Chapter 4

Upper bounds on random partitions

In this section, we will prove the existence of random partitions with the separation and padding properties that were stated in the Introduction.

4.1 Proof of Theorem 75 and the upper bound on $\mathsf{PAD}_\delta(\mathbf{X})$ in Theorem 69

Theorem 122 below asserts that every normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ admits a random partition that simultaneously has desirable padding and separation properties. In the literature, such properties are obtained for different random partitions: separating partitions of normed spaces use iterative ball partitioning with deterministic radii, while padded partitions also rely on randomizing the radii. At present, we do not have in mind an application in which good padding and separation properties are needed simultaneously for the same random partition, so it is worthwhile to note this feature for potential future use but in what follows we will use Theorem 122 to obtain two standalone conclusions that yield upper bounds on the moduli of padded and separated decomposability (in fact, the separation profile of Theorem 75).

Theorem 122. *Fix $n \in \mathbb{N}$ and a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. For every $\Delta \in (0,\infty)$ there exists a Δ-bounded random partition $\mathcal{P}_\Delta$ of $\mathbf{X}$ such that for every $x, y \in \mathbb{R}^n$ and every $\delta \in (0,1)$ we have*

$$\mathbf{Prob}\big[\mathcal{P}_\Delta(x) \neq \mathcal{P}_\Delta(y)\big] \asymp \min\left\{1, \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp}(B_{\mathbf{X}})\big)}{\Delta\,\mathrm{vol}_n(B_{\mathbf{X}})}\|x-y\|_{\ell_2^n}\right\} \tag{4.1}$$

and

$$\mathbf{Prob}\left[\mathcal{P}_\Delta(x) \supseteq \frac{1-\sqrt[n]{\delta}}{1+\sqrt[n]{\delta}}\cdot\frac{\Delta}{2}B_{\mathbf{X}}\right] = \delta.$$

By the conventions of Remark 62, the Δ-boundedness of Theorem 122 is with respect to the norm $\|\cdot\|_{\mathbf{X}}$, i.e., the clusters of the random partition $\mathcal{P}_\Delta$ have $\mathbf{X}$-diameter at most Δ. By the definitions in Section 1.7.1, the notion of random partition implies that each of the clusters of $\mathcal{P}_\Delta$ is strongly measurable, but we will see that they are also standard (recall Definition 113).

Remark 123. For every $M > 0$, consider the metric space $L_1^{\leqslant M} = (L_1, d_M)$ that is given by

$$\forall\, f, \in L_1, \quad d_M(f,g) \stackrel{\mathrm{def}}{=} \min\big\{M, \|f-g\|_{L_1}\big\}.$$

A useful property [211, Lemma 5.4] of this truncated L_1 metric is $c_{L_1}(L_1^{\leqslant M}) \lesssim 1$, i.e., $L_1^{\leqslant M}$ embeds back into L_1 with bi-Lipschitz distortion $O(1)$. Theorem 122 gives a different proof of this since if $\mathbf{X} = \ell_\infty^n$, then by (1.38) the right-hand side of (4.1) is equal to $\min\{2\Delta, \|x-y\|_1\}/(2\Delta)$. At the same time, if $\mathcal{P}_\Delta^\omega = \{\Gamma_\Delta^k(\omega)\}_{k=1}^\infty$, then the left-hand side of (4.1) embeds isometrically into an $L_1(\mu)$ space via the embedding

$$(f \in L_1) \mapsto \left(\omega \mapsto \left(\mathbf{1}_{\Gamma^k(\omega)}(f)\right)_{k=1}^\infty\right) \in L_1(\mathbf{Prob}; \ell_1).$$

By (1.30), the right-hand side of (4.1) equals $\min\{\Delta, \|x-y\|_{\Pi^*\mathbf{X}}\}/\Delta$. But, by [41] the class of finite dimensional normed spaces whose unit ball is a polar projection body coincides with those finite dimensional normed spaces that embed isometrically into L_1, so this does not give a new embedding result.

We will first describe the construction that leads to the random partition whose existence is asserted in Theorem 122. This construction is a generalization of the construction that appears in the proof of [173, Lemma 3.16], which itself combines a coloring argument with a generalization of the iterated ball partitioning technique that was used in the Euclidean setting in [76, 152].

In the rest of this section we will work under the assumptions and notation of Theorem 122. Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice such that $\{z + B_\mathbf{X}\}_{z\in\Lambda}$ have pairwise disjoint interiors (equivalently, $\|z - z'\|_\mathbf{X} \geqslant 2$ for distinct $z, z' \in \Lambda$) and $\bigcup_{z\in\Lambda}(z + 3B_\mathbf{X}) = \mathbb{R}^n$ (i.e., for every $x \in \mathbb{R}^n$ there is $z \in \Lambda$ such that $\|x - z\|_\mathbf{X} \leqslant 3$). The existence of such a lattice follows from the work of Rogers [273] (see [315, Remark 6]). The constant 3 here is not the best-known (see [70, 315]); we prefer to work with an explicit constant only for notational convenience despite the fact that its value is not important in the present context.

Denote the $\mathbf{X}$-Voronoi cell of Λ, i.e., the set of points in $\mathbb{R}^n$ whose closest lattice point is the origin, by

$$\mathcal{V} \stackrel{\mathrm{def}}{=} \left\{x \in \mathbb{R}^n : \|x\|_\mathbf{X} = \min_{z\in\Lambda} \|x - z\|_\mathbf{X}\right\}.$$

Then $\mathcal{V} \subseteq 3B_\mathbf{X}$ and the translates $\{z + \mathcal{V}\}_{z\in\Lambda}$ cover $\mathbb{R}^n$ and have pairwise disjoint interiors.

Remark 124. Our choice of the above lattice is natural since it is adapted to the intrinsic geometry of $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_\mathbf{X})$ and it leads to a simpler probability space in the construction below. Nevertheless, for the present purposes this choice is not crucial, and one could also work with any other lattice, including $\mathbb{Z}^n$. In that case, one could carry out the ensuing reasoning while adapting it to geometric characteristics of the lattice in question (its packing radius, covering radius and the diameter of its Voronoi cell, all of which are measured with respect to the metric induced by $\|\cdot\|_\mathbf{X}$). This requires several changes in the ensuing discussion, resulting in slightly more cumbersome computations that incorporate these geometric characteristics of the lattice. All of these quantities are universal constants for our choice of Λ.

Define graph $\mathsf{G} = (\Lambda, \mathsf{E}_\mathsf{G})$ whose vertex set is the lattice Λ and whose edge set E_G is given by

$$\forall w, z \in \Lambda, \quad \{w, z\} \in \mathsf{E}_\mathsf{G} \iff w \neq z \wedge \inf_{\substack{a \in w+\mathcal{V}\\ b \in z+\mathcal{V}}} \|a-b\|_\mathbf{X} \leqslant 10.$$

So, if $\{w, z\} \in \mathsf{E}_\mathsf{G}$ and $x \in B_\mathbf{X}$ then there are $u, v \in \mathcal{V}$ such that

$$\|(w+u)-(z+v)\|_\mathbf{X} \leqslant 10$$

and therefore, since $\mathcal{V} \subseteq 3B_\mathbf{X}$, we have

$$\|w-(z+x)\|_\mathbf{X} \leqslant \|(w+u)-(z+v)\|_\mathbf{X} + \|u\|_\mathbf{X} + \|v\|_\mathbf{X} + \|x\|_\mathbf{X} \leqslant 17.$$

Hence $z + B_\mathbf{X} \subseteq w + 17B_\mathbf{X}$. It follows that if $w \in \Lambda$ and $z_1, \ldots, z_m \in \Lambda$ are the distinct neighbors of w in the graph G then the balls $\{z_i + B_\mathbf{X}\}_{i=1}^m$ have disjoint interiors (since distinct elements of the lattice Λ are at $\mathbf{X}$-distance at least 2), yet they are all contained in the ball $w + 17B_\mathbf{X}$. By comparing volumes, this implies that $m \leqslant 17^n$. In other words, the degree of the graph G is at most 17^n, and therefore (by applying the greedy algorithm, see, e.g., [59]) its chromatic number is at most $17^n + 1 \leqslant 5^{2n}$, i.e., there is $\chi : \Lambda \to \{1, \ldots, 5^{2n}\}$ such that

$$\forall w, z \in \Lambda, \quad w \neq z \wedge \inf_{\substack{a \in w+\mathcal{V}\\ b \in z+\mathcal{V}}} \|a-b\|_\mathbf{X} \leqslant 10 \implies \chi(w) \neq \chi(z). \tag{4.2}$$

Consider the Polish space $\mathcal{Z} \stackrel{\mathrm{def}}{=} \mathcal{V}^\mathbb{N} \times \{1, \ldots, 5^{2n}\}^\mathbb{N}$. In what follows, every $\omega \in \mathcal{Z}$ will be written as $\omega = (\vec{x}, \vec{\gamma})$, where $\vec{x} = (x_1, x_2, \ldots) \in \mathcal{V}^\mathbb{N}$ and $\vec{\gamma} = (\gamma_1, \gamma_2, \ldots) \in \{1, \ldots, 5^{2n}\}^\mathbb{N}$. Denote by μ the normalized Lebesgue measure on $\mathcal{V}$ and by ν the normalized counting measure on $\{1, \ldots, 5^{2n}\}$, i.e., for every Lebesgue measurable $A \subseteq \mathbb{R}^n$ and every $F \subseteq \{1, \ldots, 5^{2n}\}^\mathbb{N}$ we have

$$\mu(A) \stackrel{\mathrm{def}}{=} \frac{\mathrm{vol}_n(A \cap \mathcal{V})}{\mathrm{vol}_n(\mathcal{V})} \quad \text{and} \quad \nu(F) \stackrel{\mathrm{def}}{=} \frac{|F|}{5^{2n}}.$$

Henceforth, the product probability measure $\mu^\mathbb{N} \times \nu^\mathbb{N}$ on $\mathcal{Z}$ will be denoted by **Prob**.

