 2

Lower bounds

In this section we will prove the impossibility results that were stated in the Introduction. Throughout what follows, all Banach spaces will be tacitly assumed to be separable. Given a Banach space $\mathbf{X}$, its Banach–Mazur distance to a Hilbert space will be denoted $\mathsf{d}_{\mathbf{X}} \in [1,\infty]$, i.e., $\mathsf{d}_{\mathbf{X}} = d_{\mathrm{BM}}(\mathbf{X},\mathbf{H})$ where $\mathbf{H}$ is a Hilbert space with either $\dim(\mathbf{H}) = \dim(\mathbf{X})$ when $\dim(\mathbf{X}) < \infty$, or $\mathbf{H} = \ell_2$ when $\mathbf{X}$ is infinite dimensional. By a classical result of Enflo [93, Theorem 6.3.3] (see also [36, Corollary 7.10]) we have $\mathsf{d}_{\mathbf{X}} = \mathsf{c}_2(\mathbf{X})$.

2.1 Proof of Theorem 13

Recall that the (Gaussian) type 2 and cotype 2 constants of a Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$, denoted $T_2(\mathbf{X})$ and $C_2(\mathbf{X})$, respectively, are the infimum over those $T \in [1,\infty]$ and $C \in [1,\infty]$, respectively, for which the following inequalities hold for every $m \in \mathbb{N}$ and every $x_1,\ldots,x_m \in \mathbf{X}$:

$$\frac{1}{C^2}\sum_{j=1}^{m}\|x_j\|_{\mathbf{X}}^2 \leqslant \mathbb{E}\left[\left\|\sum_{j=1}^{m}\mathsf{g}_j x_j\right\|_{\mathbf{X}}^2\right] \leqslant T^2\sum_{j=1}^{m}\|x_j\|_{\mathbf{X}}^2, \tag{2.1}$$

where henceforth $\mathsf{g}_1,\mathsf{g}_2,\ldots$ will always denote i.i.d. standard Gaussian random variables. The following theorem of Kwapień [162] is fundamental (see also [261, Theorem 3.3] or [305, Theorem 13.15]).

Theorem 92. *Every Banach space* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *satisfies* $\mathsf{d}_{\mathbf{X}} \leqslant T_2(\mathbf{X})C_2(\mathbf{X})$.

We will use Theorem 92 to estimate the following quantity, which in turn will be used to get the best bound that we currently have on the constant c that appears in the lower bound on $\mathsf{e}(\mathbf{X})$ of Theorem 13.

Definition 93 (Lindenstrauss–Tzafriri constant). Suppose that $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a Banach space. Define $\mathsf{LT}(\mathbf{X})$ to be the infimum over those $K \in [1,\infty]$ such that for every closed linear subspace $\mathbf{V} \subseteq \mathbf{X}$ there exists a projection $\mathsf{Proj}: \mathbf{X} \twoheadrightarrow \mathbf{V}$ from $\mathbf{X}$ onto $\mathbf{V}$ whose operator norm satisfies $\|\mathsf{Proj}\|_{\mathbf{X}\to\mathbf{X}} \leqslant K$.

So, the Lindenstrauss–Tzafriri constant of a Hilbert space equals 1, and Sobczyk proved [290] that

$$\forall n \in \mathbb{N}, \quad \mathsf{LT}\big(\ell_1^n\big) \asymp \mathsf{LT}\big(\ell_\infty^n\big) \asymp \sqrt{n}. \tag{2.2}$$

We chose the nomenclature of Definition 93 in reference to the famous solution [180] by Lindenstrauss and Tzafriri of the *complemented subspace problem*, which asserts that if $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a Banach space for which $\mathsf{LT}(\mathbf{X}) < \infty$, then $\mathbf{X}$ is isomorphic to a Hilbert space, i.e., $\mathsf{d}_{\mathbf{X}} < \infty$. Moreover, if $\mathbf{X}$ is infinite dimensional, then it was shown in [180] that

$$\mathsf{d}_{\mathbf{X}} \lesssim \mathsf{LT}(\mathbf{X})^4.$$

This dependence was improved in [147] by Kadec and Mitjagin, who established the following theorem, which is the currently best-known bound in the Lindenstrauss–Tzafriri theorem (see also [3, 97, 150, 262, 264] for subsequent improvements of the implicit universal constant factor and further generalizations).

Theorem 94. *Every infinite dimensional Banach space* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *satisfies*

$$\mathsf{d}_{\mathbf{X}} \lesssim \mathsf{LT}(\mathbf{X})^2.$$

When $\dim(\mathbf{X}) < \infty$ the question of bounding $\mathsf{d}_{\mathbf{X}}$ by a function of $\mathsf{LT}(\mathbf{X})$ was left open in [180]. This question, which was eventually solved by Figiel, Lindenstrauss and Milman [99, Theorem 6.7], turned out to be significantly more subtle than its infinite dimensional counterpart. The currently best-known estimate is due to Tomczak-Jaegermann [305, Theorem 29.4], who proved the following theorem.

Theorem 95. *Every finite dimensional Banach space* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *satisfies*

$$\mathsf{d}_{\mathbf{X}} \lesssim \mathsf{LT}(\mathbf{X})^5.$$

The proof of Theorem 95 is achieved in [305] through an interesting combination of the *proof of* the Lindenstrauss–Tzafriri theorem [180] with the finite dimensional machinery of [99] and Milman's Quotient of Subspace Theorem [216].

The following theorem is a link between the Lindenstrauss–Tzafriri constant and Lipschitz extension.

Theorem 96. *Every Banach space* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *satisfies* $\mathsf{e}(\mathbf{X}) \geqslant \mathsf{LT}(\mathbf{X})$.

Proof. By Remark 98, if $\dim(\mathbf{X}) = \infty$, then $\mathsf{e}(\mathbf{X}) = \infty$, so we may assume that $\dim(\mathbf{X}) < \infty$. Fix $L > \mathsf{e}(\mathbf{X})$ and let $\mathbf{V} \subseteq \mathbf{X}$ be a linear subspace of $\mathbf{X}$. Then, the identity mapping from $\mathbf{V}$ to $\mathbf{V}$ can be extended to an L-Lipschitz mapping $\rho : \mathbf{X} \to \mathbf{V}$. In other words, ρ is an L-Lipschitz retraction from $\mathbf{X}$ onto $\mathbf{V}$. By a classical theorem of Lindenstrauss [179] (see also its elegant alternative proof by Pełczyńsky in [247, p. 61]), there is a projection of norm at most L from $\mathbf{X}$ onto $\mathbf{V}$. This proves that $\mathsf{LT}(\mathbf{X}) \leqslant L$. ∎

The following theorem is the lower bound $\mathsf{e}(\ell_2^n) \gtrsim \sqrt[4]{n}$ of [210] that we already quoted in (1.22), in combination with the bi-Lipschitz invariance of the Lipschitz extension modulus.

Theorem 97. *For every $n \in \mathbb{N}$, any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies*

$$\mathsf{e}(\mathbf{X}) \gtrsim \frac{\sqrt[4]{n}}{\mathsf{d}_{\mathbf{X}}}.$$

Remark 98. The question whether $\mathsf{e}(\ell_2)$ is finite or infinite was open for quite some time: it was first stated in print in [140, p. 137], and it was also posed by Ball in [23, p. 170] (Ball conjectured that $\mathsf{e}(\ell_2) = \infty$). We answered it in [224] by proving that $\lim_{n\to\infty} \mathsf{e}(\ell_2^n) = \infty$. Due to Dvoretzky's theorem [90] this implies that $\mathsf{e}(\mathbf{X})$ is at least an unbounded function of $\dim(\mathbf{X})$ for *any* normed space $\mathbf{X}$, and in particular $\mathsf{e}(\mathbf{X}) = \infty$ if $\dim(\mathbf{X}) = \infty$. A rate at which $\mathsf{e}(\ell_2^n)$ tends to ∞ was not specified in [224], but the reasoning of [224] was inspected quantitatively in [173, Remark 5.3], yielding an explicit lower bound that depends on an auxiliary parameter, and it was noted in [62] that an optimization over this parameter yields the estimate $\mathsf{e}(\ell_2^n) \gtrsim \sqrt[8]{n}$. A further improvement from [210] (whose proof refines ideas of Kalton [149, 151]) was the aforementioned estimate $\mathsf{e}(\ell_2^n) \gtrsim \sqrt[4]{n}$ (a different proof of this bound follows from [231]), which is the currently best-known lower bound on $\mathsf{e}(\ell_2^n)$. By Milman's sharpening [215] of Dvoretzky's theorem [90], it follows that every normed space $\mathbf{X}$ satisfies $\mathsf{e}(\mathbf{X}) \gtrsim \sqrt[4]{\log n}$. As we explained in Section 1.3, the bound $\mathsf{e}(\ell_\infty^n) \gtrsim \sqrt{n}$ is classical (specifically, by substituting (2.2) into Theorem 96). In combination with the Alon–Milman theorem [5] (see also [299]), the fact that both $\mathsf{e}(\ell_2^n) = n^{\Omega(1)}$ and $\mathsf{e}(\ell_\infty^n) = n^{\Omega(1)}$ formally implies that

$$\mathsf{e}(\mathbf{X}) \geqslant e^{\eta\sqrt{\log n}}$$

for some universal constant $\eta > 0$ and every n-dimensional normed space $\mathbf{X}$, which was the best-known general lower bound on the Lipschitz extension modulus prior to Theorem 1.

The above results imply as follows the lower bound on $\mathsf{e}(\mathbf{X})$ of Theorem 13. By combining Theorems 95 and 96, we have $\mathsf{e}(\mathbf{X}) \gtrsim \sqrt[5]{\mathsf{d}_{\mathbf{X}}}$. In combination with Theorem 97, it therefore follows that

$$\mathsf{e}(\mathbf{X}) \gtrsim \max\left\{\frac{\sqrt[4]{n}}{\mathsf{d}_{\mathbf{X}}}, \sqrt[5]{\mathsf{d}_{\mathbf{X}}}\right\} \geqslant \sqrt[24]{n}, \tag{2.3}$$

where the last step follows from elementary calculus and holds as equality when

$$\mathsf{d}_{\mathbf{X}} = n^{\frac{5}{24}}.$$

We will derive a better lower bound on $\mathsf{e}(\mathbf{X})$ than (2.3) through the following theorem which improves over the power of $\mathsf{LT}(\mathbf{X})$ in Theorem 95, showing that in the finite dimensional setting one can come close (up to logarithmic factors) to the infinite dimensional bound of Theorem 94; see also Remark 103 below.

Theorem 99. *For every integer $n \geqslant 2$, any n-dimensional Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ satisfies*

$$\mathsf{d}_{\mathbf{X}} \lesssim \mathsf{LT}(\mathbf{X})^2 (\log n)^3. \tag{2.4}$$

Assuming Theorem 2.3, reason analogously to (2.3) while using (2.4) in place of Theorem 95 to get

$$\mathsf{e}(\mathbf{X}) \gtrsim \max\left\{ \frac{\sqrt[4]{n}}{\mathsf{d}_{\mathbf{X}}}, \frac{\sqrt{\mathsf{d}_{\mathbf{X}}}}{(\log n)^3} \right\} \geqslant \frac{n^{\frac{1}{12}}}{(\log n)^2}, \tag{2.5}$$

where equality holds in the final step of (2.5) if and only if $\mathsf{d}_{\mathbf{X}} = \sqrt[6]{n}(\log n)^2$.

Prior to proving Theorem 99, we will record the following two standard lemmas that will be used in its proof; both will be established in correct generality that also treats infinite dimensional Banach spaces even though here we will need them only in the finite dimensional setting (the infinite dimensional formulations are relevant to the discussion in Remark 103).

