MEMOIRS OF THE EUROPEAN MATHEMATICAL SOCIETY

Assaf Naor

Extension, Separation and Isomorphic Reverse Isoperimetry

MEMS Vol. 11 / 2024

EMS PRESS

EMS PRESS

Memoirs of the European Mathematical Society

The *Memoirs of the European Mathematical Society* publish outstanding research contributions in individual volumes, in all areas of mathematics and with a particular focus on works that are longer and more comprehensive than usual research articles.

The collection's editorial board consists of the editors-in-chief of the *Journal of the European Mathematical Society* and the *EMS Surveys in Mathematical Sciences*, along with editors of book series of the publishing house of the EMS as well as other distinguished mathematicians.

All submitted works go through a highly selective peer-review process.

Previously published in this series:

J.-M. Delort, N. Masmoudi, *Long-Time Dispersive Estimates for Perturbations of a Kink Solution of One-Dimensional Cubic Wave Equations*
G. Cotti, *Cyclic Stratum of Frobenius Manifolds, Borel–Laplace* (α,β)*-Multitransforms, and Integral Representations of Solutions of Quantum Differential Equations*
A. Kostenko, N. Nicolussi, *Laplacians on Infinite Graphs*
A. Carey, F. Gesztesy, G. Levitina, R. Nichols, F. Sukochev, D. Zanin, *The Limiting Absorption Principle for Massless Dirac Operators, Properties of Spectral Shift Functions, and an Application to the Witten Index of Non-Fredholm Operators*
J. Kigami, *Conductive Homogeneity of Compact Metric Spaces and Construction of p-Energy*
A. Buium, L. E. Miller, *Purely Arithmetic PDEs Over a p-Adic Field: δ-Characters and δ-Modular Forms*
M. Duerinckx, A. Gloria, *On Einstein's Effective Viscosity Formula*
R. Willett, G. Yu, *The Universal Coefficient Theorem for C*-Algebras with Finite Complexity*
B. Janssens, K.-H. Neeb, *Positive Energy Representations of Gauge Groups I. Localization*
S.Dipierro, G. Giacomin, E. Valdinoci, *The Lévy Flight Foraging Hypothesis in Bounded Regions*

Assaf Naor

Extension, Separation and Isomorphic Reverse Isoperimetry

EMS PRESS

Author

Assaf Naor
Princeton University
Department of Mathematics
Fine Hall, Washington Road
Princeton, NJ 08544-1000, USA

Email: naor@math.princeton.edu

Each volume of the *Memoirs of the European Mathematical Society* is available individually or as part of an annual subscription. It may be ordered from your bookseller, subscription agency, or directly from the publisher via subscriptions@ems.press.

ISSN 2747-9080, eISSN 2747-9099
ISBN 978-3-98547-069-3, eISBN 978-3-98547-569-8, DOI 10.4171/MEMS/11

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Published by EMS Press, an imprint of the

European Mathematical Society – EMS – Publishing House GmbH
Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin, Germany

https://ems.press

Typesetting: AfricanType, Cairo, Egypt
Printed in Germany
♾ Printed on acid free paper

Abstract

The Lipschitz extension modulus $\mathsf{e}(\mathfrak{M})$ of a metric space $\mathfrak{M}$ is the infimum over those $L \in [1,\infty]$ such that for any Banach space $\mathbf{Z}$ and any $\mathcal{C} \subseteq \mathfrak{M}$, any 1-Lipschitz function $f : \mathcal{C} \to \mathbf{Z}$ can be extended to an L-Lipschitz function $F : \mathfrak{M} \to \mathbf{Z}$. Johnson, Lindenstrauss and Schechtman proved (1986) that if $\mathbf{X}$ is an n-dimensional normed space, then $\mathsf{e}(\mathbf{X}) \lesssim n$. In the reverse direction, we prove that every n-dimensional normed space $\mathbf{X}$ satisfies $\mathsf{e}(\mathbf{X}) \gtrsim n^c$, where $c > 0$ is a universal constant. Our core technical contribution is a geometric structural result on stochastic clustering of finite dimensional normed spaces which implies upper bounds on their Lipschitz extension moduli using an extension method of Lee and the author (2005). The separation modulus of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ is the infimum over those $\sigma \in (0,\infty]$ such that for any $\Delta > 0$ there is a distribution over random partitions of $\mathfrak{M}$ into clusters of diameter at most Δ such that for every two points $x, y \in \mathfrak{M}$ the probability that they belong to different clusters is at most $\sigma d_{\mathfrak{M}}(x,y)/\Delta$. We obtain upper and lower bounds on the separation moduli of finite dimensional normed spaces that relate them to well-studied volumetric invariants (volume ratios and projection bodies). Using these connections, we determine the asymptotic growth rate of the separation moduli of various normed spaces. If $\mathbf{X}$ is an n-dimensional normed space with enough symmetries, then our bounds imply that its separation modulus is equal to $\mathrm{vr}(\mathbf{X}^*)\sqrt{n}$ up to factors of lower order, where $\mathrm{vr}(\mathbf{X}^*)\sqrt{n}(\mathbf{X}^*)$ is the volume ratio of the unit ball of the dual of $\mathbf{X}$. We formulate a conjecture on isomorphic reverse isoperimetric properties of symmetric convex bodies (akin to Ball's reverse isoperimetric theorem (1991), but permitting a non-isometric perturbation in addition to the choice of position) that can be used with our volumetric bounds on the separation modulus to obtain many more exact asymptotic evaluations of the separation moduli of normed spaces. Our estimates on the separation modulus imply asymptotically improved upper bounds on the Lipschitz extension moduli of various classical spaces. In particular, we deduce an improved upper bound on $\mathsf{e}(\ell_p^n)$ when $p > 2$ that resolves a conjecture of Brudnyi and Brudnyi (2005), and we prove that $\mathsf{e}(\ell_\infty^n) \asymp \sqrt{n}$, which is the first time that the growth rate of $\mathsf{e}(\mathbf{X})$ has been evaluated (as $\dim(\mathbf{X}) \to \infty$) for *any* finite dimensional normed space $\mathbf{X}$.

Dedicated with awe to the memory of Jean Bourgain.

Keywords. Lipschitz extension, randomized clustering, convex geometry, local theory of Banach spaces, projection bodies, volume ratios, Wasserstein spaces, spectral geometry, Dirichlet eigenvalues, Cheeger sets, reverse isoperimetry

Mathematics Subject Classification (2020). Primary 51F30; Secondary 46B07, 52A23, 46B06, 46B20, 46B85, 54C20

Acknowledgments. I am grateful to Moses Charikar and Piotr Indyk for a helpful conversation on the erroneous optimality assertion of [76] in the range $p \in (2, \infty]$, and to Piotr Indyk for subsequent conversations and pointers to the literature. I thank Gideon Schechtman for multiple discussions that led in particular to Lemma 173 and (1.85). I thank Emmanuel Breuillard for discussions regarding Remark 41 and for sharing with me his proofs of statements on the comparison between the notions of being canonically positioned and having enough symmetries. I thank Erwin Lutwak for showing me his proof of Proposition 87. I thank Bo'az Klartag and Emanuel Milman for showing me the proof of Proposition 61. I am grateful to Keith Ball, Franck Barthe, Károly Böröczky, Ronen Eldan, Charles Fefferman, Apostolos Giannopoulos, David Jerison, Grigoris Paouris, Gilles Pisier, Oded Regev, Carsten Schütt, Ramon van Handel and Elisabeth Werner for helpful discussions and suggestions. Finally, I am grateful to the anonymous referee for valuable corrections and recommendations.

Funding. Supported by NSF grant DMS-2054875, BSF grants 2010021 and 2018223, the Packard Foundation and the Simons Foundation. Part of this work was conducted under the auspices of the Simons Algorithms and Geometry (A&G) Think Tank. An extended abstract [227] titled "Probabilistic clustering of high dimensional norms" that announces discrete and algorithmic aspects of parts of this work appeared in the proceedings of the 28th annual ACM–SIAM Symposium on Discrete Algorithms.

Contents

Chapter 1

Introduction

Our core technical contribution is a geometric structural result (stochastic clustering) for subsets of finite dimensional normed spaces. It provides new links between nonlinear questions in metric geometry and volumetric issues in convex geometry. An unexpected aspect of our statement is that it contradicts an impossibility result of the well-known work [76] by Charikar, Chekuri, Goel, Guha and Plotkin in the computer science literature, thus leading to bounds that were previously thought to be impossible. This is reconciled in Section 1.7, where we explain the source of the error in [76].

The aforementioned link opens up a vista that allows one to apply the extensive literature on the linear theory to important and well-studied nonlinear questions. It also raises new fundamental issues within the linear theory that we will only begin to address here. So, in order to fully explain both the history and the ideas and their consequences, we will start with a quick overview of some of our main results that assumes familiarity with standard concepts in the respective areas. We will then present a gradual and complete introduction to our work that specifies all of the necessary background.

1.1 Brief highlights of main results

Associate to every separable complete metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ two bi-Lipschitz invariants $\mathsf{e}(\mathfrak{M}), \mathsf{SEP}(\mathfrak{M}) \in (0, \infty]$ called, respectively, the *Lipschitz extension modulus of* $\mathfrak{M}$ and the *separation modulus of* $\mathfrak{M}$, that are defined as follows. The Lipschitz extension modulus of $\mathfrak{M}$ is the infimum over those $L \in (0, \infty]$ such that for every Banach space $\mathbf{Z}$ and every subset $\mathcal{C} \subseteq \mathfrak{M}$, every 1-Lipschitz function $f : \mathcal{C} \to \mathbf{Z}$ can be extended to a $\mathbf{Z}$-valued L-Lipschitz function that is defined on all of $\mathfrak{M}$. The separation modulus of $\mathfrak{M}$ is the infimum over those $\sigma \in (0, \infty]$ such that for any $\Delta > 0$ there is a distribution over random partitions[1] of $\mathfrak{M}$ into clusters of diameter at most Δ such that for every two points $x, y \in \mathfrak{M}$ the probability that they belong to different clusters is at most $\sigma d_{\mathfrak{M}}(x, y)/\Delta$.

The question of estimating the Lipschitz extension modulus received great scrutiny over the past century; see Section 1.3 for an indication of (a small part of) the

[1] We are suppressing here measurability issues that are addressed in Section 1.7 and Section 3.1.

extensive knowledge on this topic. The separation modulus was introduced by Bartal in the mid-1990s and received a lot of attention in the computer science literature due to its algorithmic applications; see Section 1.7.3 for the history. Its connection to Lipschitz extension was found by Lee and the author [171, 173], who proved that

$$\mathsf{e}(\mathfrak{M}) \lesssim \mathsf{SEP}(\mathfrak{M}).$$

By a well-known theorem of Johnson, Lindenstrauss and Schechtman [140], every normed space $\mathbf{X}$ satisfies $\mathsf{e}(\mathbf{X}) = O(\dim(\mathbf{X}))$. Here we obtain a power-type lower bound on $\mathsf{e}(\mathbf{X})$ in terms of $\dim(\mathbf{X})$.

Theorem 1. *There is a universal constant $c > 0$ such that $\mathsf{e}(\mathbf{X}) \geqslant \dim(\mathbf{X})^c$ for every normed space $\mathbf{X}$.*

Theorem 1 improves over the previously best-available bound

$$\mathsf{e}(\mathbf{X}) \geqslant e^{c\sqrt{\log \dim(\mathbf{X})}};$$

see Remark 98 for the history of this question. Despite substantial efforts, the asymptotic growth rate (as $\dim(\mathbf{X}) \to \infty$) of $\mathsf{e}(\mathbf{X})$ was not previously known (even up to lower order factors) for *any* sequence of normed spaces.

Theorem 2. *For every $n \in \mathbb{N}$ we have*[2] $\mathsf{e}(\ell_\infty^n) \asymp \sqrt{n}$.

The previously best-known upper bound on $\mathsf{e}(\ell_\infty^n)$ was nothing better than the aforementioned general $O(n)$ bound of [140]. Theorem 2 is just one instance of our asymptotically improved upper bounds on the Lipschitz extension moduli of many normed spaces of interest; we also get, e.g., the best-known bound when $\mathbf{X} = \ell_p^n$ for any $p > 2$. Nevertheless, currently ℓ_∞^n is essentially[3] the only normed space whose Lipschitz extension modulus is known up to lower order factors (by Theorem 2), and the same question even for the Euclidean space ℓ_2^n remains a well-known longstanding open problem; see Section 1.3 for more on this.

All of the upper bounds on the Lipschitz extension modulus that we obtain herein use the upper bound on the separation modulus that appears in Theorem 3 below. This theorem also contains a new lower bound on the separation modulus, which we

[2] We use the following conventions for asymptotic notation, in addition to the usual $O(\cdot), o(\cdot), \Omega(\cdot)$ notation. Given $a, b > 0$, by writing $a \lesssim b$ or $b \gtrsim a$ we mean that $a \leqslant Cb$ for some universal constant $C > 0$, and $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$. If we need to allow for dependence on parameters, we indicate it by subscripts. For example, in the presence of an auxiliary parameter q, the notation $a \lesssim_q b$ means that $a \leqslant C(q)b$, where $C(q) > 0$ may depend only on q, and similarly for $a \gtrsim_q b$ and $a \asymp_q b$.

[3] The proof of Theorem 2 artificially gives more such spaces, e.g., $\ell_\infty^n \oplus \ell_2^n$, or $\ell_\infty^n \oplus \mathbf{X}$ for any normed space $\mathbf{X}$ with $\dim(\mathbf{X}) \leqslant \sqrt{n}$.

will see shows that in several cases of interest our results are a sharp evaluation of the asymptotic growth rate of the separation modulus.[4]

Theorem 3. *Let* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ *and* $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ *be normed spaces whose unit balls satisfy* $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$*. Then*

$$\mathrm{vr}(\mathbf{X}^*)\sqrt{n} \lesssim \mathsf{SEP}(\mathbf{X}) \lesssim \frac{\mathrm{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})}. \tag{1.1}$$

In the left-hand side of (1.1), $\mathrm{vr}(\mathbf{X}^*)$ is the *volume ratio* [293, 294] of the dual $\mathbf{X}^*$, i.e., it is the nth root of the ratio of the volume of $B_{\mathbf{X}^*}$ and maximal volume of an ellipsoid that is contained in $B_{\mathbf{X}^*}$. In the right-hand side of (1.1), $\Pi B_{\mathbf{Y}}$ is the *projection body* [251] of $B_{\mathbf{Y}}$, and $\mathrm{diam}_{\mathbf{X}^*}(\cdot)$ denotes diameter with respect to the metric on $\mathbb{R}^n$ that is induced by $\mathbf{X}^*$. We will recall the definition of a projection body later[5] and it suffices to mention now that the mapping $K \mapsto \Pi K$, which is of central importance in convex geometry (see [47, 102, 190, 282] for an indication of the extensive literature on this topic), associates to every convex body $K \subseteq \mathbb{R}^n$ a convex body $\Pi K \subseteq \mathbb{R}^n$ that encodes isoperimetric properties of K.

A key contribution of Theorem 3 is the role of the auxiliary normed space $\mathbf{Y}$, which appears despite the fact that we are interested in the separation modulus of $\mathbf{X}$. By substituting $\mathbf{Y} = \mathbf{X}$ into the right-hand side of (1.1) one *does* get a meaningful estimate, and in particular the resulting bound is $O(n)$, i.e., (1.1) implies the bound of [140]. However, we will see that by introducing a suitable perturbation $\mathbf{Y}$ of $\mathbf{X}$, the second inequality in (1.1) can sometimes be significantly stronger than the special case $\mathbf{Y} = \mathbf{X}$. We will exploit this powerful degree of freedom heavily; its geometric significance is discussed in Section 1.4.

The previously best-known upper and lower estimates on the separation moduli of normed spaces are due to [76], where it was proved that

$$\mathsf{SEP}(\ell_1^n) \asymp n \quad \text{and} \quad \mathsf{SEP}(\ell_2^n) \asymp \sqrt{n}.$$

By bi-Lipschitz invariance, this implies that any n-dimensional normed space $\mathbf{X}$ satisfies

$$\frac{n}{d_{\mathrm{BM}}(\ell_1^n, \mathbf{X})} \lesssim \mathsf{SEP}(\mathbf{X}) \lesssim d_{\mathrm{BM}}\big(\ell_2^n, \mathbf{X}\big)\sqrt{n}, \tag{1.2}$$

[4] Our approach also pertains to subsets of normed spaces, e.g., we will prove that for any $p \in [1, \infty]$, $n \in \mathbb{N}$ and $r \in \{1, \ldots, n\}$, the separation modulus of the set of n-by-n matrices of rank at most r, equipped with the Schatten–von Neumann-p norm, is equal up to lower order factors to $\max\{\sqrt{r}, r^{1/p}\}\sqrt{n}$, which is new even in the Euclidean (Hilbert–Schmidt) setting $p = 2$. However, for the purpose of this initial overview we will restrict attention to bounds for the entire space $\mathbf{X}$.

[5] By [187, 188] the mapping that assigns a convex body $K \subseteq \mathbb{R}^n$ to its projection body ΠK is characterized axiomatically as the unique (up to scaling) translation-invariant $\mathsf{SL}_n(\mathbb{R})$-contravariant Minkowski valuation.

where $d_{\mathrm{BM}}(\cdot,\cdot)$ denotes the Banach–Mazur distance. Both of the bounds in (1.2) can be inferior to those that follow from Theorem 3. For example, suppose that $n = m^2$ for some $m \in \mathbb{N}$ and consider $\mathbf{X} = \ell_\infty^m(\ell_1^m)$. Then,

$$d_{\mathrm{BM}}(\mathbf{X}, \ell_1^n) \asymp d_{\mathrm{BM}}(\mathbf{X}, \ell_2^n) \asymp \sqrt{n}$$

by the work [163] of Kwapień and Schütt. Therefore, in this case (1.2) becomes the estimates $\sqrt{n} \lesssim \mathsf{SEP}(\mathbf{X}) \lesssim n$, while we will see that (1.1) implies that $\mathsf{SEP}(\mathbf{X}) \asymp n^{3/4}$.

The following corollary collects examples of applications of Theorem 3 that we will deduce herein.

Corollary 4 (Examples of consequences of Theorem 3). *The following statements hold for any $n \in \mathbb{N}$.*

- *For any $p \geqslant 1$, the separation modulus of ℓ_p^n satisfies*

$$\mathsf{SEP}\big(\ell_p^n\big) \asymp n^{\max\{\frac{1}{2}, \frac{1}{p}\}}. \tag{1.3}$$

  More generally, let $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ be any n-dimensional normed space with a 1-symmetric basis $e_1, \ldots, e_n$. Then, $\mathsf{SEP}(\mathbf{E})$ is equal to the following quantity up to lower order factors:

$$\|e_1 + \cdots + e_n\|_{\mathbf{E}} \bigg(\max_{k \in \{1,\ldots,n\}} \frac{\sqrt{k}}{\|e_1 + \cdots + e_k\|_{\mathbf{E}}} \bigg).$$

- *For any $p \geqslant 1$, the separation modulus of the Schatten–von Neumann trace class S_p^n on $\mathsf{M}_n(\mathbb{R})$ is*

$$\mathsf{SEP}\big(\mathsf{S}_p^n\big) = n^{\max\{1, \frac{1}{2} + \frac{1}{p}\} + o(1)} = \dim\big(\mathsf{S}_p^n\big)^{\max\{\frac{1}{2}, \frac{1}{4} + \frac{1}{2p}\} + o(1)}. \tag{1.4}$$

  More generally, let $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ be any n-dimensional normed space with a 1-symmetric basis $e_1, \ldots, e_n$ and denote its unitary ideal by $\mathsf{S}_{\mathbf{E}} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathsf{S}_{\mathbf{E}}})$. Then, $\mathsf{SEP}(\mathsf{S}_{\mathbf{E}})$ is equal to the following quantity up to lower order factors:

$$\|e_1 + \cdots + e_n\|_{\mathbf{E}} \bigg(\max_{k \in \{1,\ldots,n\}} \frac{\sqrt{k}}{\|e_1 + \cdots + e_k\|_{\mathbf{E}}} \bigg) \sqrt{n}.$$

- *For any $p, q \geqslant 1$, the separation modulus of the $\ell_p^n(\ell_q^n)$ norm on $\mathsf{M}_n(\mathbb{R})$ is*

$$\begin{aligned} \mathsf{SEP}\big(\ell_p^n(\ell_q^n)\big) &\asymp n^{\max\{1, \frac{1}{p} + \frac{1}{q}, \frac{1}{2} + \frac{1}{p}, \frac{1}{2} + \frac{1}{q}\}} \\ &= \dim\big(\ell_p^n(\ell_q^n)\big)^{\max\{\frac{1}{2}, \frac{1}{2p} + \frac{1}{2q}, \frac{1}{4} + \frac{1}{2p}, \frac{1}{4} + \frac{1}{2q}\}}. \end{aligned} \tag{1.5}$$

- *For any $p, q \geqslant 1$, the separation modulus of $\mathsf{M}_n(\mathbb{R})$ equipped with the operator norm $\|\cdot\|_{\ell_p^n \to \ell_q^n}$ from ℓ_p^n to ℓ_q^n is equal to the following quantity up to lower order* *factors:*

$$\begin{cases} n^{\frac{3}{2}-\frac{1}{\min\{p,q\}}} & \text{if } p,q \geqslant 2,\\ n^{\frac{1}{2}+\frac{1}{\max\{p,q\}}} & \text{if } p,q \leqslant 2,\\ n & \text{if } p \leqslant 2 \leqslant q,\\ n^{\max\{1,\frac{1}{q}-\frac{1}{p}+\frac{1}{2}\}} & \text{if } q \leqslant 2 \leqslant p.\end{cases}$$

- *For any $p,q \geqslant 1$, the separation modulus of the projective tensor product $\ell_p^n \widehat{\otimes} \ell_q^n$, i.e., the norm on $\mathsf{M}_n(\mathbb{R})$ whose unit ball is the convex hull of the set*

$$\{(x_i y_j) \in \mathsf{M}_n(\mathbb{R});\ (x_1,\ldots,x_n) \in B_{\ell_p^n}\ \wedge\ (y_1,\ldots,y_n) \in B_{\ell_q^n}\},$$

is equal to the following quantity up to lower order factors:

$$\begin{cases} n^{\frac{3}{2}} & \text{if } \max\{p,q\} \geqslant 2,\\ n^{1+\frac{1}{\max\{p,q\}}} & \text{if } \max\{p,q\} \leqslant 2.\end{cases}$$

All of the results in Corollary 4 are new, except for the range $1 \leqslant p \leqslant 2$ of (1.3), which is due to [76]. The range $p \in (2,\infty]$ of (1.3) is $\mathsf{SEP}(\ell_p^n) \asymp \sqrt{n}$, which is incompatible with the statement $\mathsf{SEP}(\ell_p^n) \asymp n^{1-1/p}$ of [76]. We will explain the reason why the latter assertion of [76] is erroneous in Remark 78.

The wealth of knowledge that is available on the volumetric quantities that appear in (1.1) leads to new estimates that relate the separation modulus of an n-dimensional normed space $\mathbf{X}$ to classical invariants of $\mathbf{X}$. We will derive several such results herein, without attempting to be encyclopedic. As a noteworthy example, we will deduce from the first inequality in (1.1) that if $B_{\mathbf{X}}$ is a polytope with ρn vertices, then

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \frac{n}{\sqrt{\log \rho}}. \tag{1.6}$$

We will also deduce that if $T_2(\mathbf{X})$ denotes the type 2 constant of $\mathbf{X}$ (see (1.77) or the survey [203]), then

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \max\left\{\sqrt{\dim(\mathbf{X})}, T_2(\mathbf{X})^2\right\}. \tag{1.7}$$

We will see that both (1.6) and (1.7) are sharp for the entire range of the relevant parameters (e.g., in the two extremes, the case $\mathbf{X} = \ell_1^n$ corresponds to $\rho = O(1)$ and $T_2(\mathbf{X}) \asymp \sqrt{n}$ in (1.6) and (1.7), respectively, and the case when $\mathbf{X}$ is $O(1)$-isomorphic to ℓ_2^n corresponds to $\log \rho \asymp n$ and $T_2(\mathbf{X}) = O(1)$ in (1.6) and (1.7), respectively).

1.1.1 A conjectural isomorphic reverse isoperimetric phenomenon

The lower bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3 is not always sharp. Indeed, consider the space $\mathbf{X} = \ell_1^n \oplus \ell_2^n$ for which $\mathsf{SEP}(\mathbf{X}) \asymp n$ yet $\mathrm{vr}(\mathbf{X}^*)\sqrt{\dim(\mathbf{X})} \asymp n^{3/4}$. It could be, however, that the upper bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3 is optimal for every $\mathbf{X}$.

Question 5. Is the separation modulus of any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ bounded above and below by some universal constant multiples of the minimum of the quantity $\mathrm{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})/\operatorname{vol}_n(B_{\mathbf{Y}})$ over all those normed spaces $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ that satisfy $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$?

See Remark 23 for an explanation why the minimum that is described in Question 5 is affine invariant, which is necessary for Question 5 to make sense, since the separation modulus is a bi-Lipschitz invariant.

For sufficiently symmetric spaces, we expect that the lower bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3 is sharp.

Conjecture 6. Every finite dimensional normed space $\mathbf{X}$ with enough symmetries satisfies

$$\mathsf{SEP}(\mathbf{X}) \asymp \mathrm{vr}(\mathbf{X}^*)\sqrt{\dim(\mathbf{X})}. \tag{1.8}$$

The notion of having enough symmetries was introduced in [103]; its definition is recalled in Section 1.6.2. We prefer to formulate Conjecture 6 using this notion at the present introductory juncture even though weaker requirements are needed for our purposes because it is a standard assumption in Banach space theory and it suffices for all of the most pressing applications that we have in mind.

The upper bound on $\mathsf{SEP}(\mathbf{X})$ in (1.8) implies by [173] that

$$\mathsf{e}(\mathbf{X}) \lesssim \mathrm{vr}(\mathbf{X}^*)\sqrt{\dim(\mathbf{X})},$$

which would be a valuable Lipschitz extension theorem due to the fact that estimating the volume ratio is typically tractable given the variety of tools and extensive knowledge that are available in the literature. For example, Milman and Pisier [219] proved (improving by lower-order factors over a major theorem of Bourgain and Milman [49, 50]; see also [217]), that any finite dimensional normed space $\mathbf{X}$ satisfies

$$\mathrm{vr}(\mathbf{X}) \lesssim C_2(\mathbf{X})\big(1 + \log C_2(\mathbf{X})\big), \tag{1.9}$$

where $C_2(\mathbf{X})$ is the cotype 2 constant of $\mathbf{X}$ (see (1.77) or the survey [203]). Therefore, if (1.8) holds, then

$$\mathsf{e}(\mathbf{X}) \lesssim C_2(\mathbf{X})\big(1 + \log C_2(\mathbf{X})\big)\sqrt{\dim(\mathbf{X})}, \tag{1.10}$$

which would be a remarkable generalization of the bound $\mathsf{e}(\ell_2^n) \lesssim \sqrt{n}$ of [173].

We expect that Theorem 3 already implies Conjecture 6, as expressed in the following conjecture which would yield a positive answer to Question 5 for normed spaces with enough symmetries.

Conjecture 7. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space with enough symmetries, then there is a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ that satisfies

$$B_{\mathbf{Y}} \subseteq B_{\mathbf{X}} \quad \text{and} \quad \frac{\mathrm{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \lesssim \mathrm{vr}(\mathbf{X}^*)\sqrt{n}.$$

As an illustrative example of Conjecture 7, consider the space $\mathbf{X} = \ell_\infty^n$. We then have $\mathrm{vr}((\ell_\infty^n)^*) = \mathrm{vr}(\ell_1^n) = O(1)$. One can compute that $\Pi B_{\ell_\infty^n} = 2^{n-1} B_{\ell_\infty^n}$. Therefore, $\mathrm{diam}_{\ell_1^n}(\Pi B_{\ell_\infty^n}) / \mathrm{vol}_n(B_{\ell_\infty^n}) \asymp n$, so taking $\mathbf{Y} = \ell_\infty^n$ in Theorem 3 only gives the bound $\mathsf{SEP}(\ell_\infty^n) \lesssim n$. However, we will prove that there exists a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ with $B_{\mathbf{Y}} \subseteq B_{\ell_\infty^n}$ for which $\mathrm{diam}_{\ell_1^n}(\Pi B_{\mathbf{Y}}) / \mathrm{vol}_n(B_{\mathbf{Y}}) \lesssim \sqrt{n}$. More generally, we will prove that Conjecture 7 (hence also Conjecture 6, by Theorem 3) holds for any normed space for which the standard basis of $\mathbb{R}^n$ is 1-symmetric, and we will also see that Conjecture 7 holds up to a logarithmic factor for its unitary ideal.