For every $k \in \mathbb{N}$, $z \in \Lambda$ and $(\vec{x}, \vec{\gamma}) \in \mathcal{Z}$ define a subset $\Gamma^{k,z}(\vec{x}, \vec{\gamma}) \subseteq \mathbb{R}^n$ by

$$\begin{aligned} \chi(z) = \gamma_k &\implies \Gamma^{k,z}(\vec{x}, \vec{\gamma}) \stackrel{\mathrm{def}}{=} (z + x_k + B_\mathbf{X}) \smallsetminus \bigcup_{j=1}^{k-1} \bigcup_{\substack{w \in \Lambda\\ \chi(w)=\gamma_j}} (w + x_j + B_\mathbf{X}), \\ \chi(z) \neq \gamma_k &\implies \Gamma^{k,z}(\vec{x}, \vec{\gamma}) \stackrel{\mathrm{def}}{=} \varnothing. \end{aligned} \tag{4.3}$$

Lemma 125. *For every $k \in \mathbb{N}$ and $z \in \Lambda$ the set-valued mapping $\Gamma^{k,z} : \mathcal{Z} \to 2^{\mathbb{R}^n}$ is both strongly measurable and standard (where the underlying σ-algebra on $\mathcal{Z}$ is the* **Prob***-measurable sets).*

Proof. For every $\chi_1, \dots, \chi_k \in \{1, \dots, 5^{2n}\}$ consider the cylinder set

$$\mathcal{C}(\chi_1, \dots, \chi_k) \stackrel{\text{def}}{=} \left\{(\vec{x}, \vec{\gamma}) \in \mathcal{Z} : (\gamma_1, \dots, \gamma_k) = (\chi_1, \dots, \chi_k)\right\}.$$

As $\{\mathcal{C}(\chi_1, \dots, \chi_k) : (\chi_1, \dots, \chi_k) \in \{1, \dots, 5^{2n}\}^k\}$ is a partition of $\mathcal{Z}$ into finitely many measurable sets, it suffices to fix from now on a k-tuple of colors $\vec{\chi} = (\chi_1, \dots, \chi_k) \in \{1, \dots, 5^{2n}\}^k$ and to show that the restriction of $\Gamma^{k,z}$ to $\mathcal{C}(\chi_1, \dots, \chi_k)$ is both strongly measurable and standard.

Observe that for each fixed $z \in \Lambda$ and $\gamma \in \{1, \dots, 5^{2n}\}$ there is at most one $w \in \Lambda$ that satisfies $\chi(w) = \gamma$ and $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w + \mathcal{V} + B_{\mathbf{X}}) \neq \emptyset$. Indeed, if both $w \in \Lambda$ and $w' \in \Lambda$ satisfied these two requirements then we would have $\chi(w) = \gamma = \chi(w')$ and there would exist $a, a', b, b' \in \mathcal{V}$ and $u, u', v, v' \in B_{\mathbf{X}}$ such that $w + a + u = z + b + v$ and $w' + a' + u' = z + b' + v'$. Hence,

$$\begin{aligned}
\inf_{\substack{\alpha \in w + \mathcal{V} \\ \beta \in w' + \mathcal{V}}} \|\alpha - \beta\|_{\mathbf{X}} &\leqslant \|(w + a) - (w' + a')\|_{\mathbf{X}} \\
&= \|(z + b + v - u) - (z + b' + v' - u')\|_{\mathbf{X}} \\
&\leqslant \|b\|_{\mathbf{X}} + \|b'\|_{\mathbf{X}} + \|v\|_{\mathbf{X}} + \|v'\|_{\mathbf{X}} + \|u\|_{\mathbf{X}} + \|u'\|_{\mathbf{X}} \\
&\leqslant 3 + 3 + 1 + 1 + 1 + 1 = 10,
\end{aligned}$$

where we used the fact that $b, b' \in \mathcal{V} \subseteq 3B_{\mathbf{X}}$. By (4.2) this contradicts the fact that $\chi(w) = \chi(w')$.

Having checked that the above w is unique, denote it by $w(\gamma, z) \in \Lambda$. If there is no $w \in \Lambda$ that satisfies $\chi(w) = \gamma$ and $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w + \mathcal{V} + B_{\mathbf{X}}) \neq \emptyset$ then let $w(\gamma, z) \in \Lambda$ be an arbitrary (but fixed) lattice point such that $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w(\gamma, z) + \mathcal{V} + B_{\mathbf{X}}) = \emptyset$. Observe that $w(\chi(z), z) = z$. Under this notation, for every $x_1, \dots, x_k \in \mathcal{V}$ and $\gamma_1, \dots, \gamma_{k-1} \in \{1, \dots, 5^{2n}\}$ we have

$$\begin{aligned}
&(z + x_k + B_{\mathbf{X}}) \smallsetminus \bigcup_{j=1}^{k-1} \bigcup_{\substack{w \in \Lambda \\ \chi(w) = \gamma_j}} (w + x_j + B_{\mathbf{X}}) \\
&\quad = (w(\chi(z), z) + x_k + B_{\mathbf{X}}) \smallsetminus \bigcup_{j=1}^{k-1} (w(\gamma_j, z) + x_j + B_{\mathbf{X}}).
\end{aligned}$$

Equivalently, if we denote for every $\vec{\mathsf{y}} = (\mathsf{y}_1, \dots, \mathsf{y}_k) \in (\mathbb{R}^n)^k$,

$$\Theta^k(\vec{\mathsf{y}}) \stackrel{\text{def}}{=} (\mathsf{y}_k + B_{\mathbf{X}}) \smallsetminus \bigcup_{j=1}^{k-1} (\mathsf{y}_j + B_{\mathbf{X}}),$$

then the definition (4.3) can be rewritten as the assertion that the restriction of $\Gamma^{k,z}$ to $\mathcal{C}(\vec{\chi})$ is the constant function $\varnothing$ if $\chi(z) \neq \chi_k$, whereas if $\chi(z) = \chi_k$, then we define $\Gamma^{k,z}(\vec{x}, \vec{\gamma}) = \Theta^k(w(\vec{\chi}, z) + \vec{x})$ for every $(\vec{x}, \vec{\gamma}) \in \mathcal{C}(\vec{\chi})$, where we use the notation $w(\vec{\chi}, z) = (w(\chi_1, z), \ldots, w(\chi_k, z)) \in (\mathbb{R}^n)^k$. The desired measurability of the restriction of $\Gamma^{k,z}$ to $\mathcal{C}(\vec{\chi})$ now follows from Lemma 119 and Corollary 121. ■

Since the sets $\{z + \mathcal{V}\}_{z\in\Lambda}$ cover $\mathbb{R}^n$, for every rational point $q \in \mathbb{Q}^n$ we can fix from now on a lattice point $z_q \in \Lambda$ such that $q \in z_q + \mathcal{V}$. Define a subset $\Omega \subseteq \mathcal{Z} = \mathcal{V}^{\mathbb{N}} \times \{1, \ldots, 5^{2n}\}^{\mathbb{N}}$ by

$$\Omega \stackrel{\mathrm{def}}{=} \bigcap_{m=1}^{\infty} \bigcap_{q\in\mathbb{Q}^n} \bigcup_{k=1}^{\infty} \left\{(\vec{x}, \vec{\gamma}) \in \mathcal{Z} : \chi(z_q) = \gamma_k \wedge \|(z_q + x_k) - q\|_{\mathbf{X}} \leqslant \frac{1}{m}\right\}. \tag{4.4}$$

We record for ease of later use the following simple properties of Ω.

Lemma 126. *Ω is a Borel subset of $\mathcal{Z}$ that satisfies* $\mathbf{Prob}[\Omega] = 1$. *Furthermore, for every $(\vec{x}, \vec{\gamma}) \in \Omega$ the set $\{z + x_k : (k, z) \in \mathbb{N} \times \Lambda \wedge \chi(z) = \gamma_k\}$ is dense in $\mathbb{R}^n$.*

Proof. The fact that Ω is Borel is evident from its definition (4.4). Also, if $(\vec{x}, \vec{\gamma}) \in \Omega$, $u \in \mathbb{R}^n$ and $\varepsilon \in (0, 1)$, then choose $q \in \mathbb{Q}^n$ such that $\|u - q\|_{\mathbf{X}} < \varepsilon/2$. Setting $m = \lceil 2/\varepsilon\rceil \in \mathbb{N}$, it follows from (4.4) that there exists $k \in \mathbb{N}$ satisfying $\chi(z_q) = \gamma_k$ and $\|(z_q + x_k) - q\|_{\mathbf{X}} \leqslant 1/m \leqslant \varepsilon/2$. By our choice of q, it follows that

$$\big\|(z_q + x_k) - u\big\|_{\mathbf{X}} < \varepsilon.$$

Since this holds for every $\varepsilon \in (0, 1)$, the set $\{z + x_k : (k, z) \in \mathbb{N} \times \Lambda \wedge \chi(z) = \gamma_k\}$ is dense in $\mathbb{R}^n$. It remains to show that $\mathbf{Prob}[\Omega] = 1$. Indeed,

$$\begin{aligned}
&\mathbf{Prob}[\mathcal{Z} \smallsetminus \Omega] \\
&\stackrel{(4.4)}{\leqslant} \sum_{m=1}^{\infty} \sum_{q\in\mathbb{Q}^n} \mathbf{Prob}\left[\bigcap_{k=1}^{\infty} \mathcal{Z} \smallsetminus \left\{(\vec{x}, \vec{\gamma}) \in \mathcal{Z} : \chi(z_q) = \gamma_k \wedge \|(z_q + x_k) - q\|_{\mathbf{X}} \leqslant \frac{1}{m}\right\}\right] \\
&= \sum_{m=1}^{\infty} \sum_{q\in\mathbb{Q}^n} \lim_{\ell\to\infty} \left(1 - \frac{\mathrm{vol}_n\left(\left(q - z_q + \frac{1}{m}B_{\mathbf{X}}\right) \cap \mathcal{V}\right)}{5^{2n}\,\mathrm{vol}_n(\mathcal{V})}\right)^{\ell} = 0,
\end{aligned} \tag{4.5}$$

where for the penultimate step of (4.5) recall that $\mathbf{Prob} = \mu^{\mathbb{N}} \times \nu^{\mathbb{N}}$. For the final step of (4.5) note that $\mathrm{vol}_n((q - z_q + rB_{\mathbf{X}}) \cap \mathcal{V}) = \mathrm{vol}_n((q + rB_{\mathbf{X}}) \cap (z_q + \mathcal{V})) > 0$ for every fixed $q \in \mathbb{Q}^n$ and $r \in (0, \infty)$, because $z_q \in \Lambda$ was chosen so that $q \in z_q + \mathcal{V}$ (and $\mathcal{V}$ is a convex body). ■

The following lemma introduces the random partition that will be used to prove Theorem 122.