Lemma 100. *For every Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ we have $\mathsf{LT}(\mathbf{X}^*) \leqslant \mathsf{LT}(\mathbf{X}) + 1$.*

Proof. We may assume that $\mathsf{LT}(\mathbf{X}) < \infty$. Then $\mathbf{X}$ is reflexive (even isomorphic to Hilbert space), by [180]. Fix a closed linear subspace $\mathbf{W}$ of $\mathbf{X}^*$ and denote its pre-annihilator by

$$^{\perp}\mathbf{W} \stackrel{\mathrm{def}}{=} \bigcap_{x^* \in \mathbf{W}} \{x \in \mathbf{X} : x^*(x) = 0\} \subseteq \mathbf{X}.$$

Suppose that $K > \mathsf{LT}(\mathbf{X})$. By the definition of $\mathsf{LT}(\mathbf{X})$ there exists $\mathsf{Proj} : \mathbf{X} \to \mathbf{X}$ that is a projection from $\mathbf{X}$ onto $^{\perp}\mathbf{W}$ whose operator norm satisfies $\|\mathsf{Proj}\|_{\mathbf{X}\to\mathbf{X}} \leqslant K$. Observe that for every $x^* \in \mathbf{X}^*$ and $x \in {}^{\perp}\mathbf{W}$,

$$\big(x^* - \mathsf{Proj}^* x^*\big)(x) = x^*(x) - x^*(\mathsf{Proj}x) = 0,$$

since $\mathsf{Proj}x = x$. This shows that

$$\big(\mathsf{Id}_{\mathbf{X}^*} - \mathsf{Proj}^*\big)(\mathbf{X}^*) \subseteq ({}^{\perp}\mathbf{W})^{\perp} = \big\{x^* \in X^* : x^*({}^{\perp}\mathbf{W}) = \{0\}\big\} = \mathbf{W},$$

where the last step follows from the double annihilator theorem since $\mathbf{X}$ is reflexive and hence $\mathbf{W}$ is weak* closed in $\mathbf{X}^*$. If $x^* \in \mathbf{W}$, then for any $x \in \mathbf{X}$ we have $\mathsf{Proj}^* x^*(x) = x^*(\mathsf{Proj}x) = 0$, as $\mathsf{Proj}x \in {}^{\perp}\mathbf{W}$. Hence $\mathsf{Proj}^* x^* = 0$, and so $\mathsf{Id}_{\mathbf{X}^*} - \mathsf{Proj}^*$ acts as the identity when it is restricted to $\mathbf{W}$, i.e., $\mathsf{Id}_{\mathbf{X}^*} - \mathsf{Proj}^* : \mathbf{X}^* \to \mathbf{X}^*$ is a projection from $\mathbf{X}^*$ onto $\mathbf{W}$. It remains to note that

$$\|\mathsf{Id}_{\mathbf{X}^*} - \mathsf{Proj}^*\|_{\mathbf{X}^*\to\mathbf{X}^*} \leqslant 1 + \|\mathsf{Proj}^*\|_{\mathbf{X}^*\to\mathbf{X}^*} = 1 + \|\mathsf{Proj}\|_{\mathbf{X}\to\mathbf{X}} \leqslant K + 1. \qquad \blacksquare$$

The following simple lemma shows that the Lindenstrauss–Tzafriri constant is a bi-Lipschitz invariant.

Lemma 101. *Any two Banach spaces* $(\mathbf{W}, \|\cdot\|_{\mathbf{W}})$ *and* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *satisfy*

$$\mathsf{LT}(\mathbf{W}) \leqslant \mathsf{c}_{\mathbf{X}}(\mathbf{W})\mathsf{LT}(\mathbf{X}). \tag{2.6}$$

Proof. We may assume that $\mathsf{c}_{\mathbf{X}}(\mathbf{W}) < \infty$ and $\mathsf{LT}(\mathbf{X}) < \infty$. By [180], the latter assumption implies that $\mathbf{X}$ is isomorphic to a Hilbert space, and hence it is reflexive. We may therefore apply a classical differentiation argument (see e.g., [36, Corollary 7.10]) to deduce that there is a closed subspace $\mathbf{Y}$ of $\mathbf{X}$ such that

$$d_{\mathrm{BM}}(\mathbf{W}, \mathbf{Y}) = \mathsf{c}_{\mathbf{X}}(\mathbf{W}).$$

In other words, for every $D > \mathsf{c}_{\mathbf{X}}(\mathbf{W})$ there is a linear isomorphism $T : \mathbf{W} \to \mathbf{Y}$ satisfying $\|T\|_{\mathbf{W}\to\mathbf{Y}}\|T^{-1}\|_{\mathbf{Y}\to\mathbf{W}} < D$. If $\mathbf{V}$ is a closed subspace of $\mathbf{W}$ and $K > \mathsf{LT}(\mathbf{X})$, then there is a projection Proj from $\mathbf{X}$ onto $T\mathbf{V}$ with $\|\mathsf{Proj}\|_{\mathbf{X}\to T\mathbf{V}} < K$. Now, $T^{-1}\mathsf{Proj}T$ is a projection from $\mathbf{W}$ onto $\mathbf{V}$ of norm less than DK. ■

The type-2 constant of a normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is equal to its "equal norm type-2 constant," namely to the infimum over those $T > 0$ for which the second inequality in (2.1) holds for every $m \in \mathbb{N}$ and every choice of vectors $x_1, \ldots, x_m \in \mathbf{X}$ that satisfy the additional requirement

$$\|x_1\|_{\mathbf{X}} = \cdots = \|x_m\|_{\mathbf{X}};$$

this is a well-known result of Pisier, though it first appeared in James' important work [134], where it had a vital role. We will likewise need to use this result, with the twist that we require a small number of unit vectors for which the type-2 constant of $\mathbf{X}$ is almost attained. The classical proof of the aforementioned equivalence between type-2 and "equal norm type-2" [134, p. 2] increases the number of vectors potentially uncontrollably, so we will preform the analysis more carefully in the following lemma, which shows that one need not increase the number of vectors when passing from general vectors to unit vectors.

Lemma 102 (Equal norm type 2 without increasing the number of vectors). *Fix* $n \in \mathbb{N}$ *and* $0 < \beta \leqslant 1$. *Let* $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ *be a normed space and suppose that there exist vectors* $x_1, \ldots, x_n \in \mathbf{X} \smallsetminus \{0\}$ *that satisfy*

$$\left(\mathbb{E}\left[\left\|\sum_{i=1}^n \mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} \geqslant \beta T_2(\mathbf{X})\left(\sum_{i=1}^n \|x_i\|_{\mathbf{X}}^2\right)^{\frac12}. \tag{2.7}$$

Then, there also exist unit vectors $y_1, \ldots, y_n \in \{x_i/\|x_i\|_{\mathbf{X}}\}_{i=1}^n \subseteq \partial B_{\mathbf{X}}$ *that satisfy*

$$\left(\mathbb{E}\left[\left\|\sum_{i=1}^n \mathsf{g}_i y_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} \gtrsim \beta^2 T_2(\mathbf{X})\sqrt{n}.$$

Proof. We may assume without loss of generality the following normalized version of assumption (2.7):

$$\sum_{i=1}^{n} \|x_i\|_{\mathbf{X}}^2 = 1 \quad \text{and} \quad \mathbb{E}\left[\left\|\sum_{i=1}^{n} \mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right] \geqslant \beta^2 T_2(\mathbf{X})^2. \tag{2.8}$$

For every $k \in \mathbb{N}$ define a subset I_k of $\{1, \ldots, n\}$ by

$$I_k \stackrel{\text{def}}{=} \left\{i \in \{1, \ldots, n\} : \frac{1}{2^k} < \|x_i\|_{\mathbf{X}} \leqslant \frac{1}{2^{k-1}}\right\}. \tag{2.9}$$

So, $\{I_k\}_{k \in \mathbb{N}}$ is a partition of $\{1, \ldots, n\}$ as $0 < \|x_i\|_{\mathbf{X}} \leqslant 1$ for all $i \in \{1, \ldots, n\}$ by the first equation in (2.8). Write

$$m \stackrel{\text{def}}{=} \left\lceil \log_2\left(\frac{3\sqrt{n}}{\beta}\right)\right\rceil \quad \text{and} \quad U \stackrel{\text{def}}{=} \bigcup_{k=1}^{m} I_k \times \left\{1, \ldots, 2^{2(m-k)}\right\}. \tag{2.10}$$

With this notation, Lemma 102 will be proven if we show that there exists $S \subseteq U$ with $|S| = n$ such that

$$\left(\mathbb{E}\left[\left\|\sum_{(i,j)\in S} \frac{\mathsf{g}_{ij}}{\|x_i\|_{\mathbf{X}}} x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \gtrsim \beta^2 T_2(\mathbf{X})\sqrt{n}, \tag{2.11}$$

where $\{\mathsf{g}_{ij}\}_{i,j=1}^{\infty}$ are i.i.d. standard Gaussian random variables.

To prove (2.11), observe first that by the contraction principle (see, e.g., [168, Section 4.2]) we have

$$\left(\mathbb{E}\left[\left\|\sum_{(i,j)\in S} \frac{\mathsf{g}_{ij}}{\|x_i\|_{\mathbf{X}}} x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \geqslant \left(\mathbb{E}\left[\left\|\sum_{k=1}^{m} 2^{k-1} \sum_{i\in I_k} \sum_{j=1}^{2^{2(m-k)}} \mathbf{1}_{\{(i,j)\in S\}} \mathsf{g}_{ij} x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}}, \tag{2.12}$$

where we used the fact that $1/\|x_i\|_{\mathbf{X}} \geqslant 2^{k-1}$ for every $k \in \mathbb{N}$ and $i \in I_k$ (by the definition (2.9) of I_k). Also,

$$\begin{aligned} 1 \stackrel{(2.8)}{=} \sum_{i=1}^{n} \|x_i\|_{\mathbf{X}}^2 &= \sum_{k=1}^{\infty} \sum_{i\in I_k} \|x_i\|_{\mathbf{X}}^2 \\ &\stackrel{(2.9)}{\leqslant} \sum_{k=1}^{\infty} \frac{|I_k|}{2^{2k-2}} \\ &\leqslant \frac{4\sum_{k=1}^{m} 2^{2(m-k)}|I_k| + \sum_{k=m+1}^{\infty} |I_k|}{2^{2m}} \\ &\stackrel{(2.10)}{\leqslant} \frac{\beta^2(4|U| + n)}{9n}. \end{aligned}$$

This simplifies to give that $|U| \geqslant 2n/\beta^2 > n$. We can therefore average the right-hand side of (2.12) over all the n-point subsets of U to get the following estimate:

$$\begin{aligned}
&\frac{1}{\binom{|U|}{n}} \sum_{\substack{S\subseteq U\\ |S|=n}} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} 2^{k-1} \sum_{i\in I_k} \sum_{j=1}^{2^{2(m-k)}} \mathbf{1}_{\{(i,j)\in S\}} \mathsf{g}_{ij} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \\
&\geqslant \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} 2^{k-1} \sum_{i\in I_k} \sum_{j=1}^{2^{2(m-k)}} \frac{\binom{|U|-1}{n-1}}{\binom{|U|}{n}} \mathsf{g}_{ij} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \\
&= \frac{n}{2|U|} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} 2^{k} \sum_{i\in I_k} \sum_{j=1}^{2^{2(m-k)}} \mathsf{g}_{ij} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \\
&= \frac{2^{m-1}n}{|U|} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} \sum_{i\in I_k} \mathsf{g}_{i} y_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \\
&\asymp \frac{n^{\frac{3}{2}}}{\beta|U|} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} \sum_{i\in I_k} \mathsf{g}_{i} y_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}}, \qquad (2.13)
\end{aligned}$$

where the first step of (2.13) uses convexity, the penultimate step of (2.13) uses the fact that

$$\left(\left(\sum_{j=1}^{2^{2(m-k)}} \mathsf{g}_{ij}\right)_{i\in I_k}\right)_{k=1}^{m} \quad \text{and} \quad \left(\left(2^{m-k}\mathsf{g}_i\right)_{i\in I_k}\right)_{k=1}^{m}$$

have the same distribution, and for the final step of (2.13) recall the definition (2.10) of m.