The minimization in Question 5 can be viewed as a shape optimization problem [130] that could potentially be approached using calculus of variations. Given an origin-symmetric convex body $K \subseteq \mathbb{R}^n$, it asks for the minimum of the affine invariant functional $L \mapsto \mathsf{outradius}_{K^\circ}(\Pi L) / \mathrm{vol}_n(L)$ over all origin-symmetric convex bodies $L \subseteq K$, where for any two origin-symmetric convex bodies $A, B \subseteq \mathbb{R}^n$ we denote the minimum radius of a dilate of A that circumscribes B by

$$\mathsf{outradius}_A(B) = \min\{r \geqslant 0 : B \subseteq rA\}$$

and

$$K^\circ = \{y \in \mathbb{R}^n : \sup_{x \in K} \langle x, y \rangle \leqslant 1\}$$

is the polar of K. Conjecture 7 asserts that if K has enough symmetries, then this minimum is bounded above and below by universal constant multiples of $\mathrm{vr}(K^\circ)\sqrt{n}$.

The minimization problem in Question 5 also has an isoperimetric flavor. As such, its investigation led us to formulate the following conjectural twist of Ball's reverse isoperimetric phenomenon [22], which we think is a fundamental geometric open question and it would be valuable to understand it even without its consequences that we derive herein.

The *isoperimetric quotient* of a convex body $K \subseteq \mathbb{R}^n$ is defined (see [126, p. 269] or [286]) to be

$$\mathrm{iq}(K) = \frac{\mathrm{vol}_{n-1}(\partial K)}{\mathrm{vol}_n(K)^{\frac{n-1}{n}}}. \tag{1.11}$$

Using this notation, the classical Euclidean isoperimetric theorem states that

$$\mathrm{iq}(K) \geqslant \mathrm{iq}\big(B_{\ell_2^n}\big) = \frac{n\sqrt{\pi}}{\Gamma\big(\frac{n}{2}+1\big)^{\frac{1}{n}}} \asymp \sqrt{n}. \tag{1.12}$$

The following theorem of Ball [22] shows that a judicious choice of the scalar product on $\mathbb{R}^n$ ensures that the isoperimetric quotient of a convex body can also be bounded from above.

Theorem 8 (Ball's reverse isoperimetric theorem [22]). *For every $n \in \mathbb{N}$ and every origin-symmetric convex body $K \subseteq \mathbb{R}^n$ there exists a linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ such that* $\mathrm{iq}(SK) \leqslant 2n = \mathrm{iq}([-1,1]^n)$.

We expect that in the isomorphic regime (i.e., permitting non-isometric $O(1)$ perturbations), origin-symmetric convex bodies have asymptotically better reverse isoperimetric properties than what is guaranteed by Theorem 8. In fact, we conjecture that if in addition to passing from K to SK for some $S \in \mathsf{SL}_n(\mathbb{R})$, a $O(1)$-perturbation of SK is allowed, then the isoperimetric quotient can be decreased to be of the same order of magnitude as that of the Euclidean ball.

Conjecture 9 (Isomorphic reverse isoperimetry). There is a universal constant $c > 0$ with the following property. For every $n \in \mathbb{N}$ and every origin-symmetric convex body $K \subseteq \mathbb{R}^n$, there exist a linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq \mathbb{R}^n$ with $cSK \subseteq L \subseteq SK$ and $\mathrm{iq}(L) \lesssim \sqrt{n}$.

Conjecture 9 can be restated analytically as the assertion that any n-dimensional normed space is at Banach–Mazur distance $O(1)$ from a normed space whose unit ball has isoperimetric quotient $O(\sqrt{n})$. We will prove that Conjecture 9 holds when K is the unit ball of ℓ_p^n for any $p \in [1, \infty]$ and $n \in \mathbb{N}$, and we will also see that Conjecture 9 holds up to lower-order factors for any Schatten–von Neumann trace class.

The requirement $L \supseteq cSK$ of Conjecture 9 implies that $\sqrt[n]{\mathrm{vol}_n(L)} \geqslant c\sqrt[n]{\mathrm{vol}_n(K)}$. So, the following weaker conjecture is implied by Conjecture 9; we will prove it for any 1-unconditional body.

Conjecture 10 (Weak isomorphic reverse isoperimetry). For every $n \in \mathbb{N}$ and every origin-symmetric convex body $K \subseteq \mathbb{R}^n$ there exist a linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq SK$ that satisfies $\sqrt[n]{\mathrm{vol}_n(L)} \gtrsim \sqrt[n]{\mathrm{vol}_n(K)}$ and $\mathrm{iq}(L) \lesssim \sqrt{n}$.

In Section 1.6 we will elucidate the relation between the task of bounding from above the rightmost quantity in (1.1) and isomorphic reverse isoperimetry. While Conjecture 9 is the strongest version of the isomorphic reverse isoperimetric phenomenon that we expect holds in full generality, we will see that it would suffice to prove its weaker variant Conjecture 10 for the purpose of using Theorem 3. In particular, consider the following symmetric version of Conjecture 10, which we will prove in Section 1.6 implies Conjecture 7 (hence, using Theorem 3, it also implies Conjecture 6).

Conjecture 11 (Symmetric version of Conjecture 10). For every $n \in \mathbb{N}$ and every normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ with enough symmetries whose isometry group is a subgroup of the orthogonal group $\mathsf{O}_n \subseteq \mathsf{GL}_n(\mathbb{R})$, there is a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ with $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ and $\sqrt[n]{\mathrm{vol}_n(B_{\mathbf{Y}})} \gtrsim \sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})}$ such that $\mathrm{iq}(B_{\mathbf{Y}}) \lesssim \sqrt{n}$.

The only difference between Conjecture 10 and Conjecture 11 is that if we impose the further requirement that K is the unit ball of a normed space with enough symmetries whose isometry group consists only of orthogonal matrices, then we are naturally conjecturing that S can be taken to be the identity matrix, i.e., there is no need to change the standard Euclidean structure on $\mathbb{R}^n$.

We will prove Conjecture 11 for various spaces, including $\ell_p^n(\ell_q^n)$ for any $p,q \geqslant 1$ and $n \in \mathbb{N}$, and any finite dimensional space with a 1-symmetric basis. Also, we will show that Conjecture 11 holds up to a factor of $O(\sqrt{\log n})$ for any unitarily invariant norm on $\mathsf{M}_n(\mathbb{R})$. In general, an argument that was shown to us by B. Klartag and E. Milman and is included in Section 7 (see also Section 1.6.3) shows that Conjecture 10 and Conjecture 11 hold up to a factor of $O(\log n)$. We will see that these results lead to Corollary 4, and in general we will deduce that Conjecture 7, and hence, thanks to Theorem 3, also Conjecture 6, hold up to lower order factors. Thus, we will obtain the following theorem.

Theorem 12. $\mathsf{SEP}(\mathbf{X}) \asymp \mathrm{vr}(\mathbf{X}^*)\dim(\mathbf{X})^{\frac{1}{2}+o(1)}$ *for any normed space* $\mathbf{X}$ *with enough symmetries.*

Assuming Conjecture 11, it is possible to compute the exact asymptotic growth rate of the separation moduli of several important matrix spaces. For example, if Conjecture 11 holds for S_∞^n, then we will see that the $o(1)$ term in (1.4) could be removed altogether, i.e.,

$$\forall (p,n) \in [1,\infty] \times \mathbb{N}, \quad \mathsf{SEP}\big(\mathsf{S}_p^n\big) \asymp n^{\max\{1,\frac{1}{2}+\frac{1}{p}\}}. \tag{1.13}$$

Also, assuming Conjecture 11 the lower order factors in the last two statements of Corollary 4 could be removed, namely we will see that Conjecture 11 implies that the separation modulus of $\mathsf{M}_n(\mathbb{R})$ equipped with the operator norm $\|\cdot\|_{\ell_p^n \to \ell_q^n}$ from ℓ_p^n to ℓ_q^n satisfies

$$\mathsf{SEP}\big(\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\ell_p^n\to\ell_q^n}\big) \asymp \begin{cases} n^{\frac{3}{2}-\frac{1}{\min\{p,q\}}} & \text{if } p,q \geqslant 2,\\ n^{\frac{1}{2}+\frac{1}{\max\{p,q\}}} & \text{if } p,q \leqslant 2,\\ n & \text{if } p \leqslant 2 \leqslant q,\\ n^{\max\{1,\frac{1}{q}-\frac{1}{p}+\frac{1}{2}\}} & \text{if } q \leqslant 2 \leqslant p, \end{cases} \tag{1.14}$$

and the separation modulus of the projective tensor product $\ell_p^n \widehat{\otimes} \ell_q^n$ satisfies

$$\mathsf{SEP}\big(\ell_p^n \widehat{\otimes} \ell_q^n\big) \asymp \begin{cases} n^{\frac{3}{2}} & \text{if } \max\{p,q\} \geqslant 2,\\ n^{1+\frac{1}{\max\{p,q\}}} & \text{if } \max\{p,q\} \leqslant 2. \end{cases} \tag{1.15}$$

Remark 174 describes ramifications of these conjectural statements to norms of algorithmic importance.

Roadmap. The rest of the Introduction effectively restarts the description of the present work, with many more details/definitions/background/ideas of proofs, than what we have included above. We organized the introductory material in this way because this work pertains to multiple mathematical disciplines, including notably Banach spaces, convex geometry, nonlinear functional analysis, metric embeddings, extension of functions, and theoretical computer science. The backgrounds of potential readers are therefore varied, so even though the above overview achieves the goal of presenting the main results quickly, it inevitably includes terminology that is not familiar to some. The aforementioned organizational choice makes the ensuing discussion accessible. Additional background can be found in the monographs [181, 220, 305] (Banach space theory), [36] (nonlinear functional analysis), [201, 244] (metric embeddings), [64] (extension of functions), as well as the references that are cited throughout.

While the ensuing extended introductory text is not short, it achieves more than merely a description of the results, history, concepts and methods: it also contains groundwork that is needed for the subsequent sections. Thus, reading the Introduction will lead to a thorough conceptual understanding of the contents, leaving to the remaining sections considerations that are for the most part more technical.

We will start by focusing on the classical Lipschitz extension problem because it is more well known than the stochastic clustering issues that lead to most of our new results on Lipschitz extension, and also because it requires less technicalities (e.g., a suitable measurability setup) than our subsequent treatment of stochastic clustering. Throughout the Introduction (and beyond), we will formulate conjectures and questions that are valuable even without the links to Lipschitz extension and clustering that are derived herein. After the Introduction, the rest of this work will be organized thematically as follows. Section 2 is devoted to proofs of our various lower bounds, namely impossibility results that rule out the existence of extensions and clusterings with certain properties. Section 3 and Section 4 deal with positive results about random partitions. Specifically, Section 3 is of a more foundational nature as it describes the concepts, basic constructions, and proofs of measurability statements that are needed for later applications in the infinitary setting (of course, measurability can be ignored for statements about finite sets). Section 4 analyses in the case of normed spaces a periodic version of a commonly used randomized partitioning technique called *iterative ball partitioning*, and computes optimally (up to universal constant factors) the probabilities of its separation and padding events. Section 5 shows how to pass from random partitions to Lipschitz extension, by adjusting to the present setting the method that was developed in [173]. Section 5 also contains further foundational results on Lipschitz extension, as well questions and conjectures that are of independent interest. Section 6 contains a range of volume and surface area estimates that are needed in conjunction with the theorems of the preceding sections in order to deduce new Lipschitz extension and stochastic clustering results for various normed spaces and their subsets. Section 7 proves that Conjecture 10 and Conjecture 11 hold up to a factor of $O(\log n)$, and also shows that the approach that leads to this result cannot fully resolve Conjecture 11.

1.2 Basic notation

Given a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, a point $x \in \mathfrak{M}$ and a radius $r \geqslant 0$, the corresponding *closed* ball is denoted $B_{\mathfrak{M}}(x, r) = \{y \in \mathfrak{M} : d_{\mathfrak{M}}(y, x) \leqslant r\}$. If $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a Banach space (in this work, all vector spaces are over the real scalars unless stated otherwise), then denote by $B_{\mathbf{X}}$ the unit ball centered at the origin. Under this notation we have $B_{\mathbf{X}} = B_{\mathbf{X}}(0, 1)$ and $B_{\mathbf{X}}(x, r) = x + rB_{\mathbf{X}}$ for every $x \in X$ and $r \geqslant 0$.

If $(\mathfrak{M}, d_{\mathfrak{M}}), (\mathfrak{N}, d_{\mathfrak{N}})$ are metric spaces and $\psi : \mathfrak{M} \to \mathfrak{N}$, then for $\mathcal{C} \subseteq \mathfrak{M}$ the Lipschitz constant of ψ on $\mathcal{C}$ is denoted $\|\psi\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})} \in [0, \infty]$. Thus, if $\mathcal{C}$ contains at least two points, then

$$\|\psi\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})} \stackrel{\mathrm{def}}{=} \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{d_{\mathfrak{N}}\big(\psi(x), \psi(y)\big)}{d_{\mathfrak{M}}(x, y)}.$$

In the special case $\mathfrak{N} = \mathbb{R}$ we will use the simpler notation $\|\psi\|_{\mathrm{Lip}(\mathcal{C};\mathbb{R})} = \|\psi\|_{\mathrm{Lip}(\mathcal{C})}$.

If $(\mathbf{X}, \|\cdot\|_{\mathbf{X}}), (\mathbf{Y}, \|\cdot\|_{\mathbf{Y}})$ are isomorphic Banach spaces, then their Banach–Mazur distance $d_{\mathrm{BM}}(\mathbf{X}, \mathbf{Y})$ is the infimum of the products of the operator norms $\|T\|_{\mathbf{X}\to\mathbf{Y}}$ and $\|T^{-1}\|_{\mathbf{Y}\to\mathbf{X}}$ over all possible (surjective) linear isomorphisms

$$T : \mathbf{X} \to \mathbf{Y}.$$

The (bi-Lipschitz) distortion of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ into a metric space $(\mathfrak{N}, d_{\mathfrak{N}})$, denoted $\mathsf{c}_{(\mathfrak{N},d_{\mathfrak{N}})}(\mathfrak{M}, d_{\mathfrak{M}})$ or $\mathsf{c}_{\mathfrak{N}}(\mathfrak{M})$ if the underlying metrics are clear from the context, is the infimum over those $D \in [1, \infty]$ for which there exists a mapping $\phi : \mathfrak{M} \to \mathfrak{N}$ and (a scaling factor) $\lambda > 0$ such that

$$\forall x, y \in \mathfrak{M}, \quad \lambda d_{\mathfrak{M}}(x, y) \leqslant d_{\mathfrak{N}}\big(\phi(x), \phi(y)\big) \leqslant D\lambda d_{\mathfrak{M}}(x, y). \tag{1.16}$$

Fix $n \in \mathbb{N}$. Throughout what follows, $\mathbb{R}^n$ will be always be endowed with its standard Euclidean structure, i.e., with the scalar product $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$ for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Given $z \in \mathbb{R}^n \smallsetminus \{0\}$, the orthogonal projection onto its orthogonal hyperplane $z^{\perp} = \{x \in \mathbb{R}^n : \langle x, z\rangle = 0\}$ will be denoted $\mathsf{Proj}_{z^{\perp}} : \mathbb{R}^n \to \mathbb{R}^n$. For $0 < s \leqslant n$, the s-dimensional Hausdorff measure of a closed subset $A \subseteq \mathbb{R}^n$ is denoted $\mathrm{vol}_s(A)$. Integration with respect to the s-dimensional Hausdorff measure is indicated by $\mathrm{d}x$. If $0 < \mathrm{vol}_s(A) < \infty$ and $f : A \to \mathbb{R}$ is continuous, then write $\fint_A f(x)\,\mathrm{d}x = \mathrm{vol}_s(A)^{-1}\int_A f(x)\,\mathrm{d}x$.

Given a normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ and $p \in [1, \infty]$, $\ell_p^n(\mathbf{X})$ is the vector space $\mathbf{X}^n$ equipped with the norm

$$\forall x = (x_1, \ldots, x_n) \in \mathbf{X}^n, \quad \|x\|_{\ell_p^n(\mathbf{X})} = \big(\|x_1\|_{\mathbf{X}} + \cdots + \|x_n\|_{\mathbf{X}}\big)^{\frac{1}{p}},$$

where for $p = \infty$ this is understood to be $\|x\|_{\ell_\infty^n(\mathbf{X})} = \max_{j\in\{1,\ldots,n\}} \|x_j\|_{\mathbf{X}}$. It is common to use the simpler notation $\ell_p^n = \ell_p^n(\mathbb{R})$ and we write as usual $S^{n-1} = \partial B_{\ell_2^n}$.

$$
\begin{tikzcd}
\mathfrak{M} \arrow[dr, dashed, "F"] & \\
\mathcal{C} \arrow[u, hook, "\mathsf{Id}_{\mathcal{C}\to\mathfrak{M}}"] \arrow[r, "f"'] & \mathbf{Z}
\end{tikzcd}
$$

Figure 1.1. Given $K \geqslant 1$, the assertion that the Lipschitz extension modulus of a metric space $\mathfrak{M}$ satisfies $\mathsf{e}(\mathfrak{M}) < K$ means that for *all* subsets $\mathcal{C} \subseteq \mathfrak{M}$, *all* Banach spaces $\mathbf{Z}$ and *all* 1-Lipschitz mappings $f : \mathcal{C} \to \mathbf{Z}$, there is a K-Lipschitz mapping $F : \mathfrak{M} \to \mathbf{Z}$ such that the above diagram commutes, where $\mathsf{Id}_{\mathcal{C}\to\mathfrak{M}} : \mathcal{C} \to \mathfrak{M}$ is the formal inclusion.

The Schatten–von Neumann trace class S_p^n is the (n^2-dimensional) space of all n by n real matrices $\mathsf{M}_n(\mathbb{R})$, equipped with the norm that is defined by

$$\forall T \in \mathsf{M}_n(\mathbb{R}), \quad \|T\|_{\mathsf{S}_p^n} = \big(\mathbf{Tr}\big((TT^*)^{\frac{p}{2}}\big)\big)^{\frac{1}{p}} = \big(\mathbf{Tr}\big((T^*T)^{\frac{p}{2}}\big)\big)^{\frac{1}{p}},$$

where $\|T\|_{\mathsf{S}_\infty^n} = \|T\|_{\ell_2^n \to \ell_2^n}$ is the operator norm of T when it is viewed as a linear operator from ℓ_2^n to ℓ_2^n.

1.3 Lipschitz extension

As we recalled in Section 1.1, one associates to every metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ a bi-Lipschitz invariant[6], called the Lipschitz extension modulus of $(\mathfrak{M}, d_{\mathfrak{M}})$ and denoted $\mathsf{e}(\mathfrak{M}, d_{\mathfrak{M}})$ or $\mathsf{e}(\mathfrak{M})$ if the metric is clear from the context, by defining it to be the infimum over those $K \in [1, \infty]$ with the property that for *every* nonempty subset $\mathcal{C} \subseteq \mathfrak{M}$, *every* Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and *every* Lipschitz function $f : \mathcal{C} \to \mathbf{Z}$ there is a mapping $F : \mathfrak{M} \to \mathbf{Z}$ that extends f, i.e., $F(x) = f(x)$ whenever $x \in \mathcal{C}$, and $\|F\|_{\mathrm{Lip}(\mathfrak{M},\mathbf{Z})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C},\mathbf{Z})}$; see Figure 1.1. All of the ensuing extension theorems hold for a larger class of target metric spaces that need not necessarily be Banach spaces, including Hadamard spaces and Busemann nonpositively curved spaces [57], or more generally spaces that posses a conical geodesic bicombing (see, e.g., [86]). This greater generality will be discussed in Section 5, but we prefer at this introductory juncture to focus on the more classical and highly-studied setting of Banach space targets.

When $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a finite dimensional normed space, the currently best-available general bounds on the quantity $\mathsf{e}(\mathbf{X})$ in terms of $\dim(\mathbf{X})$ are contained the following theorem.

[6] The assertion that $\mathsf{e}(\mathfrak{M})$ is a bi-Lipschitz invariant refers to the fact that the definition immediately implies that if $(\mathfrak{N}, d_{\mathfrak{N}})$ is another metric space into which $(\mathfrak{M}, d_{\mathfrak{M}})$ admits a bi-Lipschitz embedding, then $\mathsf{e}(\mathfrak{M}) \leqslant \mathsf{c}_{\mathfrak{N}}(\mathfrak{M})\mathsf{e}(\mathfrak{N})$.

Theorem 13. *There is a universal constant $c > 0$ such that for any finite dimensional normed space* $\mathbf{X}$,

$$\dim(\mathbf{X})^c \lesssim \mathsf{e}(\mathbf{X}) \lesssim \dim(\mathbf{X}). \tag{1.17}$$

The bound $\mathsf{e}(\mathbf{X}) \lesssim \dim(\mathbf{X})$ in (1.17) is a famous result of Johnson, Lindenstrauss and Schechtman [140], which they proved by cleverly refining the classical extension method of Whitney [312]; different proofs of this estimate were found by Lee and the author [173] as well as by Brudnyi and Brudnyi [61] (see also the discussion in the paragraph following equation (1.37) below). It remains a major longstanding open problem to determine whether the bound of [140] could be improved to

$$\mathsf{e}(\mathbf{X}) = o\big(\dim(\mathbf{X})\big).$$

The new content of Theorem 13 is the lower bound on $\mathsf{e}(\mathbf{X})$, which improves over the previously known bound $\mathsf{e}(\mathbf{X}) \gtrsim \exp(c\sqrt{\log\dim(\mathbf{X})})$; see Remark 98 for the history of this question. It is a very interesting open problem to determine the supremum over those c for which Theorem 13 holds.[7] More generally, it is natural to aim to evaluate the precise power-type behavior of $\mathsf{e}(\mathbf{X})$ as $\dim(\mathbf{X}) \to \infty$ for specific (sequences of) finite dimensional normed spaces $\mathbf{X}$. However, prior to the present work and despite many efforts over the years, this was not achieved for *any* finite dimensional normed space whatsoever.

Theorem 14 (Restatement of Theorem 2). *For every $n \in \mathbb{N}$ we have $\mathsf{e}(\ell_\infty^n) \asymp \sqrt{n}$.*

The bound $\mathsf{e}(\ell_\infty^n) \gtrsim \sqrt{n}$ follows from a combination of [60, Theorem 4] and [62, Theorem 1.2]. The new content of Theorem 14 is the upper bound $\mathsf{e}(\ell_\infty^n) \lesssim \sqrt{n}$ (and, importantly, the extension procedure that leads to it; see below). The previously best-known upper bound on $\mathsf{e}(\ell_\infty^n)$ was the aforementioned $O(n)$ estimate of [140]. The question of evaluating the asymptotic behavior of $\mathsf{e}(\ell_p^n)$ as $n \to \infty$ for each $p \in [1,\infty]$ is natural and longstanding; it was stated in [60, Problem 2] and reiterated in [63, Section 4], [62, Problem 1.4] and [64, Problem 8.14]. Theorem 14 answers this question when $p = \infty$. The upper bound on $\mathsf{e}(\ell_\infty^n)$ of Theorem 14 is a special case of a general extension criterion that provides the best-known Lipschitz extension results in other settings (including for ℓ_p^n when $p > 2$), but we chose to state it separately because it yields the first (and currently essentially only) family of normed spaces for which the growth rate of their Lipschitz extension moduli has been determined.

Remark 15. It is also meaningful to study extension of θ-Hölder functions for any $0 < \theta \leqslant 1$. Namely, one can analogously define the θ-Hölder extension modulus of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, denoted $\mathsf{e}^\theta(\mathfrak{M})$. Alternatively, this notion falls into the above

[7] Our proof of the lower bound on $\mathsf{e}(\mathbf{X})$ of Theorem 13 shows that this supremum is at least $\frac{1}{12}$; see equation (2.5).

Lipschitz-extension framework because one can define

$$\mathsf{e}^{\theta}(\mathfrak{M}) \stackrel{\mathrm{def}}{=} \mathsf{e}\big(\mathfrak{M}, d_{\mathfrak{M}}^{\theta}\big). \tag{1.18}$$

The results that we obtain herein also yield improved estimates on θ-Hölder extension moduli; see Corollary 140. However, when $\theta < 1$ we never get a matching lower bound (the reason why we can do better in the Lipschitz regime $\theta = 1$ is essentially due to the fact that Lipschitz functions are differentiable almost everywhere). For example, in the setting of Theorem 14 we get the upper bound

$$\forall \theta \in (0,1], \quad \mathsf{e}^{\theta}\big(\ell_\infty^n\big) \lesssim n^{\frac{\theta}{2}}, \tag{1.19}$$

but the best lower bound on $\mathsf{e}^{\theta}(\ell_\infty^n)$ that we are at present able to prove is

$$\mathsf{e}^{\theta}\big(\ell_\infty^n\big) \gtrsim n^{\max\{\frac{\theta}{4}, \frac{\theta}{2}+\theta^2-1\}} = \begin{cases} n^{\frac{\theta}{4}} & \text{if } 0 \leqslant \theta \leqslant \frac{\sqrt{65}-1}{8}, \\ n^{\frac{\theta}{2}+\theta^2-1} & \text{if } \frac{\sqrt{65}-1}{8} \leqslant \theta \leqslant 1. \end{cases} \tag{1.20}$$

We conjecture that $\mathsf{e}^{\theta}(\ell_\infty^n) \asymp_\theta n^{\frac{\theta}{2}}$, but proving this for $\theta < 1$ would likely require a genuinely new idea.

Question 16. Despite its utility in many cases, the extension method that underlies Theorem 14 does not yield improved bounds for some important spaces, including notably ℓ_1^n and ℓ_2^n. Thus, determining the asymptotic behavior of $\mathsf{e}(\ell_1^n)$ and $\mathsf{e}(\ell_2^n)$ as $n \to \infty$ remains a tantalizing open question. Specifically, the currently best-known bounds on $\mathsf{e}(\ell_1^n)$ are

$$\sqrt{n} \lesssim \mathsf{e}\big(\ell_1^n\big) \lesssim n, \tag{1.21}$$

where the first inequality in (1.21) is due to Johnson and Lindenstrauss [138] and the second inequality in (1.21) is the aforementioned general upper bound of [140] on the Lipschitz extension modulus of *any* n-dimensional normed space. The currently best-known bounds in the Hilbertian setting are

$$\sqrt[4]{n} \lesssim \mathsf{e}\big(\ell_2^n\big) \lesssim \sqrt{n}, \tag{1.22}$$

where the first inequality in (1.22) is due to Mendel and the author [210] (a different proof of this lower bound on $\mathsf{e}(\ell_2^n)$ follows from [231]), and the second inequality in (1.22) is from [173].

By the bi-Lipschitz invariance of the Lipschitz extension modulus, the second inequality in (1.22) implies the following bound from [173], which holds for every finite dimensional normed space $\mathbf{X}$:

$$\mathsf{e}(\mathbf{X}) \lesssim d_{\mathrm{BM}}\big(\mathbf{X}, \ell_2^{\dim(\mathbf{X})}\big)\sqrt{\dim(\mathbf{X})}. \tag{1.23}$$

This refines the upper bound on $\mathsf{e}(\mathbf{X})$ in (1.17) because $d_{\mathrm{BM}}(\mathbf{X}, \ell_2^{\dim(\mathbf{X})}) \leqslant \sqrt{\dim(\mathbf{X})}$ by John's theorem [137].

Remark 17. In the context of the aforementioned question whether the bound $\mathsf{e}(\mathbf{X}) \lesssim \dim(\mathbf{X})$ of [140] is optimal, by (1.23) we see that $\mathsf{e}(\mathbf{X}) = o(\dim(\mathbf{X}))$ unless the Banach–Mazur distance between $\mathbf{X}$ and Euclidean space is of order $\sqrt{\dim(\mathbf{X})}$. Structural properties of such spaces of extremal distance to Euclidean space have been studied in [15, 43, 142, 221, 255]; see also [305, Chapters 6 and 7]. In particular, the Mil'man–Wolfson theorem [221] asserts that this holds if and only if $\mathbf{X}$ has a subspace of dimension $k = k(\dim(\mathbf{X}))$ whose Banach–Mazur distance to ℓ_1^k is $O(1)$, where $\lim_{n\to\infty} k(n) = \infty$.

As $d_{\mathrm{BM}}(\ell_p^n, \ell_2^n) \asymp n^{|p-2|/(2p)}$ for all $n \in \mathbb{N}$ and $p \in [1,\infty]$ (see [139, Section 8]), it follows from (1.23) that

$$\mathsf{e}(\ell_p^n) \lesssim \begin{cases} n^{\frac{1}{p}} & \text{if } p \in [1,2], \\ n^{1-\frac{1}{p}} & \text{if } p \in [2,\infty]. \end{cases} \tag{1.24}$$

(1.24) was the previously best-known upper bound on $\mathsf{e}(\ell_p^n)$, and here we improve it for every $p > 2$.