Lemma 127. *$\mathcal{P} \stackrel{\mathrm{def}}{=} \{\Gamma^{k,z}|_\Omega : \Omega \to 2^{\mathbb{R}^n}\}_{(k,z)\in\mathbb{N}\times\Lambda}$ is a 2-bounded random partition of $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$, each of whose clusters are both strongly measurable and standard set-valued mappings.*

Proof. Since Ω is a Borel subset of $\mathcal{Z}$, for each $(k,z) \in \mathbb{N}\times\Lambda$ the measurability requirements for the restriction of $\Gamma^{k,z}$ to Ω follow from Lemma 125. Fix $(\vec{x},\vec{\gamma}) \in \mathcal{Z}$. Recalling (4.3), if $\Gamma^{k,z}(\vec{x},\vec{\gamma}) \neq \varnothing$, then $\operatorname{diam}_{\mathbf{X}}(\Gamma^{k,z}(\vec{x},\vec{\gamma})) \leqslant \operatorname{diam}_{\mathbf{X}}(z + x_k + B_{\mathbf{X}}) \leqslant 2$. Note also that by (4.3) if $\Gamma^{k,z}(\vec{x},\vec{\gamma}) \neq \varnothing$, then

$$\Gamma^{k,z}(\vec{x},\vec{\gamma}) = (z + x_k + B_{\mathbf{X}}) \smallsetminus \bigcup_{j=1}^{k-1}\bigcup_{w\in\Lambda} \Gamma^{j,w}(\vec{x},\vec{\gamma}).$$

Hence $\Gamma^{k,z}(\vec{x},\vec{\gamma}) \cap \Gamma^{j,w}(\vec{x},\vec{\gamma}) = \varnothing$ for every distinct $j,k \in \mathbb{N}$ and for every $w,z \in \Lambda$. We claim that also

$$\Gamma^{k,z}(\vec{x},\vec{\gamma}) \cap \Gamma^{k,w}(\vec{x},\vec{\gamma}) = \varnothing$$

for every $k \in \mathbb{N}$ and every distinct $w,z \in \Lambda$. Indeed, it suffices to check this under the assumption that $\chi(w) = \chi(z) = \gamma_k$, since otherwise $\varnothing \in \{\Gamma^{k,z}(\vec{x},\vec{\gamma}), \Gamma^{k,w}(\vec{x},\vec{\gamma})\}$. So, suppose that

$$\chi(w) = \chi(z) = \gamma_k \quad \text{yet} \quad \Gamma^{k,z}(\vec{x},\vec{\gamma}) \cap \Gamma^{k,w}(\vec{x},\vec{\gamma}) \neq \varnothing.$$

By (4.3), this implies that there are $u, v \in B_{\mathbf{X}}$ such that $w + x_k + u = z + x_k + v$. Hence, for every $\alpha, \beta \in \mathcal{V}$,

$$\begin{aligned}\|(w+\alpha) - (z+\beta)\|_{\mathbf{X}} &= \|\alpha - \beta + v - u\|_{\mathbf{X}} \\ &\leqslant \|\alpha\|_{\mathbf{X}} + \|\beta\|_{\mathbf{X}} + \|u\|_{\mathbf{X}} + \|v\|_{\mathbf{X}} \\ &\leqslant 3 + 3 + 1 + 1 < 10,\end{aligned}$$

where we used the fact that $\mathcal{V} \subseteq 3B_{\mathbf{X}}$. Since w and z are distinct and $\chi(w) = \chi(z)$, this is in contradiction to (4.2). We have thus shown that the sets $\{\Gamma^{k,z}(\vec{x},\vec{\gamma})\}_{(k,z)\in\mathbb{N}\times\Lambda}$ are pairwise disjoint.

Note that by the definition (4.3), for every $(\vec{x},\vec{\gamma}) \in \mathcal{Z}$ we have

$$\bigcup_{k=1}^{\infty}\bigcup_{z\in\Lambda} \Gamma^{k,z}(\vec{x},\vec{\gamma}) = \bigcup_{\substack{(k,z)\in\mathbb{N}\times\Lambda\\ \chi(z)=\gamma_k}} (z + x_k + B_{\mathbf{X}}). \tag{4.6}$$

Indeed, it is immediate from (4.3) that the left-hand side of (4.6) is contained in the right-hand side of (4.6). If u belongs to the right-hand side of (4.6), then let k be the minimum natural number for which there is $z \in \Lambda$ with $u \in z + x_k + B_{\mathbf{X}}$ and $\chi(z) = \gamma_k$. Consequently, for all $j \in \{1,\ldots,k-1\}$ and $w \in \Lambda$ with $\chi(w) = \gamma_j$ we have $u \notin w + x_j + B_{\mathbf{X}}$, and hence by (4.3) we have $v \in \Gamma^{k,z}(\vec{x},\vec{\gamma})$, as required. By

Lemma 126, if $(\vec{x},\vec{\gamma}) \in \Omega$, then $\{z + x_k : (k,z) \in \mathbb{N} \times \Lambda \wedge \chi(z) = \gamma_k\}$ is dense in $\mathbb{R}^n$, and therefore the right-hand side of (4.6) is equal to $\mathbb{R}^n$. Thus $\mathcal{P}$ takes values in partitions of $\mathbb{R}^n$. ∎

Definition 128 introduces convenient notation that will be used several times in what follows.

Definition 128. If $\mathcal{M} \subseteq \mathbb{R}^n$ is Lebesgue measurable and $(k,z) \in \mathbb{N} \times \Lambda$, then define $\mathsf{H}^{k,z}_{\mathcal{M}} \subseteq \Omega$ by

$$\mathsf{H}^{k,z}_{\mathcal{M}} \stackrel{\mathrm{def}}{=} \big\{(\vec{x},\vec{\gamma}) \in \Omega : \chi(z) = \gamma_k \wedge z + x_k \in \mathcal{M}\big\}. \tag{4.7}$$

If $\mathcal{S},\mathcal{T} \subseteq \mathbb{R}^n$ are Lebesgue measurable and $(k,z) \in \mathbb{N} \times \Lambda$, then define $\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}} \subseteq \Omega$ by

$$\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}} \stackrel{\mathrm{def}}{=} \mathsf{H}^{k,z}_{\mathcal{S}} \smallsetminus \bigcup_{j=1}^{k-1} \bigcup_{w \in \Lambda} \mathsf{H}^{j,w}_{\mathcal{T}}. \tag{4.8}$$

The meaning of the set in (4.8) is that it consists of all of those $(\vec{x},\vec{\gamma}) \in \Omega$ such that the kth coordinate of $\vec{\gamma} \in \{1,\ldots,5^{2n}\}^{\mathbb{N}}$ is the color of the lattice point $z \in \Lambda$, the kth coordinate of $\vec{x} \in \mathcal{V}^{\mathbb{N}}$ satisfies $x_k \in \mathcal{S} - z$, and for no $j \in \{1,\ldots,k-1\}$ and no lattice point $w \in \Lambda$ do the same assertions hold with $\mathcal{S}$ replaced by $\mathcal{T}$.

Lemma 129. *Suppose that $\mathcal{S},\mathcal{T} \subseteq \mathbb{R}^n$ are Lebesgue measurable sets of positive volume such that $\mathcal{S} \subseteq \mathcal{T}$. Suppose also that* $\mathrm{diam}_{\mathbf{X}}(\mathcal{T}) \leqslant 4$. *Then the sets*

$$\{\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}}\}_{(k,z)\in\mathbb{N}\times\Lambda}$$

are pairwise disjoint and

$$\mathbf{Prob}\Bigg[\bigcup_{k=1}^{\infty} \bigcup_{z \in \Lambda} \mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}}\Bigg] = \frac{\mathrm{vol}_n(\mathcal{S})}{\mathrm{vol}_n(\mathcal{T})}. \tag{4.9}$$

Proof. The definition of the product measure **Prob** implies that for any Lebesgue measurable $\mathcal{M} \subseteq \mathbb{R}^n$ and every $(j,w) \in \mathbb{N} \times \Lambda$ we have

$$\begin{aligned} \mathbf{Prob}\big[\mathsf{H}^{j,w}_{\mathcal{M}}\big] &= \mu(\mathcal{M} - w)\nu\big(\chi(w)\big) \\ &= \frac{\mathrm{vol}_n\big(\mathcal{V} \cap (\mathcal{M} - w)\big)}{5^{2n}\,\mathrm{vol}_n(\mathcal{V})} \\ &= \frac{\mathrm{vol}_n\big((\mathcal{V} + w) \cap \mathcal{M}\big)}{5^{2n}\,\mathrm{vol}_n(\mathcal{V})}. \end{aligned} \tag{4.10}$$

We claim if $\mathrm{diam}_{\mathbf{X}}(\mathcal{M}) \leqslant 4$, then $\{\mathsf{H}^{j,w}_{\mathcal{M}}\}_{w\in\Lambda}$ are pairwise disjoint for every fixed $j \in \mathbb{N}$. Indeed, otherwise

$$\exists (\vec{x},\vec{\gamma}) \in \mathsf{H}^{j,w}_{\mathcal{M}} \cap \mathsf{H}^{j,z}_{\mathcal{M}}$$

for some distinct lattice points $w, z \in \Lambda$. Then, $w + x_j, z + x_j \in \mathcal{M}$ and $\chi(w) = \gamma_j = \chi(z)$. Hence,

$$\|w - z\|_{\mathbf{X}} = \|(w + x_j) - (z + x_j)\|_{\mathbf{X}} \leqslant \operatorname{diam}_{\mathbf{X}}(\mathcal{M}) \leqslant 4.$$

Since $\mathcal{V} \subseteq 3B_{\mathbf{X}}$, it follows that for every $\alpha, \beta \in \mathcal{V}$ we have

$$\|(w + \alpha) - (z + \beta)\|_{\mathbf{X}} \leqslant \|w - z\|_{\mathbf{X}} + \|\alpha\|_{\mathbf{X}} + \|\beta\|_{\mathbf{X}} \leqslant 4 + 3 + 3 = 10,$$

which, by virtue of (4.2), contradicts the fact that $w \neq z$ and $\chi(w) = \chi(z)$.