It follows from (2.12) and (2.13) that there must exist $S \subseteq U$ with $|S| = n$ such that

$$\left(\mathbb{E}\left[\left\| \sum_{(i,j)\in S} \frac{\mathsf{g}_{ij}}{\|x_i\|_{\mathbf{X}}} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \gtrsim \frac{n^{\frac{3}{2}}}{\beta|U|} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} \sum_{i\in I_k} \mathsf{g}_{i} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}}. \qquad (2.14)$$

To use (2.14), we claim that $|U| \lesssim n/\beta^2$. Indeed,

$$1 \overset{(2.8)}{=} \sum_{i=1}^{n} \|x_i\|_{\mathbf{X}}^2 = \sum_{k=1}^{\infty} \sum_{i\in I_k} \|x_i\|_{\mathbf{X}}^2 \overset{(2.9)}{>} \sum_{k=1}^{m} \frac{|I_k|}{2^{2k}} \overset{(2.10)}{=} \frac{|U|}{2^{2m}} \overset{(2.10)}{\geqslant} \frac{\beta^2|U|}{81n}.$$

By combining the aforementioned upper bound on the size of U with (2.12) and (2.14), we see that

$$\left(\mathbb{E}\left[\left\| \sum_{(i,j)\in S} \frac{\mathsf{g}_{ij}}{\|x_i\|_{\mathbf{X}}} x_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \gtrsim \beta\sqrt{n} \left(\mathbb{E}\left[\left\| \sum_{k=1}^{m} \sum_{i\in I_k} \mathsf{g}_{i} y_i \right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}}.$$

From this, we deduce the desired estimate (2.11) by combining as follows the second inequality in our assumption (2.8) with the triangle inequality and the definition (2.1) of the type-2 constant $T_2(\mathbf{X})$:

$$\begin{aligned}\left(\mathbb{E}\left[\left\|\sum_{k=1}^{m}\sum_{i\in I_k}\mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} &\geqslant \left(\mathbb{E}\left[\left\|\sum_{i=1}^{\infty}\mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} - \left(\mathbb{E}\left[\left\|\sum_{k=m+1}^{\infty}\sum_{i\in I_k}\mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12}\\ &\overset{(2.1)}{\geqslant} \beta T_2(\mathbf{X}) - T_2(\mathbf{X})\left(\sum_{k=m+1}^{\infty}\sum_{i\in I_k}\|x_i\|_{\mathbf{X}}^2\right)^{\frac12}\\ &\overset{(2.9)}{\geqslant} \beta T_2(\mathbf{X}) - \frac{T_2(\mathbf{X})\sqrt{n}}{2^m}\\ &\overset{(2.10)}{\asymp} \beta T_2(\mathbf{X}).\end{aligned}$$

■

Proof of Theorem 99. We will prove that the type 2 constant of $\mathbf{X}$ satisfies

$$T_2(\mathbf{X}) \lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac32}. \tag{2.15}$$

After (2.15) will be proven, we deduce Theorem 99 as follows. We first claim that the estimate (2.15) implies the same upper bound on the cotype 2 constant of $\mathbf{X}$. Namely, we also have

$$C_2(\mathbf{X}) \lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac32}. \tag{2.16}$$

Indeed,

$$\begin{aligned}C_2(\mathbf{X}) \leqslant T_2(\mathbf{X}^*) &\lesssim \mathsf{LT}(\mathbf{X}^*)(\log n)^{\frac32}\\ &\lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac32},\end{aligned} \tag{2.17}$$

where the first step of (2.17) follows from a standard duality argument [204] (see also, e.g., [220, Section 9.10], [253, Section 4.9] or [3, Proposition 6.2.12]), the second step of (2.17) is an application of (2.15) to $\mathbf{X}^*$, and the third step of (2.17) is application of Lemma 100. The desired estimate (2.4) now follows by a substitution of (2.15) and (2.16) into Theorem 92 (Kwapień's theorem).

By [99, Lemma 6.1] (see also the exposition of this fact in [141, p. 546]) there exists an integer[1]

$$1 \leqslant m \leqslant \frac{n(n+1)}{2} \tag{2.18}$$

[1] By [303], if one does not mind losing a universal constant factor in (2.19), then one could take $m = n$ here, but for the purpose of the ensuing reasoning it suffices to use the much simpler result [99, Lemma 6.1].

and $x_1,\ldots,x_m \in \mathbf{X}\smallsetminus\{0\}$ such that

$$\left(\mathbb{E}\left[\left\|\sum_{i=1}^m \mathsf{g}_i x_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} = T_2(\mathbf{X})\left(\sum_{i=1}^m \|x_i\|_{\mathbf{X}}^2\right)^{\frac12} \tag{2.19}$$

By Lemma 102, it follows that there exist $y_1,\ldots,y_m \in \partial B_{\mathbf{X}}$ and a universal constant $0<\gamma<1$ such that

$$\mathbb{E}\left[\left\|\sum_{i=1}^m \mathsf{g}_i y_i\right\|_{\mathbf{X}}\right] \geqslant \sqrt{\frac{2}{\pi}}\left(\mathbb{E}\left[\left\|\sum_{i=1}^m \mathsf{g}_i y_i\right\|_{\mathbf{X}}^2\right]\right)^{\frac12} \geqslant \gamma T_2(\mathbf{X})\sqrt{m}, \tag{2.20}$$

where the first step in (2.20) holds by (the Gaussian version of) Kahane's inequality [148] (see, e.g., [168, Corollary 3.2] and specifically [167, Corollary 3] for the (optimal) constant that we are quoting here even though its value is of secondary importance in the present context). If we denote

$$\delta \stackrel{\mathrm{def}}{=} \frac{\gamma T_2(\mathbf{X})}{\sqrt{m}}, \tag{2.21}$$

then a different way to write (2.20) is

$$\mathbb{E}\left[\left\|\sum_{i=1}^m \mathsf{g}_i y_i\right\|_{\mathbf{X}}\right] \geqslant \delta m. \tag{2.22}$$

Because we ensured that $y_1,\ldots,y_m$ are unit vectors in $\mathbf{X}$, we may use a theorem of Rudelson and Vershynin [274, Theorem 7.4] (an improved Talagrand-style two-parameter version of Elton's theorem; see Remark 103), to deduce from (2.22) that there are two numbers $0<s\leqslant 1$ and $\delta \lesssim t \leqslant 1$ that satisfy

$$t\sqrt{s} \gtrsim \frac{\delta}{\left(\log\left(\frac{2}{\delta}\right)\right)^{\frac32}}, \tag{2.23}$$

such that there exists a subset J of $\{1,\ldots,m\}$ whose cardinality satisfies

$$|J| \geqslant sm, \tag{2.24}$$

and moreover we have

$$\forall (a_j)_{j\in J}\in\mathbb{R}^J,\quad t\sum_{j\in J}|a_j| \lesssim \left\|\sum_{j\in J} a_j y_j\right\|_{\mathbf{X}} \leqslant \sum_{j\in J}|a_j|. \tag{2.25}$$

(2.25) means that the Banach–Mazur distance between $\mathrm{span}(\{y_j\}_{j\in J})$ and $\ell_1^{|J|}$ is $O(1/t)$. Hence,

$$\mathsf{c}_{\mathbf{X}}\left(\ell_1^{|J|}\right) \lesssim \frac{1}{t}. \tag{2.26}$$

Now, the justification of (2.15), and hence also the proof of Theorem 99, can be completed as follows:

$$\mathsf{LT}(\mathbf{X}) \overset{(2.6)}{\geqslant} \frac{\mathsf{LT}\big(\ell_1^{|J|}\big)}{\mathsf{c}_{\mathbf{X}}\big(\ell_1^{|J|}\big)} \overset{(2.2)\wedge(2.26)}{\gtrsim} t\sqrt{|J|} \overset{(2.24)}{\geqslant} t\sqrt{sm}$$
$$\overset{(2.23)}{\gtrsim} \frac{\delta\sqrt{m}}{\big(\log\big(\frac{2}{\delta}\big)\big)^{\frac{3}{2}}} \overset{(2.21)}{=} \frac{\gamma T_2(\mathbf{X})}{\Big(\log\Big(\frac{2\sqrt{m}}{\gamma T_2(\mathbf{X})}\Big)\Big)^{\frac{3}{2}}} \gtrsim \frac{T_2(\mathbf{X})}{(\log n)^{\frac{3}{2}}}, \tag{2.27}$$

where the final step of (2.27) holds because $T_2(\mathbf{X}) \geqslant 1$ and $\log m \lesssim \log n$ by (2.18). ■

Remark 103. In the proof of Theorem 99 we relied on [274, Theorem 7.4], which improves (in terms of the power of the logarithm in (2.23)) Talagrand's refinement [298] of Elton's theorem [92] (which is itself a major quantitative strengthening of an important theorem from [254]). Continuing with the notation of Theorem 99, Elton's theorem is a similar statement, except that the size of the subset J is a definite proportion of m that depends only on the parameter δ for which (2.22) holds, and also the parameter t for which (2.25) holds depends only on δ. The asymptotic dependence on δ in Elton's theorem [92] was improved by Pajor [245], a further improvement was obtained in [298], and the optimal dependence on δ was found by Mendelson and Vershynin in [213]. However, plugging this sharp dependence into our proof of Theorem 99 shows that the classical formulation of Elton's theorem is insufficient for our purposes. The two-parameter formulation of Elton's theorem that was introduced in [298] allows for the subset J to have any size through the parameter s in (2.24), but imposes a relation between s and t such as (2.23), thus making it possible for us to obtain Theorem 99.

The only reason why the logarithmic factor in (2.4) occurs is our use of a Talagrand-style two-parameter version of Elton's theorem, for which the currently best-known bound [274] is (2.23). Thus, if (2.23) could be improved to $t\sqrt{s} \gtrsim \delta$, i.e., if Question 104 below has a positive answer, then the conclusion (2.4) of Theorem 99 would become $\mathsf{d}_{\mathbf{X}} \lesssim \mathsf{LT}(\mathbf{X})^2$. This would improve Theorem 95 to match the bound of Theorem 94 which is currently known only for infinite dimensional Banach spaces. Moreover, since the resulting bound is independent of the dimension of $\mathbf{X}$, this would yield a new proof of the Lindenstrauss–Tzafriri solution of the complemented subspace problem; the infinite dimensional statement follows formally from its finite dimensional counterpart (e.g., [3, Theorem 12.1.6]), though all of the steps that led to Theorem 99 work for any reflexive Banach space. Question 104 is interesting in its own right regardless of the above application to the complemented subspace problem. In particular, a positive answer to Question 104 would resolve the question that Talagrand posed in the remark right after Corollary 1.2 in [298], though we warn that he characterises this in [298] as "certainly a rather formidable question."

Question 104. Fix $0 < \delta < 1$ and $n \in \mathbb{N}$. Let $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ be a Banach space and suppose that $x_1, \ldots, x_n \in \partial B_{\mathbf{X}}$ satisfy $\mathbb{E}[\|\sum_{i=1}^m \mathsf{g}_i x_i\|_{\mathbf{X}}] \geqslant \delta n$. Does this imply that there are two numbers $0 < s, t \leqslant 1$ satisfying $t\sqrt{s} \gtrsim \delta$ and a subset $J \subseteq \{1, \ldots, n\}$ with $|J| \geqslant sn$ such that $\|\sum_{j\in J} a_j x_j\|_{\mathbf{X}} \geqslant t \sum_{j \in J} |a_j|$ for every $a_1, \ldots, a_n \in \mathbb{R}$?

2.2 Proof of (1.105)

Because by [76] we know that $\mathsf{SEP}(\ell_1^n) \asymp n$ for every $n \in \mathbb{N}$, using bi-Lipschitz invariance we see that in order to prove (1.105) it suffices to show that for any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$,

$$\exists m \in \{1, \ldots, n\}, \qquad \frac{m}{\mathsf{c}_{\mathbf{X}}(\ell_1^m)} \geqslant T_2(\mathbf{X})^2. \tag{2.28}$$

We will prove (2.28) using Talagrand's two-parameter refinement of Elton's theorem [298] that we discussed in Remark 103 (it is worthwhile to note that the aforementioned improvements over [298] in [213, 274] do not yield a better bound in the ensuing reasoning. Also, the classical formulation of Elton's theorem is insufficient for our purposes, even if one incorporates the asymptotically sharp dependence on δ from [213]). Suppose that $k \in \mathbb{N}$ and $x_1, \ldots, x_k \in B_{\mathbf{X}}$. Let $\mathsf{g}_1, \ldots, \mathsf{g}_k$ be i.i.d. standard Gaussian random variables. Denote

$$E \stackrel{\mathrm{def}}{=} \mathbb{E}\Bigg[\bigg\|\sum_{j=1}^k \mathsf{g}_j x_j\bigg\|_{\mathbf{X}}\Bigg].$$

By [298, Corollary 1.2], there exist a universal constant $C \in [1, \infty)$ and a subset $S \subseteq \{1, \ldots, k\}$ satisfying

$$m \stackrel{\mathrm{def}}{=} |S| \geqslant \frac{E^2}{Ck},$$

and such that for every $(a_j)_{j\in S} \in \mathbb{R}^S$ we have

$$\frac{E}{\sqrt{Ckm}\big(\log\big(\frac{eCkm}{E^2}\big)\big)^C}\sum_{j\in S}|a_j| \leqslant \bigg\|\sum_{j\in S} a_j x_j\bigg\|_{\mathbf{X}} \leqslant \sum_{j\in S}|a_j|.$$

Consequently,

$$\mathsf{c}_{\mathbf{X}}\big(\ell_1^m\big) \leqslant \frac{\sqrt{Ckm}}{E}\left(\log\left(\frac{eCkm}{E^2}\right)\right)^C.$$