Theorem 18. *For every $n \in \mathbb{N}$ and every $p \in [1,\infty]$ we have $\mathsf{e}(\ell_p^n) \lesssim n^{\max\{\frac{1}{2},\frac{1}{p}\}}$.*

Theorem 14 is the case $p = \infty$ of Theorem 18. We do not know if Theorem 18 is optimal (perhaps up to lower order factors) as $n \to \infty$ for fixed $p \in [2,\infty)$, but we conjecture that this is indeed the case, which would resolve [60, Problem 2]. The currently best-known lower bound on $\mathsf{e}(\ell_p^n)$ for every $p \in [1,\infty]$ is

$$\mathsf{e}(\ell_p^n) \gtrsim \begin{cases} n^{\frac{1}{p}-\frac{1}{2}} & \text{if } 1 \leqslant p \leqslant \frac{4}{3}, \\ \sqrt[4]{n} & \text{if } \frac{4}{3} \leqslant p \leqslant 2, \\ n^{\frac{1}{2p}} & \text{if } 2 \leqslant p \leqslant 3, \\ n^{\frac{1}{2}-\frac{1}{p}} & \text{if } 3 \leqslant p \leqslant \infty. \end{cases} \tag{1.25}$$

A lower bound on $\mathsf{e}(\ell_p^n)$ that coincides with (1.25) when $p \in [1, 4/3] \cup [3,\infty]$ is stated in [64, Corollary 8.12], but [64, Corollary 8.12] is weaker than (1.25) when $4/3 < p < 3$. The reason for this is that the lower bound of [210] on $\mathsf{e}(\ell_2^n)$ that appears in (1.22) was not available when [64] was written, but (1.25) for $4/3 < p < 3$ follows quickly by combining the first inequality in (1.22) with [99]; see Remark 2.4.

Remark 19. Theorem 18 resolves negatively a conjecture that A. Brudnyi and Y. Brudnyi posed as Conjecture 5 in [60]. They conducted a comprehensive study of the *linear* extension problem for real-valued Lipschitz functions, where one considers for a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ a quantity $\lambda(\mathfrak{M})$ which is defined the same as $\mathsf{e}(\mathfrak{M})$, but with the further requirements that the function f is real-valued and that the extended function F depends linearly on f. Namely, $\lambda(\mathfrak{M})$ is the infimum over those $K \in [1,\infty]$ such that for every $\mathcal{C} \subseteq \mathfrak{M}$ there is a linear operator $\mathsf{Ext}_{\mathcal{C}} : \mathrm{Lip}(\mathcal{C}) \to \mathrm{Lip}(\mathfrak{M})$ that assigns to every Lipschitz function $f : \mathcal{C} \to \mathbb{R}$ a function $\mathsf{Ext}_{\mathcal{C}} f : \mathfrak{M} \to \mathbb{R}$ satisfying $\mathsf{Ext}_{\mathcal{C}} f(s) = f(s)$ for every $s \in \mathcal{C}$, and

$$\|\mathsf{Ext}_{\mathcal{C}} f\|_{\mathrm{Lip}(\mathfrak{M})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C})}.$$

They also considered a natural variant of this quantity when $\mathfrak{M} = \mathbf{X}$ is a Banach space, denoted $\lambda_{\mathrm{conv}}(\mathbf{X})$, which is defined almost identically to $\lambda(\mathbf{X})$ except that now the subset $\mathcal{C}$ is only allowed to be any *convex* subset of $\mathbf{X}$ rather than a subset of $\mathbf{X}$ without any additional restriction. Conjecture 5 in [60] states that

$$\forall (p,n) \in [1,\infty] \times \mathbb{N}, \quad \lambda\big(\ell_p^n\big) \asymp_p \lambda_{\mathrm{conv}}\big(\ell_p^n\big)\sqrt{n}. \tag{1.26}$$

Theorem 18 implies that this conjecture is *false* for every $p \in (2,\infty]$. Indeed, the asymptotic behavior of $\lambda_{\mathrm{conv}}(\ell_p^n)$ was evaluated in [63, Theorem 2.19], where it was shown that

$$\forall p \in [1,\infty], \quad \lambda_{\mathrm{conv}}\big(\ell_p^n\big) \asymp n^{\left|\frac{1}{2}-\frac{1}{p}\right|}.$$

Consequently, $\lambda_{\mathrm{conv}}(\ell_p^n)\sqrt{n} \asymp n^{1-\frac{1}{p}}$ when $p > 2$. Next, in [62] a quantity $\nu(\mathfrak{M})$ was associated to a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ by defining it almost identically to the definition of $\mathsf{e}(\mathfrak{M})$, except that the target Banach space $\mathbf{Z}$ is allowed to be any *finite dimensional* Banach space rather than any Banach space whatsoever. By definition $\nu(\mathfrak{M}) \leqslant \mathsf{e}(\mathfrak{M})$, but actually $\lambda(\mathfrak{M}) = \nu(\mathfrak{M})$ thanks to [62, Theorem 1.2] (see the work [11] of Ambrosio and Puglisi for more on this "linearization phenomenon"). Using these results in combination with Theorem 18, we see that for every $p \in (2,\infty]$, as $n \to \infty$ we have

$$\lambda\big(\ell_p^n\big) = \nu\big(\ell_p^n\big) \leqslant \mathsf{e}\big(\ell_p^n\big) \lesssim \sqrt{n} = o\big(n^{1-\frac{1}{p}}\big).$$

Thus, $\lambda(\ell_p^n) = o\big(\lambda_{\mathrm{conv}}(\ell_p^n)\sqrt{n}\big)$ as $n \to \infty$ for any $p > 2$, in contrast to the conjecture (1.26) of [60].

Prior to passing to the general Lipschitz extension theorem that underlies the new results that were described above, we will further illustrate its utility by stating one more concrete application. For each $p \in [1,\infty]$ and $n \in \mathbb{N}$, if $k \in \{1,\ldots,n\}$, then let $(\ell_p^n)_{\leqslant k}$ denote the subset of $\mathbb{R}^n$ consisting of those vectors with at most k nonzero coordinates, equipped with the metric that is inherited from ℓ_p^n.

Theorem 20. *For every $p \in [1,\infty]$, every $n \in \mathbb{N}$ and every $k \in \{1,\ldots,n\}$ we have*

$$\mathsf{e}\big((\ell_p^n)_{\leqslant k}\big) \lesssim k^{\max\{\frac{1}{p},\frac{1}{2}\}}.$$

Theorem 18 is the special case $k = n$ and $p \geqslant 2$ of Theorem 20. If $1 \leqslant p \leqslant 2$ and $k = n$, then Theorem 20 is the estimate (1.24), which is the best-known upper bound on $\mathsf{e}(\ell_p^n)$ for p in this range. However, for general $k \in \{1,\ldots,n\}$ Theorem 20 yields a refinement of (1.24) in the entire range $p \in [1, \infty]$ which does not seem to follow from previously known results. In particular, the case $p = 2$ of Theorem 20 becomes

$$\mathsf{e}\big((\ell_2^n)_{\leqslant k}\big) \lesssim \sqrt{k}. \tag{1.27}$$

Even though (1.27) concerns a Euclidean setting, its proof relies on a construction that employs a multi-scale partitioning scheme using balls of an auxiliary metric on $\mathbb{R}^n$ that differs from the ambient Euclidean metric. The utility of such a non-Euclidean geometric reasoning despite the Euclidean nature of the question being studied is discussed further in Section 1.4.

1.4 A volumetric upper bound on the Lipschitz extension modulus

We will prove that Theorem 20 (hence also its special cases Theorem 14 and Theorem 18) is a consequence of Theorem 21 below, which is a Lipschitz extension theorem for subsets of finite dimensional normed spaces in terms of volumes of hyperplane projections of their unit balls. Throughout what follows, for dealing with volumetric notions we will adhere to the following conventions. Given $n \in \mathbb{N}$, when we say that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space we mean that the underlying vector space is $\mathbb{R}^n$ and that $\|\cdot\|_{\mathbf{X}} : \mathbb{R}^n \to [0, \infty)$ is a norm on $\mathbb{R}^n$. This is, of course, always achievable by fixing any scalar product on an n-dimensional normed space. While the ensuing statements hold in this setting, i.e., for an arbitrary identification of $\mathbf{X}$ with $\mathbb{R}^n$, a judicious choice of such an identification is beneficial; the discussion of this important matter is postponed to Section 1.6.2 because it is not needed for the initial description of the main results. We will continue using the notation

$$B_{\mathbf{X}} = \{x \in \mathbb{R}^n : \|x\|_{\mathbf{X}} \leqslant 1\}$$

for the unit ball of $\mathbf{X}$. Also, given $\mathcal{C} \subseteq \mathbb{R}^n$ we denote by $\mathcal{C}_{\mathbf{X}}$ the metric space consisting of the set $\mathcal{C}$ equipped with the metric that is inherited from $\|\cdot\|_{\mathbf{X}}$. This notation is important for us because we will crucially need to simultaneously consider more than one norm on $\mathbb{R}^n$.

Theorem 21. *Suppose that $n \in \mathbb{N}$ and that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ are two normed spaces. Then, for every $\mathcal{C} \subseteq \mathbb{R}^n$ we have the following upper bound on the Lipschitz extension modulus of $\mathcal{C}_{\mathbf{X}}$:*

$$\mathsf{e}(\mathcal{C}_{\mathbf{X}}) \lesssim \Bigg(\sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{\|x-y\|_{\mathbf{X}}}{\|x-y\|_{\mathbf{Y}}}\Bigg) \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \Bigg(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp} B_{\mathbf{Y}}\big)}{\mathrm{vol}_n(B_{\mathbf{Y}})} \cdot \frac{\|x-y\|_{\ell_2^n}}{\|x-y\|_{\mathbf{X}}}\Bigg). \tag{1.28}$$

We will next discuss the geometric meaning of Theorem 21 and derive some of its consequences, including Theorem 20. Firstly, by homogeneity the case $\mathcal{C} = \mathbb{R}^n$

of (1.28) becomes

$$\mathsf{e}(\mathbf{X}) \lesssim \Big(\sup_{y\in\partial B_{\mathbf{Y}}} \|y\|_{\mathbf{X}}\Big) \sup_{x\in\partial B_{\mathbf{X}}} \left(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp} B_{\mathbf{Y}}\big)}{\mathrm{vol}_n(B_{\mathbf{Y}})}\|x\|_{\ell_2^n}\right). \tag{1.29}$$

The quantity $\sup_{y\in\partial B_{\mathbf{Y}}}\|y\|_{\mathbf{X}}$ in (1.29) is the norm $\|\mathsf{Id}_n\|_{\mathbf{Y}\to\mathbf{X}}$ of the identity matrix $\mathsf{Id}_n \in \mathsf{M}_n(\mathbb{R})$ as an operator from $\mathbf{Y}$ to $\mathbf{X}$. Alternatively,

$$\sup_{y\in\partial B_{\mathbf{Y}}} \|y\|_{\mathbf{X}} = \frac{1}{2}\,\mathrm{diam}_{\mathbf{X}}(B_{\mathbf{Y}}),$$

where for each $\mathcal{C} \subseteq \mathbb{R}^n$ we denote its diameter with respect to the metric that $\mathbf{X}$ induces by

$$\mathrm{diam}_{\mathbf{X}}(\mathcal{C}) = \sup_{x,y\in\mathcal{C}} \|x-y\|_{\mathbf{X}}.$$

Given a convex body $K \subseteq \mathbb{R}^n$, let $\Pi^*K \subseteq \mathbb{R}^n$ be the polar of the *projection body* of K, which is defined to be the unit ball of the norm $\|\cdot\|_{\Pi^*K}$ on $\mathbb{R}^n$ that is given by setting for every $x \in \mathbb{R}^n \smallsetminus \{0\}$,

$$\|x\|_{\Pi^*K} \stackrel{\mathrm{def}}{=} \frac{1}{2}\int_{\partial K} \big|\langle x, N_K(y)\rangle\big|\,\mathrm{d}y = \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp} K\big)\|x\|_{\ell_2^n}, \tag{1.30}$$

where $N_K(y) \in S^{n-1}$ denotes the unit outer normal to ∂K at $y \in \partial K$ (which is uniquely defined almost everywhere with respect to the surface-area measure on ∂K), and the final equality in (1.30) is the Cauchy projection formula (see, e.g., [102, Appendix A]). The projection body ΠK of K is the polar of Π^*K. These important notions were introduced by Petty [251]. When $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space let $\Pi^*\mathbf{X}$ be the normed space whose unit ball is $\Pi^*B_{\mathbf{X}}$. Let $\Pi\mathbf{X} = (\Pi^*\mathbf{X})^*$ be the normed space whose unit ball is $\Pi B_{\mathbf{X}}$.

By substituting (1.30) into (1.29) we get the following interpretation of our bound on $\mathsf{e}(\mathbf{X})$ in terms of analytic and geometric properties of projection bodies; it is worthwhile to state it as a separate corollary even though it is only a matter of notation because of its intrinsic interest and also because these alternative viewpoints were useful for guiding some of the subsequent considerations.

Corollary 22. *Any two normed spaces* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$, $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ *satisfy*

$$\begin{aligned}\mathsf{e}(\mathbf{X}) &\lesssim \frac{\mathrm{diam}_{\mathbf{X}}(B_{\mathbf{Y}})\,\mathrm{diam}_{\Pi^*\mathbf{Y}}(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \asymp \frac{\|\mathsf{Id}_n\|_{\mathbf{Y}\to\mathbf{X}}\|\mathsf{Id}_n\|_{\mathbf{X}\to\Pi^*\mathbf{Y}}}{\mathrm{vol}_n(B_{\mathbf{Y}})}\\ &= \frac{\|\mathsf{Id}_n\|_{\mathbf{X}\to\mathbf{Y}}\|\mathsf{Id}_n\|_{\Pi\mathbf{Y}\to\mathbf{X}^*}}{\mathrm{vol}_n(B_{\mathbf{Y}})} \asymp \frac{\mathrm{diam}_{\mathbf{X}}(B_{\mathbf{Y}})\,\mathrm{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})}{\mathrm{vol}(B_{\mathbf{Y}})}.\end{aligned} \tag{1.31}$$

The penultimate step in (1.31) is duality (the norm of an operator equals the norm of its adjoint) and the final quantity in (1.31) relates Theorem 21 to the second estimate in Theorem 3.

Remark 23. It is worthwhile to note that Corollary 22 has the right *affine invariance*. For $S \in \mathsf{SL}_n(\mathbb{R})$ let $S\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{S\mathbf{X}})$ be the normed space whose unit ball is $SB_{\mathbf{X}}$. Equivalently, $\|x\|_{S\mathbf{X}} = \|S^{-1}x\|_{\mathbf{X}}$ for every $x \in \mathbb{R}^n$. Then $\mathbf{X}$ and $S\mathbf{X}$ are isometric, so $\mathsf{e}(S\mathbf{X}) = \mathsf{e}(\mathbf{X})$. We have $(S\mathbf{X})^* = (S^*)^{-1}\mathbf{X}^*$ (by definition), and $\Pi(SB_{\mathbf{Y}}) = (S^*)^{-1}\Pi B_{\mathbf{Y}}$ by [251]. From this we see that $\operatorname{diam}_{(S\mathbf{X})^*}(\Pi B_{S\mathbf{Y}}) = \operatorname{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})$. Thus, the minimum of the right-hand side of (1.31) over all normed spaces $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ is also invariant under the action of $\mathsf{SL}_n(\mathbb{R})$.

The special case of Theorem 21 in which the normed space $\mathbf{Y}$ coincides with the given normed space $\mathbf{X}$ is in itself a nontrivial bound on the Lipschitz extension modulus. Examining this special case first will help elucidate how the idea arose to introduce an auxiliary space $\mathbf{Y}$ that may differ from $\mathbf{X}$, and why this can yield stronger estimates. If $\mathbf{X} = \mathbf{Y}$, then the bound (1.28) becomes

$$\mathsf{e}(\mathcal{C}_{\mathbf{X}}) \lesssim \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \left(\frac{\operatorname{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp} B_{\mathbf{X}}\big)}{\operatorname{vol}_n(B_{\mathbf{X}})} \cdot \frac{\|x-y\|_{\ell_2^n}}{\|x-y\|_{\mathbf{X}}}\right). \tag{1.32}$$

Correspondingly, the bound (1.29) becomes

$$\mathsf{e}(\mathbf{X}) \lesssim \sup_{z\in\partial B_{\mathbf{X}}} \left(\frac{\operatorname{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp} B_{\mathbf{X}}\big)}{\operatorname{vol}_n(B_{\mathbf{X}})}\|z\|_{\ell_2^n}\right) = \frac{\operatorname{diam}_{\Pi^*\mathbf{X}}(B_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})}. \tag{1.33}$$

Even these weaker estimates suffice to obtain new results, e.g., we will see that this is so if $2 \leqslant p = O(1)$ and $\mathbf{X} = \ell_p^n$. However, as we will soon explain, (1.33) does not imply an upper bound on ℓ_∞^n that is better than the aforementioned general bound of [140]. Despite this shortcoming of (1.32) and (1.33) relative to (1.28), it is worthwhile to state these special cases of Theorem 21 separately because they are simpler than (1.28) and hence perhaps somewhat easier to remember. Moreover, a naïve way to enhance the applicability of (1.32) is to leverage the fact that the Lipschitz extension modulus is a bi-Lipschitz invariant, so that

$$\mathsf{e}(\mathcal{C}_{\mathbf{X}}) \leqslant \|\mathsf{Id}_n\|_{\operatorname{Lip}(\mathcal{C}_{\mathbf{Y}},\mathcal{C}_{\mathbf{X}})}\|\mathsf{Id}_n\|_{\operatorname{Lip}(\mathcal{C}_{\mathbf{X}},\mathcal{C}_{\mathbf{Y}})}\mathsf{e}(\mathcal{C}_{\mathbf{Y}}).$$

Consequently, by estimating $\mathsf{e}(\mathcal{C}_{\mathbf{Y}})$ through (1.32) we formally deduce from (1.32) that

$$\begin{aligned}\mathsf{e}(\mathcal{C}_{\mathbf{X}}) \lesssim{}& \left(\sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{\|x-y\|_{\mathbf{X}}}{\|x-y\|_{\mathbf{Y}}}\right)\left(\sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{\|x-y\|_{\mathbf{Y}}}{\|x-y\|_{\mathbf{X}}}\right)\\ &\cdot \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \left(\frac{\operatorname{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp} B_{\mathbf{Y}}\big)}{\operatorname{vol}_n(B_{\mathbf{Y}})} \cdot \frac{\|x-y\|_{\ell_2^n}}{\|x-y\|_{\mathbf{Y}}}\right).\end{aligned} \tag{1.34}$$

We do not see how to deduce Theorem 18 and Theorem 20 from (1.34). However, we will show that (1.34) suffices for proving Theorem 14 (as well as some other results that will be presented later). In summary, even the case of Theorem 21 in which the auxiliary space $\mathbf{Y}$ coincides with $\mathbf{X}$ is valuable, but Theorem 21 does not follow from merely combining its special case $\mathbf{Y} = \mathbf{X}$ with bi-Lipschitz invariance.

Given a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $z \in \mathbb{R}^n \setminus \{0\}$, the quantity

$$\frac{1}{n}\operatorname{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp} B_{\mathbf{X}}\big)\|z\|_{\ell_2^n} \tag{1.35}$$

is equal to the volume of the cone

$$\mathsf{Cone}_z(B_{\mathbf{X}}) \stackrel{\mathrm{def}}{=} \operatorname{conv}\big(\{z\} \cup \mathsf{Proj}_{z^\perp} B_{\mathbf{X}}\big) \subseteq \mathbb{R}^n \tag{1.36}$$

whose base is the $(n-1)$-dimensional convex set $\mathsf{Proj}_{z^\perp} B_{\mathbf{X}} \subseteq z^\perp$ and whose apex is z. In (1.36) and throughout what follows, $\operatorname{conv}(\cdot)$ denotes the convex hull. Thus, the estimate (1.33) can be restated as follows:

$$\mathsf{e}(\mathbf{X}) \lesssim n \sup_{z \in \partial B_{\mathbf{X}}} \frac{\operatorname{vol}_n\big(\mathsf{Cone}_z(B_{\mathbf{X}})\big)}{\operatorname{vol}_n(B_{\mathbf{X}})}. \tag{1.37}$$

Through (1.37) we see that the geometric interpretation of the “bad spaces” $\mathbf{X}$ for (1.33) is that these are the spaces that have a “pointy direction” $z \in \partial B_{\mathbf{X}}$ for which the volume of the cone $\mathsf{Cone}_z(B_{\mathbf{X}})$ is a significant fraction of the volume of $B_{\mathbf{X}}$. Examples will be presented next, but note first that a short geometric argument (see the proof of [109, Lemma 5.1]) shows that $\operatorname{vol}_n(\mathsf{Cone}_z(B_{\mathbf{X}})) \leqslant \operatorname{vol}_n(B_{\mathbf{X}})/2$, so the right-hand side of (1.37) is at most $n/2$. Hence, (1.33) is a refinement of the classical bound $\mathsf{e}(\mathbf{X}) \lesssim n$ of [140].

Nevertheless, a “vanilla” application of (1.33) does not yield an asymptotically better estimate than that of [140] even when $\mathbf{X} = \ell_\infty^n$. Indeed, $B_{\ell_\infty^n} = [-1,1]^n$ and a simple argument (see [75]) shows that

$$\forall z \in \mathbb{R}^n \setminus \{0\}, \qquad \frac{\operatorname{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp}[-1,1]^n\big)}{\operatorname{vol}_n([-1,1]^n)} = \frac{\|z\|_{\ell_1^n}}{2\|z\|_{\ell_2^n}}. \tag{1.38}$$

So, by considering the all 1’s vector $z = \mathbf{1}_{\{1,\ldots,n\}} \in \partial B_{\ell_\infty^n}$ we see that for $\mathbf{X} = \ell_\infty^n$ the right-hand side of (1.33) is at least $n/2$. The right-hand side of (1.33) is at least $n/2$ when $\mathbf{X} = \ell_1^n$, as seen by taking $z = (1,0,\ldots,0) \in \partial B_{\ell_1^n}$. Such “problematic" directions $z \in \partial B_{\mathbf{X}}$ can sometimes be the overwhelming majority of $\partial B_{\mathbf{X}}$. Consider Ball’s counterexample [21] to the Shepard Problem [287], which states that for any $n \in \mathbb{N}$ there is a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ such that $\operatorname{vol}_n(B_{\mathbf{X}}) = 1$ yet $\operatorname{vol}_{n-1}(\mathsf{Proj}_{z^\perp} B_{\mathbf{X}}) \gtrsim \sqrt{n}$ for *every* $z \in S^{n-1}$. Since $\operatorname{vol}_n(B_{\ell_2^n}) \leqslant (3/\sqrt{n})^n$ while $\operatorname{vol}_n(B_{\mathbf{X}}) = 1$, the proportion of those $z \in \partial B_{\mathbf{X}}$ for which $\|z\|_{\ell_2^n} \geqslant \sqrt{n}/4$ tends to 1 as $n \to \infty$ (exponentially fast). Any such z satisfies

$$\frac{\operatorname{vol}_{n-1}(\mathsf{Proj}_{z^\perp} B_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})}\|z\|_{\ell_2^n} \gtrsim n.$$

These obstacles can sometimes be overcome by perturbing the given normed space $\mathbf{X}$ prior to invoking (1.33), i.e., by using of Theorem 21 with a suitably chosen auxiliary normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$. In particular, $\|\cdot\|_{\ell_2^n} \leqslant n^{1/2-1/p}\|\cdot\|_{\ell_p^n}$ when $p \geqslant 2$ by Hölder's inequality, so Theorem 18 follows from a substitution of the space $\mathbf{Y}_p^n$ of Theorem 24 below into Theorem 21 (with $\mathbf{X} = \ell_p^n$), or even into (1.34).

Theorem 24. *For any $n \in \mathbb{N}$ and $p \in [1,\infty]$ there is a normed space*

$$\mathbf{Y}_p^n = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}_p^n})$$

that satisfies

$$\forall x \in \mathbb{R}^n \smallsetminus \{0\}, \quad \|x\|_{\mathbf{Y}_p^n} \asymp \|x\|_{\ell_p^n}, \quad \textit{and} \quad \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp} B_{\mathbf{Y}_p^n}\big)}{\mathrm{vol}_n\big(B_{\mathbf{Y}_p^n}\big)} \lesssim n^{\frac{1}{p}}. \tag{1.39}$$

The case $p=\infty$ of Theorem 24 implies Theorem 20 by an application of Theorem 21. Indeed, fix $p \geqslant 1$ and $n \in \mathbb{N}$. Suppose that $x, y \in (\ell_p^n)_{\leqslant k}$ for some $k \in \{1,\ldots,n\}$. Then $x-y$ has at most $2k$ nonzero coordinates. Therefore, if $\mathbf{Y}_\infty^n$ is as in Theorem 24, then by Hölder's inequality we have

$$(2k)^{-\max\{\frac{1}{2}-\frac{1}{p},0\}}\|x-y\|_{\ell_2^n} \leqslant \|x-y\|_{\ell_p^n} \leqslant (2k)^{\frac{1}{p}}\|x-y\|_{\ell_\infty^n} \asymp k^{\frac{1}{p}}\|x-y\|_{\mathbf{Y}_\infty^n}. \tag{1.40}$$

Theorem 20 follows by substituting these bounds and the case $p=\infty$ of (1.39) into (1.28). Observe that we would have obtained the weaker bound $\mathsf{e}((\ell_p^n)_{\leqslant k}) \lesssim k^{1/p+1/2}$ if we used (1.34) instead of (1.28).

If $p = O(1)$, then one can take $\mathbf{Y}_p^n = \ell_p^n$ in Theorem 24. In fact, the direction $z \in S^{n-1}$ at which

$$\max_{z \in S^{n-1}} \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp} B_{\ell_p^n}\big) \tag{1.41}$$

is attained was determined by Barthe and the author in [32]. This result implies that

$$\forall p \geqslant 1, \qquad \max_{z \in S^{n-1}} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp} B_{\ell_p^n}\big)}{\mathrm{vol}_n(B_{\ell_p^n})} \asymp n^{\frac{1}{p}}\sqrt{\min\{p,n\}}. \tag{1.42}$$

As [32] computes (1.41) exactly, the implicit constant factors in (1.42) can be evaluated, but in the present context such precision is of secondary importance. While (1.42) follows from [32] (see the deduction in [227]), we will give a self-contained proof of (1.42) in Section 6 as a special case of a more general result that we will use for other purposes as well. In the range $q \in (2,\infty)$, a different approach to computing (1.41) was found in [157]. Earlier methods for estimating (1.41) with worse lower order factors are due to [223, 286]; the latter is an adaptation of an idea (used for related purposes) in [45].

For each $k \in \{1,\ldots,n\}$, by applying (1.28) with $\mathbf{Y} = \ell_q^n$ for some $q \geqslant p$, using (1.42) with p replaced by q, and optimizing the resulting bound over q, one obtains a result that matches Theorem 20 up to unbounded lower order factors. More precisely, the best that one can get with this approach is when

$$q = \max\Big\{2\log\Big(\frac{n}{k}\Big), p\Big\}$$

if $p \leqslant \log(2k)$. If $p \geqslant \log(2k)$, then use (1.28) with $\mathbf{Y} = \ell_{\log(2k)}^n$.

Theorem 24 provides an auxiliary space $\mathbf{Y}$ for which a use of (1.28) removes the above lower order factors, and yields a sharp result when $p = \infty$ (we conjecture that it is sharp for any $p \geqslant 2$). Regardless of whether we apply (1.28) with the space $\mathbf{Y} = \mathbf{Y}_\infty^n$ of Theorem 24 or with $\mathbf{Y} = \ell_q^n$ for a suitable choice of $q \geqslant p$, we have seen that without using an auxiliary space $\mathbf{Y} \neq \ell_p^n$ in (1.28) we do not come close to such results.

Even though in Theorem 21 we are interested in extending functions that are Lipschitz in the metric that is induced by the given norm $\|\cdot\|_{\mathbf{X}}$, the underlying reason for the bounds of Theorem 21 is a partitioning scheme (to be described below) that iteratively carves out balls in the metric that is induced by the auxiliary norm $\|\cdot\|_{\mathbf{Y}}$. So, the perturbation of $\mathbf{X}$ into $\mathbf{Y}$ amounts to exhibiting a Lipschitz extension operator through the use of a multi-scale construction that utilizes geometric shapes that differ from balls in the ambient metric. This strategy is feasible because the quantity $\mathsf{e}(\mathcal{C}_{\mathbf{X}})$ in the left-hand side of (1.32) is a bi-Lipschitz invariant, while the volumes that appear in the right-hand side of (1.32) scale exponentially in n. Hence, by passing to an equivalent norm one could hope to reduce the right-hand side of (1.32) significantly, while not changing the left-hand side of (1.32) by too much.

This perturbative approach is decisively useful for $\mathbf{X} = \ell_\infty^n$. When one unravels the ensuing proofs, the upper bound on $\mathsf{e}(\ell_\infty^n)$ of Theorem 14 arises from a multi-scale construction of an extension operator (using a *gentle partition of unity* [173]) that utilizes a partition of space that is obtained by iteratively removing sets of the form $x + rB_{\mathbf{Y}_\infty^n}$, where $\mathbf{Y}_\infty^n$ is as in Theorem 24. If one carries out the same procedure while using balls of the intrinsic metric of ℓ_∞^n (namely, hypercubes $x + r[-1,1]^n$ in place of $x + rB_{\mathbf{Y}_\infty^n}$, which look like hypercubes with "rounded corners"), then only the weaker bound $\mathsf{e}(\ell_\infty^n) \lesssim n$ is obtained. We already mentioned that such a phenomenon even occurs in the proof of the Euclidean estimate (1.27).

The following two examples describe further uses of Theorem 21; we will work out several more later.