Since $\{\mathsf{H}_{\mathcal{M}}^{j,w}\}_{w\in\Lambda}$ are pairwise disjoint and $\{w + \mathcal{V}\}_{w\in\Lambda}$ cover $\mathbb{R}^n$ and have pairwise disjoint interiors,

$$\begin{aligned}\mathbf{Prob}\Big[\bigcup_{w\in\Lambda} \mathsf{H}_{\mathcal{M}}^{j,w}\Big] &= \sum_{w\in\Lambda} \mathbf{Prob}\big[\mathsf{H}_{\mathcal{M}}^{j,w}\big]\\ &\overset{(4.10)}{=} \frac{1}{5^{2n}\operatorname{vol}_n(\mathcal{V})} \sum_{w\in\Lambda} \operatorname{vol}_n\big((\mathcal{V}+w)\cap\mathcal{M}\big) = \frac{\operatorname{vol}_n(\mathcal{M})}{5^{2n}\operatorname{vol}_n(\mathcal{V})}. \end{aligned} \tag{4.11}$$

As $\mathcal{S} \subseteq \mathcal{T}$, we have $\operatorname{diam}_{\mathbf{X}}(\mathcal{S}) \leqslant \operatorname{diam}_{\mathbf{X}}(\mathcal{T}) \leqslant 4$. So, $\{\mathsf{H}_{\mathcal{S}}^{k,z}\}_{z\in\Lambda}$ are pairwise disjoint for every $k \in \mathbb{N}$ by the case $\mathcal{M} = \mathcal{S}$ of the above reasoning. Recalling (4.8), this implies that for every $k \in \mathbb{N}$ and distinct $w, z \in \Lambda$,

$$\mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,w} \cap \mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z} = \varnothing.$$

To establish that $\{\mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z}\}_{(k,z)\in\mathbb{N}\times\Lambda}$ are pairwise disjoint it therefore remains to check that

$$\mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z} \cap \mathsf{K}_{\mathcal{S},\mathcal{T}}^{j,w} = \varnothing$$

for every $j, k \in \mathbb{N}$ with $j < k$ and any $w, z \in \Lambda$. This is so because if $(\vec{x}, \vec{\gamma}) \in \mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z}$, then $(\vec{x}, \vec{\gamma}) \notin \mathsf{H}_{\mathcal{T}}^{j,w}$ by (4.8). Therefore, either $\chi(w) \neq \gamma_j$ or $w + x_j \notin \mathcal{T} \supseteq \mathcal{S}$. Consequently,

$$(\vec{x}, \vec{\gamma}) \notin \mathsf{H}_{\mathcal{S}}^{j,w} \supseteq \mathsf{K}_{\mathcal{S},\mathcal{T}}^{j,w}.$$

This concludes the verification of the disjointness of $\{\mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z}\}_{(k,z)\in\mathbb{N}\times\Lambda}$.

As for every $k \in \mathbb{N}$ and $z \in \Lambda$, the membership of $(\vec{x}, \vec{\gamma}) \in \{1+, \ldots, 5^{2n}\}^{\mathbb{N}} \times \mathcal{V}^{\mathbb{N}}$ in $\mathsf{H}_{\mathcal{S}}^{k,z}$ and $\mathsf{H}_{\mathcal{T}}^{k,z}$ depends only on the kth coordinates of $\vec{x}$ and $\vec{\gamma}$, it follows from the independence of the coordinates that

$$\begin{aligned}\mathbf{Prob}\big[\mathsf{K}_{\mathcal{S},\mathcal{T}}^{k,z}\big] &\overset{(4.8)}{=} \mathbf{Prob}\left[\mathsf{H}_{\mathcal{S}}^{k,z} \cap \left(\bigcap_{j=1}^{k-1}\Big(\Omega \smallsetminus \bigcup_{w\in\Lambda} \mathsf{H}_{\mathcal{T}}^{j,w}\Big)\right)\right]\\ &= \mathbf{Prob}\big[\mathsf{H}_{\mathcal{S}}^{k,z}\big] \prod_{j=1}^{k-1}\Big(1 - \mathbf{Prob}\Big[\bigcup_{w\in\Lambda} \mathsf{H}_{\mathcal{T}}^{j,w}\Big]\Big)\\ &\overset{(4.10)\wedge(4.11)}{=} \frac{\operatorname{vol}_n\big((\mathcal{V}+z)\cap\mathcal{S}\big)}{5^{2n}\operatorname{vol}_n(\mathcal{V})}\left(1 - \frac{\operatorname{vol}_n(\mathcal{T})}{5^{2n}\operatorname{vol}_n(\mathcal{V})}\right)^{k-1}. \end{aligned} \tag{4.12}$$

Hence, since we already checked that $\{\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}}\}_{(k,z)\in\mathbb{N}\times\Lambda}$ are pairwise disjoint,

$$\begin{aligned}
&\mathbf{Prob}\left[\bigcup_{k=1}^{\infty}\bigcup_{z\in\Lambda}\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}}\right]\\
&\quad=\sum_{k=1}^{\infty}\sum_{z\in\Lambda}\mathbf{Prob}\big[\mathsf{K}^{k,z}_{\mathcal{S},\mathcal{T}}\big]\\
&\quad\overset{(4.12)}{=}\frac{1}{5^{2n}\operatorname{vol}_n(\mathcal{V})}\Big(\sum_{z\in\Lambda}\operatorname{vol}_n\big((\mathcal{V}+z)\cap\mathcal{S}\big)\Big)\sum_{k=1}^{\infty}\left(1-\frac{\operatorname{vol}_n(\mathcal{T})}{5^{2n}\operatorname{vol}_n(\mathcal{V})}\right)^{k-1}\\
&\quad=\frac{\operatorname{vol}_n(\mathcal{S})}{\operatorname{vol}_n(\mathcal{T})},
\end{aligned}$$

where in the final step we used once more the fact that the sets $\{w+\mathcal{V}\}_{w\in\Lambda}$ cover $\mathbb{R}^n$ and have pairwise disjoint interiors. This completes the verification of the desired identity (4.9). ∎

The following lemma is a computation of the probability of the “padding event” corresponding to the random partition $\mathcal{P}$, as a consequence of Lemma 129. In [208] a similar argument was carried out for general finite metric spaces, but it relied on a different random partition in which the radius of the balls is also a random variable (namely, the partition of [71]). This subtlety is circumvented here by using properties of normed spaces that are not available in the full generality of [208].

Lemma 130. *Let $\mathcal{P}$ be the random partition of Lemma 127. For every $\rho\in(0,1)$ and $u\in\mathbb{R}^n$ we have*

$$\mathbf{Prob}\big[u+\rho B_{\mathbf{X}}\subseteq\mathcal{P}(u)\big]=\left(\frac{1-\rho}{1+\rho}\right)^n. \tag{4.13}$$

Proof. For every $k\in\mathbb{N}$, $z\in\Lambda$ and $r\in(0,\infty)$ define $\mathcal{E}^{k,z}_{u,r},\mathcal{F}^{k,z}_{u,r}\subseteq\Omega$ by

$$\mathcal{E}^{k,z}_{u,r}\overset{\mathrm{def}}{=}\mathsf{H}^{k,z}_{u+rB_{\mathbf{X}}}\quad\text{and}\quad\mathcal{F}^{k,z}_{u,r}\overset{\mathrm{def}}{=}\mathsf{K}^{k,z}_{u+(1-r)B_{\mathbf{X}},u+(1+r)B_{\mathbf{X}}}, \tag{4.14}$$

i.e., we are using here the notations of Definition 128 for the following sets:

$$\mathcal{M}=u+rB_{\mathbf{X}},\quad \mathcal{S}=u+(1-r)B_{\mathbf{X}},\quad\text{and}\quad\mathcal{T}=u+(1+r)B_{\mathbf{X}}.$$

We claim that

$$\forall(k,z)\in\mathbb{N}\times\Lambda,\quad\big\{(\vec{x},\vec{\gamma})\in\Omega:\Gamma^{k,z}(\vec{x},\vec{\gamma})\supseteq u+\rho B_{\mathbf{X}}\big\}=\mathcal{F}^{k,z}_{u,\rho}. \tag{4.15}$$

As $u+(1-\rho)B_{\mathbf{X}}\subseteq u+(1+\rho)B_{\mathbf{X}}$ and

$$\operatorname{diam}_{\mathbf{X}}(u+(1+\rho)B_{\mathbf{X}})=2(1+\rho)\leqslant 4,$$

once (4.15) is proven we could apply Lemma 129 to deduce the desired identity (4.13) as follows:

$$\begin{aligned}
&\mathbf{Prob}\big[u+\rho B_{\mathbf{X}}\subseteq \mathcal{P}(u)\big]\\
&\overset{(4.3)}{=} \mathbf{Prob}\Big[\big\{(\vec{x},\vec{\gamma})\in\Omega : \exists (k,z)\in\mathbb{N}\times\Lambda,\ \Gamma^{k,z}(\vec{x},\vec{\gamma})\supseteq u+\rho B_{\mathbf{X}}\big\}\Big]\\
&\overset{(4.15)}{=} \mathbf{Prob}\left[\bigcup_{k=1}^{\infty}\bigcup_{z\in\Lambda}\mathcal{F}^{k,z}_{u,\rho}\right] \overset{(4.9)\wedge(4.14)}{=} \frac{\mathrm{vol}_n(u+(1-\rho)B_{\mathbf{X}})}{\mathrm{vol}_n(u+(1+\rho)B_{\mathbf{X}})} = \left(\frac{1-\rho}{1+\rho}\right)^n.
\end{aligned}$$

To establish (4.15), suppose first that $(\vec{x},\vec{\gamma})\in\mathcal{F}^{k,z}_{u,\rho}$. By the definition of $\mathcal{F}^{k,z}_{u,\rho}$ we therefore know that

$$\forall (j,w)\in\{1,\dots,k-1\}\times\Lambda,\quad (\vec{x},\vec{\gamma})\in\mathcal{E}^{k,z}_{u,1-\rho}\quad\text{yet}\quad (\vec{x},\vec{\gamma})\notin\mathcal{E}^{j,w}_{u,1+\rho}.$$

Hence, by the definition of $\mathcal{E}^{j,w}_{u,1-\rho}$ we know that

$$\chi(z)=\gamma_k \quad\text{and}\quad z+x_k\in u+(1-\rho)B_{\mathbf{X}},$$

which (using the triangle inequality), implies that $z+x_k+B_{\mathbf{X}}\supseteq u+\rho B_{\mathbf{X}}$. At the same time, if $j\in\{1,\dots,k-1\}$ and $w\in\Lambda$, then by the definition of $\mathcal{E}^{j,w}_{u,1+\rho}$, the fact that $(\vec{x},\vec{\gamma})\notin\mathcal{E}^{j,w}_{u,1+\rho}$ means that if $\chi(w)=\gamma_j$ then necessarily $\|w+x_j-u\|_{\mathbf{X}}>1+\rho$, which (using the triangle inequality) implies that $(w+x_j+B_{\mathbf{X}})\cap(u+\rho B_{\mathbf{X}})=\varnothing$. Hence, the ball $u+\rho B_{\mathbf{X}}$ does not intersect the union of the balls

$$\big\{w+x_j+B_{\mathbf{X}} : (j,w)\in\{1,\dots,k-1\}\times\Lambda\wedge\chi(w)=\gamma_j\big\}.$$

Since $\chi(z)=\gamma_k$, due to (4.3), this implies that

$$\Gamma^{k,z}(\vec{x},\vec{\gamma})\cap(u+\rho B_{\mathbf{X}})=(z+x_k+B_{\mathbf{X}})\cap(u+\rho B_{\mathbf{X}})=u+\rho B_{\mathbf{X}},$$

i.e., $(\vec{x},\vec{\gamma})$ belongs to the left-hand side of (4.15).