Therefore,

$$\frac{m}{\mathsf{c}_{\mathbf{X}}(\ell_1^m)} \geqslant \frac{E\sqrt{m}}{\sqrt{Ck}\big(\log\big(\frac{eCkm}{E^2}\big)\big)^C} \geqslant \frac{e^{C-\frac{1}{2}}}{2^C C^{C+1}}\cdot\frac{E^2}{k} \asymp \frac{E^2}{k},$$

where the last step uses the fact that the function $u \mapsto \sqrt{u}/(\log(eCku/E^2))^C$ attains its minimum on the ray $[E^2/(Ck), \infty)$ at $u = e^{2C-1}E^2/(Ck)$. It remains to choose $x_1, \dots, x_k$ so that $E^2/k \asymp T_2(\mathbf{X})^2$. This is possible because the equal norm type 2 constant of $\mathbf{X}$ equals $T_2(\mathbf{X})$, so there are $x_1, \dots, x_k \in \partial B_{\mathbf{X}}$ for which

$$T_2(\mathbf{X})\sqrt{k} \asymp \left(\mathbb{E}\left[\left\|\sum_{j=1}^{k} \mathsf{g}_j x_j\right\|_{\mathbf{X}}^2\right]\right)^{\frac{1}{2}} \asymp E,$$

where the last step uses Kahane's inequality. ∎

2.3 Hölder extension

In this section we will prove the lower bound on $\mathsf{e}^{\theta}(\ell_\infty^n)$ in (1.20) for every $n \in \mathbb{N}$ and $0 < \theta \leqslant 1$. It consists of two estimates, the first of which is

$$\mathsf{e}^{\theta}(\ell_\infty^n) \gtrsim n^{\frac{\theta}{2}+\theta^2-1}, \tag{2.29}$$

and the second of which is

$$\mathsf{e}^{\theta}(\ell_\infty^n) \gtrsim n^{\frac{\theta}{4}}. \tag{2.30}$$

We will justify (2.29) and (2.30) separately.

Note that (2.29) is vacuous if $\theta/2 + \theta^2 - 1 \leqslant 0$, i.e., if $0 < \theta \leqslant (\sqrt{17} - 1)/2$. The reason for this is that (2.29) is based on a reduction to the linear theory from [233] (extending the approach of [138] to the Hölder regime), that breaks down for functions which are too far from being Lipschitz. Specifically, for a Banach space $\mathbf{X}$ and a closed subspace $\mathbf{E}$ of $\mathbf{X}$, let $\lambda(\mathbf{E}; \mathbf{X})$ be the projection constant [122] of $\mathbf{E}$ relative to $\mathbf{X}$, i.e., it is the infimum over those $\lambda \in [1, \infty]$ for which there is a projection Proj from $\mathbf{X}$ onto $\mathbf{E}$ whose operator norm satisfies $\|\mathsf{Proj}\|_{\mathbf{X}\to\mathbf{E}} \leqslant \lambda$. Also, let $\mathsf{e}^{\theta}(\mathbf{X}; \mathbf{E})$ be the infimum over those $L \in [1, \infty]$ such that for every $\mathcal{C} \subseteq \mathbf{X}$ and every $f : \mathcal{C} \to \mathbf{E}$ that is θ-Hölder with constant 1, there is $F : \mathbf{X} \to \mathbf{E}$ that extend f and is θ-Hölder with constant L. With this notation, it was proved in [233] (see equation (106) there) that

$$\mathsf{e}^{\theta}(\mathbf{X}; \mathbf{E}) \gtrsim \frac{\lambda(\mathbf{E}; \mathbf{X})^{\theta}}{\dim(\mathbf{E})^{\frac{1-\theta}{2}} \dim(\mathbf{X})^{\theta(1-\theta)} \mathsf{c}_2(\mathbf{E})^{1-\theta}}. \tag{2.31}$$

Using the bounds $\dim(\mathbf{E}) \leqslant \dim(\mathbf{X})$ and $\mathsf{c}_2(\mathbf{E}) \leqslant \sqrt{\dim(\mathbf{E})}$ (John's theorem) in (2.31), we get that

$$\mathsf{e}^{\theta}(\mathbf{X}; \mathbf{E}) \gtrsim \frac{\lambda(\mathbf{E}; \mathbf{X})^{\theta}}{\dim(\mathbf{X})^{1-\theta^2}}. \tag{2.32}$$

By [290] there is a linear subspace $\mathbf{E}$ of ℓ_∞^n with $\lambda(\mathbf{E}; \ell_\infty^n) \asymp \sqrt{n}$, using which (2.32) implies (2.29).

Remark 105. In [233] it was deduced from (2.31) that

$$\mathsf{e}^{\theta}(\ell_1^n) \gtrsim n^{\theta^2-\frac{1}{2}}. \tag{2.33}$$

Specifically, by [153] there is a linear subspace $\mathbf{E}$ of ℓ_1^n with $\mathsf{c}_2(\mathbf{E}) \lesssim 1$ and $\dim(\mathbf{E}) = \lfloor n/2 \rfloor$; call such $\mathbf{E}$ a Kašin subspace of ℓ_1^n. By [275] we have $\lambda(\mathbf{E}; \ell_1^n) \asymp \sqrt{n}$, so (2.33) follows by substituting these parameters into (2.31). For $\mathbf{X} = \ell_\infty^n$, the poorly complemented subspace that we used above can be taken to be the orthogonal complement of any Kašin subspace of ℓ_1^n. Such a subspace of ℓ_∞^n has pathological properties [98]; in particular its Banach–Mazur distance to a Euclidean space is of order $\sqrt{n}$. So, a "vanilla" use of (2.31) leads at best to (2.29). However, we expect that it should be possible to improve (2.29) to

$$\mathsf{e}^{\theta}(\ell_\infty^n) \gtrsim n^{\theta^2-\frac{1}{2}}. \tag{2.34}$$

If (2.34) holds, then (1.20) improves to

$$\mathsf{e}^{\theta}(\ell_\infty^n) \gtrsim n^{\max\{\frac{\theta}{4}, \theta^2-\frac{1}{2}\}} = \begin{cases} n^{\frac{\theta}{4}} & \text{if } 0 \leqslant \theta \leqslant \frac{1+\sqrt{33}}{8}, \\ n^{\theta^2-\frac{1}{2}} & \text{if } \frac{1+\sqrt{33}}{8} \leqslant \theta \leqslant 1. \end{cases}$$

For (2.34), it would suffice to prove the following variant of Conjecture 7 for random subspaces of ℓ_∞^n. Let $\mathbf{E}$ be a subspace of $\mathbb{R}^n$ of dimension $m = \lfloor n/2 \rfloor$ that is chosen from the Haar measure on the Grassmannian. We conjecture that there is a universal constant $D \geqslant 1$ such that with high probability there is an origin-symmetric convex body $L \subseteq B_{\mathbf{E}}$ that satisfies $\mathrm{MaxProj}(L)/\operatorname{vol}_m(L) \lesssim 1$. If this indeed holds, then by using it in the *proof of* (2.31) in [233] we can deduce (2.34) (specifically, replace in [233, Lemma 20] the averaging over $B_{\ell_2^m}$ by averaging over L; we omit the details of this adaptation of [233]).

Proof of (2.30). Fix $k, m \in \mathbb{N}$ satisfying $k \leqslant 2m \leqslant n/2$ whose value will be specified later so as to optimize the ensuing reasoning (see (2.48) below). Denote $\ell = \lfloor (4m/k) \rfloor$ and define $\mathcal{C} = \mathcal{C}(k, m, n) \subseteq \ell_\infty^n(\mathbb{C})$ by

$$\mathcal{C} \stackrel{\mathrm{def}}{=} \{E_m(ks) : s \in \{1, \ldots, \ell\}^n\},$$

where for every $s = (s_1, \ldots, s_n) \in \mathbb{R}^n$ we define $E_m(s) \in \mathbb{C}^n$ by

$$E_m(s) \stackrel{\mathrm{def}}{=} \sum_{j=1}^{n} e^{\frac{\pi i}{2m} s_j} e_j.$$

Denote the standard basis (delta masses) of $\mathbb{R}^{\mathcal{C}}$ by $\{\boldsymbol{\delta}_s\}_{s \in \mathcal{C}}$. Let $\mathbb{R}_0^{\mathcal{C}}$ be the hyperplane of $\mathbb{R}^{\mathcal{C}}$ consisting of those $(a_s)_{s \in \mathcal{C}} = \sum_{s \in \mathcal{C}} a_s \boldsymbol{\delta}_s$ with $\sum_{s \in \mathcal{C}} a_s = 0$. Suppose that $\mathbf{X}_\theta = (\mathbb{R}_0^{\mathcal{C}}, \|\cdot\|_{\mathbf{X}_\theta})$ is a normed space that satisfies

$$\forall x, y \in \mathcal{C}, \quad \|\boldsymbol{\delta}_x - \boldsymbol{\delta}_y\|_{\mathbf{X}_\theta} = \|x - y\|_{\ell_\infty^n(\mathbb{C})}^{\theta} \tag{2.35}$$

and

$$\forall \mu \in \mathbb{R}_0^{\mathcal{C}}, \quad \left(\frac{k}{m}\right)^\theta \|\mu\|_{\ell_1(\mathcal{C})} \lesssim \|\mu\|_{\mathbf{X}_\theta} \lesssim \|\mu\|_{\ell_1(\mathcal{C})}. \tag{2.36}$$

For this, $\mathbf{X}_\theta$ can be taken to be the normed space whose unit ball is

$$B_{\mathbf{X}_\theta} = \mathrm{conv}\left\{\frac{1}{\|x-y\|^\theta_{\ell_\infty^n(\mathbb{C})}}(\boldsymbol{\delta}_x - \boldsymbol{\delta}_y) : x, y \in \mathcal{C},\ x \neq y\right\} \subseteq \mathbb{R}_0^{\mathcal{C}}, \tag{2.37}$$

which is the maximal norm on $\mathbb{R}_0^{\mathcal{C}}$ satisfying (2.35). To check that (2.36) holds for the choice (2.37), note that, as $1 \leqslant k \leqslant 2m$, distinct $x, y \in \mathcal{C}$ satisfy

$$\frac{k}{m} \lesssim \|x-y\|_{\ell_\infty^n(\mathbb{C})} \lesssim 1.$$

It is simple to deduce (2.36) from this, as done in [233, Lemma 7]. The choice (2.37) makes $\mathbf{X}_\theta$ be the Wasserstein-1 space over $(\mathcal{C}, d_\theta)$, where d_θ is the θ-snowflake of the $\ell_\infty^n(\mathbb{C})$ metric, i.e., $d_\theta(x,y) = \|x-y\|^\theta_{\ell_\infty^n(\mathbb{C})}$ for $x, y \in \ell_\infty^n(\mathbb{C})$; see Section 5.1.