Example 25. In the forthcoming work [234], the author and Schechtman prove (for an application to metric embedding theory) the following asymptotic evaluation of the maximal volumes of hyperplane projections of the unit balls of the Schatten–von Neumann trace classes:

$$\forall q \geqslant 1, \qquad \max_{A \in \mathsf{M}_n(\mathbb{R}) \setminus \{0\}} \frac{\mathrm{vol}_{n^2-1}\big(\mathsf{Proj}_{A^\perp} B_{\mathsf{S}_q^n}\big)}{\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)} \asymp n^{\frac{1}{2}+\frac{1}{q}} \sqrt{\min\{q,n\}}. \tag{1.43}$$

Upon substitution into Theorem 21, this yields the following new estimates on the Lipschitz extension moduli of Schatten–von Neumann trace classes, which holds for every $p \geqslant 1$ and every integer $n \geqslant 2$:

$$\mathsf{e}\big(\mathsf{S}_p^n\big) \lesssim \begin{cases} n^{\frac{1}{2}+\frac{1}{p}} & \text{if } p \in [1,2], \\ n\sqrt{\min\{p, \log n\}} & \text{if } p \in [2,\infty]. \end{cases} \tag{1.44}$$

Indeed, by Hölder's inequality

$$\|\cdot\|_{\mathsf{S}_2^n} \leqslant n^{\max\{0,\frac{1}{2}-\frac{1}{p}\}}\|\cdot\|_{\mathsf{S}_p^n},$$

so (1.44) for $p \leqslant \log n$ follows from a substitution of these point-wise bounds and (1.43) when $q = p$ into the case $\mathbf{X} = \mathbf{Y} = \mathsf{S}_p^n$ of Theorem 21. The case $p \geqslant \log n$ of (1.44) follows from the same reasoning using (1.43) when $q = \log n$ and Theorem 21 for $\mathbf{X} = \mathsf{S}_p^n$ and $\mathbf{Y} = \mathsf{S}_q^n$, since in this case $d_{\mathrm{BM}}(\mathsf{S}_p^n, \mathsf{S}_q^n) \lesssim 1$. Note that, since $\dim(\mathsf{S}_p^n) = n^2$, for every $p \in [1,\infty]$ the bound on $\mathsf{e}(\mathsf{S}_p^n)$ in (1.44) is $o(\dim(\mathsf{S}_p^n))$, i.e., it is asymptotically better than what follows from [140].

More generally, given $p \geqslant 1$, an integer $n \geqslant 2$ and $r \in \{3,\ldots,n\}$, let $(\mathsf{S}_p^n)_{\leqslant r}$ be the set of n by n matrices of rank at most r, equipped with the metric inherited from S_p^n. Then, (1.44) has the following strengthening:

$$\mathsf{e}\big((\mathsf{S}_p^n)_{\leqslant r}\big) \lesssim r^{\max\{\frac{1}{p},\frac{1}{2}\}}\sqrt{n}\cdot \begin{cases} \sqrt{\max\{\log(\frac{n}{r}), p\}} & \text{if } p \leqslant \log r, \\ \sqrt{\log n} & \text{if } p \geqslant \log r. \end{cases} \tag{1.45}$$

To justify (1.45), apply Theorem 21 with $\mathbf{X} = \mathsf{S}_p^n$ and $\mathbf{Y} = \mathsf{S}_q^n$ for some $q \geqslant p$ while using (1.43), and optimize the resulting bound over q. Specifically, as for $A, B \in (\mathsf{S}_p^n)_{\leqslant r}$ the matrix $A - B$ has at most $2r$ nonzero singular values, by Hölder's inequality we have

$$\|A-B\|_{\mathsf{S}_2^n} \leqslant (2r)^{\max\{0,\frac{1}{2}-\frac{1}{p}\}}\|A-B\|_{\mathsf{S}_p^n}$$

and

$$\|A-B\|_{\mathsf{S}_p^n} \leqslant (2r)^{\frac{1}{p}-\frac{1}{q}}\|A-B\|_{\mathsf{S}_q^n}.$$

In combination with (1.43), we therefore get the following bound from (1.28):

$$\begin{aligned}\mathsf{e}\big((\mathsf{S}_p^n)_{\leqslant r}\big) &\lesssim \Bigg(\sup_{\substack{A,B\in(\mathsf{S}_p^n)_{\leqslant r}\\ A\neq B}} \frac{\|A-B\|_{\mathsf{S}_p^n}}{\|A-B\|_{\mathsf{S}_q^n}}\Bigg) \sup_{\substack{A,B\in(\mathsf{S}_p^n)_{\leqslant r}\\ A\neq B}} \Bigg(n^{\frac{1}{2}+\frac{1}{q}}\sqrt{q}\frac{\|A-B\|_{\mathsf{S}_2^n}}{\|A-B\|_{\mathsf{S}_p^n}}\Bigg) \\ &\lesssim r^{\frac{1}{p}-\frac{1}{q}} n^{\frac{1}{2}+\frac{1}{q}}\sqrt{q}\, r^{\max\{\frac{1}{2}-\frac{1}{p},0\}}. \end{aligned} \tag{1.46}$$

The $q \geqslant p$ that minimizes the right-hand side of (1.46) is $\max\{2\log(n/r), p\}$, yielding (1.45) when $p \leqslant \log r$. If $p \geqslant \log r$, then $\|A-B\|_{\mathsf{S}_p^n} \asymp \|A-B\|_{\mathsf{S}_{\log r}^n}$ for every $A, B \in (\mathsf{S}_p^n)_{\leqslant r}$, so (1.45) reduces to its special case $p = \log r$.

We conjecture that it is possible to replace the logarithmic factor in (1.45) by a universal constant, i.e.,

$$\mathsf{e}\big((\mathsf{S}_p^n)_{\leqslant r}\big) \lesssim r^{\max\{\frac{1}{p},\frac{1}{2}\}}\sqrt{n}. \tag{1.47}$$

As we will see in Section 1.6, Conjecture 26 below is equivalent to the symmetric isomorphic reverse isoperimetry conjecture (see Conjecture 47) for $\mathsf{M}_n(\mathbb{R})$ equipped with the operator norm, which is an especially interesting special case of this much more general conjectural phenomenon; by reasoning as we did in the above deduction of Theorem 20 from (the special case $p=\infty$ of) Theorem 24 (recall the discussion immediately following (1.40)), a positive answer to Conjecture 26 would imply (1.47).

Conjecture 26. For every $n \in \mathbb{N}$ there exists a normed space

$$\mathbf{Y} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathbf{Y}})$$

such that for every nonzero n by n matrix $A \in \mathsf{M}_n(\mathbb{R}) \smallsetminus \{0\}$ we have $\|A\|_{\mathbf{Y}} \asymp \|A\|_{\mathsf{S}_\infty^n}$ and

$$\mathrm{vol}_{n^2-1}(\mathsf{Proj}_{A^\perp} B_{\mathbf{Y}}) \lesssim \mathrm{vol}_{n^2}(B_{\mathbf{Y}})\sqrt{n}.$$

Example 27. Since the $\ell_\infty^n(\ell_\infty^n)$ norm on $\mathsf{M}_n(\mathbb{R})$ is isometric to $\ell_\infty^{n^2}$, by Theorem 24 there is a normed space $\mathbf{Y} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathbf{Y}})$ that satisfies

$$\|A\|_{\ell_\infty^n(\ell_\infty^n)} \leqslant \|A\|_{\mathbf{Y}} \lesssim \|A\|_{\ell_\infty^n(\ell_\infty^n)}$$

for every $A \in \mathsf{M}_n(\mathbb{R})$, and

$$\max_{A\in \mathsf{M}_n(\mathbb{R})\smallsetminus\{0\}} \frac{\mathrm{vol}_{n^2-1}\big(\mathsf{Proj}_{A^\perp} B_{\mathbf{Y}}\big)}{\mathrm{vol}_{n^2}(B_{\mathbf{Y}})} = O(1).$$

By Hölder's inequality, for every $p,q \in [1,\infty]$ and $A \in \mathsf{M}_n(\mathbb{R})$ we have

$$\|A\|_{\ell_p^n(\ell_q^n)} \leqslant n^{\frac{1}{p}+\frac{1}{q}}\|A\|_{\ell_\infty^n(\ell_\infty^n)} \leqslant n^{\frac{1}{p}+\frac{1}{q}}\|A\|_{\mathbf{Y}}$$

and

$$\|A\|_{\ell_2^n(\ell_2^n)} \leqslant n^{\max\{\frac{1}{2}-\frac{1}{p},0\}+\max\{\frac{1}{2}-\frac{1}{q},0\}}\|A\|_{\ell_p^n(\ell_q^n)}.$$

Therefore, Theorem 21 gives the Lipschitz extension bound

$$\mathsf{e}\big(\ell_p^n(\ell_q^n)\big) \lesssim n^{\frac{1}{p}+\frac{1}{q}+\max\{\frac{1}{2}-\frac{1}{p},0\}+\max\{\frac{1}{2}-\frac{1}{q},0\}} = n^{\max\{1,\frac{1}{p}+\frac{1}{q},\frac{1}{2}+\frac{1}{p},\frac{1}{2}+\frac{1}{q}\}}. \tag{1.48}$$

As in the case of ℓ_p^n, we get (1.48) if $p,q=O(1)$ by using Theorem 21 with $\mathbf{Y} = \mathbf{X} = \ell_p^n(\ell_p^n)$, but otherwise we need to work with an auxiliary space $\mathbf{Y} \neq \mathbf{X}$ as above. Specifically, in Section 6 we will prove the following asymptotic evaluation of the maximal volume of hyperplane projections of the unit ball of $\ell_p^n(\ell_q^n)$:

$$\max_{A\in \mathsf{M}_n(\mathbb{R})\setminus\{0\}} \frac{\mathrm{vol}_{n^2-1}\big(\mathsf{Proj}_{A^\perp} B_{\ell_p^n(\ell_q^n)}\big)}{\mathrm{vol}_{n^2}\big(B_{\ell_p^n(\ell_q^n)}\big)} \asymp \begin{cases} n & \text{if } n \leqslant \min\{\sqrt{p},q\},\\ \sqrt{q}\,n^{\frac12+\frac1q} & \text{if } q \leqslant n \leqslant \sqrt{p},\\ \sqrt{p} & \text{if } \sqrt{p} \leqslant n \leqslant \min\{p,q\},\\ \sqrt{pq}\,n^{\frac1q-\frac12} & \text{if } \max\{\sqrt{p},q\} \leqslant n \leqslant p,\\ n^{\frac12+\frac1p} & \text{if } p \leqslant n \leqslant q,\\ \sqrt{q}\,n^{\frac1p+\frac1q} & \text{if } n \geqslant \max\{p,q\}. \end{cases} \tag{1.49}$$

The intricacy of (1.49) is perhaps unexpected, though it is nonetheless sharp in all of the six ranges (depending on the relative locations of p, q, n and, somewhat curiously, $\sqrt{p}$) that appear in (1.49). By reasoning analogously to the discussion following (1.42), one can prove a bound on $\mathsf{e}(\ell_p^n(\ell_q^n))$ that matches (1.48) up to lower order factors by applying Theorem 21 with $\mathbf{Y} = \ell_r^n(\ell_s^n)$ and then optimizing over $r, s \geqslant 1$. For the sole purpose of this application, only the range $n \geqslant \max\{p,q\}$ of (1.49) is needed. However, results such as (1.49) have geometric interest in their own right for all of the possible values of the relevant parameters. We will actually prove a version of (1.49) for $\ell_p^n(\ell_q^m)$ even when $n \neq m$; the case of rectangular matrices is independently interesting, but we will also use it elsewhere (see Remark 56 below).

Problem 28. Determine the exact maximizers of volumes of hyperplane projections of the unit balls of S_p^n and $\ell_p^n(\ell_q^n)$, i.e., for which $A \in \mathsf{M}_n(\mathbb{R}) \setminus \{0\}$ are the maxima in (1.43) and (1.49) attained.

1.5 A dimension-independent extension theorem

In the preceding sections we stated all of the extension theorems using the traditional setup that aims to extend a Lipschitz function to a function that is Lipschitz with respect to the given metric. However, all of our new (positive) extension theorems are a consequence of Theorem 29 below, which is a nonstandard Lipschitz extension theorem.

Theorem 29 asserts that if $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space and f is a 1-Lipschitz function from a subset of $\mathbb{R}^n$ to a Banach space $\mathbf{Z}$, then f can be extended to a $\mathbf{Z}$-valued function that is defined on all of $\mathbb{R}^n$ and is $O(1)$-Lipschitz with respect to the metric that is induced on $\mathbb{R}^n$ by the norm $|||\cdot||| = 2\|\cdot\|_{\Pi^*\mathbf{X}}/\,\mathrm{vol}_n(B_{\mathbf{X}})$, i.e., a suitable rescaling of the norm whose unit ball is the polar projection body of $B_{\mathbf{X}}$. This rescaling ensures that $|||\cdot|||$ dominates $\|\cdot\|_{\mathbf{X}}$; indeed, by an elementary geometric argument (see Remark 112),

$$\forall x \in \mathbb{R}^n, \quad \|x\|_{\mathbf{X}} \leqslant \frac{2\|x\|_{\Pi^*\mathbf{X}}}{\mathrm{vol}_n(B_{\mathbf{X}})} \leqslant n\|x\|_{\mathbf{X}}. \tag{1.50}$$

Thus, the conclusion of Theorem 29 that the extended function is Lipschitz with respect to $|||\cdot|||$ is less stringent than the traditional requirement that it should be Lipschitz with respect to $\|\cdot\|_{\mathbf{X}}$, but Theorem 29 has the feature that the upper bound on the Lipschitz constant is independent of the dimension.

Theorem 29. *Fix $n \in \mathbb{N}$ and a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. Fix also a Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$. Suppose that $\mathcal{C} \subseteq \mathbb{R}^n$ and $f : \mathcal{C} \to \mathbf{Z}$ is 1-Lipschitz with respect to the metric that is induced by $\|\cdot\|_{\mathbf{X}}$, i.e., $\|f(x) - f(y)\|_{\mathbf{Z}} \leqslant \|x - y\|_{\mathbf{X}}$ for every $x, y \in \mathcal{C}$. Then, there exists $F : \mathbb{R}^n \to \mathbf{Z}$ that coincides with f on $\mathcal{C}$ and satisfies*

$$\forall x, y \in \mathbb{R}^n, \quad \|F(x) - F(y)\|_{\mathbf{Z}} \lesssim \frac{\|x - y\|_{\Pi^*\mathbf{X}}}{\mathrm{vol}_n(B_{\mathbf{X}})}.$$

To see how Theorem 29 implies Theorem 21, denote (in the setting of the statement of Theorem 21):

$$M = \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \left(\frac{\|x-y\|_{\mathbf{X}}}{\|x-y\|_{\mathbf{Y}}}\right) \tag{1.51}$$

and

$$M' = \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \left(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^{\perp}} B_{\mathbf{Y}}\big)}{\mathrm{vol}_n(B_{\mathbf{Y}})} \cdot \frac{\|x-y\|_{\ell_2^n}}{\|x-y\|_{\mathbf{X}}}\right). \tag{1.52}$$

Thus, every $x, y \in \mathcal{C}$ satisfy $\|x - y\|_{\mathbf{X}} \leqslant M\|x - y\|_{\mathbf{Y}}$ and, recalling (1.30), also

$$\frac{\|x-y\|_{\Pi^*\mathbf{Y}}}{\mathrm{vol}_n(B_{\mathbf{Y}})} \leqslant M'\|x-y\|_{\mathbf{X}}.$$

Let $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ be a Banach space. Consider an arbitrary subset $\mathcal{C}' \subseteq \mathcal{C}$. If $f : \mathcal{C}' \to \mathbf{Z}$ is 1-Lipschitz with respect to the metric that is induced by $\|\cdot\|_{\mathbf{X}}$, then the function f/M is 1-Lipschitz with respect to the metric that is induced by $\mathbf{Y}$. By Theorem 29 (with $\mathbf{X}$ replaced by $\mathbf{Y}$, $\mathcal{C}$ replaced by $\mathcal{C}'$, f replaced by f/M) we therefore see that there exists $F : \mathbb{R}^n \to \mathbf{Z}$ (for Theorem 21 we only need F to be defined on $\mathcal{C}$) that extends F and satisfies $\|F(x) - F(y)\|_{\mathbf{Z}} \lesssim M\|x - y\|_{\Pi^*\mathbf{Y}}/\mathrm{vol}_n(B_{\mathbf{Y}}) \leqslant MM'\|x - y\|_{\mathbf{X}}$ for all $x, y \in \mathcal{C}$. This coincides with (1.28).

Remark 30. Given $p \geqslant 1$, consider what happens when we apply Theorem 29 to the space $\mathbf{Y}_p^n$ of Theorem 24. We get that for any $\mathcal{C} \subseteq \mathbb{R}^n$ and any Banach space $\mathbf{Z}$, if $f : \mathcal{C} \to \mathbf{Z}$ is 1-Lipschitz with respect to the ℓ_p^n metric, then f can be extended to $F : \mathbb{R}^n \to \mathbf{Z}$ that is $O(n^{1/p})$-Lipschitz with respect to the Euclidean metric. When

$p < 2$, the Lipschitz assumption on f is less stringent than requiring it to be $O(1)$-Lipschitz with respect to the Euclidean metric, but we then get an extension F that is $O(n^{1/p})$-Lipschitz with respect to the Euclidean metric; this upper bound on the Lipschitz constant of F is asymptotically larger than the $O(\sqrt{n})$ bound that we would get if f were assumed to be 1-Lipschitz with respect to the Euclidean metric and we applied the second inequality in (1.22), but we get it while requiring less from f. In particular, when $p = 1$ we see that any $\mathbf{Z}$-valued function on a subset of $\mathbb{R}^n$ that is 1-Lipschitz with respect to the ℓ_1^n metric can be extended to a $\mathbf{Z}$-valued function defined on all of $\mathbb{R}^n$ whose Lipschitz constant with respect to the Euclidean metric is $O(n)$, while an application of [140] will give an extension that is $O(n)$-Lipschitz with respect to the ℓ_1^n metric. On the other hand, if $p > 2$, then the Lipschitz assumption on f is more stringent than requiring it to be $O(1)$-Lipschitz with respect to the Euclidean metric, but we then get an extension F that is $O(n^{1/p})$-Lipschitz with respect to the Euclidean metric, which is asymptotically better than the $O(\sqrt{n})$ bound from (1.22). In particular, when $p = \infty$ we see that any $\mathbf{Z}$-valued function on a subset of $\mathbb{R}^n$ that is 1-Lipschitz with respect to the ℓ_∞^n metric can be extended to a $\mathbf{Z}$-valued function on all of $\mathbb{R}^n$ whose Lipschitz constant with respect to the Euclidean metric is $O(1)$.

1.6 Isomorphic reverse isoperimetry

All of the applications that we found for Theorem 21 proceed by bounding the volumes of hyperplane projections of $B_{\mathbf{Y}}$ that appear in right-hand side of (1.28) by

$$\mathrm{MaxProj}(B_{\mathbf{Y}}) \stackrel{\mathrm{def}}{=} \max_{z \in S^{n-1}} \mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z^\perp} B_{\mathbf{Y}}\big). \tag{1.53}$$

Thus, it follows from (1.29) that for any two normed spaces $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ with $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ we have

$$\mathsf{e}(\mathbf{X}) \lesssim \frac{\mathrm{MaxProj}(B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}}). \tag{1.54}$$

Even though there could conceivably be an application of (1.29) that is more refined than (1.54), in this section we will investigate the ramifications of bounding $\mathrm{MaxProj}(B_{\mathbf{X}})$ as a way to use Theorem 21. This will relate to the isomorphic reverse isoperimetric phenomena that we conjectured in Section 1.1.1.

Any origin-symmetric convex body $L \subseteq \mathbb{R}^n$ satisfies

$$\mathrm{MaxProj}(L) \gtrsim \frac{\mathrm{vol}_{n-1}(\partial L)}{\sqrt{n}}. \tag{1.55}$$

Indeed, this follows immediately from the following classical *Cauchy surface area formula* (see, e.g., [282, equation (5.73)]) by bounding the integrand by its maximum:

$$\begin{aligned}\mathrm{vol}_{n-1}(\partial L) &= \frac{2\sqrt{\pi}\,\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \fint_{S^{n-1}} \mathrm{vol}_{n-1}\left(\mathsf{Proj}_{z^\perp} L\right) \mathrm{d}z \\ &\asymp \sqrt{n} \fint_{S^{n-1}} \mathrm{vol}_{n-1}\left(\mathsf{Proj}_{z^\perp} L\right) \mathrm{d}z.\end{aligned}$$

Remark 31. Using (1.55), Theorem 24 implies that Conjecture 9 (isomorphic reverse isoperimetry) holds (with S the identity mapping) when $K = B_{\ell_p^n}$ for any $p \geqslant 1$ and $n \in \mathbb{N}$. Indeed, let $\mathbf{Y}_p^n$ be the normed space from Theorem 24. By the first inequality in (1.40) we have

$$\mathrm{vol}_n\left(B_{\mathbf{Y}_p^n}\right)^{\frac{1}{n}} \asymp \mathrm{vol}_n\left(B_{\ell_p^n}\right)^{\frac{1}{n}} \asymp n^{-\frac{1}{p}}, \tag{1.56}$$

where the last equivalence in (1.56) is a standard computation (e.g., [263, p. 11]). By (1.55) and (1.56), the second inequality in (1.40) implies that the isoperimetric quotient of $B_{\mathbf{Y}_p^n}$ is $O(\sqrt{n})$. So, Conjecture 9 holds for $K = B_{\ell_p^n}$ if we take L to be a rescaling by a universal constant factor of $B_{\mathbf{Y}_p^n}$ so that $L \subseteq K$.

Thanks to (1.55), if we set $K = B_{\mathbf{X}}$ and $L = B_{\mathbf{Y}}$ in (1.54), then the right-hand side of (1.54) satisfies

$$\begin{aligned}\frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_n(L)} \mathrm{diam}_{\ell_2^n}(K) &\gtrsim \frac{\mathrm{vol}_{n-1}(\partial L)}{\sqrt{n}\,\mathrm{vol}_n(L)} \mathrm{diam}_{\ell_2^n}(K) \\ &= \frac{\mathrm{iq}(L)}{\sqrt{n}} \cdot \frac{\mathrm{diam}_{\ell_2^n}(K)}{\mathrm{vol}_n(L)^{\frac{1}{n}}} \gtrsim \frac{\mathrm{diam}_{\ell_2^n}(K)}{\mathrm{vol}_n(K)^{\frac{1}{n}}},\end{aligned} \tag{1.57}$$

where we recall notation (1.11) for the isoperimetric quotient $\mathrm{iq}(\cdot)$ and the last step uses the isoperimetric theorem (1.12) and the assumption $L \subseteq K$. The following proposition explains what it would entail for one to be able to reverse (1.57) after an application of a suitable linear transformation; in particular, it shows that one can find $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq SK$ such that

$$\frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_n(L)} \mathrm{diam}_{\ell_2^n}(SK) \lesssim \frac{\mathrm{diam}_{\ell_2^n}(SK)}{\mathrm{vol}_n(K)^{\frac{1}{n}}}$$

if and only if Conjecture 10 on weak isomorphic reverse isoperimetry holds for K.

Proposition 32. *The following two statements are equivalent for every $n \in \mathbb{N}$, every origin-symmetric convex body $K \subseteq \mathbb{R}^n$ and every $\alpha > 0$.*

(1) *There exist a linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq SK$ with*

$$\frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_n(L)} \mathrm{vol}_n(K)^{\frac{1}{n}} \lesssim \alpha. \tag{1.58}$$

(2) *There exist a linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq SK$ that satisfies $\sqrt[n]{\mathrm{vol}_n(L)} \geqslant \beta \sqrt[n]{\mathrm{vol}_n(K)}$ and $\mathrm{iq}(L) \leqslant \gamma\sqrt{n}$ for some $\beta \gtrsim 1/\alpha$ and $\gamma \lesssim \alpha$ with $\gamma/\beta \lesssim \alpha$.*

Proof. For the implication (1)⇒(2) we introduce the notations $\gamma = \mathrm{iq}(L)/\sqrt{n}$ and $\beta = \sqrt[n]{\mathrm{vol}_n(L)}/\sqrt[n]{\mathrm{vol}_n(K)}$. Then,

$$\alpha \overset{(1.58)}{\gtrsim} \frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_n(L)} \mathrm{vol}_n(K)^{\frac{1}{n}} \overset{(1.55)}{\gtrsim} \frac{\mathrm{vol}_{n-1}(\partial L)}{\mathrm{vol}_n(L)\sqrt{n}} \mathrm{vol}_n(K)^{\frac{1}{n}} = \frac{\gamma}{\beta}.$$

Since by the isoperimetric theorem (1.12) we have $\gamma \gtrsim 1$, it follows from this that $\beta \gtrsim 1/\alpha$, and since $L \subseteq SK$ and $S \in SL_n(\mathbb{R})$, we have $\mathrm{vol}_n(L) \leqslant \mathrm{vol}_n(K)$, so $\beta \leqslant 1$ and it also follows from this that $\gamma \lesssim \alpha$.

For the implication (2)⇒(1), fix $T \in \mathsf{SL}_n(\mathbb{R})$ that satisfies

$$\mathrm{vol}_{n-1}(\partial TL) = \min\big\{\mathrm{vol}_{n-1}(\partial T'L) : T' \in \mathsf{SL}_n(\mathbb{R})\big\},$$

i.e., TL is in its *minimum surface area position* [250]. So, $\mathrm{vol}_{n-1}(\partial TL) \leqslant \mathrm{vol}_{n-1}(\partial L)$ by the definition of T, and by Proposition 3.1 in the work [104] of Giannopoulos and Papadimitrakis combined with (1.55) we have

$$\mathrm{MaxProj}(TL) \asymp \frac{\mathrm{vol}_{n-1}(\partial TL)}{\sqrt{n}}.$$

Consequently, if L satisfies part (2) of Proposition 32, then

$$\begin{aligned} \frac{\mathrm{MaxProj}(TL)}{\mathrm{vol}_n(TL)} \mathrm{vol}_n(K)^{\frac{1}{n}} &\asymp \frac{\mathrm{vol}_{n-1}(\partial TL)}{\mathrm{vol}_n(TL)\sqrt{n}} \mathrm{vol}_n(K)^{\frac{1}{n}} \\ &\leqslant \frac{\mathrm{vol}_{n-1}(\partial L)}{\mathrm{vol}_n(TL)\sqrt{n}} \mathrm{vol}_n(K)^{\frac{1}{n}} \\ &= \frac{\mathrm{iq}(L)}{\sqrt{n}} \left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(L)}\right)^{\frac{1}{n}} \leqslant \frac{\gamma}{\beta} \lesssim \alpha. \end{aligned}$$

Hence, (1) holds with S replaced by $TS \in \mathsf{SL}_n(\mathbb{R})$ and L replaced by $TL \subseteq TSK$. ■

Since when $\alpha \lesssim 1$ in Proposition 32 the assertion of its part (2) coincides with Conjecture 10, it follows that Conjecture 10, and a fortiori Conjecture 9, imply that for any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ there is $S \in \mathsf{SL}_n(\mathbb{R})$ such that $\mathsf{e}(\mathbf{X})$ is at most a universal constant multiple of $\mathrm{diam}_{\ell_2^n}(SB_{\mathbf{X}})/\sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})}$. Indeed, this follows by applying Theorem 21 to the normed spaces $\mathbf{X}' = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}'})$ and $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ whose unit balls are $SB_{\mathbf{X}}$ and L, respectively, where S and L are as in part (1) of Proposition 32 for $K = B_{\mathbf{X}}$, while noting that $\mathsf{e}(\mathbf{X}') = \mathsf{e}(\mathbf{X})$ since $\mathbf{X}'$ is isometric to $\mathbf{X}$. We record this conclusion as the following corollary.

Corollary 33. *If Conjecture 10 holds for a normed space* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$*, then there is* $S \in \mathsf{SL}_n(\mathbb{R})$ *such that*

$$\mathsf{e}(\mathbf{X}) \lesssim \frac{\operatorname{diam}_{\ell_2^n}(SB_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}}. \tag{1.59}$$

The upshot of Corollary 33 is that the right-hand side of (1.59) involves only Euclidean diameters and nth roots of volumes, which are typically much easier to estimate than extremal volumes of hyperplane projections. This comes at the cost of having to find the auxiliary linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$, but we expect that in concrete settings it will be simple to determine S. Moreover, in all of the specific examples of spaces for which we are interested (at least initially) in estimating their Lipschitz extension modulus, S should be the identity mapping. We will discuss this matter and its consequences in Section 1.6.2.