To establish the reverse inclusion, suppose that $\Gamma^{k,z}(\vec{x},\vec{\gamma})\supseteq u+\rho B_{\mathbf{X}}$. The definition (4.3) implies in particular that $\Gamma^{k,z}(\vec{x},\vec{\gamma})\subseteq z+x_k+B_{\mathbf{X}}$ and that for $\Gamma^{k,z}(\vec{x},\vec{\gamma})$ to be nonempty we must have $\chi(z)=\gamma_k$. So, we know that $z+x_k+B_{\mathbf{X}}\supseteq u+\rho B_{\mathbf{X}}$ and $\chi(z)=\gamma_k$. Assuming first that $z+x_k\neq u$, consider the vector

$$v=u+\frac{\rho}{\|u-z-x_k\|_{\mathbf{X}}}(u-z-x_k).$$

Then, $v\in u+\rho B_{\mathbf{X}}$ and hence also $v\in z+x_k+B_{\mathbf{X}}$, i.e.,

$$1\geqslant\|v-z-x_k\|_{\mathbf{X}}=\|u-z-x_k\|_{\mathbf{X}}+\rho.$$

This shows that $\|z+x_k-u\|_{\mathbf{X}}\leqslant 1-\rho$, i.e., $z+x_k\in u+(1-\rho)B_{\mathbf{X}}$. We obtained this conclusion under the assumption that $z+x_k\neq u$, but it of course holds trivially also when $z+x_k=u$. We have thus shown that $(\vec{x},\vec{\gamma})\in\mathcal{E}^{k,z}_{u,1-\rho}$.

By the definition of $\mathcal{F}^{k,z}_{u,\rho}$, it remains to check that

$$\forall (j,w) \in \{1,\dots,k-1\} \times \Lambda, \quad \big(\vec{x},\vec{\gamma}\big) \notin \mathcal{E}^{j,w}_{u,1+\rho}. \tag{4.16}$$

Assume for the purpose of obtaining a contradiction that (4.16) does not hold. Then, let $j_{\min}$ be the minimum $j \in \{1,\dots,k-1\}$ for which $(\vec{x},\vec{\gamma}) \in \mathcal{E}^{j,w}_{u,1+\rho}$ for some $w \in \Lambda$. Hence, $\chi(w) = \gamma_{j_{\min}}$ and $w + x_{j_{\min}} \in u + (1+\rho)B_{\mathbf{X}}$. If $w + x_{j_{\min}} \neq u$, then the vector

$$u + \frac{\rho}{\|w + x_{j_{\min}} - u\|_{\mathbf{X}}}(w + x_{j_{\min}} - u)$$

is at $\mathbf{X}$-distance ρ from u and also at $\mathbf{X}$-distance $|\rho - \|w + x_{j_{\min}} - u\|_{\mathbf{X}}| \leqslant 1$ from $w + x_{j_{\min}}$, where we used the fact that $\|w + x_{j_{\min}} - u\|_{\mathbf{X}} \leqslant 1+\rho$. We have thus shown that $(w + x_{j_{\min}} + B_{\mathbf{X}}) \cap (u + \rho B_{\mathbf{X}}) \neq \varnothing$ under the assumption $w + x_{j_{\min}} \neq u$, and this assertion trivially holds also if $w + x_{j_{\min}} = u$. By the minimality of the index $j_{\min}$, for every $j \in \{1,\dots,j_{\min}-1\}$ and every $w' \in \Lambda$ with $\chi(w') = \gamma_j$ we have $w' + x_j \notin u + (1+\rho)B_{\mathbf{X}}$, i.e., $\|w' + x_j - u\|_{\mathbf{X}} > 1+\rho$. Hence, by the triangle inequality $(w' + x_j + B_{\mathbf{X}}) \cap (u + \rho B_{\mathbf{X}}) = \varnothing$. The definition of $\Gamma^{j_{\min},w}(\vec{x},\vec{\gamma})$ now shows that $(u + \rho B_{\mathbf{X}}) \cap \Gamma^{j_{\min},w}(\vec{x},\vec{\gamma}) \neq \varnothing$, and since by Lemma 127 we know that $\Gamma^{j_{\min},w}(\vec{x},\vec{\gamma})$ and $\Gamma^{k,z}(\vec{x},\vec{\gamma})$ are disjoint (as $j_{\min} < k$), this contradicts the premise $\Gamma^{k,z}(\vec{x},\vec{\gamma}) \supseteq u + \rho B_{\mathbf{X}}$. ∎

The probability of the "separation event" corresponding to the random partition $\mathcal{P}$ is estimated in the following lemma by using Lemma 129, together with input from Brunn–Minkowski theory.

Lemma 131. *Let $\mathcal{P}$ be the random partition of Lemma 127. For every $u, v \in \mathbb{R}^n$ we have*

$$\mathbf{Prob}\big[\mathcal{P}(u) \neq \mathcal{P}(v)\big] \asymp \min\left\{1, \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(u-v)^\perp}(B_{\mathbf{X}})\big)}{\mathrm{vol}_n(B_{\mathbf{X}})}\|u-v\|_{\ell_2^n}\right\}. \tag{4.17}$$

More precisely, if we denote $\psi(0) = 0$ and

$$\forall w \in \mathbb{R}^n \smallsetminus \{0\}, \quad \psi(w) \stackrel{\mathrm{def}}{=} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{w^\perp}(B_{\mathbf{X}})\big)}{\mathrm{vol}_n(B_{\mathbf{X}})}\|w\|_{\ell_2^n} = \frac{\|w\|_{\Pi^*\mathbf{X}}}{\mathrm{vol}_n(B_{\mathbf{X}})}, \tag{4.18}$$

then for every $u, v \in \mathbb{R}^n$ we have

$$\frac{2e^{\psi(u-v)} - 2}{2e^{\psi(u-v)} - 1} \leqslant \mathbf{Prob}\big[\mathcal{P}(u) \neq \mathcal{P}(v)\big] \leqslant \frac{2\psi(u-v)}{1+\psi(u-v)}. \tag{4.19}$$

In particular, (4.19) *implies the following more precise version of* (4.17)*:*

$$\frac{2e-2}{2e-1}\min\{1, \psi(u-v)\} \leqslant \mathbf{Prob}\big[\mathcal{P}(u) \neq \mathcal{P}(v)\big] \leqslant 2\min\{1, \psi(u-v)\}.$$

Moreover, (4.19) *shows that* $\mathbf{Prob}[\mathcal{P}(u) \neq \mathcal{P}(v)] = 2\psi(u-v) + O(\psi(u-v)^2)$ *as* $u \to v$.

Proof. If $\|u-v\|_{\mathbf{X}} > 2$, then $\mathbf{Prob}[\mathcal{P}(u) \neq \mathcal{P}(v)] = 1$ because $\mathcal{P}$ is 2-bounded. Since $(2e^{\psi(u-v)}-2)/(2e^{\psi(u-v)}-1) < 1$, the first inequality in (4.19) holds. By (1.50) we have $\psi(u-v) \geqslant \|u-v\|_{\mathbf{X}}/2 > 1$, so $2\psi(u-v)/(\psi(u-v)+1) > 1$ and hence the second inequality in (4.19) holds. We will therefore assume from now on that $\|u-v\|_{\mathbf{X}} \leqslant 2$.

Denote $\mathcal{I}(u,v) = (u+B_{\mathbf{X}}) \cap (v+B_{\mathbf{X}})$ and $\mathcal{U}(u,v) = (u+B_{\mathbf{X}}) \cup (v+B_{\mathbf{X}})$. We claim that

$$\forall (k,z) \in \mathbb{N} \times \Lambda, \quad \big\{(\vec{x},\vec{\gamma}) \in \Omega : \{u,v\} \subseteq \Gamma^{k,z}(\vec{x},\vec{\gamma})\big\} = \mathsf{K}^{k,z}_{\mathcal{I}(u,v),\mathcal{U}(u,v)}, \tag{4.20}$$

where we recall the notation that was introduced in Definition 128. Assuming (4.20) for the moment, we will next explain how to conclude the proof of Lemma 131.