By virtue of (2.35), if we define $f : \mathcal{C} \to \mathbf{X}_\theta$ by setting

$$\forall x \in \mathcal{C}, \quad f(x) \stackrel{\mathrm{def}}{=} \boldsymbol{\delta}_x - \frac{1}{|\mathcal{C}|}\sum_{y\in\mathcal{C}} \boldsymbol{\delta}_y,$$

then f is θ-Hölder with constant 1. We claim that if $m \geqslant \pi\sqrt{n}$, then by (2.35) every $F : \ell_\infty^n(\mathbb{C}) \to \mathbf{X}_\theta$ satisfies

$$\begin{aligned} &\frac{1}{(4m)^n}\sum_{j=1}^n \sum_{s\in\{1,\ldots,4m\}^n} \left\|F\big(E_m(s+2me_j)\big) - F\big(E_m(s)\big)\right\|_{\mathbf{X}_\theta} \\ &\quad\lesssim \frac{m^{2+\theta}}{k^\theta(12m)^n} \sum_{\varepsilon\in\{-1,0,1\}^n}\ \sum_{s\in\{1,\ldots,4m\}^n} \left\|F\big(E_m(s+\varepsilon)\big) - F\big(E_m(s)\big)\right\|_{\mathbf{X}_\theta}. \end{aligned} \tag{2.38}$$

Indeed, (2.38) follows from a substitution of (2.35) into the following inequality from [209, Remark 7.5]:

$$\begin{aligned} &\frac{1}{(4m)^n}\sum_{j=1}^n \sum_{s\in\{1,\ldots,4m\}^n} \left\|F\big(E_m(s+2me_j)\big) - F\big(E_m(s)\big)\right\|_{\ell_1(\mathcal{C})} \\ &\quad\lesssim \frac{m^{2}}{(12m)^n} \sum_{\varepsilon\in\{-1,0,1\}^n}\ \sum_{s\in\{1,\ldots,4m\}^n} \left\|F\big(E_m(s+\varepsilon)\big) - F\big(E_m(s)\big)\right\|_{\ell_1(\mathcal{C})}. \end{aligned}$$

Suppose that $F : \{1,\ldots,4m\}^n \to \mathbf{X}_\theta$ is θ-Hölder with constant $L \geqslant 1$ on the set $(\{1,\ldots,4m\}^n, \|\cdot\|_{\ell_\infty^n(\mathbb{C})})$, i.e.,

$$x, y \in \{1,\ldots,4m\}^n, \quad \|F(x) - F(y)\|_{\mathbf{X}_\theta} \leqslant L\|x-y\|^\theta_{\ell_\infty^n(\mathbb{C})}.$$

Then, each of the summands that appear in the right-hand side of (2.38) is at most $2L/m^\theta$. Consequently,

$$\frac{1}{n(4m)^n}\sum_{j=1}^{n}\sum_{s\in\{1,\ldots,4m\}^n}\big\|F\big(E_m(s+2me_j)\big)-F\big(E_m(s)\big)\big\|_{\mathbf{X}_\theta}\lesssim\frac{Lm^2}{k^\theta n}. \tag{2.39}$$

If F also extends f, then $F(E_m(s)) = f(E_m(s'))$ for every $s \in \mathbb{N}^n$, where we use the notation $s' = (s'_1,\ldots,s'_n)$ and for each $u \in \mathbb{N}$ we let u' be an element α of $\{k, 2k,\ldots,\ell k\}$ for which $|\alpha - u \mod (4m)|$ is minimized, so that $s' \in \mathcal{C}$ and

$$\forall s\in\mathbb{N}^n,\quad \|E_m(s)-E_m(s')\|_{\ell_\infty^n(\mathbb{C})}\lesssim\frac{k}{m}. \tag{2.40}$$

Hence, for any $j \in \{1,\ldots,n\}$ and $s \in \{1,\ldots,4m\}^n$ we have

$$\begin{aligned}
2^\theta &= \big\|-2e^{\frac{\pi i}{2m}s_j}e_j\big\|_{\ell_\infty^n(\mathbb{C})}^\theta\\
&= \|E_m(s+2me_j)-E_m(s)\|_{\ell_\infty^n(\mathbb{C})}^\theta \qquad (2.41)\\
&\leqslant \|E_m((s+2me_j)')-E_m(s')\|_{\ell_\infty^n(\mathbb{C})}^\theta\\
&\quad+\|E_m((s+2me_j)')-E_m(s+2me_j)\|_{\ell_\infty^n(\mathbb{C})}^\theta\\
&\quad+\|E_m(s')-E_m(s)\|_{\ell_\infty^n(\mathbb{C})}^\theta\\
&\leqslant \|E_m((s+2me_j)')-E_m(s')\|_{\ell_\infty^n(\mathbb{C})}^\theta+\frac{2k^\theta}{m^\theta} \qquad (2.42)\\
&= \big\|\boldsymbol{\delta}_{E_m((s+2me_j)')}-\boldsymbol{\delta}_{E_m(s')}\big\|_{\mathbf{X}_\theta}+\frac{2k^\theta}{m^\theta} \qquad (2.43)\\
&= \big\|f\big(E_m((s+2me_j)')\big)-f\big(E_m(s')\big)\big\|_{\mathbf{X}_\theta}+\frac{2k^\theta}{m^\theta} \qquad (2.44)\\
&= \big\|F\big(E_m((s+2me_j)')\big)-F\big(E_m(s')\big)\big\|_{\mathbf{X}_\theta}+\frac{2k^\theta}{m^\theta} \qquad (2.45)\\
&\leqslant \big\|F\big(E_m(s+2me_j)\big)-F\big(E_m(s)\big)\big\|_{\mathbf{X}_\theta}\\
&\quad+\big\|F\big(E_m((s+2me_j)')\big)-F\big(E_m(s+2me_j)\big)\big\|_{\mathbf{X}_\theta}\\
&\quad+\big\|F\big(E_m(s')\big)-F\big(E_m(s)\big)\big\|_{\mathbf{X}_\theta}+\frac{2k^\theta}{m^\theta}\\
&\leqslant \big\|F\big(E_m(s+2me_j)\big)-F\big(E_m(s)\big)\big\|_{\mathbf{X}_\theta}\\
&\quad+L\|E_m((s+2me_j)')-E_m(s+2me_j)\|_{\ell_\infty^n(\mathbb{C})}^\theta\\
&\quad+L\|E_m(s')-E_m(s)\|_{\ell_\infty^n(\mathbb{C})}^\theta+\frac{2k^\theta}{m^\theta} \qquad (2.46)\\
&\leqslant \big\|F\big(E_m(s+2me_j)\big)-F\big(E_m(s)\big)\big\|_{\mathbf{X}_\theta}+\frac{2(L+1)k^\theta}{m^\theta}, \qquad (2.47)
\end{aligned}$$

where for (2.41) recall the definition of E_m, in (2.42) and (2.47) we used (2.40), in (2.43) we used (2.35), for (2.44) recall the definition of f, in (2.45) we used the fact that F extends f and $\{(s+2me_j)', s'\} \subseteq \mathcal{C}$, and in (2.46) we used the fact that F is θ-Hölder with constant L. By averaging this inequality over (j,s) chosen uniformly at random from $\{1,\ldots,n\} \times \{1,\ldots,4m\}^n$ and applying (2.39), we conclude that

$$1 \lesssim \left(\frac{m^2}{k^\theta n} + \frac{k^\theta}{m^\theta}\right) L. \tag{2.48}$$

This holds whenever $k, m \in \mathbb{N}$ satisfy $k \leqslant 2m \leqslant n/2$ and $m \geqslant \pi\sqrt{n}$, so choose $m \asymp \sqrt{n}$ and $k \asymp \sqrt[4]{n}$ to minimize (up to constants) the right-hand side of (2.48) and deduce the desired lower bound $L \gtrsim n^{\theta/4}$. ■

By [210, Lemma 6.5], for every $\theta \in (0,1]$ and $n \in \mathbb{N}$ we have

$$\mathsf{e}^\theta(\ell_2^n) \gtrsim n^{\frac{\theta}{4}}. \tag{2.49}$$

In combination with (2.30) and [5], this implies that there is a universal constant $c > 0$ such that

$$\mathsf{e}^\theta(\mathbf{X}) \geqslant e^{c\theta\sqrt{\log n}} \tag{2.50}$$

for every n-dimensional normed space $\mathbf{X}$ and every $\theta \in (0,1]$.

Conjecture 106. For any $\theta \in (0,1]$ there is $c(\theta) > 0$ such that $\mathsf{e}^\theta(\mathbf{X}) \geqslant \dim(\mathbf{X})^{c(\theta)}$ for every normed space $\mathbf{X}$.

Conjecture 106 has a positive answer when the Hölder exponent is close enough to 1. Specifically, if

$$0.9307777\ldots = \frac{\sqrt{193}+1}{16} < \theta \leqslant 1, \tag{2.51}$$

then

$$\mathsf{e}^\theta(\mathbf{X}) \gtrsim \frac{n^{\frac{\theta(8\theta^2-\theta-6)}{20\theta-8}}}{(\log n)^{\frac{3\theta^2}{5\theta-2}}}. \tag{2.52}$$

Indeed, by bi-Lipschitz invariance, (2.49) implies the following generalization of Theorem 97:

$$\mathsf{e}^\theta(\mathbf{X}) \gtrsim \frac{n^{\frac{\theta}{4}}}{\mathsf{d}_{\mathbf{X}}^\theta}.$$

Also,

$$\mathsf{e}^\theta(\mathbf{X}) \overset{(2.31)}{\gtrsim} \frac{\mathsf{LT}(\mathbf{X})^\theta}{n^{(1-\theta)(\theta+\frac{1}{2})}\mathsf{d}_{\mathbf{X}}^{1-\theta}} \overset{(2.4)}{\gtrsim} \frac{\mathsf{d}_{\mathbf{X}}^{\frac{\theta}{2}}/(\log n)^{\frac{3\theta}{2}}}{n^{(1-\theta)(\theta+\frac{1}{2})}\mathsf{d}_{\mathbf{X}}^{1-\theta}} = \frac{\mathsf{d}_{\mathbf{X}}^{\frac{3\theta}{2}-1}}{n^{(1-\theta)(\theta+\frac{1}{2})}(\log n)^{\frac{3\theta}{2}}}.$$

Therefore, in analogy to (2.5) we see that

$$\mathsf{e}^{\theta}(\mathbf{X}) \gtrsim \max\left\{\frac{n^{\frac{\theta}{4}}}{\mathsf{d}_{\mathbf{X}}^{\theta}}, \frac{\mathsf{d}_{\mathbf{X}}^{\frac{3\theta}{2}-1}}{n^{(1-\theta)(\theta+\frac{1}{2})}(\log n)^{\frac{3\theta}{2}}}\right\}. \tag{2.53}$$

Elementary calculus shows that (2.53) implies (2.52) in the range (2.51). If θ does not satisfy (2.51), then (2.53) does not imply a lower bound $\mathsf{e}^{\theta}(\mathbf{X})$ that depends only on n and grows to ∞ with n; for such θ the best lower bound that we know is (2.50). The application of (2.38) in the above proof of (2.30) can be mimicked using other bi-Lipschitz invariants to prove Conjecture 106 for various normed spaces, such as $\ell_2^n(\ell_1^n)$ or S_1^n, using [237] and [235], respectively. We do not know if Conjecture 106 holds even when, say, $\mathbf{X} = \ell_1^n$.

2.4 Justification of (1.25)

In the range $p \in [1, 4/3] \cup \{2\} \cup [3, \infty]$ the bound in (1.25) is a combination of [64, Corollary 8.12] and [210, Theorem 1.17]. We only need to justify (1.25) in the range $p \in (4/3, 3) \smallsetminus \{2\}$ because it was not previously stated in the literature. Suppose first that $p \in (4/3, 2)$. By [99], there is $k \in \{1, \ldots, n\}$ with $k \asymp n$ such that $\mathsf{c}_{\ell_p^n}(\ell_2^k) \asymp 1$. Hence,

$$\mathsf{e}(\ell_p^n) \gtrsim \mathsf{e}(\ell_2^k) \gtrsim \sqrt[4]{k} \asymp \sqrt[4]{n},$$

where the penultimate inequality follows from [210, Theorem 1.17]. Analogously, if $q \in (2, 3)$, then by [99] there is $m \in \{1, \ldots, n\}$ with $m \asymp n^{2/q}$ such that $\mathsf{c}_{\ell_q^n}(\ell_2^m) \asymp 1$. We therefore have

$$\mathsf{e}(\ell_q^n) \gtrsim \mathsf{e}(\ell_2^m) \gtrsim \sqrt[4]{m} \asymp n^{\frac{1}{2q}}.$$

2.5 Proof of the lower bound on SEP(X) in Theorem 3

Thanks to (1.71), the first part of Theorem 107 below coincides with the lower bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3, except that in (2.54) below we also specify the constant factor that our proof provides (there is no reason to expect that this constant is optimal; due to the fundamental nature of this randomized clustering problem it would be interesting to find the optimal constant here). The second part of Theorem 107 relates to dimension reduction by controlling the cardinality of a finite subset $\mathcal{C}$ of $\mathbf{X}$ on which the lower bound is attained. We conjecture that the first part of (2.55) below could be improved to $|\mathcal{C}|^{1/n} = O(1)$; an inspection of the ensuing proof suggests that a possible route towards this improved bound is to incorporate a proportional Dvoretzky–Rogers factorization [51, 106, 297] in place of our use of the "vanilla" Dvoretzky–Rogers lemma [91].