Remark 34. There is a degree of freedom that the above discussion does not exploit. Let $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ be a normed space. By (1.31), we know that $\mathsf{e}(\mathbf{X})$ is bounded from above by a constant multiple of the minimum of $\operatorname{diam}_{\Pi^*\mathbf{Y}}(B_{\mathbf{X}})/\operatorname{vol}_n(B_{\mathbf{Y}})$ over all the normed spaces $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ for which $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$. By (1.54), to control this minimum it suffices to estimate the minimum of $\operatorname{MaxProj}(\mathrm{B}_{\mathbf{Y}})/\operatorname{vol}_n(B_{\mathbf{Y}})$ over all such $\mathbf{Y}$, which relates to isomorphic reverse isoperimetric phenomena. But, we could also take a normed space $\mathbf{W} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{W}})$ for $m \geqslant n$ such that $B_{\mathbf{W}} \cap \mathbb{R}^n = B_{\mathbf{X}}$ (we need that $\mathbf{W}$ contains an isometric copy of $\mathbf{X}$), estimate either of the two minima above for the super-space $\mathbf{W}$, and then use $\mathsf{e}(\mathbf{X}) \leqslant \mathsf{e}(\mathbf{W})$. Thus, it would suffice to embed $\mathbf{X}$ into a larger normed space that exhibits good isomorphic reverse isoperimetry. Our conjectures imply that such an embedding step is not needed, namely we expect that the desired isomorphic reverse isoperimetric property holds for $\mathbf{X}$. Nevertheless, it could be that by finding a suitable super-space $\mathbf{W}$ one could bound $\mathsf{e}(\mathbf{X})$ while circumventing the difficulty of proving Conjecture 10 for $\mathbf{X}$. For example, if $\mathbf{X}$ is a subspace of ℓ_∞^m for some $m = O(n)$, then by Theorem 14 we know that $\mathsf{e}(\mathbf{X}) \lesssim \sqrt{n}$, but this is because we know that ℓ_∞^m has the desired isomorphic reverse isoperimetric property, and it is not clear how to prove it for $\mathbf{X}$ itself. It is also unclear how to construct for a given normed $\mathbf{X}$ a super-space $\mathbf{W}$ that could be used as above. We leave the exploration of this possibility for future research.

1.6.1 A spectral interpretation, reverse Faber–Krahn and the Cheeger space of a normed space

We will henceforth quantify the extent to which Conjecture 10 holds through the following condition:

$$\frac{\operatorname{iq}(L)}{\sqrt{n}}\left(\frac{\operatorname{vol}_n(K)}{\operatorname{vol}_n(L)}\right)^{\frac{1}{n}} = \frac{\operatorname{vol}_n(K)^{\frac{1}{n}}}{\sqrt{n}}\left(\frac{\operatorname{vol}_{n-1}(\partial L)}{\operatorname{vol}_n(L)}\right) \leqslant \alpha. \tag{1.60}$$

The factors $\mathrm{iq}(L)/\sqrt{n}$ and $(\mathrm{vol}_n(K)/\mathrm{vol}_n(L))^{1/n}$ that appear in the left-hand side of (1.60) are at least a positive universal constant (by, respectively, the isoperimetric theorem and the assumed inclusion $L \subseteq K$), so (1.60) implies that

$$\mathrm{iq}(L) \leqslant \alpha\sqrt{n} \quad \text{and} \quad \sqrt[n]{\mathrm{vol}_n(L)} \gtrsim \alpha^{-1}\sqrt[n]{\mathrm{vol}_n(K)}.$$

Thus, if $\alpha = O(1)$, then (1.60) is equivalent to the conclusion of Conjecture 10. However, even though Conjecture 10 expresses our expectation that (1.60) is always achievable with $\alpha = O(1)$ upon a judicious choice of the Euclidean structure on $\mathbb{R}^n$, in lieu of Conjecture 10 it would still be valuable to obtain (1.60) with α unbounded but slowly growing. In such a situation, the bi-parameter quantification that we used in part (2) of Proposition 32 contains more geometric information than (1.60), but below we will work with (1.60) in order to simplify the ensuing discussion; this suffices for our purposes because (1.60) is what shows up in all of the applications herein (per the proof Proposition 32) since they all proceed by bounding the right-hand side of (1.54) from above.

Alter and Caselles proved [7] that for every convex body $K \subseteq \mathbb{R}^n$ there is a *unique* measurable set $A \subseteq K$, which we call the *Cheeger body* of K and denote $\mathrm{Ch}\,K$, satisfying $\mathrm{Per}(A)/\mathrm{vol}_n(A) \leqslant \mathrm{Per}(B)/\mathrm{vol}_n(B)$ for every measurable $B \subseteq K$ with $\mathrm{vol}_n(B) > 0$, where $\mathrm{Per}(\cdot)$ denotes perimeter in the sense of Caccioppoli and de Giorgi; this notion is covered in [9] but we do not need to recall its definition here since the perimeter of a convex body coincides with the $(n-1)$-dimensional Hausdorff measure of its boundary. It was proved in [7] that $\mathrm{Ch}\,K$ is convex and its boundary is $C^{1,1}$. Further information on this remarkable theorem can be found in [7], where $\mathrm{Ch}\,K$ is characterized in terms of the mean curvature of its boundary through the work [8] of Alter, Caselles and Chambolle (see also the precursor [74] by Caselles, Chambolle and Novaga which obtained these statements under stronger assumptions on K).

Beyond the fact that it allows us to use the notation $\mathrm{Ch}\,K$ and call it *the* Cheeger body of K, the aforementioned uniqueness statement will be used substantially in the ensuing reasoning. It implies in particular that if K is origin-symmetric, then so is $\mathrm{Ch}\,K$. Consequently, if $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space, then $\mathrm{Ch}\,B_{\mathbf{X}}$ is the unit ball of a normed space that we denote by $\mathrm{Ch}\,\mathbf{X}$ and call the *Cheeger space* of $\mathbf{X}$.

For a convex body $K \subseteq \mathbb{R}^n$, let $\lambda(K)$ be the smallest Dirichlet eigenvalue of the Laplacian on K, namely it is the smallest $\lambda > 0$ for which there is a nonzero function

$$\varphi : K \to \mathbb{R}$$

that is smooth on the interior of K, vanishes on the boundary of K, and satisfies $\Delta\varphi = -\lambda\varphi$ on the interior of K; see, e.g., [77,81,265] for background on this classical topic. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space, then we denote

$$\lambda(\mathbf{X}) = \lambda(B_{\mathbf{X}}).$$

The quantity $h(K) = \mathrm{vol}_{n-1}(\partial\,\mathrm{Ch}\,K)/\,\mathrm{vol}_n(\mathrm{Ch}\,K)$ is called the Cheeger constant of K; it relates to $\lambda(K)$ by

$$\frac{2}{\pi}\sqrt{\lambda(K)} \leqslant h(K) = \frac{\mathrm{vol}_{n-1}(\partial\,\mathrm{Ch}\,K)}{\mathrm{vol}_n(\mathrm{Ch}\,K)} \leqslant 2\sqrt{\lambda(K)}. \tag{1.61}$$

It is important for our purposes that the constants appearing in (1.61) are independent of the dimension n. The second inequality in (1.61) is the Cheeger inequality for the Dirichlet Laplacian on Euclidean domains. Cheeger's proof of it for compact Riemannian manifolds without boundary appears in [78] and that proof works mutatis mutandis in the present setting; see its derivation in, e.g., the appendix of [174]. The first inequality in (1.61) can be called the Buser inequality for the Dirichlet Laplacian on convex Euclidean domains, since Buser proved [69] its analogue for compact Riemannian manifolds without boundary that have a lower bound on their Ricci curvature. In our setting, this reverse Cheeger inequality is more recent, namely it was noted for planar convex sets by Parini [246] and in any dimension by Brasco [53]. It can be justified quickly using the convexity of K and its Cheeger body $\mathrm{Ch}\,K$ as follows. By a classical theorem of Pólya we have $\lambda(K) \leqslant \pi^2(\mathrm{vol}_{n-1}(\partial K)/\,\mathrm{vol}_n(K))^2/4$ (Pólya proved this for planar convex sets, but in [144] Joó and Stachó carried out Pólya's approach for convex bodies in $\mathbb{R}^n$ for any $n \in \mathbb{N}$). Therefore,

$$\lambda(K) \leqslant \lambda(\mathrm{Ch}\,K) \leqslant \left(\frac{\pi\,\mathrm{vol}_{n-1}(\partial\,\mathrm{Ch}\,K)}{2\,\mathrm{vol}_n(\mathrm{Ch}\,K)}\right)^2 = \frac{\pi^2}{4}h(K)^2,$$

since $\mathrm{Ch}\,K$ is convex.

Let $j_{n/2-1,1}$ be the smallest positive zero of the Bessel function $J_{n/2-1}$; see [14, Chapter 4] for a treatment of Bessel functions and their zeros, though here we will only need to know that $j_{n/2-1,1} \asymp n$ (see [306] for more precise asymptotics). By classical computations (see, e.g., [129, equation (1.29)]),

$$\lambda\big(B_{\ell_2^n}\big) = j^2_{\frac{n}{2}-1,1}.$$

The Faber–Krahn inequality [95, 159] (see also, e.g., [77, 265]) asserts that $\lambda(K)$ is at least the first Dirichlet eigenvalue of a Euclidean ball whose volume is the same as the volume of K. Thus,

$$\lambda(K)\,\mathrm{vol}_n(K)^{\frac{2}{n}} \geqslant \lambda\big(B_{\ell_2^n}\big)\,\mathrm{vol}_n\big(B_{\ell_2^n}\big)^{\frac{2}{n}} = j^2_{\frac{n}{2}-1,1}\,\mathrm{vol}_n\big(B_{\ell_2^n}\big)^{\frac{2}{n}} \asymp n,$$

where we used the straightforward fact that $\lambda(rK) = \lambda(K)/r^2$ for every $r > 0$.

Observe that (1.61) can be rewritten as follows for every convex body $K \subseteq \mathbb{R}^n$:

$$\frac{2}{\pi}\left(\frac{\lambda(K)\,\mathrm{vol}_n(K)^{\frac{2}{n}}}{n}\right)^{\frac{1}{2}} \leqslant \frac{\mathrm{iq}(\mathrm{Ch}\,K)}{\sqrt{n}}\left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(\mathrm{Ch}\,K)}\right)^{\frac{1}{n}} \leqslant 2\left(\frac{\lambda(K)\,\mathrm{vol}_n(K)^{\frac{2}{n}}}{n}\right)^{\frac{1}{2}}.$$

Hence, for every $\alpha > 0$ we have

$$\frac{\mathrm{iq}(\mathrm{Ch}\,K)}{\sqrt{n}}\left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(\mathrm{Ch}\,K)}\right)^{\frac{1}{n}} \lesssim \alpha \iff \lambda(K)\,\mathrm{vol}_n(K)^{\frac{2}{n}} \lesssim \alpha^2 n. \tag{1.62}$$

Since $\mathrm{Ch}\,K$ is convex, the convex body $L \subseteq K$ that minimizes the left-hand side of (1.60) is equal to $\mathrm{Ch}\,K$. We therefore see that Conjecture 35 below is equivalent to Conjecture 10. Furthermore, if one of these two conjectures hold for a matrix $S \in \mathsf{SL}_n(\mathbb{R})$, then the same matrix would work for the other conjecture.

Conjecture 35 (Reverse Faber–Krahn). For any origin-symmetric convex body $K \subseteq \mathbb{R}^n$ there exists a volume-preserving linear transformation $S \in \mathsf{SL}_n(\mathbb{R})$ such that

$$\lambda(SK)\,\mathrm{vol}(K)^{\frac{2}{n}} \asymp n.$$

Remark 36. One can also wonder about exact maximizers in the context of Conjecture 35. Specifically, Bucur and Fragalà stated in [67, p. 389] that they expect that for any origin-symmetric convex body $K \subseteq \mathbb{R}^n$ with $\mathrm{vol}_n(K) = 1$ there exists $S \in \mathsf{SL}_n(\mathbb{R})$ such that $\lambda(SK) \leqslant \lambda([0,1]^n) = \pi^2 n$. If true, then this would be a beautiful statement even though it does not have substantial impact on Conjecture 10 and its implications herein (it would only influence the value of the implicit constant factors in our statements, which incur further losses that are most likely not sharp in other steps of their derivations). The only available evidence for the aforementioned (speculative) exact statement is the partial result of [67] in the planar case $n = 2$, which proves that it indeed holds when $K \subseteq \mathbb{R}^2$ is a convex axisymmetric octagon that has four of its vertices lying on the axes at the same distance from the origin; see specifically [67, Proposition 10], whose proof involves delicate reasoning that incorporate computer-assisted steps. A complete result for $n = 2$ has been subsequently obtained by the same authors in [68] for the analogous question in which one replaces the Dirichlet eigenvalue of the Laplacian by the Cheeger constant. Namely, [68, Theorem 1.1] states that for every origin-symmetric convex body $K \subseteq \mathbb{R}^2$ with $\mathrm{vol}_2(K) = 1$ there exists $S \in \mathsf{SL}_2(\mathbb{R})$ such that $h(SK) \leqslant h([0,1]^2) = 2 + \sqrt{\pi}$ (furthermore, in this case S can be taken to be the matrix that puts K in John position, i.e., the ellipse of maximal area that is contained in SK is a circle).

This above spectral interpretation of Conjecture 10 is useful for multiple purposes, including the following lemma whose proof appears in Section 6.1. For its statement, as well as throughout the ensuing discussion, recall that a basis $x_1, \ldots, x_n$ of an n-dimensional normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is a 1-unconditional basis of $\mathbf{X}$ if

$$\|\varepsilon_1 a_1 x_1 + \cdots + \varepsilon_n a_n x_n\|_{\mathbf{X}} = \|a_1 x_1 + \cdots + a_n x_n\|_{\mathbf{X}}$$

for every choice of scalars $a_1, \ldots, a_n \in \mathbb{R}$ and signs $\varepsilon_1, \ldots, \varepsilon_n \in \{-1, 1\}$. When we say that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is an unconditional normed space, we mean that the standard (coordinate) basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$ is a 1-unconditional basis of $\mathbf{X}$.

Lemma 37 (Closure of Conjecture 10 under unconditional composition). *Fix $n \in \mathbb{N}$ and $m_1, \ldots, m_n \in \mathbb{N}$. Let $\mathbf{X}_1 = (\mathbb{R}^{m_1}, \|\cdot\|_{\mathbf{X}_1}), \ldots, \mathbf{X}_n = (\mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}_n})$ be normed spaces. Also, let $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ be an unconditional normed space. Define a normed space $\mathbf{X} = (\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}})$ by*

$$\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}, \qquad \|x\|_{\mathbf{X}} \stackrel{\mathrm{def}}{=} \big\|\big(\|x_1\|_{\mathbf{X}_1}, \ldots, \|x_n\|_{\mathbf{X}_n}\big)\big\|_{\mathbf{E}}.$$

Suppose that there exist $\alpha > 0$, linear transformations $S_1 \in \mathsf{SL}_{m_1}(\mathbb{R}), \ldots, S_n \in \mathsf{SL}_{m_n}(\mathbb{R})$, and normed spaces $\mathbf{Y}_1 = (\mathbb{R}^{m_1}, \|\cdot\|_{\mathbf{Y}_1}), \ldots, \mathbf{Y}_n = (\mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{Y}_n})$ such that

$$B_{\mathbf{Y}_k} \subseteq S_k B_{\mathbf{X}_k} \quad \text{and} \quad \frac{\mathrm{iq}\big(B_{\mathbf{Y}_k}\big)}{\sqrt{m_k}} \left(\frac{\mathrm{vol}_{m_k}\big(B_{\mathbf{X}_k}\big)}{\mathrm{vol}_{m_k}\big(B_{\mathbf{Y}_k}\big)}\right)^{\frac{1}{m_k}} \leqslant \alpha, \tag{1.63}$$

for every $k \in \{1, \ldots, n\}$. Then, there exist a normed space

$$\mathbf{Y} = (\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}})$$

and $S \in \mathsf{SL}(\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n})$ such that

$$B_{\mathbf{Y}} \subseteq S B_{\mathbf{X}} \quad \text{and} \quad \frac{\mathrm{iq}(B_{\mathbf{Y}})}{\sqrt{m_1 + \cdots + m_n}} \left(\frac{\mathrm{vol}_{m_1+\cdots+m_n}(B_{\mathbf{X}})}{\mathrm{vol}_{m_1+\cdots+m_n}(B_{\mathbf{Y}})}\right)^{\frac{1}{m_1+\cdots+m_n}} \lesssim \alpha. \tag{1.64}$$

As (1.63) with $\alpha = O(1)$ is immediate when $n_0 = 1$, Lemma 37 establishes Conjecture 10 for when K is the unit ball of an unconditional normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. This holds, in particular, for $\mathbf{X} = \ell_p^n$, though we will prove in Section 6.1 that the stronger conclusion of Conjecture 9 holds in this case (recall Remark 31). Lemma 37 also shows that Conjecture 10 holds for, say, $\mathbf{X} = \ell_p^n(\ell_q^m)$; we expect that the reasoning of Section 6.1 could be adapted to yield Conjecture 9 for these spaces as well, but we did not attempt to carry this out. Other spaces that satisfy (1.63) with α slowly growing will be presented in Section 1.6.2; upon their substitution into Lemma 37, more examples for which Conjecture 10 holds up to lower-order factors are obtained (of course, we are conjecturing here that it holds for *any* space).

Remark 38. Say that a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is in *Cheeger position* if

$$\forall S \in \mathsf{SL}_n(\mathbb{R}), \qquad \frac{\mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, B_{\mathbf{X}})}{\mathrm{vol}_n(\mathrm{Ch}\, B_{\mathbf{X}})} \leqslant \frac{\mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, SB_{\mathbf{X}})}{\mathrm{vol}_n(\mathrm{Ch}\, SB_{\mathbf{X}})}.$$

Observe that if $\mathbf{X}$ is in Cheeger position, then its Cheeger space $\mathrm{Ch}\,\mathbf{X}$ is in minimum surface area position, namely, $\mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, B_{\mathbf{X}}) \leqslant \mathrm{vol}_{n-1}(\partial S \,\mathrm{Ch}\, B_{\mathbf{X}})$ for every $S \in \mathsf{SL}_n(\mathbb{R})$. Indeed, $S \,\mathrm{Ch}\, B_{\mathbf{X}} \subseteq SB_{\mathbf{X}}$, so by the definition of the Cheeger body of $SB_{\mathbf{X}}$ we have $\mathrm{vol}_{n-1}(\partial S \,\mathrm{Ch}\, B_{\mathbf{X}}) / \mathrm{vol}_n(\mathrm{Ch}\, B_{\mathbf{X}}) \geqslant \mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, SB_{\mathbf{X}}) / \mathrm{vol}_n(\mathrm{Ch}\, SB_{\mathbf{X}})$. At the same time, $\mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, SB_{\mathbf{X}}) / \mathrm{vol}_n(\mathrm{Ch}\, SB_{\mathbf{X}}) \geqslant \mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, B_{\mathbf{X}}) / \mathrm{vol}_n(\mathrm{Ch}\, B_{\mathbf{X}})$ by the definition of the Cheeger position, so $\mathrm{vol}_{n-1}(\partial S \,\mathrm{Ch}\, B_{\mathbf{X}}) \geqslant \mathrm{vol}_{n-1}(\partial \,\mathrm{Ch}\, B_{\mathbf{X}})$. This shows that in the proof of the implication (2)⇒(1) of Proposition 32, if we worked with $L = \mathrm{Ch}\, SK$, then there would be no need to introduce the additional linear transformation $T \in \mathsf{SL}_n(\mathbb{R})$. It would be worthwhile to study the Cheeger position for its own sake even if it were not for its connection to reverse isoperimetry. In particular, we do not know if the converse of the above deduction holds, namely whether it is true that if $\mathrm{Ch}\, \mathbf{X}$ is in minimum surface area position, then $\mathbf{X}$ is in Cheeger position. We also do not know if the Cheeger position is unique up to orthogonal transformation (as is the case for the minimum surface area position [104]); we did not investigate these matters since they are not needed for the present purposes, but we expect that the characterisations of the Cheeger body in [7] would be relevant here. One could also define that a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is in *Dirichlet position* if $\lambda(\mathbf{X}) \leqslant \lambda(S\mathbf{X})$ for every $S \in \mathsf{SL}_n(\mathbb{R})$. It is unclear how the Cheeger position relates to the Dirichlet position and it would be also worthwhile to study the Dirichlet position for its own sake. By (1.61), working with either the Cheeger position or the Dirichlet position would be equally valuable for the reverse isoperimetric questions in which we are interested here.

1.6.2 Symmetries and positions

Thus far we considered an arbitrary scalar product on an n-dimensional normed space through which we identified its underlying vector space structure with $\mathbb{R}^n$. However, the Lipschitz extension modulus is insufficiently understood for "very nice" normed spaces (including even the Euclidean space ℓ_2^n) that belong to a natural class of normed spaces that have a canonical identification with $\mathbb{R}^n$. It therefore makes sense to first focus on this class.

For a finite dimensional normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$, let $\mathsf{Isom}(\mathbf{X})$ be the group of all of the isometric automorphism of $\mathbf{X}$, i.e., all the linear operators $U : \mathbf{X} \to \mathbf{X}$ that satisfy $\|Ux\|_{\mathbf{X}} = \|x\|_{\mathbf{X}}$ for every $x \in \mathbf{X}$. We will denote the Haar probability measure on the compact group $\mathsf{Isom}(\mathbf{X})$ by $\hbar_{\mathbf{X}}$.

Definition 39. We say that a finite dimensional normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is *canonically positioned* if any two $\mathsf{Isom}(\mathbf{X})$-invariant scalar products on $\mathbf{X}$ are proportional to each other. In other words, if $\langle\cdot,\cdot\rangle : \mathbf{X}\times\mathbf{X} \to \mathbb{R}$ and $\langle\cdot,\cdot\rangle' : \mathbf{X}\times\mathbf{X} \to \mathbb{R}$ are scalar products on $\mathbf{X}$ such that $\langle Ux, Uy\rangle = \langle x, y\rangle$ and $\langle Ux, Uy\rangle' = \langle x, y\rangle'$ for every $x, y \in \mathbf{X}$ and every $U \in \mathsf{Isom}(\mathbf{X})$, then there necessarily exists $\lambda \in \mathbb{R}$ such that $\langle\cdot,\cdot\rangle' = \lambda\langle\cdot,\cdot\rangle$.

On any finite dimensional normed space $\mathbf{X}$ there exists at least one scalar product $\langle\cdot,\cdot\rangle : \mathbf{X}\times\mathbf{X} \to \mathbb{R}$ that is invariant under $\mathsf{Isom}(\mathbf{X})$, as seen, e.g., by averaging any given scalar product $\langle\cdot,\cdot\rangle_0$ on $\mathbf{X}$ with respect $\hbar_{\mathbf{X}}$, i.e., defining

$$\forall x, y \in \mathbf{X}, \quad \langle x, y\rangle \stackrel{\mathrm{def}}{=} \int_{\mathsf{Isom}(\mathbf{X})} \langle Sx, Sy\rangle_0 \,\mathrm{d}\hbar_{\mathbf{X}}(S).$$

Definition 39 concerns those spaces $\mathbf{X}$ for which such an invariant scalar product is unique up to rescaling, so there is (essentially, i.e., up to rescaling) no arbitrariness when we identify $\mathbf{X}$ with $\mathbb{R}^{\dim(\mathbf{X})}$.

Example 40. The class of n-dimensional canonically positioned spaces includes those normed spaces $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ that have a basis $e_1, \ldots, e_n$ such that for any distinct $i, j \in \{1, \ldots, n\}$ there are a permutation $\pi \in S_n$ with $\pi(i) = j$ and a sign vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, 1\}^n$ with $\varepsilon_i = -\varepsilon_j$ such that $T_\pi, S_\varepsilon \in \mathsf{Isom}(\mathbf{X})$, where we denote $T_\pi x = \sum_{i=1}^n a_{\pi(i)} e_i$ and $S_\varepsilon x = \sum_{i=1}^n \varepsilon_i a_i e_i$ for $x = \sum_{i=1}^n a_i e_i \in \mathbf{X}$ with $a_1, \ldots, a_n \in \mathbb{R}$. Indeed, let $\langle \cdot, \cdot \rangle$ be a scalar product on $\mathbf{X}$ that is $\mathsf{Isom}(\mathbf{X})$-invariant. For every distinct $i, j \in \{1, \ldots, n\}$, if $\pi \in S_n$ and $\varepsilon \in \{-1, 1\}^n$ are as above, then $\langle e_i, e_i \rangle = \langle e_{\pi(i)}, e_{\pi(i)} \rangle = \langle e_j, e_j \rangle$ while $\langle e_i, e_j \rangle = \langle \varepsilon_i e_i, \varepsilon_j e_j \rangle = -\langle e_i, e_j \rangle$, so $\langle e_i, e_j \rangle = 0$.

Example 40 covers all of the spaces for which we think that it is most pressing (given the current state of knowledge) to understand their Lipschitz extension modulus, including normed spaces $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ that have a 1-symmetric basis, i.e., a basis $e_1, \ldots, e_n \in \mathbf{E}$ such that $\|\sum_{i=1}^n \varepsilon_i a_{\pi(i)} e_i\|_{\mathbf{E}} = \|\sum_{i=1}^n a_i e_i\|_{\mathbf{E}}$ for every $(\varepsilon, \pi) \in \{-1, 1\}^n \times S_n$. In particular, ℓ_p^n, and more generally Orlicz and Lorentz spaces (see, e.g., [181]), are canonically positioned. We will use below the common convention that a normed space $(\mathbb{R}^n, \|\cdot\|)$ is said to be symmetric if it is 1-symmetric with respect to the standard (coordinate) basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$.

Example 40 also includes matrix norms

$$\mathbf{X} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathbf{X}})$$

that remain unchanged if one transposes a pair of rows or columns, or changes the sign of an entire row or a column, such as S_p^n. More generally, if $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ is a symmetric normed space, then its unitary ideal $\mathsf{S}_{\mathbf{E}} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathsf{S}_{\mathbf{E}}})$ is canonically positioned (see, e.g., [37]), where for $T \in \mathsf{M}_n(\mathbb{R})$ one denotes its singular values by $s_1(T) \geqslant \cdots \geqslant s_n(T)$ and defines $\|T\|_{\mathsf{S}_{\mathbf{E}}} = \|(s_1(T), \ldots, s_n(T))\|_{\mathbf{E}}$. More examples of such matrix norms are projective and injective tensor products (see, e.g., [276]) of symmetric spaces, where if $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{Y}})$ are normed spaces, then their projective tensor product $\mathbf{X} \widehat{\otimes} \mathbf{Y}$ is the norm on $M_{n\times m}(\mathbb{R}) = \mathbb{R}^n \otimes \mathbb{R}^m$ whose unit ball is the convex hull of $\{x \otimes y : (x, y) \in B_{\mathbf{X}} \times B_{\mathbf{Y}}\}$, and their injective tensor product $\mathbf{X} \check{\otimes} \mathbf{Y}$ is the dual of $\mathbf{X}^* \widehat{\otimes} \mathbf{Y}^*$ (equivalently, $\mathbf{X} \check{\otimes} \mathbf{Y}$ is isometric to the operator norm from $\mathbf{X}^*$ to $\mathbf{Y}$; see, e.g., [87, Section 1.1]).

Henceforth, when we will say that a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is canonically positioned it will always be tacitly assumed that the standard scalar product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^n$ is $\mathsf{Isom}(\mathbf{X})$-invariant, i.e., $\mathsf{Isom}(\mathbf{X})$ is a subgroup of the orthogonal group $\mathsf{O}_n \subseteq \mathsf{M}_n(\mathbb{R})$. This is equivalent to the requirement that for every symmetric positive definite matrix $T \in \mathsf{M}_n(\mathbb{R})$, if $TU = UT$ for every $U \in \mathsf{Isom}(\mathbf{X})$, then there is $\lambda \in (0, \infty)$

such that $T = \lambda \mathsf{Id}_n$. Indeed, any scalar product $\langle\cdot,\cdot\rangle' : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is of the form $\langle x, y\rangle' = \langle Tx, y\rangle$ for some symmetric positive definite $T \in \mathsf{M}_n(\mathbb{R})$ and all $x, y \in \mathbb{R}^n$, and using the $\mathsf{Isom}(\mathbf{X})$-invariance of $\langle\cdot,\cdot\rangle$ we see that $\langle\cdot,\cdot\rangle'$ is $\mathsf{Isom}(\mathbf{X})$-invariant if and only if T commutes with all of the elements of $\mathsf{Isom}(\mathbf{X})$.

Remark 41. A symmetry assumption that is common in the literature is *enough symmetries*. A normed space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ is said [103] to have enough symmetries if any linear transformation $T : \mathbf{X} \to \mathbf{X}$ must be a scalar multiple of the identity if T commutes with every element of $\mathsf{Isom}(\mathbf{X})$. By the above discussion, if $\mathbf{X}$ has enough symmetries, then $\mathbf{X}$ is canonically positioned. The converse implication does not hold, i.e., there exist normed spaces that are canonically positioned but do not have enough symmetries. For example, let $\mathsf{Rot}_{\pi/2} \in \mathsf{O}_2$ be the rotation by 90 degrees and let G be the subgroup of O_2 that is generated by $\mathsf{Rot}_{\pi/2}$. Thus, G is cyclic of order 4. Suppose that

$$\mathbf{X} = (\mathbb{R}^2, \|\cdot\|_{\mathbf{X}})$$

is a normed space with $\mathsf{Isom}(\mathbf{X}) = G$; the fact that there is such a normed space follows from the general result [118, Theorem 3.1] of Gordon and Loewy on existence of norms with a specified group of isometries, though in this particular case it is simple to construct such an example (e.g., the unit ball of $\mathbf{X}$ can be taken to be a suitable non-regular octagon). Since $\mathsf{Isom}(\mathbf{X})$ is Abelian, the matrix $\mathsf{Rot}_{\pi/2}$ commutes with all of the elements of $\mathsf{Isom}(\mathbf{X})$ yet it is not a multiple of the identity matrix, so $\mathbf{X}$ does not have enough symmetries. Nevertheless, $\mathbf{X}$ is canonically positioned. Indeed, suppose that $T \in \mathsf{M}_2(\mathbb{R})$ is a symmetric matrix that commutes with $\mathsf{Rot}_{\pi/2}$. Then, $\mathsf{Rot}_{\pi/2}$ preserves any eigenspace of T, which means that any such eigenspace must be $\{0\}$ or $\mathbb{R}^2$. But T is diagonalizable over $\mathbb{R}$, so it follows that for some $\lambda \in \mathbb{R}$ we have $T = \lambda \mathsf{Id}_2$. If n is even, then one obtains such an n-dimensional example by considering $\ell_\infty^{n/2}(\mathbf{X})$. However, a representation-theoretic argument due to Emmanuel Breuillard (private communication; details omitted) shows that if n is odd, then any n-dimensional normed space has enough symmetries if and only if it is canonically positioned.