Note that $\mathcal{I}(u,v) \subseteq \mathcal{U}(u,v)$ and $\operatorname{diam}_{\mathbf{X}}(\mathcal{U}(u,v)) \leqslant \|u-v\|_{\mathbf{X}} + 2\operatorname{diam}_{\mathbf{X}}(B_{\mathbf{X}}) \leqslant 4$. Consequently, by Lemma 129,

$$\begin{aligned}
&\mathbf{Prob}\big[\mathcal{P}(u) = \mathcal{P}(v)\big] \\
&\qquad \overset{(4.3)}{=} \mathbf{Prob}\Big[\big\{(\vec{x},\vec{\gamma}) \in \Omega : \exists (k,z) \in \mathbb{N} \times \Lambda,\ \{u,v\} \subseteq \Gamma^{k,z}\big(\vec{x},\vec{\gamma}\big)\big\}\Big] \\
&\qquad \overset{(4.20)}{=} \mathbf{Prob}\left[\bigcup_{k=1}^{\infty} \bigcup_{z\in\Lambda} \mathsf{K}^{k,z}_{\mathcal{I}(u,v),\mathcal{U}(u,v)}\right] \\
&\qquad \overset{(4.9)}{=} \frac{\operatorname{vol}_n\big(\mathcal{I}(u,v)\big)}{\operatorname{vol}_n\big(\mathcal{U}(u,v)\big)} \\
&\qquad = \frac{\operatorname{vol}_n\big((u+B_{\mathbf{X}}) \cap (v+B_{\mathbf{X}})\big)}{2\operatorname{vol}_n(B_{\mathbf{X}}) - \operatorname{vol}_n\big((u+B_{\mathbf{X}}) \cap (v+B_{\mathbf{X}})\big)}.
\end{aligned}$$

Hence,

$$\mathbf{Prob}\big[\mathcal{P}(u) \neq \mathcal{P}(v)\big] = \frac{2 - 2\frac{\operatorname{vol}_n((u+B_{\mathbf{X}})\cap(v+B_{\mathbf{X}}))}{\operatorname{vol}_n(B_{\mathbf{X}})}}{2 - \frac{\operatorname{vol}_n((u+B_{\mathbf{X}})\cap(v+B_{\mathbf{X}}))}{\operatorname{vol}_n(B_{\mathbf{X}})}}. \tag{4.21}$$

Now, by the work [280, Corollary 1] of Schmuckenschläger we have the following general estimates:

$$1 - \psi(u-v) \leqslant \frac{\operatorname{vol}_n\big((u+B_{\mathbf{X}}) \cap (v+B_{\mathbf{X}})\big)}{\operatorname{vol}_n(B_{\mathbf{X}})} \leqslant e^{-\psi(u-v)}, \tag{4.22}$$

where $\psi(\cdot)$ is defined in (4.18). The mapping $t \mapsto (2-2t)/(2-t)$ is decreasing on $[0,1]$, so (4.19) is consequence of (4.21) and (4.22). The remaining assertions of Lemma 131 (in particular the asymptotic evaluation (4.17) of the separation probability) follow from (4.19) by elementary calculus. Observe that for the purpose of bounding the separation modulus of $\mathbf{X}$ from above, we need only the first inequality in (4.22); since it is stated in [280] but not proved there, for completeness we will include its elementary proof in Section 4.1.1 below. The second inequality in (4.22) is used here only to show that our bounds are sharp; its proof in [280] relies on a more substantial use of Brunn–Minkowski theory.

It remains to verify (4.20). Fix $(k,z) \in \mathbb{N} \times \Lambda$. Suppose first that $(\vec{x},\vec{\gamma})$ is an element of the right-hand side of (4.20). Recalling the definitions (4.7) and (4.8), this implies that $\chi(z) = \gamma_k$ and $z + x_k \in (u + B_{\mathbf{X}}) \cap (v + B_{\mathbf{X}})$, while for every $j \in \{1,\ldots,k-1\}$ and $w \in \Lambda$ with $\chi(w) = \gamma_j$ we have $w + x_j \notin (u + B_{\mathbf{X}}) \cup (v + B_{\mathbf{X}})$. By the triangle inequality these facts imply that $z + x_k + B_{\mathbf{X}} \supseteq \{u, v\}$ and the union of the balls

$$\left\{w + x_j + B_{\mathbf{X}} : (j,w) \in \{1,\ldots,k-1\} \times \Lambda \wedge \chi(w) = \gamma_j\right\}$$

contains neither of the vectors u, v. The definition (4.3) of $\Gamma^{k,z}(\vec{x},\vec{\gamma})$ now shows that $\{u,v\} \subseteq \Gamma^{k,z}(\vec{x},\vec{\gamma})$.

For the reverse inclusion, assume that $\{u,v\} \subseteq \Gamma^{k,z}(\vec{x},\vec{\gamma})$. Then $\chi(z) = \gamma_k$ and $\{u,v\} \subseteq z + x_k + B_{\mathbf{X}}$ by (4.3), which implies that

$$z + x_k \in (u + B_{\mathbf{X}}) \cap (v + B_{\mathbf{X}}) = \mathcal{I}(u,v).$$

If there were $j \in \{1,\ldots,k-1\}$ and $w \in \Lambda$ such that $(w + x_j + B_{\mathbf{X}}) \cap \{u,v\} \neq \varnothing$ and $\chi(w) = \gamma_j$, then when one subtracts $w + x_j + B_{\mathbf{X}}$ from $z + x_k + B_{\mathbf{X}}$ one removes at least one of the vectors u, v, which by (4.3) would mean that one of these two vectors is not an element of $\Gamma^{k,z}(\vec{x},\vec{\gamma})$, in contradiction to our assumption. Hence for all $j \in \{1,\ldots,k-1\}$ and $w \in \Lambda$ with $\chi(w) = \gamma_j$ we have $u \notin w + x_j + B_{\mathbf{X}}$ and $v \notin w + x_j + B_{\mathbf{X}}$, i.e., $w + x_j \notin (u + B_{\mathbf{X}}) \cup (v + B_{\mathbf{X}}) = \mathcal{U}(u,v)$. This shows that $(\vec{x},\vec{\gamma})$ belongs to the right-hand side of (4.20), thus completing the proof of Lemma 131. ■

Proof of Theorem 122. By rescaling, namely considering the norm $(2/\Delta)\|\cdot\|_{\mathbf{X}}$, it suffices to treat the case $\Delta = 2$. The desired random partition will then be the partition $\mathcal{P}$ of Lemma 127 and the conclusions of Theorem 122 follow from Lemma 130 and Lemma 131. ■

4.1.1 Proof of the first inequality in (4.22)

The proof of the first inequality in (4.22) is a simple and elementary application of standard reasoning using Fubini's theorem. Denote

$$t \stackrel{\mathrm{def}}{=} \|v - u\|_{\ell_2^n} \quad \text{and} \quad x \stackrel{\mathrm{def}}{=} \frac{1}{t}(v - u) \in S^{n-1}.$$

Then,

$$\mathrm{vol}_n\big((u + B_{\mathbf{X}}) \cap (v + B_{\mathbf{X}})\big) = \mathrm{vol}_n\big(B_{\mathbf{X}} \cap (tx + B_{\mathbf{X}})\big),$$

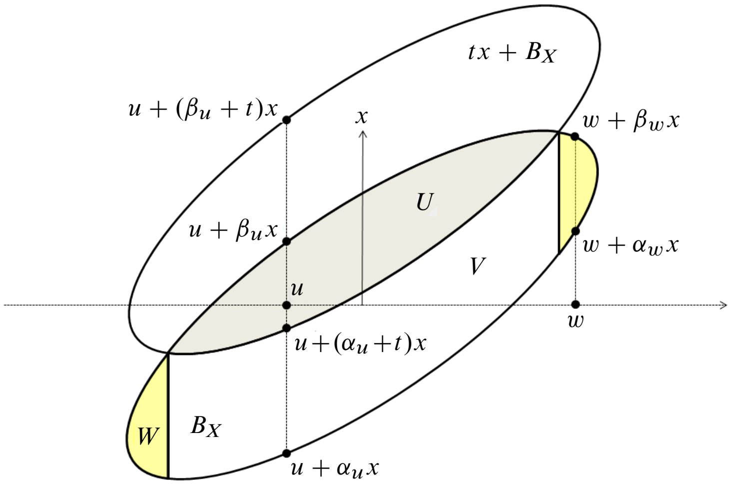

Figure 4.1. A schematic depiction of the partition of $B_{\mathbf{X}}$ into the sets U, V, W (with the sets U, W shaded), as well as the line segments parallel to x that are used in the justification of the estimate (4.23).

The desired estimate is therefore equivalent to the following assertion:

$$\mathrm{vol}_n(B_{\mathbf{X}}) \leqslant \mathrm{vol}_n\big(B_{\mathbf{X}} \cap (tx + B_{\mathbf{X}})\big) + t \cdot \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})\big). \tag{4.23}$$

To prove (4.23), partition $B_{\mathbf{X}}$ into the following three sets:

$$U \stackrel{\mathrm{def}}{=} B_{\mathbf{X}} \cap (tx + B_{\mathbf{X}}), \tag{4.24}$$

$$V \stackrel{\mathrm{def}}{=} \big\{y \in B_{\mathbf{X}} \smallsetminus (tx + B_{\mathbf{X}}) : \mathsf{Proj}_{x^\perp}(y) \in \mathsf{Proj}_{x^\perp}(U)\big\}, \tag{4.25}$$

$$W \stackrel{\mathrm{def}}{=} B_{\mathbf{X}} \smallsetminus (U \cup V) = \big\{y \in B_{\mathbf{X}} : \mathsf{Proj}_{x^\perp}(y) \notin \mathsf{Proj}_{x^\perp}(U)\big\}. \tag{4.26}$$

A schematic depiction of this partition, as well as the notation of ensuing discussion, appears in Figure 4.1. We recommend examining Figure 4.1 while reading the following reasoning because it consists of a formal justification of a situation that is clear when one keeps the geometric picture in mind.

For $z \in \mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})$ let $\alpha_z \in \mathbb{R}$ be the smallest real number such that $z + \alpha_z x \in B_{\mathbf{X}}$ and let $\beta_z \in \mathbb{R}$ be the largest real number such that $z + \beta_z x \in B_{\mathbf{X}}$. Thus the intersection of the line $z + \mathbb{R}x$ with $B_{\mathbf{X}}$ is the segment $w + [\alpha_z, \beta_z]x \subseteq \mathbb{R}^n$. Since $\|x\|_{\ell_2^n} = 1$, by Fubini's theorem we have

$$\begin{aligned}\mathrm{vol}_n(B_{\mathbf{X}}) &= \int_{\mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})} (\beta_z - \alpha_z)\,\mathrm{d}z \\ &= \int_{\mathsf{Proj}_{x^\perp}(U)} (\beta_u - \alpha_u)\,\mathrm{d}u + \int_{\mathsf{Proj}_{x^\perp}(W)} (\beta_w - \alpha_w)\,\mathrm{d}w.\end{aligned} \tag{4.27}$$

To see why the final step of (4.27) holds, simply observe that by (4.26) we have $\mathsf{Proj}_{x^\perp}(B_\mathbf{X}) = \mathsf{Proj}_{x^\perp}(U) \cup \mathsf{Proj}_{x^\perp}(W)$, and the sets $\mathsf{Proj}_{x^\perp}(U), \mathsf{Proj}_{x^\perp}(W)$ have disjoint interiors (in the subspace $x^\perp$).