Theorem 107. *For every $n \in \mathbb{N}$, any n-dimensional normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ satisfies*

$$\mathsf{SEP}(\mathbf{X}) \geqslant \mathrm{evr}(\mathbf{X}) \frac{2(n!)^{\frac{1}{2n}}\Gamma\left(1+\frac{n}{2}\right)^{\frac{1}{n}}}{\sqrt{\pi n}} = \frac{\sqrt{2}+o(1)}{e\sqrt{\pi}}\mathrm{evr}(\mathbf{X})\sqrt{n}. \tag{2.54}$$

Furthermore, there exists a finite subset $\mathcal{C}$ of $\mathbf{X}$ satisfying

$$|\mathcal{C}|^{\frac{1}{n}} \lesssim \frac{\sqrt{n}}{\mathrm{evr}(\mathbf{X})} \quad \textit{and} \quad \mathsf{SEP}(\mathcal{C}_{\mathbf{X}}) \gtrsim \mathrm{evr}(\mathbf{X})\sqrt{n}. \tag{2.55}$$

Our proof of Theorem 107 builds upon the strategy that was used in [76] to treat ℓ_1^n. A combinatorial fact on which it relies is Lemma 108 below, which is implicit in the proof of [76, Lemma 3.1]. After proving Theorem 107 while using Lemma 108, we will present a proof of Lemma 108 which is a quick application of the Loomis–Whitney inequality [185]; the proof in [76] uses a result of [4] which is proved in [4] via information-theoretic reasoning through the use of Shearer's inequality [80]; the relation between the Loomis–Whitney inequality and Shearer's inequality is well known (see, e.g., [64]), so our proof of Lemma 108 is in essence a repackaging of the classical ideas.

Lemma 108. *Fix $n, M \in \mathbb{N}$ and a nonempty finite subset Ω of $\mathbb{Z}^n$. Suppose that $\mathcal{P}$ is a random partition of Ω that is supported on partitions into subsets of cardinality at most M, i.e.,*

$$\mathbf{Prob}\Big[\max_{\Gamma\in\mathcal{P}} |\Gamma| \leqslant M\Big] = 1.$$

Then, there exists $i \in \{1, \ldots, n\}$ and $x \in \Omega \cap (\Omega - e_i)$ for which

$$\mathbf{Prob}\big[\mathcal{P}(x) \neq \mathcal{P}(x+e_i)\big] \geqslant \frac{1}{\sqrt[n]{M}} - \frac{1}{n}\sum_{i=1}^{n}\frac{|\Omega \smallsetminus (\Omega - e_i)|}{|\Omega|}. \tag{2.56}$$

Proof of Theorem 107 assuming Lemma 108. By suitably choosing the identification of $\mathbf{X}$ with $\mathbb{R}^n$, we may assume without loss of generality that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $B_{\ell_2^n}$ is the Löwner ellipsoid of $B_{\mathbf{X}}$. Then,

$$\mathrm{evr}(\mathbf{X}) = \left(\frac{\mathrm{vol}_n(B_{\ell_2^n})}{\mathrm{vol}_n(B_{\mathbf{X}})}\right)^{\frac{1}{n}} = \frac{\sqrt{\pi}}{\Gamma\left(1+\frac{n}{2}\right)^{\frac{1}{n}}\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}}. \tag{2.57}$$

By the Dvoretzky–Rogers lemma [91], there exist contact points

$$x_1, \ldots, x_n \in S^{n-1} \cap \partial B_{\mathbf{X}}$$

that satisfy

$$\forall k \in \{1, \ldots, n\}, \qquad \left\|\mathsf{Proj}_{\mathrm{span}(x_1,\ldots,x_{k-1})^\perp}(x_k)\right\|_{\ell_2^n} \geqslant \sqrt{\frac{n-k+1}{n}}. \tag{2.58}$$

Let $\Lambda = \Lambda(x_1, \ldots, x_n) \subseteq \mathbb{R}^n$ denote the lattice that is generated by $x_1, \ldots, x_n$, namely

$$\Lambda = \sum_{i=1}^{n} \mathbb{Z} x_i = \left\{ \sum_{i=1}^{n} k_i x_i : k_1, \ldots, k_n \in \mathbb{Z} \right\}.$$

By (2.58), Λ is a full-rank lattice. Denote the fundamental parallelepiped of Λ by $Q = Q(x_1, \ldots, x_n)$, i.e.,

$$Q = \sum_{i=1}^{n} [0,1) x_i = \left\{ \sum_{i=1}^{n} s_i x_i : 0 \leqslant s_1, \ldots, s_n < 1 \right\}.$$

Since $x_1, \ldots, x_n \in B_{\mathbf{X}}$, we have $Q - Q \subseteq n B_{\mathbf{X}}$ and by (2.58) the volume of Q (the determinant of Λ) satisfies

$$\begin{aligned} \det(\Lambda) = \mathrm{vol}_n(Q) &= \prod_{k=1}^{n} \left\| \mathsf{Proj}_{\mathrm{span}\{x_1, \ldots, x_{k-1}\}^{\perp}}(x_k) \right\|_{\ell_2^n} \\ &\overset{(2.58)}{\geqslant} \prod_{k=1}^{n} \sqrt{\frac{n-k+1}{n}} = \frac{\sqrt{n!}}{n^{\frac{n}{2}}}. \end{aligned} \tag{2.59}$$

Fix $m \in \mathbb{N}$ and $\sigma, \Delta > 0$. Denote

$$\begin{aligned} \mathcal{C}_m = \mathcal{C}_m(x_1, \ldots, x_n) &= \Lambda \cap (mQ) \\ &= \left\{ \sum_{i=1}^{n} k_i x_i : k_1, \ldots, k_n \in \{0, \ldots, m-1\} \right\}, \end{aligned}$$

and suppose that $\mathcal{P}$ is σ-separating Δ-bounded random partition of $\mathcal{C}_m$. The fact that $\mathcal{P}$ is Δ-bounded means that $\Gamma - \Gamma \subseteq \Delta B_{\mathbf{X}}$ for every $\Gamma \subseteq \mathcal{C}_m$ with $\mathbf{Prob}[\Gamma \in \mathcal{P}] > 0$. Recalling that $Q - Q \subseteq n B_{\mathbf{X}}$, this implies that

$$B_{\mathbf{X}} \supseteq \frac{1}{\Delta + n} \big((\Gamma + Q) - (\Gamma + Q) \big). \tag{2.60}$$

Now,

$$\begin{aligned} \frac{\sqrt{\pi}}{\Gamma\left(1 + \frac{n}{2}\right)^{\frac{1}{n}} \mathrm{evr}(\mathbf{X})} &= \mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}} \geqslant \frac{2}{\Delta + n} \mathrm{vol}_n(\Gamma + Q)^{\frac{1}{n}} \\ &= \frac{2}{\Delta + n} \big(|\Gamma| \, \mathrm{vol}_n(Q) \big)^{\frac{1}{n}} \geqslant \frac{2(n!)^{\frac{1}{2n}}}{(\Delta + n)\sqrt{n}} |\Gamma|^{\frac{1}{n}}, \end{aligned} \tag{2.61}$$

where the first step of (2.61) is (2.57), the second step of (2.61) uses (2.60) and the Brunn–Minkowski inequality, the third step of (2.61) holds because the parallelepipeds $\{\gamma + Q : \gamma \in \Gamma\}$ are disjoint, and the final step of (2.61) is (2.59). If

$T \in \mathsf{GL}_n(\mathbb{R})$ is given by $Te_i = x_i$, then by (2.61) the random partition

$$T^{-1}\mathcal{P} \stackrel{\mathrm{def}}{=} \{T^{-1}\Gamma : \Gamma \in \mathcal{P}\}$$

of $T^{-1}\mathcal{C}_m = \{0, \dots, m-1\}^n$ satisfies the assumptions of Lemma 108 with

$$M = \frac{(\pi n)^{\frac{n}{2}}(\Delta + n)^n}{2^n \Gamma\big(1 + \frac{n}{2}\big)\sqrt{n!}} \cdot \frac{1}{\mathrm{evr}(\mathbf{X})^n}.$$

If we choose $\Omega = \{0, \dots, m-1\}^n = T^{-1}\mathcal{C}_m$ in Lemma 108, then we have $|\Omega| = m^n$ and

$$|\Omega \smallsetminus (\Omega - e_i)| = m^{n-1}$$

for every $i \in \{1, \dots, n\}$, so it follows from Lemma 108 that there exist $i \in \{1, \dots, n\}$ and $x \in \mathcal{C}_m$ such that

$$\mathbf{Prob}\big[\mathcal{P}(x) \neq \mathcal{P}(x + e_i)\big] \geqslant \mathrm{evr}(\mathbf{X}) \frac{2(n!)^{\frac{1}{2n}} \Gamma\big(1 + \frac{n}{2}\big)^{\frac{1}{n}}}{(\Delta + n)\sqrt{\pi n}} - \frac{1}{m}. \tag{2.62}$$

At the same time, the left-hand side of (2.62) is at most σ/Δ, since $\mathcal{P}$ is σ-separating and $\|x_i\|_{\mathbf{X}} \leqslant 1$. Thus,

$$\sigma \geqslant \mathrm{evr}(\mathbf{X}) \frac{2\Delta(n!)^{\frac{1}{2n}} \Gamma(1 + \frac{n}{2})^{\frac{1}{n}}}{(\Delta + n)\sqrt{\pi n}} - \frac{\Delta}{m}. \tag{2.63}$$

By letting $m \to \infty$ in (2.63) and then letting $\Delta \to \infty$ in the resulting estimate, we get (2.54). Also, if we set $\Delta = n$ in (2.63), then for sufficiently large $m \asymp \sqrt{n}/\,\mathrm{evr}(\mathbf{X})$ we have

$$\mathsf{SEP}(\mathcal{C}_m) \gtrsim \mathrm{evr}(\mathbf{X})\sqrt{n},$$

giving (2.55). ■

We will next provide a proof of Lemma 108 whose main ingredient is the following lemma.

Lemma 109 (Application of Loomis–Whitney). *Fix an integer $n \geqslant 2$ and a finite subset Γ of $\mathbb{Z}^n$. For $x \in \mathbb{Z}^n$ and $i \in \{1, \dots, n\}$, let $d_i(x;\Gamma) \in \mathbb{N} \cup \{0\}$ be the number of times that the oriented discrete axis-parallel line $x + \mathbb{Z}e_i$ transitions from Γ to $\mathbb{Z}^n \smallsetminus \Gamma$, and let $\mathsf{g}(x;\Gamma)$ be the geometric mean of $d_1(x;\Gamma), \dots, d_n(x;\Gamma)$. Thus*

$$\forall i \in \{1, \dots, n\}, \quad d_i(x;\Gamma) \stackrel{\mathrm{def}}{=} \big|\{k \in \mathbb{Z} : x + ke_i \in \Gamma \wedge x + (k+1)e_i \notin \Gamma\}\big|$$

and

$$g(x;\Gamma) \stackrel{\mathrm{def}}{=} \sqrt[n]{d_1(x;\Gamma) \cdots d_n(x;\Gamma)}.$$

Then,

$$\frac{1}{n}\sum_{i=1}^{n} |\Gamma \smallsetminus (\Gamma - e_i)| \geqslant \Big(\sum_{x \in \mathbb{Z}^n} g(x;\Gamma)^{\frac{n}{n-1}}\Big)^{\frac{n-1}{n}} \geqslant |\Gamma|^{\frac{n-1}{n}}. \tag{2.64}$$

Proof. The second inequality in (2.64) holds because $d_1(x;\Gamma),\ldots,d_n(x;\Gamma)\geqslant 1$ for every $x\in\Gamma$ (as $|\Gamma|<\infty$), and hence $g(\cdot;\Gamma)\geqslant \mathbf{1}_\Gamma(\cdot)$ point-wise. For the first inequality in (2.64), observe that for each $i\in\{1,\ldots,n\}$,

$$
\begin{aligned}
|\Gamma\smallsetminus(\Gamma-e_i)| &= \sum_{x\in\mathbb{Z}^n}\mathbf{1}_\Gamma(x)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}(x+e_i)\\
&= \sum_{y\in \mathsf{Proj}_{e_i^\perp}\Gamma}\Big(\sum_{k\in\mathbb{Z}}\mathbf{1}_\Gamma(y+ke_i)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}\big(y+(k+1)e_i\big)\Big)\\
&= \sum_{y\in \mathsf{Proj}_{e_i^\perp}\mathbb{Z}^n} d_i(y;\Gamma).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^n|\Gamma\smallsetminus(\Gamma-e_i)| &= \frac{1}{n}\sum_{i=1}^n\big\|d_i(\cdot;\Gamma)^{\frac{1}{n-1}}\big\|^{n-1}_{\ell_{n-1}(\mathsf{Proj}_{e_i^\perp}\mathbb{Z}^n)}\\
&\geqslant \prod_{i=1}^n\big\|d_i(\cdot;\Gamma)^{\frac{1}{n-1}}\big\|^{\frac{n-1}{n}}_{\ell_{n-1}(\mathsf{Proj}_{e_i^\perp}\mathbb{Z}^n)}\\
&\geqslant \sum_{x\in\mathbb{Z}^n}\prod_{i=1}^n d_i(\mathsf{Proj}_{e_i^\perp}x)^{\frac{1}{n-1}},
\end{aligned}
$$

where the second step is an application of the arithmetic-mean/geometric-mean inequality and the final step is an application of the Loomis–Whitney inequality [185] (see [288, Theorem 3] for the functional version of the Loomis–Whitney inequality that they are using here); we note that even though this inequality is commonly stated for functions on $\mathbb{R}^n$ rather than for functions on $\mathbb{Z}^n$, its proof for functions on $\mathbb{Z}^n$ is identical (in fact, [185] proves the continuous inequality by first proving its discrete counterpart). ∎

Note that when $n=1$ Lemma 109 holds trivially if we interpret (2.64) as the estimate $|\Gamma\smallsetminus(\Gamma-1)|\geqslant\max_{x\in\mathbb{Z}}g(x;\Gamma)\geqslant 1$, since in this case

$$g(x;\Gamma)=|\Gamma\smallsetminus(\Gamma-1)|$$

for every $x\in\mathbb{Z}$.