The following lemma is important for us even though it is an immediate consequence of the (major) theorem of [7] that the Cheeger body of a given convex body in $\mathbb{R}^n$ is unique (recall Section 1.6.1).

Lemma 42. *Let $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ be a normed space such that $\mathsf{Isom}(\mathbf{X}) \leqslant \mathsf{O}_n$ is a subgroup of the orthogonal group. Then the isometry group of its Cheeger space* $\mathrm{Ch}\,\mathbf{X}$ *satisfies*

$$\mathsf{Isom}(\mathrm{Ch}\,\mathbf{X}) \supseteq \mathsf{Isom}(\mathbf{X}).$$

Consequently, if $\mathbf{X}$ is canonically positioned, then also $\mathrm{Ch}\,\mathbf{X}$ *is canonically positioned.*

Proof. For any $U \in \mathsf{Isom}(\mathbf{X})$ we have

$$\frac{\mathrm{vol}_{n-1}(\partial U \operatorname{Ch} B_{\mathbf{X}})}{\mathrm{vol}_n(U \operatorname{Ch} B_{\mathbf{X}})} = \frac{\mathrm{vol}_{n-1}(\partial \operatorname{Ch} B_{\mathbf{X}})}{\mathrm{vol}_n(\operatorname{Ch} B_{\mathbf{X}})},$$

and also $U \operatorname{Ch} B_{\mathbf{X}} \subseteq UB_{\mathbf{X}} = B_{\mathbf{X}}$, since $U \in \mathsf{O}_n$. Consequently, (by definition), $U \operatorname{Ch} B_{\mathbf{X}}$ is a Cheeger body of $B_{\mathbf{X}}$. The uniqueness of the Cheeger body now implies that $U \operatorname{Ch} B_{\mathbf{X}} = \operatorname{Ch} B_{\mathbf{X}}$. Therefore, $U \in \mathsf{Isom}(\operatorname{Ch} \mathbf{X})$. ∎

The following corollary is a quick consequence of Lemma 42.

Corollary 43. *Let* $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ *be a symmetric normed space. Then, its Cheeger space* $\operatorname{Ch} \mathbf{E}$ *is also symmetric and there exists a (unique) symmetric normed space* $\chi\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\chi\mathbf{E}})$ *such that the Cheeger space of the unitary ideal* $\mathsf{S}_{\mathbf{E}}$ *is the unitary ideal of* $\chi\mathbf{E}$*, i.e.,* $\operatorname{Ch} \mathsf{S}_{\mathbf{E}} = \mathsf{S}_{\chi\mathbf{E}}$.

Proof. The assertion that $\operatorname{Ch}\mathbf{E}$ is symmetric coincides with requiring that $\mathsf{Isom}(\operatorname{Ch}\mathbf{E})$ contains the group $\{-1,1\}^n \rtimes S_n = \{T_\varepsilon S_\pi : (\varepsilon, \pi) \in \{-1,1\}^n \times S_n\} \leqslant \mathsf{O}_n$, where we recall the notation of Example 40. We are assuming that $\mathsf{Isom}(\mathbf{E}) \supseteq \{-1,1\}^n \rtimes S_n$, so this follows from Lemma 42. For $U, V \in \mathsf{O}_n$ define $R_{U,V} : \mathsf{M}_n(\mathbb{R}) \to \mathsf{M}_n(\mathbb{R})$ by $(A \in M_n(\mathbb{R})) \mapsto UAV$. Since $\mathsf{Isom}(\mathsf{S}_{\mathbf{E}}) \supseteq \{R_{U,V} : U, V \in \mathsf{O}_n\}$, by Lemma 42 so does $\mathsf{Isom}(\operatorname{Ch} S_{\mathbf{E}})$. A normed space $(\mathsf{M}_n(\mathbb{R}), \|\cdot\|)$ that is invariant under $R_{U,V}$ for all $U, V \in \mathsf{O}_n$ is the unitary ideal of a symmetric normed space $\mathbf{F} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{F}})$; see, e.g., [37, Theorem IV.2.1]. This $\mathbf{F}$ is unique (consider the values of $\|\cdot\|_{S_{\mathbf{F}}}$ on diagonal matrices), so we can introduce the notation $\mathbf{F} = \chi\mathbf{E}$. ∎

The same reasoning as in the proof of Corollary 43 shows that if

$$\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$$

is an unconditional normed space, then so is $\operatorname{Ch}\mathbf{E}$. Thus, the space $\mathbf{Y}$ in Lemma 37 when

$$\mathbf{X}_1 = \cdots = \mathbf{X}_n = \mathbb{R}$$

that satisfies (1.64) can be taken to unconditional, as seen by an inspection of the proof of Lemma 37 (specifically, the operator S in (1.64) that arises in this case is diagonal, so $S\mathbf{E}$ is also unconditional and we can take $\mathbf{Y} = \operatorname{Ch} S\mathbf{E}$).

Problem 44. We associated above to every symmetric normed space

$$\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$$

two symmetric normed spaces $\operatorname{Ch}\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\operatorname{Ch}\mathbf{E}})$ and $\chi\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\chi\mathbf{E}})$. It would be valuable to understand these auxiliary norms on $\mathbb{R}^n$, and in particular how they relate to each other. By the definition of the Cheeger body, its convexity and uniqueness, $\mathrm{Ch}\,\mathbf{E}$ is the unique minimizer of the functional

$$\mathbf{F} \mapsto \frac{\mathrm{vol}_{n-1}\big(\partial B_{\mathbf{F}}\big)}{\mathrm{vol}_n\big(B_{\mathbf{F}}\big)} = \frac{\int_{\partial B_{\mathbf{F}}} 1\,\mathrm{d}x}{\int_{B_{\mathbf{F}}} 1\,\mathrm{d}x} \tag{1.65}$$

over all symmetric normed spaces $\mathbf{F} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{F}})$ with $B_{\mathbf{F}} \subseteq B_{\mathbf{E}}$; denote the set of all such $\mathbf{F}$ by $\mathcal{S}\mathfrak{ym}(\subseteq B_{\mathbf{E}})$. In contrast to (1.65), $\chi\mathbf{E}$ is the unique minimizer of the functional

$$\mathbf{F} \mapsto \frac{\int_{\partial B_{\mathbf{F}}} \prod_{1\leqslant i<j\leqslant n} |x_i^2 - x_j^2|\,\mathrm{d}x}{\int_{B_{\mathbf{F}}} \prod_{1\leqslant i<j\leqslant n} |x_i^2 - x_j^2|\,\mathrm{d}x} \tag{1.66}$$

over the same domain $\mathcal{S}\mathfrak{ym}(\subseteq B_{\mathbf{E}})$. To justify (1.66), observe first that by Corollary 43 we know that $\chi\mathbf{E}$ is the unique minimizer of the following functional over $\mathcal{S}\mathfrak{ym}(\subseteq B_{\mathbf{E}})$:

$$\mathbf{F} \mapsto \frac{\mathrm{vol}_{n^2-1}\big(\partial B_{\mathsf{S}_{\mathbf{F}}}\big)}{\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{F}}}\big)} = \lim_{\varepsilon\to 0+} \frac{\int_{\left(B_{\mathsf{S}_{\mathbf{F}}}+\varepsilon B_{\mathsf{S}_2^n}\right)\smallsetminus B_{\mathsf{S}_{\mathbf{F}}}} 1\,\mathrm{d}x}{\varepsilon \int_{B_{\mathbf{F}}} 1\,\mathrm{d}x}. \tag{1.67}$$

We claim that for every $\mathbf{F} \in \mathcal{S}\mathfrak{ym}(\subseteq B_{\mathbf{E}})$ and $\varepsilon > 0$,

$$\begin{aligned}&\big(B_{\mathsf{S}_{\mathbf{F}}} + \varepsilon B_{\mathsf{S}_2^n}\big) \smallsetminus B_{\mathsf{S}_{\mathbf{F}}}\\ &= \Big\{A \in M_n(\mathbb{R}) : s(A) \stackrel{\mathrm{def}}{=} \big(s_1(A),\ldots,s_n(A)\big) \in \big(B_{\mathbf{F}} + \varepsilon B_{\ell_2^n}\big) \smallsetminus B_{\mathbf{F}}\Big\},\end{aligned} \tag{1.68}$$

where we denote the singular values of $A \in \mathsf{M}_n(\mathbb{R})$ by $s_1(A) \geqslant \cdots \geqslant s_n(A)$. Indeed, if A belongs to the right-hand side of (1.68), then $\|s(A)\|_{\mathbf{F}} > 1$ and $s(A) = x + y$ for $x, y \in \mathbb{R}^n$ that satisfy $\|x\|_{\mathbf{F}} \leqslant 1$ and $\|y\|_{\ell_2^n} \leqslant \varepsilon$. Write $A = UDV$, where $D \in \mathsf{M}_n(\mathbb{R})$ is the diagonal matrix whose diagonal is the vector $s(A) \in \mathbb{R}^n$, and $U, V \in \mathsf{O}_n$. Let $D(x), D(y) \in \mathsf{M}_n(\mathbb{R})$ be the diagonal matrices whose diagonals equal x, y, respectively. By noting that $\|A\|_{\mathsf{S}_{\mathbf{F}}} = \|s(A)\|_{\mathbf{F}} > 1$ and $A = UD_xV + UD_yV$, where $\|UD(x)V\|_{\mathsf{S}_{\mathbf{F}}} \leqslant 1$ and $\|UD(y)V\|_{\mathsf{S}_2^n} \leqslant \varepsilon$, we conclude that A belongs to the left-hand side of (1.68). The reverse inclusion is less straightforward. If A belongs to the left-hand side of (1.68), then $\|A\|_{\mathsf{S}_{\mathbf{F}}} > 1$ and $A = B + C$, where $B, C \in \mathsf{M}_n(\mathbb{R})$ satisfy $\|B\|_{\mathsf{S}_{\mathbf{F}}} = \|s(B)\|_{\mathbf{F}} \leqslant 1$ and $\|C\|_{\mathsf{S}_2^n} \leqslant \varepsilon$. By an inequality of Mirsky [222] we have $\|s(A) - s(B)\|_{\ell_2^n} \leqslant \|A - B\|_{\mathsf{S}_2^n} = \|C\|_{\mathsf{S}_2^n} \leqslant \varepsilon$. Hence $s(A) = s(B) + (s(A) - s(B)) \in (B_{\mathbf{F}} + \varepsilon B_{\ell_2^n}) \smallsetminus B_{\mathbf{F}}$, i.e., A belongs to the right-hand side of (1.68). With (1.68) established, since membership of a matrix A in either $B_{\mathbf{F}}$ or $(B_{\mathbf{F}} + \varepsilon B_{\ell_2^n}) \smallsetminus B_{\mathbf{F}}$ depends only on $s(A)$, by the Weyl integration formula [311] (see [12, Proposition 4.1.3] for the formulation that we are using),

$$\frac{\int_{(B_{\mathsf{S}_{\mathbf{F}}}+\varepsilon B_{\mathsf{S}_2^n})\smallsetminus B_{\mathsf{S}_{\mathbf{F}}}} 1\,\mathrm{d}x}{\int_{B_{\mathbf{F}}} 1\,\mathrm{d}x} = \frac{\int_{(B_{\mathbf{F}}+\varepsilon B_{\ell_2^n})\smallsetminus B_{\mathbf{F}}} \prod_{1\leqslant i<j\leqslant n} |x_i^2 - x_j^2|\,\mathrm{d}x}{\int_{B_{\mathbf{F}}} \prod_{1\leqslant i<j\leqslant n} |x_i^2 - x_j^2|\,\mathrm{d}x}.$$

Thus (1.66) follows from (1.67). Analysing the functional in (1.66) seems nontrivial but likely tractable using ideas from random matrix theory. It would be especially interesting to treat the case $\mathbf{E} = \ell_\infty^n$. While we have a reasonably good understanding of the (isomorphic) geometry space $\mathrm{Ch}\,\ell_\infty^n$, its noncommutative counterpart $\chi\ell_\infty^n$ is still mysterious and understanding its geometry is closely related to Conjecture 10 (and likely also Conjecture 9) in the important special case of the operator norm S_∞^n; see also Remark 172.

If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is canonically positioned and μ is a Borel measure on $\mathbb{R}^n$ that is $\mathsf{Isom}(\mathbf{X})$-invariant, i.e., $\mu(UA) = \mu(A)$ for every $U \in \mathsf{Isom}(\mathbf{X})$ and every Borel subset $A \subseteq \mathbb{R}^n$, then consider the scalar product

$$\forall x, y \in \mathbb{R}^n, \qquad \langle x, y\rangle' \stackrel{\mathrm{def}}{=} \int_{\mathbb{R}^n} \langle x, z\rangle\langle y, z\rangle \,\mathrm{d}\mu(z).$$

For every $U \in \mathsf{Isom}(\mathbf{X})$ and $x, y \in \mathbb{R}^n$ we have

$$\begin{aligned}\langle Ux, Uy\rangle' &= \int_{\mathbb{R}^n} \langle Ux, z\rangle\langle Uy, z\rangle \,\mathrm{d}\mu(z) = \int_{\mathbb{R}^n} \langle x, U^{-1}z\rangle\langle y, U^{-1}z\rangle \,\mathrm{d}\mu(z)\\ &= \int_{\mathbb{R}^n} \langle x, z\rangle\langle y, z\rangle \,\mathrm{d}\mu(z) = \langle x, y\rangle',\end{aligned}$$

where the second step uses the $\mathsf{Isom}(\mathbf{X})$-invariance of $\langle\cdot,\cdot\rangle$, and the third step uses the $\mathsf{Isom}(\mathbf{X})$-invariance of μ. Hence $\langle x, y\rangle' = \lambda\langle x, y\rangle$ for some $\lambda \in \mathbb{R}$ and every $x, y \in \mathbb{R}^n$. By considering the case $x = y$ of this identity and integrating over $x \in S^{n-1}$ one sees that necessarily $n\lambda = \int_{\mathbb{R}^n} \|z\|_{\ell_2^n}^2 \,\mathrm{d}\mu(z)$. Hence,

$$\forall x, y \in \mathbb{R}^n, \qquad \int_{\mathbb{R}^n} \langle x, z\rangle\langle y, z\rangle \,\mathrm{d}\mu(z) = \frac{\int_{\mathbb{R}^n} \|z\|_{\ell_2^n}^2 \,\mathrm{d}\mu(z)}{n}\langle x, y\rangle. \tag{1.69}$$

By establishing (1.69) we have shown that if

$$\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$$

is a canonically positioned normed space, then any $\mathsf{Isom}(\mathbf{X})$-invariant Borel measure on $\mathbb{R}^n$ is *isotropic* [55, 107] (the converse also holds, i.e., $\mathbf{X}$ is canonically positioned if and only if every $\mathsf{Isom}(\mathbf{X})$-invariant Borel measure on $\mathbb{R}^n$ is isotropic). In particular, let $\sigma_{\mathbf{X}}$ be the measure on S^{n-1} that is given by

$$\sigma_{\mathbf{X}}(A) = \mathrm{vol}_{n-1}(\{x \in \partial B_{\mathbf{X}} : N_{\mathbf{X}}(x) \in A\})$$

for every measurable $A \subseteq S^{n-1}$, where for $x \in \partial B_{\mathbf{X}}$ the vector $N_{\mathbf{X}}(x) \in S^{n-1}$ is the (almost-everywhere uniquely defined) unit outer normal to $\partial B_{\mathbf{X}}$ at x, i.e., recalling (1.30), we use the simpler notation $N_{B_{\mathbf{X}}} = N_{\mathbf{X}}$. In other words, $\sigma_{\mathbf{X}}$ is the image under the Gauss map of the $(n-1)$-dimensional Hausdorff measure on $\partial B_{\mathbf{X}}$. Then,

$\sigma_{\mathbf{X}}$ is $\mathsf{Isom}(\mathbf{X})$-invariant because every $U \in \mathsf{Isom}(\mathbf{X})$ is an orthogonal transformation and $N_{\mathbf{X}} \circ U = U \circ N_{\mathbf{X}}$ almost everywhere on $\partial B_{\mathbf{X}}$. By [250], this implies that $\mathbf{X}$ is in its minimum surface area position (recall the proof of Proposition 32), so $\mathrm{MaxProj}(B_{\mathbf{X}}) \asymp \mathrm{vol}_{n-1}(\partial B_{\mathbf{X}})/\sqrt{n}$ by [104, Proposition 3.1].

The following corollary follows by substituting the above conclusion into Theorem 21.

Corollary 45. *Suppose that $n \in \mathbb{N}$ and that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ are two n-dimensional normed spaces. Suppose also that $\mathbf{Y}$ is canonically positioned and $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$. Then,*

$$\mathsf{e}(\mathbf{X}) \lesssim \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Y}})\,\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{Y}})\sqrt{n}}.$$

The assumption in Corollary 45 that $\mathbf{Y}$ is canonically positioned can be replaced by the requirement $\mathrm{MaxProj}(B_{\mathbf{Y}}) \lesssim \mathrm{vol}_{n-1}(\partial B_{\mathbf{Y}})/\sqrt{n}$, which is much less stringent. In particular, by [104, Proposition 3.1] it is enough to assume here that $B_{\mathbf{Y}}$ is in its minimum surface area position; see also Section 6.2.

We will denote the John and Löwner ellipsoids of a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ by $\mathcal{J}_{\mathbf{X}}$ and $\mathcal{L}_{\mathbf{X}}$, respectively; see [128]. Thus, $\mathcal{J}_{\mathbf{X}} \subseteq \mathbb{R}^n$ is the ellipsoid of maximum volume that is contained in $B_{\mathbf{X}}$ and $\mathcal{L}_{\mathbf{X}} \subseteq \mathbb{R}^n$ is the ellipsoid of minimum volume that contains $B_{\mathbf{X}}$. Both of these ellipsoids are unique [137]. The *volume ratio* $\mathrm{vr}(\mathbf{X})$ of $\mathbf{X}$ and *external volume ratio* $\mathrm{evr}(\mathbf{X})$ of $\mathbf{X}$ are defined by

$$\mathrm{vr}(\mathbf{X}) \stackrel{\mathrm{def}}{=} \left(\frac{\mathrm{vol}_n(B_{\mathbf{X}})}{\mathrm{vol}_n(\mathcal{J}_{\mathbf{X}})}\right)^{\frac{1}{n}} \quad \text{and} \quad \mathrm{evr}(\mathbf{X}) \stackrel{\mathrm{def}}{=} \left(\frac{\mathrm{vol}_n(\mathcal{L}_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})}\right)^{\frac{1}{n}}. \tag{1.70}$$

By the Blaschke–Santaló inequality [39, 278] and the Bourgain–Milman inequality [50],

$$\mathrm{evr}(\mathbf{X}) \asymp \mathrm{vr}(\mathbf{X}^*). \tag{1.71}$$

By the above discussion, we can quickly deduce the following theorem that relates the Lipschitz extension modulus of a canonically positioned space to volumetric and spectral properties of its unit ball.

Theorem 46. *Suppose that $n \in \mathbb{N}$ and that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a canonically positioned normed space. Then,*

$$\begin{aligned}\mathsf{e}(\mathbf{X}) \lesssim \frac{\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\sqrt{n}}\sqrt{\lambda(\mathbf{X})} &\asymp \mathrm{evr}(\mathbf{X})\sqrt{\lambda(\mathbf{X})\,\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{2}{n}}} \\ &\asymp \mathrm{vr}(\mathbf{X}^*)\sqrt{\lambda(\mathbf{X})\,\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{2}{n}}}.\end{aligned} \tag{1.72}$$

In fact, the minimum of the right-hand side of (1.54) over all those normed spaces $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ for which $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ is bounded above and below by universal constant multiples of $\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})\sqrt{\lambda(\mathbf{X})/n}$.

Proof. By Lemma 42 the Cheeger space Ch $\mathbf{X}$ is canonically positioned. So, by Corollary 45 with $\mathbf{Y} = \mathrm{Ch}\,\mathbf{X}$,

$$\mathsf{e}(\mathbf{X}) \lesssim \frac{\mathrm{vol}_{n-1}(\partial\,\mathrm{Ch}\,B_{\mathbf{Y}})\,\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\mathrm{vol}_n(\mathrm{Ch}\,B_{\mathbf{Y}})\sqrt{n}} \overset{(1.61)}{\lesssim} \frac{\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\sqrt{n}}\sqrt{\lambda(\mathbf{X})}.$$

This proves the first inequality in (1.72). The final equivalence in (1.72) is (1.71). To prove the rest of (1.72), let $r_{\min} = \min\{r > 0 : rB_{\ell_2^n} \supseteq B_{\mathbf{X}}\}$ denote the radius of the circumscribing Euclidean ball of $B_{\mathbf{X}}$. We claim that $r_{\min}B_{\ell_2^n} = \mathcal{L}_{\mathbf{X}}$. Indeed, for every $U \in \mathsf{Isom}(\mathbf{X}) \subseteq \mathsf{O}_n$ the ellipsoid $U\mathcal{L}_{\mathbf{X}}$ contains $B_{\mathbf{X}}$ and has the same volume as $\mathcal{L}_{\mathbf{X}}$, so because the minimum volume ellipsoid that contains $B_{\mathbf{X}}$ is unique [137], it follows that $U\mathcal{L}_{\mathbf{X}} = \mathcal{L}_{\mathbf{X}}$. Hence, the scalar product that corresponds to $\mathcal{L}_{\mathbf{X}}$ is $\mathsf{Isom}(\mathbf{X})$-invariant and since $\mathbf{X}$ is canonically positioned, this means that $\mathcal{L}_{\mathbf{X}}$ is a multiple of $B_{\ell_2^n}$. Now,

$$\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}\,\mathrm{evr}(\mathbf{X}) \overset{(1.70)}{=} \mathrm{vol}_n\big(r_{\min}B_{\ell_2^n}\big)^{\frac{1}{n}} \asymp \frac{r_{\min}}{\sqrt{n}} = \frac{\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{2\sqrt{n}}.$$

The above reasoning shows that the minimum of the right-hand side of (1.54) over all the normed spaces $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ with $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ is at most a universal constant multiple of $\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})\sqrt{\lambda(\mathbf{X})/n}$ (take $\mathbf{Y} = \mathrm{Ch}\,\mathbf{X}$). In the reverse direction, for any such $\mathbf{Y}$ by (1.55) with $L = B_{\mathbf{Y}}$ we have

$$\frac{\mathrm{MaxProj}(B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \gtrsim \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})\sqrt{n}} \geqslant \frac{\mathrm{vol}_{n-1}(\partial\,\mathrm{Ch}\,B_{\mathbf{X}})}{\mathrm{vol}_n(\mathrm{Ch}\,B_{\mathbf{X}})\sqrt{n}} \overset{(1.61)}{\geqslant} \frac{2\sqrt{\lambda(\mathbf{X})}}{\pi\sqrt{n}},$$

where the penultimate step follows from the definition of the Cheeger body Ch $B_{\mathbf{X}}$. ∎

It is natural to expect that if $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a canonically positioned normed space, then in Conjecture 9 for $K = B_{\mathbf{X}}$ holds with S the identity matrix and with L being the unit ball of a canonically positioned normed space. We formulate this refined special case of Conjecture 9 as the following conjecture.

Conjecture 47. Fix $n \in \mathbb{N}$ and a canonically positioned normed space

$$\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}}).$$

Then, there exists a canonically positioned normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ that satisfies $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\mathbf{X}}$ and $\mathrm{iq}(B_{\mathbf{Y}}) \lesssim \sqrt{n}$.

Theorem 48 below shows that Conjecture 47 holds if $\mathbf{X} = \ell_p^n$ for any $p \geqslant 1$ and infinitely many dimensions $n \in \mathbb{N}$; specifically, it holds if n satisfies the mild arithmetic (divisibility) requirement (1.73) below. An obvious question that this leaves is to prove Conjecture 47 for $\mathbf{X} = \ell_p^n$ and arbitrary $(p, n) \in [1, \infty] \times \mathbb{N}$. We expect that this question is tractable by (likely nontrivially) adapting the approach herein, but we did not make a major effort to do so since obtaining Conjecture 47 for such a dense set of dimensions n suffices for our purposes (the bi-Lipschitz invariants that we consider can be estimated from above for any $n \in \mathbb{N}$ since the requirement (1.73) holds for some $N \in \mathbb{N} \cap [n, O(n)]$ and ℓ_p^n embeds isometrically into ℓ_p^N). In Section 6 we will prove Theorem 48, and deduce Theorem 24 from it. Recall Remark 31, which explains that Conjecture 9 when K is the unit ball of ℓ_p^n follows (with S the identity matrix) from Theorem 24. Thus, we *do* know that a body L as in Conjecture 9 exists for all the possible choices of $p \geqslant 1$ and $n \in \mathbb{N}$, and (1.73) is only relevant to ensure that L is the unit ball of a canonically positioned normed space.

Theorem 48. *Fix $n \in \mathbb{N}$ and $p \geqslant 1$. Conjecture 47 holds for $\mathbf{X} = \ell_p^n$ if the following condition is satisfied:*

$$\exists m \in \mathbb{N}, \quad m \mid n \quad \textit{and} \quad \max\{p, 2\} \leqslant m \leqslant e^p. \tag{1.73}$$

The following conjecture is a variant of Conjecture 11.

Conjecture 49. Fix $n \in \mathbb{N}$ and suppose that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a canonically positioned normed space. Then, there exists a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ with $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ yet $\sqrt[n]{\mathrm{vol}_n(B_{\mathbf{Y}})} \gtrsim \sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})}$ such that $\mathrm{iq}(B_{\mathbf{Y}}) \lesssim \sqrt{n}$.

Conjecture 47 requires $\mathbf{Y}$ to be canonically positioned while Conjecture 49 does not. The reason for this is that if any normed space $\mathbf{Y}$ satisfies the conclusion of Conjecture 49, then also the Cheeger space $\mathrm{Ch}\,\mathbf{X}$ of $\mathbf{X}$ satisfies it (this is so because the convex body L that minimizes the second quantity in (1.60) is, by definition, the Cheeger body of $K = B_{\mathbf{X}}$), and by Lemma 42 the Cheeger space of $\mathbf{X}$ inherits from $\mathbf{X}$ the property of being canonically positioned. This use of the uniqueness of the Cheeger body will be important below. By (1.62), Conjecture 49 is equivalent to the following symmetric version of Conjecture 35.

Conjecture 50. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a canonically positioned normed space, then $\lambda(\mathbf{X})\,\mathrm{vol}(B_{\mathbf{X}})^{\frac{2}{n}} \asymp n$.

The following corollary is a substitution of Conjecture 50 into Theorem 46.

Corollary 51. *If Conjecture 49 (equivalently, Conjecture 50) holds for a canonically positioned normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$, then the right-hand side of (1.54) when $\mathbf{Y} = \mathrm{Ch}\,\mathbf{X}$ is $O(\mathrm{evr}(\mathbf{X})\sqrt{n})$. Consequently,*

$$\mathsf{e}(\mathbf{X}) \lesssim \mathrm{evr}(\mathbf{X})\sqrt{n} \asymp \mathrm{vr}(\mathbf{X}^*)\sqrt{n}. \tag{1.74}$$

It is worthwhile to note that by [19], the rightmost quantity in (1.74) is maximized (over all possible n-dimensional normed spaces) when $\mathbf{X} = \ell_1^n$, in which case we have $\mathrm{evr}(\ell_1^n)\sqrt{n} \asymp n$.