Since $U = B_\mathbf{X} \cap (tx + B_\mathbf{X})$ is convex, for every u in the interior of $\mathsf{Proj}_{x^\perp}(U)$ the line $u + \mathbb{R}x$ intersects U in an interval, say $(u + \mathbb{R}x) \cap U = u + [\gamma_u, \delta_u]x$ with $\gamma_u, \delta_u \in \mathbb{R}$ satisfying $\gamma_u < \delta_u$ such that $u + \gamma_u x, u + \delta_u x \in \partial U$ and $u + sx \in \mathbf{int}(U)$ for every $s \in (\gamma_u, \delta_u)$. Also,

$$(u + \mathbb{R}x) \cap B_\mathbf{X} = u + [\alpha_u, \beta_u]x$$

with $u + \alpha_u x, u + \beta_u x \in \partial B_\mathbf{X}$. Thus $[\gamma_u, \delta_u] \subseteq [\alpha_u, \beta_u]$. Since $u + \gamma_u x \in U \subseteq tx + B_\mathbf{X}$, it follows that $\gamma_w - t \in [\alpha_w, \beta_w]$. But $\gamma_u \in [\alpha_u, \beta_u]$, so $\beta_u - \alpha_u \geqslant t$ and therefore $\alpha_u + t, \beta_u - t \in [\alpha_u, \beta_u]$, or equivalently $u + (\alpha_u + t)x, u + (\beta_u - t)x \in B_\mathbf{X}$. As $u + \alpha_u x, u + \beta_u x \in \partial B_\mathbf{X}$, we have $u + (\alpha_u + t)x \in B_\mathbf{X} \cap (tx + \partial B_\mathbf{X}) \subseteq \partial U$ and $u + \beta_u x \in (\partial B_\mathbf{X}) \cap (tx + B_\mathbf{X}) \subseteq \partial U$. Hence $\gamma_u = \alpha_u + t$ and $\delta_u = \beta_u$, from which we conclude that

$$u \in \mathsf{Proj}_{x^\perp}(U) \implies (u + \mathbb{R}x) \cap U = u + [\alpha_u + t, \beta_u]x, \tag{4.28}$$

and therefore also

$$\begin{aligned} u \in \mathsf{Proj}_{x^\perp}(U) \implies (u + \mathbb{R}x) \cap V &\overset{(4.25)}{=} B_\mathbf{X} \smallsetminus \big((u + \mathbb{R}x) \cap U\big) \\ &\overset{(4.28)}{=} u + [\alpha_u, \alpha_u + t]x. \end{aligned} \tag{4.29}$$

Another application of Fubini's theorem now implies that

$$\begin{aligned} &\int_{\mathsf{Proj}_{x^\perp}(U)} (\beta_u - \alpha_u)\, \mathrm{d}u \\ &= \int_{\mathsf{Proj}_{x^\perp}(U)} \mathrm{vol}_1\big((u + \mathbb{R}x) \cap U\big)\, \mathrm{d}u + \int_{\mathsf{Proj}_{x^\perp}(U)} t\, \mathrm{d}u \\ &= \mathrm{vol}_n(U) + t\, \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(U)\big) \\ &= \mathrm{vol}_n(U) + t\Big(\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(B_\mathbf{X})\big) - \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(W)\big)\Big), \end{aligned} \tag{4.30}$$

where the first step of (4.30) uses (4.28) and (4.29) and for the last step of (4.30) recall the definition (4.26).

Observe next that

$$w \in \mathsf{Proj}_{x^\perp}(W) \implies \beta_w - \alpha_w \leqslant t. \tag{4.31}$$

Indeed, if $w \in \mathsf{Proj}_{x^\perp}(W)$ yet $\beta_w - \alpha_w > t$ then $w + (\beta_w - t)x$ belongs to the interval joining $w + \alpha_w x$ and $w + \beta_w x$. We therefore have $w + (\beta_w - t)x \in B_\mathbf{X}$ by the convexity of $B_\mathbf{X}$, or equivalently $w + \beta_w x \in tx + B_\mathbf{X}$. Recalling that $w + \beta_w x \in B_\mathbf{X}$, this means that $w + \beta_w x \in B_\mathbf{X} \cap (tx + B_\mathbf{X})$. By the definition (4.24) of U, it follows that $w \in \mathsf{Proj}_{x^\perp}(U)$. By the definition (4.26) of W, this means that $w \notin \mathsf{Proj}_{x^\perp}(W)$, a contradiction.

Having established (4.31) we see that

$$\int_{\mathsf{Proj}_{x^\perp}(W)} (\beta_w - \alpha_w)\,\mathrm{d}w \overset{(4.31)}{\leqslant} t\,\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(W)\big). \tag{4.32}$$

The estimate (4.23) now follows from a substitution of (4.30) and (4.32) into (4.27). ∎

4.2 Proof of Theorem 81

For any $m \in \mathbb{N}$, because $\mathrm{evr}(\ell_1^m) \asymp \sqrt{m}$, by the second part (2.55) of Theorem 107 there exists $\mathcal{C} \subseteq \mathbb{R}^m$ with $|\mathcal{C}| \leqslant e^{\beta m}$ for some universal constant $\beta > 0$ such that $\mathsf{SEP}(\mathcal{C}_{\ell_1^m}) \gtrsim m$ (as we are considering here ℓ_1^m rather than more general normed spaces, this statement is due [76]). Fix an integer $n \geqslant 2$ and $1 \leqslant p \leqslant 2$. Let m be the largest integer such that $e^{\beta m} \leqslant n$. Thus $m \asymp \log n$ and

$$\mathsf{SEP}^n(\ell_p) \geqslant \mathsf{SEP}\big(\mathcal{C}_{\ell_p^m}\big) \geqslant \frac{\mathsf{SEP}\big(\mathcal{C}_{\ell_1^m}\big)}{d_{\mathrm{BM}}\big(\ell_1^m,\ell_p^m\big)} \gtrsim \frac{m}{d_{\mathrm{BM}}\big(\ell_1^m,\ell_p^m\big)} = m^{\frac{1}{p}} \asymp (\log n)^{\frac{1}{p}}.$$

This proves the lower bound on $\mathsf{SEP}^n(\ell_p)$ in Theorem 81.

It remains to prove the upper bound on $\mathsf{SEP}^n(\ell_p)$ in Theorem 81, i.e., that for all $x_1,\ldots,x_n \in \ell_p$,

$$\mathsf{SEP}\big(\{x_1,\ldots,x_n\},\|\cdot\|_{\ell_p}\big) \lesssim \frac{(\log n)^{\frac{1}{p}}}{p-1}. \tag{4.33}$$

The proof of (4.33) will refer to the following technical probabilistic lemma.

Lemma 132. *Suppose that $p \in (1,\infty)$ and let X be a nonnegative random variable, defined on some probability space $(\Omega, \mathbf{Prob})$, that satisfies the following Laplace transform identity:*

$$\forall u \in [0,\infty), \quad \mathbb{E}\big[e^{-u\mathsf{X}^2}\big] = e^{-u^{\frac{p}{2}}}. \tag{4.34}$$

Then

$$\mathbb{E}[\mathsf{X}] = \frac{\Gamma\big(1-\frac{1}{p}\big)}{\sqrt{\pi}} \asymp \frac{p}{p-1}. \tag{4.35}$$

Moreover, we have

$$\forall t \in (0,\infty), \quad \mathbf{Prob}\big[\mathsf{X} \leqslant t\big] \leqslant \exp\left(-\frac{\big(\frac{p}{2}\big)^{\frac{p}{2-p}}\big(1-\frac{p}{2}\big)}{t^{\frac{2p}{2-p}}}\right). \tag{4.36}$$

Proof. Suppose that $\alpha \in (0,1)$. Then every $x \in (0,\infty)$ satisfies

$$\int_0^\infty \frac{1-e^{-ux}}{u^{1+\alpha}}\,\mathrm{d}x = x^\alpha \int_0^\infty \frac{1-e^{-v}}{v^{1+\alpha}}\,\mathrm{d}x = \frac{\Gamma(1-\alpha)}{\alpha}x^\alpha, \tag{4.37}$$

where the first step of (4.37) is a straightforward change of variable and the last step of (4.37) follows by integration by parts. The case $\alpha = 1/2$ of (4.37) implies (4.35) as follows:

$$\begin{aligned}\mathbb{E}[\mathsf{X}] = \mathbb{E}\bigg[\frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-e^{-u\mathsf{X}^2}}{u^{\frac{3}{2}}}\,\mathrm{d}u\bigg] &= \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-\mathbb{E}\big[e^{-u\mathsf{X}^2}\big]}{u^{\frac{3}{2}}}\,\mathrm{d}u\\ \overset{(4.34)}{=} \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{1-e^{-u^{\frac{p}{2}}}}{u^{\frac{3}{2}}}\,\mathrm{d}u &= \frac{1}{p\sqrt{\pi}}\int_0^\infty \frac{1-e^{-v}}{v^{1+\frac{1}{p}}}\,\mathrm{d}v \overset{(4.37)}{=} \frac{\Gamma\big(1-\frac{1}{p}\big)}{\sqrt{\pi}}.\end{aligned}$$

The small ball probability estimate (4.36) is a consequence of the following standard use of Markov's inequality. For every $u,t \in (0,\infty)$ we have

$$\mathbf{Prob}\big[\mathsf{X} \leqslant t\big] = \mathbf{Prob}\big[e^{-u\mathsf{X}^2} \geqslant e^{-ut^2}\big] \leqslant e^{ut^2}\mathbb{E}\big[e^{-u\mathsf{X}^2}\big] = e^{ut^2-u^{\frac{p}{2}}}. \tag{4.38}$$

The value of $u \in (0,\infty)$ that minimizes the right-hand side of (4.38) is

$$u = u(p,t) \overset{\mathrm{def}}{=} \Big(\frac{p}{2t^2}\Big)^{\frac{2}{2-p}}.$$

A substitution of this value of u into (4.38) simplifies to give the estimate (4.36). ■

Proof of (4.33). Fix distinct $x_1,\ldots,x_n \in \ell_p$. It suffices to prove (4.33) when $p \in (1,2)$, since the quantity that appears in the right-hand side of (4.33) remains bounded as $p \to 2^-$, and every finite subset of ℓ_2 embeds isometrically into ℓ_p for every $p \in [1,2]$ (see, e.g., [314, Chapter III.A]). We will therefore assume in the remainder of the proof of (4.33) that $p \in (1,2)$.