The following corollary of Lemma 109 is a deterministic counterpart of Lemma 108.

Corollary 110. *Fix $n,M\in\mathbb{N}$ and a nonempty finite subset Ω of $\mathbb{Z}^n$. Suppose that $\mathcal{P}$ is a partition of Ω with*

$$\max_{\Gamma\in\mathcal{P}}|\Gamma|\leqslant M. \tag{2.65}$$

Then,

$$\frac{1}{n}\sum_{i=1}^{n}|\{x\in\Omega\cap(\Omega-e_i):\mathcal{P}(x)\neq\mathcal{P}(x+e_i)\}|$$
$$\geqslant \frac{|\Omega|}{\sqrt[n]{M}}-\frac{1}{n}\sum_{i=1}^{n}|\Omega\smallsetminus(\Omega-e_i)|. \tag{2.66}$$

Proof. Observe that for each fixed $i\in\{1,\ldots,n\}$ we have

$$\begin{aligned}
&|\Omega\smallsetminus(\Omega-e_i)|+\sum_{x\in\Omega\cap(\Omega-e_i)}\mathbf{1}_{\mathcal{P}(x)\neq\mathcal{P}(x+e_j)}\\
&=|\Omega\smallsetminus(\Omega-e_i)|+\sum_{x\in\Omega\cap(\Omega-e_i)}\Big(\sum_{\Gamma\in\mathcal{P}}\mathbf{1}_\Gamma(x)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}(x+e_i)\Big)\\
&=\sum_{x\in\mathbb{Z}^n}\sum_{\Gamma\in\mathcal{P}}\mathbf{1}_\Gamma(x)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}(x+e_i)\\
&=\sum_{\Gamma\in\mathcal{P}}|\Gamma\smallsetminus(\Gamma-e_i)|,
\end{aligned}\tag{2.67}$$

where the first step of (2.67) holds because $\mathcal{P}$ is a partition of Ω and the second step of (2.67) holds because

$$\mathbf{1}_\Gamma(x)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}(x+e_i)=0$$

for every $\Gamma\subseteq\Omega$ if $x\in\mathbb{Z}^n\smallsetminus\Omega$, and if $x\in\Omega\smallsetminus(\Omega-e_i)$, then

$$\mathbf{1}_\Gamma(x)\mathbf{1}_{\mathbb{Z}^n\smallsetminus\Gamma}(x+e_i)=1$$

for exactly one $\Gamma\in\mathcal{P}$ (specifically, this is satisfied only for $\Gamma=\mathcal{P}(x)$ because we have $x+e_i\in\mathbb{Z}^n\smallsetminus\Omega\subseteq\mathbb{Z}^n\smallsetminus\mathcal{P}(x)$). Now,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}|\{x\in\Omega\cap(\Omega-e_i):\mathcal{P}(x)\neq\mathcal{P}(x+e_i)\}|+\frac{1}{n}\sum_{i=1}^{n}|\Omega\smallsetminus(\Omega-e_i)|\\
&\overset{(2.67)}{=}\sum_{\Gamma\in\mathcal{P}}\frac{1}{n}\sum_{i=1}^{n}|\Gamma\smallsetminus(\Gamma-e_i)|\\
&\overset{(2.64)}{\geqslant}\sum_{\Gamma\in\mathcal{P}}|\Gamma|^{\frac{n-1}{n}}\overset{(2.65)}{\geqslant}\frac{1}{\sqrt[n]{M}}\sum_{\Gamma\in\mathcal{P}}|\Gamma|=\frac{|\Omega|}{\sqrt[n]{M}},
\end{aligned}$$

where the last step holds because $\mathcal{P}$ is a partition of Ω. ∎

Proof of Lemma 108. Denoting

$$p=\max_{i\in\{1,\ldots,n\}}\max_{x\in\Omega\cap(\Omega-e_i)}\mathbf{Prob}[\mathcal{P}(x)\neq\mathcal{P}(x+e_i)],$$

the goal is to show that p is at least the right-hand side of (2.56). This follows from Corollary 110 because

$$\begin{aligned}
p|\Omega| &\geqslant \frac{p}{n}\sum_{i=1}^{n}|\Omega\cap(\Omega-e_i)|\\
&\geqslant \frac{1}{n}\sum_{i=1}^{n}\sum_{x\in\Omega\cap(\Omega-e_i)}\mathbf{Prob}\big[\mathcal{P}(x)\neq\mathcal{P}(x+e_i)\big]\\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{x\in\Omega\cap(\Omega-e_i)}\mathbb{E}\big[\mathbf{1}_{\mathcal{P}(x)\neq\mathcal{P}(x+e_i)}\big]\\
&= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}|\{x\in\Omega\cap(\Omega-e_i):\mathcal{P}(x)\neq\mathcal{P}(x+e_i)\}|\right]\\
&\overset{(2.66)}{\geqslant} \frac{|\Omega|}{\sqrt[n]{M}}-\frac{1}{n}\sum_{i=1}^{n}|\Omega\smallsetminus(\Omega-e_i)|. \qquad \blacksquare
\end{aligned}$$

2.6 Proof of the lower bound on $\mathsf{PAD}_\delta(\mathbf{X})$ in Theorem 69

Fixing $n\in\mathbb{N}$, a normed space $\mathbf{X}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{X}})$, and $\delta\in(0,1)$, recalling the notation in Definition 65 we will prove here that

$$\mathsf{PAD}_\delta(\mathbf{X})\geqslant\sup_{m\in\mathbb{N}}\mathsf{PAD}_\delta^m(\mathbf{X})\geqslant\frac{2}{1-\sqrt[n]{\delta}}, \tag{2.68}$$

which gives the first inequality in (1.102).

Proof of (2.68). Suppose that $0<\varepsilon<1$ and $r>2$. Let $\mathfrak{N}_\varepsilon$ be any ε-net of $rB_{\mathbf{X}}$. Then, $\log|\mathfrak{N}_\varepsilon|\asymp n\log(r/\varepsilon)$ (see, e.g., [244, Lemma 9.18]). Fix a (disjoint) Voronoi tessellation $\{V_x\}_{x\in\mathfrak{N}_\varepsilon}$ of $rB_{\mathbf{X}}$ that is induced by $\mathfrak{N}_\varepsilon$. Thus, $\{V_x\}_{x\in\mathfrak{N}_\varepsilon}$ is a partition of $rB_{\mathbf{X}}$ into Borel subsets such that $x\in V_x\subseteq x+\varepsilon B_{\mathbf{X}}$ for every $x\in\mathfrak{N}_\varepsilon$. So, for every $w\in rB_{\mathbf{X}}$ there is a unique net point $x(w)\in\mathfrak{N}_\varepsilon$ such that $w\in V_{x(w)}$.

Fix $\mathfrak{p}>\sup_{m\in\mathbb{N}}\mathsf{PAD}_\delta^m(\mathbf{X})\geqslant\mathsf{PAD}_\delta(\mathfrak{N}_\varepsilon)$. Assume from now that $0<\varepsilon<1/(2\mathfrak{p})$ and $r>1/\mathfrak{p}-2\varepsilon$ (eventually we will consider the limits $\varepsilon\to0$ and $r\to\infty$). By the definition of $\mathsf{PAD}_\delta(\mathfrak{N}_\varepsilon)$, there exists a probability distribution $\mathcal{P}$ over 1-bounded partitions of $\mathfrak{N}_\varepsilon$ such that

$$\forall y\in\mathfrak{N}_\varepsilon,\quad \mathbf{Prob}\left[\left(y+\frac{1}{\mathfrak{p}}B_{\mathbf{X}}\right)\cap\mathfrak{N}_\varepsilon\subseteq\mathcal{P}(y)\right]\geqslant\delta. \tag{2.69}$$

For every $y\in\mathfrak{N}_\varepsilon$ define

$$\mathcal{P}^*(y)\overset{\mathrm{def}}{=}\bigcup_{z\in\mathcal{P}(y)}V_z=\big\{w\in rB_{\mathbf{X}}:x(w)\in\mathcal{P}(y)\big\}.$$

Then $\{\mathcal{P}^*(y)\}_{y\in\mathfrak{N}_\varepsilon}$ is a (finitely supported) random partition of $rB_{\mathbf{X}}$ into Borel subsets.

We claim that for every $y \in \mathfrak{N}_\varepsilon$ the following inclusion of events holds:

$$\left\{w \in \mathbb{R}^n : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*(y)\right\} + \frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\big(\mathcal{P}^*(y) - \mathcal{P}^*(y)\big) \subseteq \mathcal{P}^*(y). \tag{2.70}$$

Indeed, take any $w \in \mathbb{R}^n$ such that

$$w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*(y),$$

and also take any $u, v \in \mathcal{P}^*(y)$. By the definition of $\mathcal{P}^*$ we have $\mathfrak{x}(u), \mathfrak{x}(v) \in \mathcal{P}(y)$. As $\mathcal{P}$ is 1-bounded, we have $\|\mathfrak{x}(u) - \mathfrak{x}(v)\|_{\mathbf{X}} \leqslant 1$. Therefore,

$$\|u - v\|_{\mathbf{X}} \leqslant \|u - \mathfrak{x}(u)\|_{\mathbf{X}} + \|\mathfrak{x}(u) - \mathfrak{x}(v)\|_{\mathbf{X}} + \|v - \mathfrak{x}(v)\|_{\mathbf{X}} \leqslant 1 + 2\varepsilon.$$

Hence,

$$\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}(u - v) \in \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}},$$

so the assumption on w implies that

$$w + \frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}(u - v) \in \mathcal{P}^*(y).$$

This is precisely the assertion in (2.70). By the Brunn–Minkowski inequality, (2.70) gives

$$\begin{aligned}\mathrm{vol}_n\big(\mathcal{P}^*(y)\big)^{\frac{1}{n}} \geqslant 2&\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\,\mathrm{vol}_n\big(\mathcal{P}^*(y)\big)^{\frac{1}{n}} \\ &+ \mathrm{vol}_n\left(\left\{w \in \mathbb{R}^n : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*(y)\right\}\right)^{\frac{1}{n}}.\end{aligned}$$

This simplifies to give the following estimate:

$$\begin{aligned}\mathrm{vol}_n&\left(\left\{w \in \mathbb{R}^n : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*(y)\right\}\right) \\ &\leqslant \left(1 - 2\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\right)^n \mathrm{vol}_n\big(\mathcal{P}^*(y)\big).\end{aligned} \tag{2.71}$$

Now,

$$\begin{aligned}\mathrm{vol}_n&\left(\left\{w \in rB_{\mathbf{X}} : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(\mathfrak{x}(w)\big)\right\}\right) \\ &= \sum_{y\in\mathfrak{N}_\varepsilon} \mathrm{vol}_n\left(\left\{w \in \mathcal{P}^*(y) : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}}B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(\mathfrak{x}(w)\big)\right\}\right)\end{aligned} \tag{2.72}$$

$$= \sum_{y\in\mathfrak{N}_\varepsilon} \mathrm{vol}_n\Big(\Big\{w \in \mathcal{P}^*(y) : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} B_{\mathbf{X}} \subseteq \mathcal{P}^*(y)\Big\}\Big) \tag{2.73}$$

$$\leqslant \left(1 - 2\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\right)^n \sum_{y\in\mathfrak{N}_\varepsilon} \mathrm{vol}_n\big(\mathcal{P}^*(y)\big) \tag{2.74}$$

$$= \left(1 - 2\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\right)^n r^n \,\mathrm{vol}_n(B_{\mathbf{X}}). \tag{2.75}$$

Here (2.72) holds because $\{\mathcal{P}^*(y)\}_{y\in\mathfrak{N}_\varepsilon}$ is a partition of $rB_{\mathbf{X}}$. The identity (2.73) holds because, since by the definition of $\mathcal{P}^*$ we have $w \in \mathcal{P}^*(x(w))$ for every $w \in rB_{\mathbf{X}}$ and the sets $\{\mathcal{P}^*(y)\}_{y\in\mathfrak{N}_\varepsilon}$ are pairwise disjoint, if $w \in \mathcal{P}^*(y)$ for some $y \in \mathfrak{N}_\varepsilon$ then necessarily $\mathcal{P}^*(x(w)) = \mathcal{P}^*(y)$. The estimate (2.74) uses (2.71). The identity (2.75) uses once more that $\{\mathcal{P}^*(y)\}_{y\in\mathfrak{N}_\varepsilon}$ is a partition of $rB_{\mathbf{X}}$.