Remark 52. We currently do not have any example of a normed space

$$\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$$

for which (1.74) provably does not hold. If (1.74) were true in general, or even if it were true for a restricted class of normed spaces that is affine invariant and closed under direct sums, such as spaces that embed into ℓ_1 with distortion $O(1)$, then it would be an excellent result. When one leaves the realm of canonically positioned spaces, (1.74) acquires a self-improving property[8] as follows. Suppose that $\mathbf{X}$ is in Löwner position, i.e., $\mathcal{L}_{\mathbf{X}} = B_{\ell_2^n}$. Fix $m \in \mathbb{N}$ and consider the $(n+m)$-dimensional space $\mathbf{X}' = \mathbf{X} \oplus_\infty \ell_2^m$. If (1.74) holds for $\mathbf{X}'$, then

$$\begin{aligned}
\mathsf{e}(\mathbf{X}) &\leqslant \mathsf{e}(\mathbf{X}') \\
&\lesssim \mathrm{evr}(\mathbf{X}')\sqrt{\dim(\mathbf{X}')} \\
&\lesssim \left(\frac{\mathrm{vol}_{n+m}\big(B_{\ell_2^{n+m}}\big)}{\mathrm{vol}_n(B_{\mathbf{X}})\,\mathrm{vol}_m\big(B_{\ell_2^m}\big)}\right)^{\frac{1}{n+m}}\sqrt{n+m} \\
&= \left(\frac{\mathrm{vol}_n(\mathcal{L}_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})}\right)^{\frac{1}{n+m}}\left(\frac{\mathrm{vol}_{n+m}\big(B_{\ell_2^{n+m}}\big)}{\mathrm{vol}_n\big(\ell_2^n\big)\,\mathrm{vol}_m\big(B_{\ell_2^m}\big)}\right)^{\frac{1}{n+m}}\sqrt{n+m} \\
&\asymp \mathrm{evr}(\mathbf{X})^{\frac{n}{n+m}}\, n^{\frac{n}{2(n+m)}}\, m^{\frac{m}{2(n+m)}}.
\end{aligned} \tag{1.75}$$

The value of m that minimizes the right-hand side of (1.75) is

$$m \asymp n\log(\mathrm{evr}(\mathbf{X})+1),$$

for which (1.75) becomes

$$\mathsf{e}(\mathbf{X}) \lesssim \sqrt{n\log\big(\mathrm{evr}(\mathbf{X})+1\big)}. \tag{1.76}$$

As $\mathrm{evr}(\mathbf{X}) \leqslant \sqrt{n}$ by John's theorem, (1.76) gives $\mathsf{e}(\mathbf{X}) \lesssim \sqrt{n\log n}$, which would be an improvement of [140]. Also, by (1.9) the bound (1.76) gives

$$\mathsf{e}(\mathbf{X}) \lesssim \sqrt{n\log(C_2(\mathbf{X})+1)},$$

which is better than the conjectural bound (1.10). Here and throughout what follows, for $1 \leqslant p \leqslant 2 \leqslant q$ the (Gaussian) type-p and cotype-q constants [204] of a Banach space $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$, denoted $T_p(\mathbf{X})$ and $C_q(\mathbf{X})$, respectively, are the infimum over those

[8]We recommend checking that the analogous stabilization argument does not lead to a similar self-improvement phenomenon in Conjecture 9, Conjecture 10 and Corollary 33; the computations in Section 4 of [198] are relevant for this purpose.

$T \in [1,\infty]$ and $C \in [1,\infty]$, respectively, for which the following inequalities hold for every $m \in \mathbb{N}$ and every $x_1,\ldots,x_m \in \mathbf{X}$, where the expectation is with respect to i.i.d. standard Gaussian random variables $\mathsf{g}_1,\ldots,\mathsf{g}_m$:

$$\frac{1}{C}\Bigg(\sum_{j=1}^{m}\|x_j\|_{\mathbf{X}}^{q}\Bigg)^{\frac{1}{q}} \leqslant \Bigg(\mathbb{E}\Bigg[\Bigg\|\sum_{j=1}^{m}\mathsf{g}_j x_j\Bigg\|_{\mathbf{X}}^{2}\Bigg]\Bigg)^{\frac{1}{2}} \leqslant T\Bigg(\sum_{j=1}^{m}\|x_j\|_{\mathbf{X}}^{p}\Bigg)^{\frac{1}{p}}. \tag{1.77}$$

This observation indicates that it might be too optimistic to expect that (1.74) holds in full generality, but it would be very interesting to understand the extent to which it does. Obvious potential counterexamples are $\ell_1^n \oplus \ell_2^m$; if (1.74) holds for these spaces, then $\mathsf{e}(\ell_1^n) \lesssim \sqrt{n\log n}$ by the above reasoning (with $m \asymp n\log n$), which would be a big achievement because the best-known bound remains $\mathsf{e}(\ell_1^n) \lesssim n$ from [140].

Lemma 53 below, whose proof appears in Section 6.1, shows that Conjecture 49 holds for a class of normed space that includes any normed spaces with a 1-symmetric basis, as well as, say, $\ell_p^n(\ell_q^m)$ for any $n,m \in \mathbb{N}$ and $p,q \geqslant 1$. Other (related) examples of such spaces arise from Lemma 151 below.

Lemma 53. *Let $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ be an unconditional normed space. Suppose that for any $j,k \in \{1,\ldots,n\}$ there is a permutation $\pi \in S_n$ with $\pi(j) = k$ such that $\|\sum_{i=1}^n a_{\pi(i)}e_i\|_{\mathbf{X}} = \|\sum_{i=1}^n a_i e_i\|_{\mathbf{X}}$ for every $a_1,\ldots,a_n \in \mathbb{R}$. Then, Conjecture 49 holds for $\mathbf{X}$. Therefore, we have $\lambda(\mathbf{X})\,\mathrm{vol}_n(B_{\mathbf{X}})^{2/n} \asymp n$ and $\mathsf{e}(\mathbf{X}) \lesssim \mathrm{evr}(\mathbf{X})\sqrt{n}$.*

By [293, Theorem 2.1], any unconditional normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies $\mathrm{vr}(\mathbf{X}) \lesssim C_2(\mathbf{X})\sqrt{n}$, where $C_2(\mathbf{X})$ is the cotype-2 constant of $\mathbf{X}$ (this is an earlier special case of (1.9) in which the logarithmic term is known to be redundant). Hence, if $\mathbf{X}$ satisfies the assumptions of Lemma 53, then we know that

$$\mathsf{e}(\mathbf{X}) \lesssim C_2(\mathbf{X}^*)\sqrt{n}. \tag{1.78}$$

By combining [22, Theorem 6] and (1.71), for any $p \in [1,\infty]$, if a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is isometric to a quotient of L_p (equivalently, the dual of $\mathbf{X}$ is isometric to a subspace of $L_{p/(p-1)}$), then

$$\mathrm{evr}(\mathbf{X}) \lesssim \mathrm{evr}\big(\ell^n_{\frac{p}{p-1}}\big) \asymp \min\big\{n^{\frac{1}{p}-\frac{1}{2}}, 1\big\}.$$

Consequently, if $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies the assumptions of Lemma 53 and is also a quotient of L_p, then

$$\mathsf{e}(\mathbf{X}) \lesssim n^{\max\{\frac{1}{2},\frac{1}{p}\}}. \tag{1.79}$$

Both (1.78) and (1.79) are generalizations of Theorem 18.

Lemma 54 below, whose proof appears in Section 6.3, shows that the unitary ideal of any n-dimensional normed space with a 1-symmetric basis (in particular, any Schatten–von Neumann trace class), satisfies Conjecture 49 up to a factor of $O(\sqrt{\log n})$. Upon its substitution into Lemma 151 below, more such examples are obtained.

Lemma 54. *Let* $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ *be a symmetric normed space. Conjecture 49 holds up to lower order factors for its unitary ideal* $\mathsf{S}_{\mathbf{E}}$*. More precisely, there is a normed space* $\mathbf{Y} = (\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\mathbf{Y}})$ *such that* $B_{\mathbf{Y}} \subseteq B_{\mathsf{S}_{\mathbf{E}}}$ *and*

$$\operatorname{vol}_{n^2}(B_{\mathbf{Y}})^{\frac{1}{n^2}} \asymp \operatorname{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{E}}}\big)^{\frac{1}{n^2}} \quad \textit{and} \quad n \lesssim \operatorname{iq}(B_{\mathbf{Y}}) \lesssim n\sqrt{\log n}. \tag{1.80}$$

Therefore, we have

$$n^2 \lesssim \lambda\big(\mathsf{S}_{\mathbf{E}}\big) \operatorname{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{E}}}\big)^{\frac{2}{n^2}} \lesssim n^2 \log n \quad \textit{and} \quad \mathsf{e}(\mathsf{S}_{\mathbf{E}}) \lesssim \operatorname{evr}(\mathsf{S}_{\mathbf{E}})n \asymp \operatorname{evr}(\mathbf{E})n.$$

For the final assertion of Lemma 54, the fact that $\operatorname{evr}(\mathsf{S}_{\mathbf{E}}) \asymp \operatorname{evr}(\mathbf{E})$ follows by combining Proposition 2.2 in [285], which states that $\operatorname{vr}(\mathsf{S}_{\mathbf{E}}) \asymp \operatorname{vr}(\mathbf{E})$, with (1.71) and the duality $\mathsf{S}_{\mathbf{E}}^* = \mathsf{S}_{\mathbf{E}^*}$ (e.g., [289, Theorem 1.17]).

The proof of Lemma 54 also shows (see Remark 172 below) that if we could prove Conjecture 49 for S_∞^n, then it would follow that $\mathsf{S}_{\mathbf{E}}$ satisfies Conjecture 49 for any symmetric normed space $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$, i.e., the logarithmic factor in (1.80) could be replaced by a universal constant.

By substituting Lemma 54 into Corollary 51 and using volume ratio computations of Schütt [285], we will derive in Section 6.3 the following proposition.

Proposition 55. *If* $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ *is a symmetric normed space, then*

$$\mathsf{e}(\mathbf{E}) \lesssim \operatorname{diam}_{\ell_2^n}\big(B_{\mathbf{E}}\big)\|e_1 + \cdots + e_n\|_{\mathbf{E}}$$

and

$$\mathsf{e}\big(\mathsf{S}_{\mathbf{E}}\big) \lesssim \operatorname{diam}_{\ell_2^n}\big(B_{\mathbf{E}}\big)\|e_1 + \cdots + e_n\|_{\mathbf{E}}\sqrt{n \log n}.$$

The following remark sketches an alternative approach towards Conjecture 9 when K is the hypercube $[-1,1]^n$ that differs from how we will prove Theorem 24. It yields the desired result up to a lower order factor that grows extremely slowly. Specifically, it constructs an origin-symmetric convex body $L \subseteq [-1,1]^n$ with

$$\operatorname{iq}(L) = e^{O(\log^* n)} \quad \text{and} \quad [-1,1]^n \subseteq e^{O(\log^* n)} L.$$

Here, for each $x \geqslant 1$ the quantity $\log^* x$ is defined to be the $k \in \mathbb{N}$ such that

$$\mathsf{tower}(k-1) \leqslant x < \mathsf{tower}(k)$$

for the sequence $\{\mathsf{tower}(i)\}_{i=0}^\infty$ that is defined by $\mathsf{tower}(0) = 1$ and $\mathsf{tower}(i+1) = \exp(\mathsf{tower}(i))$. We think that this approach is worthwhile to describe despite the fact that it falls slightly short of fully establishing Conjecture 9 for $[-1,1]^n$ due to its flexibility that could be used for other purposes, as well as due to its intrinsic interest.

Remark 56. Fix $n \in \mathbb{N}$ and $q \geqslant 1$. Since the nth root of the volume of the unit ball of ℓ_q^n is of order $n^{-1/q}$ and ℓ_q^n is in minimum surface area position, we can restate (1.42) as

$$\mathrm{iq}\big(B_{\ell_q^n}\big) \asymp \min\{\sqrt{qn}, n\}. \tag{1.81}$$

In particular, for $\mathbf{Y} = \ell_q^n$ with $q = \log n$, we have $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\ell_\infty^n}$ and

$$\mathrm{iq}(\mathbf{Y}) \lesssim \sqrt{n \log n},$$

which already comes close to the conclusion of Conjecture 9. We can do better using the following evaluation of the isoperimetric quotient of the unit ball of $\ell_p^n(\ell_q^m)$, which holds for every $n, m \in \mathbb{N}$ and $p, q \geqslant 1$:

$$\mathrm{iq}\big(B_{\ell_p^n(\ell_q^m)}\big) \asymp \begin{cases} nm & m \leqslant \min\{\frac{p}{n}, q\}, \\ n\sqrt{qm} & q \leqslant m \leqslant \frac{p}{n}, \\ \sqrt{pnm} & \frac{p}{n} \leqslant m \leqslant \min\{p, q\}, \\ \sqrt{pqn} & \max\{\frac{p}{n}, q\} \leqslant m \leqslant p, \\ m\sqrt{n} & p \leqslant m \leqslant q, \\ \sqrt{qnm} & m \geqslant \max\{p, q\}. \end{cases} \tag{1.82}$$

We will prove (1.82) in Section 6. Note that when $m = 1$ this yields (1.81). The case $n = m$ of (1.82) is equivalent to (1.49) since $\ell_p^n(\ell_q^m)$ is canonically positioned (it belongs to the class of spaces in Example 40) and using a simple evaluation of the volume of its unit ball (see (6.6) below). The range of (1.82) that is most pertinent for the present context is $m \geqslant \max\{p, q\}$, which has the feature that the factor that multiplies the quantity

$$\sqrt{nm} = \sqrt{\dim\big(\ell_p^n(\ell_q^m)\big)}$$

is $O(\sqrt{q})$ and there is no dependence on p. This can be used as follows. Suppose that $n = ab$ for $a, b \in \mathbb{N}$ satisfying $a \asymp n/\log n$ and $b \asymp \log n$. Identify ℓ_∞^n with $\ell_\infty^a(\ell_\infty^b)$. If we set $\mathbf{Y} = \ell_p^a(\ell_q^b)$ for $p = \log a \asymp \log n$ and $q = \log b \asymp \log\log n$, then $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\ell_\infty^n}$, while

$$\mathrm{iq}(B_{\mathbf{Y}}) \asymp \sqrt{n \log\log n}$$

by (1.82). By iterating we get that for infinitely many $n \in \mathbb{N}$ there is a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ for which

$$\|\cdot\|_{\mathbf{Y}} \leqslant \|\cdot\|_{\ell_\infty^n} \lesssim e^{O(\log^* n)}\|\cdot\|_{\mathbf{Y}} \quad \text{and} \quad \mathrm{iq}(B_{\mathbf{Y}}) = e^{O(\log^* n)}.$$

Even though the set of $n \in \mathbb{N}$ for which this works is not all of $\mathbb{N}$, it is quite dense in $\mathbb{N}$ per Lemma 163 below. This will allow us to deduce that a space $\mathbf{Y}$ with the above properties exists for every $n \in \mathbb{N}$; see Section 6.1 for the details.

Remark 57. Recalling Remark 38, Conjecture 10 is equivalent to the assertion that if a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is in Cheeger position, then $\mathrm{iq}(\mathrm{Ch}\, B_{\mathbf{X}}) \lesssim \sqrt{n}$ and $\mathrm{vol}_n(\mathrm{Ch}\, B_{\mathbf{X}})^{1/n} \gtrsim \mathrm{vol}_n(B_{\mathbf{X}})^{1/n}$. Since $\mathrm{Ch}\,\mathbf{X}$ is in minimum surface area position when $\mathbf{X}$ is in Cheeger position (as explained in Remark 38), the proof of Proposition 32 shows that Conjecture 10 implies that if $\mathbf{X}$ is in Cheeger position, then

$$\mathsf{e}(\mathbf{X}) \lesssim \frac{\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}}. \tag{1.83}$$

In fact, the right-hand side of (1.54) is at most the right-hand side of (1.83) for a suitable choice of normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$, specifically for $\mathbf{Y} = \mathrm{Ch}\,\mathbf{X}$. The discussion in Section 1.6.2 was about establishing (1.83) when $\mathbf{X}$ is canonically positioned (conceivably that assumption implies that $\mathbf{X}$ is in Cheeger position or close to it, which would be a worthwhile to prove, if true). Even though, as we explained earlier, given the current state of knowledge, understanding the Lipschitz extension problem for canonically positioned spaces is the most pressing issue for future research, it would be very interesting to study if (1.83) holds in other situations. For examples, we pose the following two natural questions.

Question 58. Does (1.83) hold if the normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is in minimum surface area position?

The extent to which $\Pi\mathbf{X}$ is close to being in minimum surface area position when $\mathbf{X}$ is in minimum surface area position seems to be unknown. Therefore, the connection between Question 59 below and Question 58 is unclear, but even if there is no formal link between these two questions, both are natural next steps beyond the setting of canonically positioned normed spaces.

Question 59. Let $\mathbf{Z} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Z}})$ be a normed space in minimum surface area position. Does (1.83) hold for the normed space $\mathbf{X} = \Pi\mathbf{Z}$ whose unit ball is the projection body of $B_{\mathbf{X}}$?

If $\mathbf{Z} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Z}})$ is a normed space in minimum surface area position, then

$$\frac{\mathrm{diam}_{\ell_2^n}(\Pi B_{\mathbf{Z}})}{\mathrm{vol}_n(\Pi B_{\mathbf{Z}})^{\frac{1}{n}}} \asymp \sqrt{n}. \tag{1.84}$$

Indeed, because $\mathbf{Z}$ is in minimum surface area position, by [104, Corollary 3.4] we have

$$\mathrm{vol}_n(\Pi B_{\mathbf{Z}})^{\frac{1}{n}} \asymp \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})}{n},$$

and also by combining [104, Proposition 3.1] and (1.55) we have

$$\mathrm{MaxProj}(B_{\mathbf{Z}}) \asymp \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})}{\sqrt{n}}.$$

We can therefore justify (1.84) using these results from [104] and duality as follows:

$$\frac{\mathrm{diam}_{\ell_2^n}(\Pi B_{\mathbf{Z}})}{\mathrm{vol}_n(\Pi B_{\mathbf{Z}})^{\frac{1}{n}}} \asymp \frac{n\|\mathsf{Id}_n\|_{\Pi\mathbf{Z}\to\ell_2^n}}{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})} = \frac{n\|\mathsf{Id}_n\|_{\ell_2^n\to\Pi^*\mathbf{Z}}}{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})}$$
$$= \frac{n\max_{z\in S^{n-1}}\|z\|_{\Pi^*\mathbf{Z}}}{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})} \overset{(1.30)}{=} \frac{n\mathrm{MaxProj}(B_{\mathbf{Z}})}{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Z}})} \asymp \sqrt{n}.$$

By this observation, a positive answer to Question 59 would show that $\mathsf{e}(\Pi\mathbf{Z}) \lesssim \sqrt{n}$ for any normed space $\mathbf{Z} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Z}})$. Indeed, if we take $S \in \mathsf{SL}_n(\mathbb{R})$ such that $S\mathbf{Z}$ is in minimum surface area position, then by [251] we know that $\Pi\mathbf{Z}$ and $\Pi S\mathbf{Z}$ are isometric, so $\mathsf{e}(\Pi\mathbf{Z}) = \mathsf{e}(\Pi S\mathbf{Z})$. As the class of projection bodies coincides with the class of zonoids [41, 283], which coincides with the class of convex bodies whose polar is the unit ball of a subspace of L_1, we have thus shown that a positive answer to Question 59 would imply the following conjecture (which would simultaneously improve (1.23) and generalize Theorem 18).

Conjecture 60. For any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ we have

$$\mathsf{e}(\mathbf{X}) \lesssim \mathsf{c}_{L_1}(\mathbf{X}^*)\sqrt{n}.$$

Note that Conjecture 60 is consistent with the estimate $\mathsf{e}(\mathbf{X}) \lesssim \mathrm{evr}(\mathbf{X})\sqrt{n}$ that has been arising thus far. Indeed, if $\mathbf{X}^*$ is isometric to a subspace of L_1 (it suffices to consider only this case in Conjecture 60 by a well-known differentiation argument; see, e.g., [36, Corollary 7.10]), then we have the bound $\mathrm{evr}(\mathbf{X}) \lesssim 1$ which can be seen to hold by combining (1.71) with (1.9), since $C_2(\mathbf{X}^*) \leqslant C_2(L_1) \lesssim 1$.[9]

Relating $\mathsf{e}(\mathbf{X})$ to $\mathrm{evr}(\mathbf{X})$ is valuable since the Lipschitz extension modulus is for the most part shrouded in mystery, while the literature contains extensive knowledge on volume ratios (we have already seen several examples of such consequences above, and we will derive more later). Section 6.3 contains examples of volume ratio evaluations for various canonically positioned normed spaces. Through their substitution into Corollary 51, they illustrate how our work yields a range of new Lipschitz extension results, some of which are currently conjectural because they hold assuming Conjecture 49 for the respective spaces; specifically, consider the Lipschitz extension bounds that correspond to using (1.14) and (1.15) with [173].

[9] Alternatively, $\mathrm{evr}(\mathbf{X}) \lesssim 1$ can be justified by writing $\mathbf{X} = \Pi\mathbf{Z}$ for some normed space $\mathbf{Z} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Z}})$ (using [41, 283]), and then applying the bound (1.84) that we derived above (this even demonstrates that the external volume ratio of $\Pi\mathbf{Z}$ is $O(1)$ when $\mathbf{Z}$ is in minimum surface area position rather when $\mathbf{Z}$ is in Löwner position). Actually, the sharp bound $\mathrm{evr}(\mathbf{X}) \leqslant \mathrm{evr}(\ell_\infty^n)$ holds, as seen by combining [22, Theorem 6] with Reisner's theorem [271] that the Mahler conjecture [193] holds for zonoids.

1.6.3 Intersection with a Euclidean ball

Fix an integer $n \geqslant 2$ and a canonically positioned normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. A natural first attempt to prove Conjecture 49 for $\mathbf{X}$ is to consider the normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ such that $B_{\mathbf{Y}} = B_{\mathbf{X}} \cap r B_{\ell_2^n}$ for a suitably chosen $r > 0$ (equivalently, we have $\|x\|_{\mathbf{Y}} = \max\{\|x\|_{\mathbf{X}}, \|x\|_{\ell_2^n}/r\}$ for every $x \in \mathbb{R}^n$). However, we checked with G. Schechtman that this fails even when $\mathbf{X} = \ell_\infty^n$. Specifically, if the nth root of the volume of $B_{\ell_\infty^n} \cap (r B_{\ell_2^n})$ is at least a universal constant, then necessarily $r \gtrsim \sqrt{n}$, but

$$\forall s > 0, \quad \mathrm{iq}\big(B_{\ell_\infty^n} \cap (s\sqrt{n} B_{\ell_2^n})\big) \gtrsim_s n. \tag{1.85}$$

A justification of (1.85) appears in Section 7 below. In terms of the quantification (1.60) of Conjecture 49 that is pertinent to the applications that we study herein, we will also show in Section 7 that

$$\min_{r>0} \frac{\mathrm{iq}\big(B_{\ell_\infty^n} \cap (r B_{\ell_2^n})\big)}{\sqrt{n}} \left(\frac{\mathrm{vol}_n(B_{\ell_\infty^n})}{\mathrm{vol}_n\big(B_{\ell_\infty^n} \cap (r B_{\ell_2^n})\big)} \right)^{\frac{1}{n}} \asymp \sqrt{\log n}, \tag{1.86}$$

where the minimum in the right-hand side of (1.86) is attained at some $r > 0$ that satisfies $r \asymp \sqrt{n/\log n}$.

Even though the above bounds demonstrate that it is impossible to resolve Conjecture 49 by intersecting with a Euclidean ball, this approach cannot fail by more than a lower-order factor; the reasoning that proves this assertion was shown to us by B. Klartag and E. Milman in unpublished private communication that is explained with their permission in Section 7. Specifically, we have the following proposition.

Proposition 61. *For any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ there exist a matrix $S \in \mathsf{SL}_n(\mathbb{R})$ and a radius $r > 0$ such that for $L = (SB_{\mathbf{X}}) \cap (r B_{\ell_2^n}) \subseteq SB_{\mathbf{X}}$ we have $\mathrm{iq}(L) \lesssim \sqrt{n}$ and $\sqrt[n]{\mathrm{vol}_n(L)} \gtrsim \sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})}/K(\mathbf{X})$, where $K(\mathbf{X})$ is the K-convexity constant of $\mathbf{X}$. If $\mathbf{X}$ is canonically positioned, then this holds when S is the identity matrix.*

For Proposition 61, the K-convexity constant of $\mathbf{X}$ is an isomorphic invariant that was introduced by Maurey and Pisier [204]; we defer recalling its definition to Section 7 since for the discussion here it suffices to state the following bounds that relate $K(\mathbf{X})$ to quantities that we already encountered. Firstly,

$$K(\mathbf{X}) \lesssim \log\big(d_{\mathrm{BM}}(\ell_2^n, \mathbf{X}) + 1\big) \lesssim \log n, \tag{1.87}$$

The first inequality in (1.87) is a useful theorem of Pisier [256, 257]. The second inequality in (1.87) follows from John's theorem [137], though for this purpose it suffices to use the older Auberbach lemma (see [27, p. 209] and [83, 300]). By [257] (see also, e.g., [143, Lemma 17]) the rightmost quantity in (1.87) can be reduced if $\mathbf{X}$ is a subspace of L_1, namely we have

$$K(\mathbf{X}) \lesssim \mathsf{c}_{L_1}(\mathbf{X})\sqrt{\log n}. \tag{1.88}$$

Secondly, $K(\mathbf{X})$ relates to the notion of type that we recalled in (1.77) through the following bounds:

$$T_{1+\frac{c}{K(\mathbf{X})^2}}(\mathbf{X})^{\frac{1}{2}} \lesssim K(\mathbf{X}) \leqslant \inf_{p\in(1,2]} e^{(CT_p(\mathbf{X}))^{\frac{p}{p-1}}}, \tag{1.89}$$

where $c, C > 0$ are universal constants. The qualitative meaning of (1.89) is that the K-convexity constant of a Banach space is finite if and only if it has type p for some $p > 1$; this is a landmark theorem of Pisier (the 'if' direction is due to [259] and the 'only if' direction is due to [254]). Since in our setting $\mathbf{X}$ is finite dimensional ($\dim(\mathbf{X}) = n \geqslant 2$), such a qualitative statement is vacuous without its quantitative counterpart (1.89). The first inequality in (1.89) can be deduced from [260] (together with the computation of the implicit dependence on p in [260] that was carried out in [131, Lemma 32]). The second inequality in (1.89) follows from an examination of the proof in [259]. We omit the details of both deductions as they would result in a (quite lengthy and tedious) digression. It would be very interesting to determine the best bounds in the context of (1.89).

Proposition 61 combined with (1.87) implies that Conjecture 10 holds up to a logarithmic factor in the sense that for every integer $n \geqslant 2$, any origin-symmetric convex body $K \subseteq \mathbb{R}^n$ admits a matrix $S \in \mathsf{SL}_n(\mathbb{R})$ and an origin-symmetric convex body $L \subseteq SK$ such that

$$\frac{\mathrm{iq}(L)}{\sqrt{n}}\left(\frac{\mathrm{vol}_n(K)}{\mathrm{vol}_n(L)}\right)^{\frac{1}{n}} \lesssim \log n. \tag{1.90}$$

Furthermore, by (1.88) the $\log n$ in (1.90) can be replaced by $\sqrt{\log n}$ if K is the unit ball of a subspace of L_1 (equivalently, the polar of K is a zonoid), and by the second inequality in (1.89) if $p > 1$, then the $\log n$ in (1.90) can be replaced by a dimension-independent quantity that depends only on p and the type-p constant of the norm whose unit ball is K. Also, Corollary 33 holds with the right-hand side of (1.59) multiplied by $\log n$, and the reverse Faber–Krahn inequality of Conjecture 35 holds up to a factor of $(\log n)^2$, i.e., for any origin-symmetric convex body $K \subseteq \mathbb{R}^n$ there is $S \in \mathsf{SL}_n(\mathbb{R})$ such that $\lambda(SK)\,\mathrm{vol}(K)^{2/n} \lesssim n(\log n)^2$. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a canonically positioned normed space, then it follows that for a suitable choice of normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ the right-hand side of (1.28), and hence also $\mathsf{e}(\mathbf{X})$ by Theorem 21, is at most a universal constant multiple of $\mathrm{evr}(\mathbf{X})\sqrt{n}\log n$, and also $n \lesssim \lambda(\mathbf{X})\,\mathrm{vol}_n(B_{\mathbf{X}})^{2/n} \lesssim n(\log n)^2$.

1.7 Randomized clustering

All of the new upper bounds on Lipschitz extension moduli that we stated above rely on a geometric structural result for finite dimensional normed spaces (and subsets thereof). Beyond the application to Lipschitz extension, this result is of value in its own right because it yields an improvement of a basic randomized clustering method from the computer science literature.

The link between random partitions of metric spaces and Lipschitz extension was found in [173]. We will adapt the methodology of [173] to deduce the aforementioned Lipschitz extension theorems from our new bound on randomized partitions of normed spaces. In order to formulate the corresponding definitions and results, one must first set some groundwork for a notion of a random partition of a metric space, whose subsequent applications necessitate certain measurability requirements.

A framework for reasoning about random partitions of metric spaces was developed in [173], but we will formulate a different approach. The reason for this is that the definitions of [173] are in essence the minimal requirements that allow one to use at once several different types of random partitions for Lipschitz extension, which leads to definitions that are more cumbersome than the approach that we take below. Greater simplicity is not the only reason why we chose to formulate a foundation that differs from [173]. The approach that we take is easier to implement, and, importantly, it yields a bi-Lipschitz invariant, while we do not know if the corresponding notions in [173] are bi-Lipschitz invariants (we suspect that they are *not*, but we did not attempt to construct examples that demonstrate this). The Lipschitz extension theorem of [173] is adapted accordingly in Section 5, thus making the present article self-contained, and also yielding simplification and further applications. Nevertheless, the key geometric ideas that underly this use of random partitions are the same as in [173].