Marcus and Pisier proved [197, Section 2] the following statement, relying on a structural result for p-stable processes; its deduction from the formulation in [197] appears in [169, Lemma 2.1]. There exists a probability space $(\Omega, \mathbf{Prob})$ for which there is a **Prob**-to-Borel measurable mapping $(\omega \in \Omega) \mapsto T_\omega \in \mathfrak{L}(\ell_p,\ell_2)$ (we denote by $\mathfrak{L}(\ell_p,\ell_2)$ the space of bounded operators from ℓ_p to ℓ_2, equipped with the strong operator topology) such that for every $\omega \in \Omega$ and $x \in \ell_p \smallsetminus \{0\}$ the random variable

$$(\omega \in \Omega) \mapsto \frac{\|T_\omega(x)\|_{\ell_2}}{\|x\|_{\ell_p}}$$

has the same distribution as the random variable X of Lemma 132 (in particular, its distribution is independent of the choice of $x \in \ell_p \smallsetminus \{0\}$). Consequently, for every $i,j \in \{1,\ldots,n\}$ we have

$$\int_\Omega \|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2}\,\mathrm{d}\mathbf{Prob}(\omega) = \|x_i - x_j\|_{\ell_p}\mathbb{E}[\mathsf{X}] \overset{(4.35)}{\asymp} \frac{\|x_i - x_j\|_{\ell_p}}{p-1}. \tag{4.39}$$

It also follows from the above discussion and Lemma 132 that for every $t \in (0, \infty)$ we have

$$\begin{aligned}
&\mathbf{Prob}\Big[\bigcup_{i,j\in\{1,\ldots,n\}} \{\omega \in \Omega : \|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2} \geqslant t\|x_i - x_j\|_{\ell_p}\}\Big] \\
&\leqslant \sum_{\substack{i,j\in\{1,\ldots,n\}\\ i\neq j}} \mathbf{Prob}\left[\left\{\omega \in \Omega : \frac{\|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2}}{\|x_i - x_j\|_{\ell_p}} < t\right\}\right] \\
&\overset{(4.36)}{\leqslant} \binom{n}{2} \exp\left(-\frac{\left(\frac{p}{2}\right)^{\frac{p}{2-p}}\left(1-\frac{p}{2}\right)}{t^{\frac{2p}{2-p}}}\right).
\end{aligned} \tag{4.40}$$

If we choose

$$t = t(n,p) \overset{\text{def}}{=} \sqrt{\frac{p}{2}}\left(\frac{2-p}{4\log n}\right)^{\frac{1}{p}-\frac{1}{2}},$$

then the right-hand side of (4.40) becomes less than $1/2$. In other words, this shows that there exists a measurable subset $A \subseteq \Omega$ with $\mathbf{Prob}[A] \geqslant 1/2$ such that for every $\omega \in A$ and $i, j \in \{1, \ldots, n\}$,

$$\begin{aligned}
\|x_i - x_j\|_{\ell_p} &\leqslant \sqrt{\frac{2}{p}}\left(\frac{4\log n}{2-p}\right)^{\frac{1}{p}-\frac{1}{2}} \|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2} \\
&\leqslant 4(\log n)^{\frac{1}{p}-\frac{1}{2}} \|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2}.
\end{aligned} \tag{4.41}$$

The last step of (4.41) uses the elementary inequality

$$\left(\frac{2}{2-p}\right)^{\frac{2-p}{2p}} \sqrt{\frac{2}{p}} \leqslant 4,$$

which holds (with room to spare) for every $p \in [1, 2)$.

$\{T_\omega(x_1), \ldots, T_\omega(x_n)\} \subseteq \ell_2$ is a subset of Hilbert space of size at most n, so by the Johnson–Lindenstrauss dimension reduction lemma [138] there is $k \in \mathbb{N}$ with $k \lesssim \log n$ such that for every $\omega \in \Omega$ there is a linear operator $Q_\omega : \ell_2 \to \mathbb{R}^k$ such that for all $i, j \in \{1, \ldots, n\}$,

$$\begin{aligned}
\|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2} &\leqslant \|Q_\omega T_\omega(x_i) - Q_\omega T_\omega(x_j)\|_{\ell_2^k} \\
&\leqslant 2\|T_\omega(x_i) - T_\omega(x_j)\|_{\ell_2}.
\end{aligned} \tag{4.42}$$

An examination of the proof in [138] reveals that the mapping

$$\omega \mapsto Q_\omega$$

can be taken to be **Prob**-to-Borel measurable, but actually Q_ω can be chosen from a finite set of operators (see, e.g., [2]).

Fix $\Delta \in (0,\infty)$. Since by [76] we have $\mathsf{SEP}(\ell_2^k) \lesssim \sqrt{k}$, there exists a probability space (Θ,μ) and a mapping $\theta \mapsto \mathcal{R}^\theta$ that is a random partition of $\mathbb{R}^k$ for which

$$\forall(\omega,\theta,i) \in \Omega\times\Theta\times\{1,\ldots,n\}, \quad \mathrm{diam}_{\ell_2^k}\left(\mathcal{R}^\theta\big(Q_\omega T_\omega(x_i)\big)\right) \leqslant \frac{\Delta}{4(\log n)^{\frac{1}{p}-\frac{1}{2}}}, \tag{4.43}$$

and also every $\omega \in \Omega$ and $i,j \in \{1,\ldots,n\}$ satisfy

$$\begin{aligned}
&\mu\Big(\big\{\theta \in \Theta : \mathcal{R}^\theta\big(Q_\omega T_\omega(x_i)\big) \neq \mathcal{R}^\theta\big(Q_\omega T_\omega(x_j)\big)\big\}\Big)\\
&\lesssim \frac{\sqrt{k}}{\Delta/\big(4(\log n)^{\frac{1}{p}-\frac{1}{2}}\big)}\big\|Q_\omega T_\omega(x_i)-Q_\omega T_\omega(x_i)\big\|_{\ell_2^k}\\
&\lesssim \frac{(\log n)^{\frac{1}{p}}}{\Delta}\big\|T_\omega(x_i)-T_\omega(x_i)\big\|_{\ell_2},
\end{aligned}\tag{4.44}$$

where the last step of (4.44) uses the right-hand inequality in (4.42) and the fact that $k \lesssim \log n$.

Recalling the set $A \subseteq \Omega$ on which (4.41) holds for every $i,j \in \{1,\ldots,n\}$, let ν be the probability measure on A defined by

$$\nu[E] = \frac{\mathbf{Prob}[E]}{\mathbf{Prob}[A]}$$

for every **Prob**-measurable $E \subseteq A$ (recall that $\mathbf{Prob}[A] \geqslant 1/2$). For every $(\omega,\theta) \in A\times\Theta$ define a partition $\mathcal{P}^{(\omega,\theta)}$ of $\{x_1,\ldots,x_n\}$ by setting for every $i \in \{1,\ldots,n\}$,

$$\mathcal{P}^{(\omega,\theta)}(x_i) \stackrel{\mathrm{def}}{=} \Big\{x \in \{x_1,\ldots,x_n\} : Q_\omega T_\omega(x) \in \mathcal{R}^\theta\big(Q_\omega T_\omega(x_i)\big)\Big\}. \tag{4.45}$$

Then, for every $(\omega,\theta) \in A\times\Theta$ and every $i \in \{1,\ldots,n\}$ we have

$$\begin{aligned}
&\mathrm{diam}_{\ell_p}\big(\mathcal{P}^{(\omega,\theta)}(x_i)\big)\\
&= \max_{\substack{u,v\in\{1,\ldots,n\}\\ Q_\omega T_\omega(x_u),Q_\omega T_\omega(x_v)\in\mathcal{R}^\theta(Q_\omega T_\omega(x_i))}} \|x_u-x_v\|_{\ell_p}\\
&\leqslant 4(\log n)^{\frac{1}{p}-\frac{1}{2}} \max_{\substack{u,v\in\{1,\ldots,n\}\\ Q_\omega T_\omega(x_u),Q_\omega T_\omega(x_v)\in\mathcal{R}^\theta(Q_\omega T_\omega(x_i))}} \|T_\omega(x_u)-T_\omega(x_v)\|_{\ell_2}\\
&\leqslant 4(\log n)^{\frac{1}{p}-\frac{1}{2}} \max_{\substack{u,v\in\{1,\ldots,n\}\\ Q_\omega T_\omega(x_u),Q_\omega T_\omega(x_v)\in\mathcal{R}^\theta(Q_\omega T_\omega(x_i))}} \|Q_\omega T_\omega(x_u)-Q_\omega T_\omega(x_v)\|_{\ell_2^k}\\
&\leqslant 4(\log n)^{\frac{1}{p}-\frac{1}{2}}\,\mathrm{diam}_{\ell_2^k}\left(\mathcal{R}^\theta\big(Q_\omega T_\omega(x_i)\big)\right) \leqslant \Delta,
\end{aligned}\tag{4.46}$$

where the first step of (4.46) uses (4.45), the second step of (4.46) uses (4.41), the third step of (4.46) uses (4.42), and the final step of (4.46) uses (4.43). Also, every distinct $i, j \in \{1, \ldots, n\}$ satisfy

$$
\begin{aligned}
&\nu \times \mu\big(\{(\omega,\theta) \in A \times \Theta : \mathcal{P}^{(\omega,\theta)}(x_i) \neq \mathcal{P}^{(\omega,\theta)}(x_j)\}\big) \\
&\quad = \int_A \mu\big(\{\theta \in \Theta : \mathcal{R}^\theta\big(Q_\omega T_\omega(x_i)\big) \neq \mathcal{R}^\theta\big(Q_\omega T_\omega(x_j)\big)\}\big)\,\mathrm{d}\nu(\omega) \\
&\quad \lesssim \frac{1}{\mathbf{Prob}[A]} \int_A \frac{(\log n)^{\frac{1}{p}}}{\Delta} \big\|T_\omega(x_i) - T_\omega(x_i)\big\|_{\ell_2} \mathrm{d}\mathbf{Prob}(\omega) \\
&\quad \leqslant \frac{2(\log n)^{\frac{1}{p}}}{\Delta} \int_\Omega \big\|T_\omega(x_i) - T_\omega(x_i)\big\|_{\ell_2}\,\mathrm{d}\mathbf{Prob}(\omega) \\
&\quad \lesssim \frac{(\log n)^{\frac{1}{p}}}{p-1} \cdot \frac{\|x_i - x_j\|_{\ell_p}}{\Delta},
\end{aligned}
\tag{4.47}
$$

where the first step of (4.47) uses (4.45), the second step of (4.47) uses (4.44), the third step of (4.47) uses $\mathbf{Prob}[A] \geqslant \frac{1}{2}$, and the last step of (4.47) uses (4.39). By (4.46) and (4.47), the proof of (4.33) is complete. ■