We next claim that for every $w \in (r + 2\varepsilon - 1/\mathfrak{p})B_{\mathbf{X}}$ the following inclusion of events holds:

$$\Big\{\Big(x(w) + \frac{1}{\mathfrak{p}}B_{\mathbf{X}}\Big) \cap \mathfrak{N}_\varepsilon \subseteq \mathcal{P}\big(x(w)\big)\Big\} \subseteq \Big\{w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(x(w)\big)\Big\}. \tag{2.76}$$

Indeed, suppose that $w \in \mathbf{X}$ satisfies $(x(w) + (1/\mathfrak{p})B_{\mathbf{X}}) \cap \mathfrak{N}_\varepsilon \subseteq \mathcal{P}(x(w))$ and also $\|w\|_{\mathbf{X}} \leqslant r + 2\varepsilon - 1/\mathfrak{p}$. Fix any $z \in \mathbf{X}$ such that $\|w - z\|_{\mathbf{X}} \leqslant (1 - 2\varepsilon\mathfrak{p})/\mathfrak{p}$. Then we have $\|z\|_{\mathbf{X}} \leqslant \|w\|_{\mathbf{X}} + \|w - z\|_{\mathbf{X}} \leqslant r$, so $z \in rB_{\mathbf{X}}$ and therefore $x(z) \in \mathfrak{N}_\varepsilon$ is well defined. Now,

$$\|x(w) - x(z)\|_{\mathbf{X}} \leqslant \|x(w) - w\|_{\mathbf{X}} + \|w - z\|_{\mathbf{X}} + \|z - x(z)\|_{\mathbf{X}} \leqslant \varepsilon + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} + \varepsilon = \frac{1}{\mathfrak{p}}.$$

So, our assumption on w implies that $x(z) \in \mathcal{P}(x(w))$. By the definition of $\mathcal{P}^*(x(w))$, this means that $z \in \mathcal{P}^*(x(w))$, thus completing the verification of (2.76). Due to (2.69) and (2.76) we conclude that

$$\forall w \in \Big(r + 2\varepsilon - \frac{1}{\mathfrak{p}}\Big)B_{\mathbf{X}}, \quad \mathbf{Prob}\Big[w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(x(w)\big)\Big] \geqslant \delta. \tag{2.77}$$

Finally,

$$\begin{aligned}
&\delta\Big(r + 2\varepsilon - \frac{1}{\mathfrak{p}}\Big)^n \mathrm{vol}_n(B_{\mathbf{X}}) \\
&\overset{(2.77)}{\leqslant} \int_{(r+2\varepsilon-\frac{1}{\mathfrak{p}})B_{\mathbf{X}}} \mathbf{Prob}\Big[w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(x(w)\big)\Big]\,\mathrm{d}w \\
&= \mathbb{E}\Big[\mathrm{vol}_n\Big(\Big\{w \in \Big(r + 2\varepsilon - \frac{1}{\mathfrak{p}}\Big)B_{\mathbf{X}} : w + \frac{1-2\varepsilon\mathfrak{p}}{\mathfrak{p}} B_{\mathbf{X}} \subseteq \mathcal{P}^*\big(x(w)\big)\Big\}\Big)\Big] \\
&\overset{(2.75)}{\leqslant} \left(1 - 2\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}\right)^n r^n \,\mathrm{vol}_n(B_{\mathbf{X}}).
\end{aligned}$$

This simplifies to give the estimate

$$\sqrt[n]{\delta}\left(1-\frac{1}{\mathfrak{p}r}+\frac{2\varepsilon}{r}\right)\leqslant 1-2\frac{1-2\varepsilon\mathfrak{p}}{(1+2\varepsilon)\mathfrak{p}}.$$

By letting $r\to\infty$, then $\varepsilon\to 0$, and then

$$\mathfrak{p}\to\sup_{m\in\mathbb{N}}\mathsf{PAD}_\delta^m(\mathbf{X}),$$

the desired bound (2.68) follows. ∎

2.7 Proof of Proposition 87

The final lower bound from the Introduction that remains to be proven is Proposition 87. The ensuing reasoning is a restructuring of a proof that was shown to us by Lutwak.

Lemma 111. *Every origin-symmetric convex body $K\subseteq\mathbb{R}^n$ satisfies*

$$\int_{S^{n-1}}\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,du\geqslant\frac{n^2\Gamma\left(\frac{n}{2}\right)}{2\sqrt{\pi}\,\Gamma\left(\frac{n+1}{2}\right)}\mathrm{vol}_n(K)^2. \tag{2.78}$$

Equality in (2.78) *holds if and only if K is an ellipsoid.*

Before proving Lemma 111, we will explain how it implies Proposition 87.

Proof of Proposition 87 assuming Lemma 111. The following standard identity follows from integration in polar coordinates (its quick derivation can be found, for example, on [263, p. 91]):

$$\mathrm{vol}_n(K)=\frac{1}{n}\int_{S^{n-1}}\frac{\mathrm{d}u}{\|u\|_K^n}. \tag{2.79}$$

Hence,

$$\begin{aligned}\int_{S^{n-1}}\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,du&\leqslant\left(\int_{S^{n-1}}\frac{\mathrm{d}u}{\|u\|_K^n}\right)\max_{u\in S^{n-1}}\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K}\\&\overset{(2.79)}{=}n\,\mathrm{vol}_n(K)\max_{z\in\partial K}\Big(\|z\|_{\ell_2^n}\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp}(K)\big)\Big)\\&=n^2\,\mathrm{vol}_n(K)\max_{z\in\partial K}\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big).\end{aligned} \tag{2.80}$$

The desired inequality (1.126) follows by contrasting (2.80) with (2.78). Consequently, if there is equality in (1.126), then (2.78) must hold as equality as well, so the characterization of the equality case in Proposition 87 follows from the characterization of the quality case in Lemma 111. ∎

The important *Petty projection inequality* [252] (see also [194, 281] for different proofs, as well as the survey [190]) states that for every convex body $K \subseteq \mathbb{R}^n$, the affine invariant quantity

$$\mathrm{vol}_n(K)^{n-1}\,\mathrm{vol}_n(\Pi^* K) \tag{2.81}$$

is maximized when K is an ellipsoid, and ellipsoids are the only maximizers of (2.81). Recall that the polar projection body $\Pi^* K$ is given by (1.30), which shows in particular that $\mathrm{vol}_{n-1}(B_{\ell_2^{n-1}})\Pi^* B_{\ell_2^n} = B_{\ell_2^n}$. Hence,

$$\begin{aligned}\mathrm{vol}_n(K)^{n-1}\,\mathrm{vol}_n(\Pi^* K) &\leqslant \mathrm{vol}_n(B_{\ell_2^n})^{n-1}\,\mathrm{vol}_n(\Pi^* B_{\ell_2^n})\\ &= \left(\frac{\mathrm{vol}_n\big(B_{\ell_2^n}\big)}{\mathrm{vol}_{n-1}\big(B_{\ell_2^{n-1}}\big)}\right)^n = \left(\frac{2\sqrt{\pi}\,\Gamma\big(\frac{n+1}{2}\big)}{n\Gamma\big(\frac{n}{2}\big)}\right)^n.\end{aligned}$$

At the same time, by combining (1.30) and (2.79) we have

$$\mathrm{vol}_n(\Pi^* K) = \frac{1}{n}\int_{S^{n-1}} \frac{\mathrm{d}u}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^n}.$$

Consequently, Petty's projection inequality can be restated as the following estimate:

$$\int_{S^{n-1}} \frac{\mathrm{d}u}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^n} \leqslant \left(\frac{2\sqrt{\pi}\,\Gamma\big(\frac{n+1}{2}\big)}{n\Gamma\big(\frac{n}{2}\big)}\right)^n \frac{n}{\mathrm{vol}_n(K)^{n-1}}, \tag{2.82}$$

together with the assertion that (2.82) holds as an equality if and only if K is an ellipsoid.

Proof of Lemma 111. Observe that

$$\mathrm{vol}_n(K)$$

$$= \frac{1}{n}\int_{S^{n-1}} \left(\frac{1}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^{\frac{n}{n+1}}}\right)\left(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^{\frac{n}{n+1}}}{\|u\|_K^n}\right)\mathrm{d}u \tag{2.83}$$

$$\leqslant \frac{1}{n}\left(\int_{S^{n-1}} \frac{\mathrm{d}u}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^n}\right)^{\frac{1}{n+1}}\left(\int_{S^{n-1}} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,\mathrm{d}u\right)^{\frac{n}{n+1}} \tag{2.84}$$

$$\leqslant \frac{1}{n}\left(\frac{2\sqrt{\pi}\,\Gamma\big(\frac{n+1}{2}\big)}{n\Gamma\big(\frac{n}{2}\big)}\right)^{\frac{n}{n+1}} \frac{n^{\frac{1}{n+1}}}{\mathrm{vol}_n(K)^{\frac{n-1}{n+1}}}\left(\int_{S^{n-1}} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,\mathrm{d}u\right)^{\frac{n}{n+1}}, \tag{2.85}$$

where (2.83) is (2.79), in (2.84) we used Hölder's inequality with the conjugate exponents $1+\frac{1}{n}$ and $n+1$, and (2.85) is an application of (2.82). This simplifies to give the desired inequality (2.78). ■

Remark 112. Fix $n \in \mathbb{N}$, a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $x \in S^{n-1}$. Both of the bounds in (1.50) follow from elementary geometric reasoning (convexity and Fubini's theorem). Recalling (1.30), the second inequality in (1.50) is $\mathrm{vol}_{n-1}(\mathsf{Proj}_{x^\perp} B_{\mathbf{X}}) \leqslant n\|x\|_{\mathbf{X}} \,\mathrm{vol}_n(B_{\mathbf{X}})/2$; its justification can be found in the proof of [109, Lemma 5.1] (this was not included in the version of [109] that appeared in the journal, but it appears in the arxiv version of [109]). The rest of (1.50) is

$$\mathrm{vol}_n(B_{\mathbf{X}})\|x\|_{\mathbf{X}} \leqslant 2\,\mathrm{vol}_{n-1}(\mathsf{Proj}_{x^\perp} B_{\mathbf{X}});$$

since we did not find a reference for the derivation of this simple lower bound on hyperplane projections, we will now quickly justify it. For every $u \in \mathsf{Proj}_{x^\perp} B_{\mathbf{X}}$ let $s(u) = \inf\{s \in \mathbb{R} : u + sx \in B_{\mathbf{X}}\}$ and $t(u) = \sup\{t \in \mathbb{R} : u + tx \in B_{\mathbf{X}}\}$. For every $u \in \mathsf{Proj}_{x^\perp} B_{\mathbf{X}}$ we have $u + t(u)x \in B_{\mathbf{X}}$, and by symmetry also $-u - s(u)x \in B_{\mathbf{X}}$. Hence, by convexity

$$\frac{1}{2}\big(u + t(u)x\big) + \frac{1}{2}\big(-u - s(u)x\big) = \frac{t(u) - s(u)}{2}x \in B_{\mathbf{X}}.$$

By the definition of $t(0)$, this means that $(t(u) - s(u))/2 \leqslant t(0) = 1/\|x\|_{\mathbf{X}}$. Consequently, using Fubini's theorem (recall that $x \in S^{n-1}$) we conclude that

$$\begin{aligned}\mathrm{vol}_n(B_{\mathbf{X}}) &= \int_{\mathsf{Proj}_{x^\perp} B_{\mathbf{X}}} \big(t(u) - s(u)\big)\,\mathrm{d}u \\ &\leqslant \int_{\mathsf{Proj}_{x^\perp} B_{\mathbf{X}}} \frac{2}{\|x\|_{\mathbf{X}}}\,\mathrm{d}u = \frac{2}{\|x\|_{\mathbf{X}}}\,\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp} B_{\mathbf{X}}\big).\end{aligned}$$