Obviously, there are no measurability issues when one considers finite metric spaces (in our setting, finite subsets of normed spaces). The ensuing measurability discussions can therefore be ignored in the finitary setting. In particular, the computer science literature on random partitions focuses exclusively on finite objects. So, for the purpose of algorithmic clustering, one does not need the more general treatment below, but it is needed for the purpose of Lipschitz extension.

1.7.1 Basic definitions related to random partitions

Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space. Suppose that $\mathcal{P} \subseteq 2^{\mathfrak{M}}$ is a partition of $\mathfrak{M}$. For $x \in \mathfrak{M}$, denote by $\mathcal{P}(x) \subseteq \mathfrak{M}$ the unique element of $\mathcal{P}$ to which x belongs. The sets $\{\mathcal{P}(x)\}_{x\in\mathfrak{M}}$ are often called the *clusters* of $\mathcal{P}$. Given $\Delta > 0$, one says that $\mathcal{P}$ is Δ-bounded if $\operatorname{diam}_{\mathfrak{M}}(\mathcal{P}(x)) \leqslant \Delta$ for every $x \in \mathfrak{M}$, where

$$\operatorname{diam}_{\mathfrak{M}}(S) = \sup\{d_{\mathfrak{M}}(x, y) : x, y \in S\}$$

denotes the diameter of $\varnothing \neq S \subseteq \mathfrak{M}$.

Suppose that $(\mathcal{Z}, \mathcal{F})$ is a measurable space, i.e., $\mathcal{Z}$ is a set and $\mathcal{F} \subseteq 2^{\mathcal{Z}}$ is a σ-algebra of subsets of $\mathcal{Z}$. Recall (see [133] or the convenient survey [309]) that if

$(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space, then a set-valued mapping

$$\Gamma : \mathcal{Z} \to 2^{\mathfrak{M}}$$

is said to be strongly measurable if for every closed subset $E \subseteq \mathfrak{M}$ we have

$$\Gamma^{-}(E) \stackrel{\mathrm{def}}{=} \big\{z \in \mathcal{Z} : E \cap \Gamma(z) \neq \varnothing\big\} \in \mathcal{F}. \tag{1.91}$$

Throughout what follows, when we say that $\mathcal{P}$ is a *random partition* of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, we mean the following (formally, the objects that we will be considering are random *ordered* partitions into countably many clusters). There is a probability space $(\Omega, \mathbf{Prob})$ and a sequence of set-valued mappings

$$\big\{\Gamma^k : \Omega \to 2^{\mathfrak{M}}\big\}_{k=1}^{\infty}.$$

We write $\mathcal{P}^\omega = \{\Gamma^k(\omega)\}_{k=1}^\infty$ for each $\omega \in \Omega$ and require that the mapping $\omega \mapsto \mathcal{P}^\omega$ takes values in partitions of $\mathfrak{M}$. We also require that for every fixed $k \in \mathbb{N}$, the set-valued mapping $\Gamma^k : \Omega \to 2^{\mathfrak{M}}$ is strongly measurable, where the σ-algebra on Ω is the **Prob**-measurable sets. Given $\Delta > 0$, we say that $\mathcal{P}$ is a Δ-bounded random partition of $(\mathfrak{M}, d_{\mathfrak{M}})$ if $\mathcal{P}^\omega$ is a Δ-bounded partition of $(\mathfrak{M}, d_{\mathfrak{M}})$ for every $\omega \in \Omega$.

Remark 62. Recall that when we say that $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is a normed space we mean that the underlying vector space is $\mathbb{R}^n$, equipped with a norm $\|\cdot\|_{\mathbf{X}} : \mathbb{R}^n \to [0, \infty)$. By doing so, we introduce a second metric on $\mathbf{X}$, i.e., $\mathbb{R}^n$ is also endowed with the standard Euclidean structure that corresponds to the norm $\|\cdot\|_{\ell_2^n}$. This leads to ambiguity when we discuss Δ-bounded partitions of $\mathbf{X}$ for some $\Delta > 0$, as there are two possible metrics with respect to which one could bound the diameters of the clusters. In fact, a key aspect of our work is that it can be beneficial to consider another auxiliary norm $\|\cdot\|_{\mathbf{Y}}$ on $\mathbb{R}^n$, as in, e.g., Theorem 21, thus leading to three possible interpretations of Δ-boundedness of a partition of $\mathbb{R}^n$. To avoid any confusion, we will adhere throughout to the convention that when we say that a partition $\mathcal{P}$ of $\mathbf{X}$ is Δ-bounded we mean exclusively that all the clusters of $\mathcal{P}$ have diameter at most Δ with respect to the norm $\|\cdot\|_{\mathbf{X}}$.

1.7.2 Iterative ball partitioning

Fix $\Delta \in (0, \infty)$. Iterative ball partitioning is a common procedure to construct a Δ-bounded random partition of a metric probability space. We will next describe it to clarify at the outset the nature of the objects that we investigate, and because our new positive partitioning results are solely about this type of partition. Thus, our contribution to the theory of random partitions is a sharp understanding of the performance of iterative ball partitioning of normed spaces, and, importantly, the demonstration of the utility of its implementation using balls that are induced by a suitably chosen auxiliary norm rather than the given norm that we aim to study. On the other hand, our impossibility results rule out the existence of any random partition whatsoever with certain desirable properties.

The iterative ball partitioning method is a ubiquitous tool in metric geometry and algorithm design. To the best of our knowledge, it was first used by Karger, Motwani and Sudan [152] and the aforementioned work [76] in the context of normed spaces, and it has become very influential in the context of general metric spaces due to its use in that setting (with the important twist of randomizing the radii) by Calinescu, Karloff and Rabani [71]. To describe it, suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space and that μ is a Borel probability measure on $\mathfrak{M}$. Let $\{\mathsf{X}_k\}_{k=1}^{\infty}$ be a sequence of i.i.d. points sampled μ. Define inductively a sequence $\{\Gamma^k\}_{k=1}^{\infty}$ of random subsets of $\mathfrak{M}$ by setting $\Gamma^1 = B_{\mathfrak{M}}(\mathsf{X}_1, \Delta/2)$ and

$$\forall k \in \{2, 3, \ldots, \}, \qquad \Gamma^k \stackrel{\mathrm{def}}{=} B_{\mathfrak{M}}\Big(\mathsf{X}_k, \frac{\Delta}{2}\Big) \smallsetminus \bigcup_{j=1}^{k-1} B_{\mathfrak{M}}\Big(\mathsf{X}_j, \frac{\Delta}{2}\Big).$$

By design, $\mathrm{diam}_{\mathfrak{M}}(\Gamma^k) \leqslant \Delta$. Under mild assumptions on $\mathfrak{M}$ and μ that are simple to check, Γ^k will have the measurability properties that we require below and $\mathcal{P} = \{\Gamma^k\}_{k=1}^{\infty}$ will be a partition of $\mathfrak{M}$ almost-surely. While initially the clusters of $\mathcal{P}$ are quite "tame," e.g., they start out as balls in $\mathfrak{M}$, as the iteration proceeds and we discard the balls that were used thus far, the resulting sets become increasingly "jagged." In particular, even when the underlying metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ is very "nice," the clusters of $\mathcal{P}$ need not be connected; see Figure 1.2. Nevertheless, we will see that such a simple procedure results in a random partition with probabilistically small boundaries in sense that will be described rigorously below.

In the present setting, the metric space that we wish to partition is a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$, so it is natural to want to use the Lebesgue measure on $\mathbb{R}^n$ in the above construction. Since this measure is not a probability measure, we cannot use the above framework directly. For this reason, we will in fact use a periodic variant of iterative ball partitioning of $\mathbf{X}$ by adapting a construction that was used in [173].

1.7.3 Separation and padding

Fix $\Delta > 0$. Let $\mathcal{P}$ be a Δ-bounded random partition of a metric space $\mathfrak{M}$. As a random "clustering" of $\mathfrak{M}$ into pieces of small diameter, $\mathcal{P}$ yields a certain "simplification" of $\mathfrak{M}$. For such a simplification to be useful, one must add a requirement that it "mimics" the geometry of $\mathfrak{M}$ in a meaningful way. The literature contains multiple definitions that achieve this goal, leading to applications in both algorithms and pure mathematics. We will not attempt to survey the literature on this topic, quoting only the definitions of *separating* and *padded* random partitions, which are the simplest

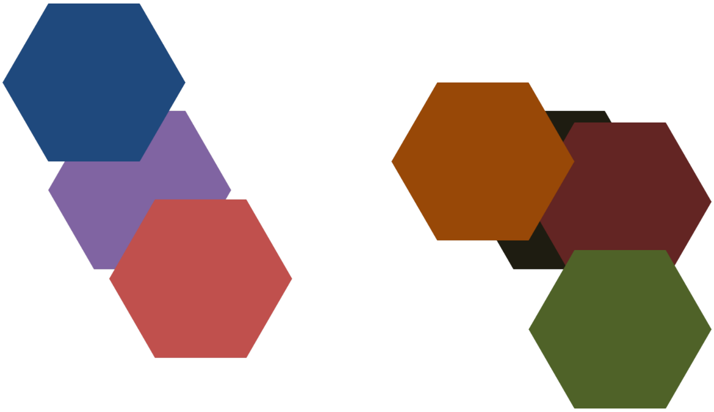

Figure 1.2. A schematic depiction of (randomized) iterative ball partitioning of a bounded subset of $\mathbb{R}^2$, where $\mathbb{R}^2$ is equipped with a norm whose unit ball is a regular hexagon. The centers of the above hexagons are chosen independently and uniformly at random from a large region that contains the given subset of $\mathbb{R}^2$. At each step of the iteration, a new hexagon appears, and it carves out a new cluster which consists of the part of the hexagon that does not intersect any of the clusters that have been formed in the previous stages of the iteration. The first few clusters that are formed by this procedure are typically hexagons, but at later stages the clusters become more complicated and less "round." In particular, they can eventually become disconnected, as exhibited by the region that is shaded black above.

and most popular notions of random partitions of metric spaces among those that have been introduced.

Definition 63 (Separating random partition and separation modulus). Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space. For $\sigma, \Delta > 0$, a Δ-bounded random partition $\mathcal{P}$ of $(\mathfrak{M}, d_{\mathfrak{M}})$ is *σ-separating* if

$$\forall x, y \in \mathfrak{M}, \quad \mathbf{Prob}\big[\mathcal{P}(x) \neq \mathcal{P}(y)\big] \leqslant \frac{\sigma}{\Delta} d_{\mathfrak{M}}(x, y). \tag{1.92}$$

The *separation modulus*[10] of $(\mathfrak{M}, d_{\mathfrak{M}})$, denoted $\mathsf{SEP}(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{SEP}(\mathfrak{M})$ if the metric is clear from the context, is the infimum over those $\sigma > 0$ such that for every $\Delta > 0$ there exists a σ-separating Δ-bounded random partition of $(\mathfrak{M}, d_{\mathfrak{M}})$. If no such σ exists, then write $\mathsf{SEP}(\mathfrak{M}, d_{\mathfrak{M}}) = \infty$. Similarly, for $n \in \mathbb{N}$, the *size-n separation modulus* of $(\mathfrak{M}, d_{\mathfrak{M}})$, denoted $\mathsf{SEP}^n(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{SEP}^n(\mathfrak{M})$ if the metric is clear from the context, is the infimum over those $\sigma > 0$ such that for every $S \subseteq \mathfrak{M}$ with $|S| \leqslant n$ and every $\Delta > 0$ there exists a σ-separating Δ-bounded random

[10] In [227] we called the same quantity the "modulus of separated decomposability."

partition of $(S, d_{\mathfrak{M}})$. In other words,

$$\mathsf{SEP}^n(\mathfrak{M}, d_{\mathfrak{M}}) \stackrel{\mathrm{def}}{=} \sup_{\substack{S \subseteq \mathfrak{M} \\ |S| \leqslant n}} \mathsf{SEP}(S, d_{\mathfrak{M}}).$$

While the notions that we presented in Definition 63 are standard (see below for the history), it will be beneficial for us (e.g., for proving Theorem 29) to introduce the following terminology.

Definition 64 (Separation profile). Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space. We say that a metric $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ on $\mathfrak{M}$ is a *separation profile* of $(\mathfrak{M}, d_{\mathfrak{M}})$ if for every $\Delta > 0$ there exists a Δ-bounded random partition $\mathcal{P}_\Delta$ of $(\mathfrak{M}, d_{\mathfrak{M}})$ that is defined on some probability space $(\Omega_\Delta, \mathbf{Prob}_\Delta)$ such that

$$\forall x, y \in \mathfrak{M}, \quad \mathfrak{d}(x, y) \geqslant \sup_{\Delta \in (0,\infty)} \Delta \mathbf{Prob}_\Delta\big[\mathcal{P}_\Delta(x) \neq \mathcal{P}_\Delta(y)\big]. \tag{1.93}$$

So, the separation modulus of $(\mathfrak{M}, d_{\mathfrak{M}})$ is the infimum over those $\sigma > 0$ for which $\sigma d_{\mathfrak{M}}$ is a separation profile of $(\mathfrak{M}, d_{\mathfrak{M}})$. Definition 64 would make sense for functions $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ that need not be metrics on $\mathfrak{M}$, but we prefer to deal only with separation profiles of $(\mathfrak{M}, d_{\mathfrak{M}})$ that are metrics on $\mathfrak{M}$ so as to be able to discuss the Lipschitz condition with respect to them; observe that the right-hand side of (1.93) is a metric on $\mathfrak{M}$, so any such function is always at least (pointwise) a metric that is a separation profile of $(\mathfrak{M}, d_{\mathfrak{M}})$. If $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ is a separation profile of $(\mathfrak{M}, d_{\mathfrak{M}})$, then $\mathfrak{d}(x, y) \geqslant d_{\mathfrak{M}}(x, y)$ for all $x, y \in \mathfrak{M}$ because $\mathrm{diam}_{\mathfrak{M}}(\mathcal{P}_{d_{\mathfrak{M}}(x,y)-\varepsilon}(x)) \leqslant d_{\mathfrak{M}}(x, y) - \varepsilon < d_{\mathfrak{M}}(x, y)$ for any $0 < \varepsilon < d_{\mathfrak{M}}(x, y)$, so we necessarily have $y \notin \mathcal{P}_{d_{\mathfrak{M}}(x,y)-\varepsilon}(x)$ (deterministically) and therefore

$$\begin{aligned}\mathfrak{d}(x, y) &\geqslant (d_{\mathfrak{M}}(x, y) - \varepsilon)\mathbf{Prob}_{d_{\mathfrak{M}}(x,y)-\varepsilon}\big[\mathcal{P}_{d_{\mathfrak{M}}(x,y)-\varepsilon}(x) \neq \mathcal{P}_{d_{\mathfrak{M}}(x,y)-\varepsilon}(y)\big]\\ &= d_{\mathfrak{M}}(x, y) - \varepsilon.\end{aligned} \tag{1.94}$$

Definition 65 (Padded random partition and padding modulus). Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space. For $\delta, \mathfrak{p}, \Delta > 0$, a Δ-bounded random partition $\mathcal{P}$ of $(\mathfrak{M}, d_{\mathfrak{M}})$ is $(\mathfrak{p}, \delta)$-*padded* if

$$\forall x \in \mathfrak{M}, \quad \mathbf{Prob}\Big[B_{\mathfrak{M}}\Big(x, \frac{\Delta}{\mathfrak{p}}\Big) \subseteq \mathcal{P}(x)\Big] \geqslant \delta. \tag{1.95}$$

Denote by $\mathsf{PAD}_\delta(\mathfrak{M}, d_{\mathfrak{M}})$, or simply $\mathsf{PAD}_\delta(\mathfrak{M})$ if the metric is clear from the context, the infimum over those $\mathfrak{p} > 0$ such that for every $\Delta > 0$ there exists a $(\mathfrak{p}, \delta)$-padded Δ-bounded random partition $\mathcal{P}$ of $(\mathfrak{M}, d_{\mathfrak{M}})$. If no such $\mathfrak{p}$ exists, then write $\mathsf{PAD}_\delta(\mathfrak{M}, d_{\mathfrak{M}}) = \infty$. For every $n \in \mathbb{N}$, denote

$$\mathsf{PAD}^n_\delta(\mathfrak{M}, d_{\mathfrak{M}}) \stackrel{\mathrm{def}}{=} \sup_{\substack{S \subseteq \mathfrak{M} \\ |S| \leqslant n}} \mathsf{PAD}_\delta(S, d_{\mathfrak{M}}).$$

See Section 3 for a quick justification why the above definition of random partition implies that the events that appear in (1.92) and (1.95) are indeed **Prob**-measurable.

Qualitatively, condition (1.92) says that despite the fact that $\mathcal{P}$ decomposes $\mathfrak{M}$ into clusters of small diameter, any two nearby points are likely to belong to the same cluster. Condition (1.95) says that every point in $\mathfrak{M}$ is likely to be "well within" its cluster (its distance to the complement of its cluster is at least a definite proportion of the assumed upper bound on the diameter of that cluster). Both of these requirements express the (often nonintuitive) property that the "boundaries" that the random partition induces are "thin" in a certain distributional sense, despite the fact that each realization of the partition consists only of small diameter clusters that can sometimes be very jagged. Neither of the above two definitions implies the other, but it follows from [170] that if $\mathcal{P}$ is a $(\mathfrak{p}, \delta)$-padded Δ-bounded random partition of $(\mathfrak{M}, d_{\mathfrak{M}})$, then there exists a random partition $\mathcal{P}'$ of $(\mathfrak{M}, d_{\mathfrak{M}})$ that is (2Δ)-bounded and $(4\mathfrak{p}/\delta)$-separating.

Separating and padded random partitions were introduced in the articles [29, 30] of Bartal, which contained decisive algorithmic applications and influenced a flurry of subsequent works that obtained many more applications in several directions. Other works considered such partitions implicitly, with a variety of applications; see the works of Leighton–Rao [175], Awerbuch–Peleg [18], Linial–Saks [184], Alon–Karp–Peleg–West [4], Klein–Plotkin–Rao [156] and Rao [269]. The nomenclature of Definition 63 and Definition 65 comes from [124, 160, 170, 171, 173].

By [29], for every metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ and every integer $n \geqslant 2$, we have the bound $\mathsf{SEP}^n(\mathfrak{M}) \lesssim \log n$. It was observed by Gupta, Krauthgamer and Lee [124] that [29] also implicitly yields the padding bound $\mathsf{PAD}^n_{0.5}(\mathfrak{M}) \lesssim \log n$. It was proved in [29] that both of these estimates are sharp.

Random partitions of normed spaces were first studied by Peleg and Reshef [248] for applications to network routing and distributed computing. The aforementioned work [76] improved and generalized the bounds of [248], and influenced later works; see, e.g., [173], and the work [13] of Andoni and Indyk. Similar partitioning schemes appeared implicitly in earlier work [152] on algorithms for graph colorings based on semidefinite programming.

1.7.4 From separation to Lipschitz extension

As we already explained, the connection between random partitions and Lipschitz extension was found in [173]. Here we will use the following theorem to deduce Theorem 29. It implies in particular the bound

$$\mathsf{e}(\mathfrak{M}) \lesssim \mathsf{SEP}(\mathfrak{M}) \tag{1.96}$$

of [173] and its proof is an adaptation of the ideas of [173] to both the present setup (extension to a function that is Lipschitz with respect to a different metric) and our different measurability requirements from the random partitions; we stress, however, that even though we cannot apply [173] directly as a "black box," the geometric ideas that underly the proof of Theorem 66 are the same as those of [173].

Theorem 66. *Suppose that $\mathfrak{d}$ is a separation profile of a locally compact metric space $(\mathcal{M}, d_{\mathcal{M}})$. For every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and every subset $\mathcal{C} \subseteq \mathcal{M}$, if $f : \mathcal{C} \to \mathbf{Z}$ is 1-Lipschitz with respect to the metric $d_{\mathcal{M}}$, i.e., $\|f(x) - f(y)\|_{\mathbf{Z}} \leqslant d_{\mathcal{M}}(x, y)$ for every $x, y \in \mathcal{M}$, then there is $F : \mathcal{M} \to \mathbf{Z}$ that extends f and is $O(1)$-Lipschitz with respect to the metric $\mathfrak{d}$, i.e., $\|F(x) - F(y)\|_{\mathbf{Z}} \lesssim \mathfrak{d}(x, y)$ for every $x, y \in \mathcal{M}$.*

1.7.5 Bounds on the separation and padding moduli of normed spaces

To facilitate the ensuing discussion of upper and lower bounds on the separation and padding moduli of (subsets of) normed spaces, we will first record two of their rudimentary properties. Firstly, the following lemma formally expresses the aforementioned advantage of the definitions in Section 1.7.3 over those of [173], namely that the moduli $\mathsf{SEP}(\cdot)$ and $\mathsf{PAD}_\delta(\cdot)$ are bi-Lipschitz invariants; its straightforward proof appears in Section 3.

Lemma 67 (Bi-Lipschitz invariance of separation and padding moduli). *Let $(\mathcal{M}, d_{\mathcal{M}})$ be a complete metric space that admits a bi-Lipschitz embedding into a metric space $(\mathcal{N}, d_{\mathcal{N}})$. Then*

$$\mathsf{SEP}(\mathcal{M}, d_{\mathcal{M}}) \leqslant \mathsf{c}_{(\mathcal{N}, d_{\mathcal{N}})}(\mathcal{M}, d_{\mathcal{M}})\mathsf{SEP}(\mathcal{N}, d_{\mathcal{N}}) \tag{1.97}$$

and

$$\forall \delta \in (0,1), \quad \mathsf{PAD}_\delta(\mathcal{M}, d_{\mathcal{M}}) \leqslant \mathsf{c}_{(\mathcal{N}, d_{\mathcal{N}})}(\mathcal{M}, d_{\mathcal{M}})\mathsf{PAD}_\delta(\mathcal{N}, d_{\mathcal{N}}). \tag{1.98}$$

Secondly, we have the following tensorization property of the separation and padding moduli, whose simple proof appears in Section 3. For $s \in [1, \infty]$ and metric spaces $(\mathcal{M}_1, d_{\mathcal{M}_1}), (\mathcal{M}, d_{\mathcal{M}_2})$, the metric $d_{\mathcal{M}_1 \oplus_s \mathcal{M}_2} : \mathcal{M}_1 \times \mathcal{M}_2 \to [0, \infty)$ on the Cartesian product $\mathcal{M}_1 \times \mathcal{M}_2$ is defined by setting for every $(x_1, x_2), (y_1, y_2) \in \mathcal{M}_1 \times \mathcal{M}_2$,

$$d_{\mathcal{M}_1 \oplus_s \mathcal{M}_2}\big((x_1, x_2), (y_1, y_2)\big) \stackrel{\mathrm{def}}{=} \big(d_{\mathcal{M}}(x_1, y_1)^s + d_{\mathcal{N}}(x_2, y_2)^s\big)^{\frac{1}{s}}. \tag{1.99}$$

With the usual convention that when $s = \infty$ the right-hand side of (1.99) is equal to the maximum of $d_{\mathcal{M}}(x_1, y_1)$ and $d_{\mathcal{N}}(x_2, y_2)$. The metric space $(\mathcal{M}_1 \times \mathcal{M}_2, d_{\mathcal{M}_1 \oplus_s \mathcal{M}_2})$ is will be denoted $\mathcal{M}_1 \oplus_s \mathcal{M}_2$.

Lemma 68 (Tensorization of separation and padding moduli). *For any $s \in [1,\infty]$ and $\delta_1,\delta_2 \in (0,1)$, any two metric spaces $(\mathfrak{M}_1, d_{\mathfrak{M}_1})$ and $(\mathfrak{M}_2, d_{\mathfrak{M}_2})$ satisfy*

$$\mathsf{SEP}(\mathfrak{M}_1 \oplus_s \mathfrak{M}_2) \leqslant \mathsf{SEP}(\mathfrak{M}_1) + \mathsf{SEP}(\mathfrak{M}_2), \tag{1.100}$$

and

$$\mathsf{PAD}_{\delta_1\delta_2}(\mathfrak{M}_1 \oplus_s \mathfrak{M}_2) \leqslant \big(\mathsf{PAD}_{\delta_1}(\mathfrak{M}_1)^s + \mathsf{PAD}_{\delta_2}(\mathfrak{M}_2)^s\big)^{\frac{1}{s}}. \tag{1.101}$$

The following theorem shows that the bi-Lipschitz invariant $\mathsf{PAD}_\delta(\cdot)$ is not sufficiently sensitive to distinguish substantially between normed spaces, as its value is essentially independent of the norm.

Theorem 69. *For every $n \in \mathbb{N}$, every normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies*

$$\forall \delta \in (0,1), \qquad \frac{1}{1-\sqrt[n]{\delta}} \leqslant \frac{1}{2}\mathsf{PAD}_\delta(\mathbf{X}) \leqslant \frac{1+\sqrt[n]{\delta}}{1-\sqrt[n]{\delta}}. \tag{1.102}$$

Therefore, $\mathsf{PAD}_\delta(\mathbf{X}) \asymp \max\{1, \frac{\dim(\mathbf{X})}{\log(1/\delta)}\}$ for every finite dimensional normed space $\mathbf{X}$ and $\delta \in (0,1)$.

As we explained above, in the setting of Theorem 69 the fact that

$$\mathsf{PAD}_{\frac{1}{2}}(\mathbf{X}) = O(n)$$

is well known. We will prove the upper bound on $\mathsf{PAD}_\delta(\mathbf{X})$ that appears in (1.102), i.e., with sharp dependence on both n and δ, in Section 4.1. The fact that $\mathsf{PAD}_{0.5}(\mathbf{X})$ is at least a universal constant multiple of n was proved in the manuscript [170]. Because [170] is not intended for publication, we will prove the lower bound on $\mathsf{PAD}_\delta(\mathbf{X})$ that appears in (1.102) in Section 2.6, by following the reasoning of [170] while taking more care than we did in [170] in order to obtain sharp dependence on δ in addition to sharp dependence on n.

In contrast to Theorem 69, the separation modulus of a finite dimensional normed space can have different asymptotic dependencies on its dimension. For example, we have $\mathsf{SEP}(\ell_2^n) \asymp \sqrt{n}$ and $\mathsf{SEP}(\ell_1^n) \asymp n$ by [76]. Using Lemma 67, we see from this that every normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies the a priori bounds

$$\frac{n}{d_{\mathrm{BM}}(\ell_1^n, \mathbf{X})} \lesssim \mathsf{SEP}(\mathbf{X}) \lesssim d_{\mathrm{BM}}(\ell_2^n, \mathbf{X})\sqrt{n}, \tag{1.103}$$

which we already quoted in the above overview as (1.2).

Giannopoulos proved [105] that every n-dimensional normed space $\mathbf{X}$ satisfies $d_{\mathrm{BM}}(\ell_1^n, \mathbf{X}) \lesssim n^{5/6}$, so the first inequality in (1.103) implies that $\mathsf{SEP}(\mathbf{X}) \gtrsim \sqrt[6]{n}$. Alternatively, the fact that $\mathsf{SEP}(\mathbf{X}) \geqslant n^c$ for some universal constant $c > 0$ follows from by combining Theorem 1 with (1.96). Actually, we always have

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \sqrt{n}, \tag{1.104}$$

which coincides with the first half of (1.7). Observe that (1.104) cannot follow from a "vanilla" application of the first inequality in (1.103) by Szarek's work [295]. In fact, the first inequality of (1.103) must sometimes yield a worse power type dependence on n than in (1.104), because Tikhomirov proved in [302] that there is a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ that satisfies $d_{\mathrm{BM}}(\ell_1^n, \mathbf{X}) \geqslant n^a$ for some universal constant $a > 1/2$.

Nevertheless, we can prove (1.104) by the following a "hereditary" application of (1.103). Bourgain–Szarek [51] and independently Ball (see [51, Remark 7], [296, Remark 7], [305, p. 138]) proved (relying on the Bourgain–Tzafriri restricted invertibility principle [52]) that there is $m \in \{1, \ldots, n\}$ with $m \asymp n$ such that $\mathsf{c}_{\mathbf{X}}(\ell_1^m) \lesssim \sqrt{n}$ (in fact, by [51] any $2n$-dimensional normed space has Banach–Mazur distance $O(\sqrt{n})$ from $\ell_1^n \oplus \ell_2^n$). Hence, $\mathsf{SEP}(\mathbf{X}) \gtrsim \mathsf{SEP}(\ell_1^m)/\mathsf{c}_{\mathbf{X}}(\ell_1^m) \asymp m/\mathsf{c}_{\mathbf{X}}(\ell_1^m) \gtrsim \sqrt{n}$, by (1.97).

The second half of (1.7) is the following lower bound on $\mathsf{SEP}(\mathbf{X})$ in terms of the type 2 constant of $\mathbf{X}$:

$$\mathsf{SEP}(\mathbf{X}) \gtrsim T_2(\mathbf{X})^2. \tag{1.105}$$

We will prove (1.105) in Section 2.2 using Talagrand's refinement [298] of Elton's theorem [92], by the same hereditary use of (1.103), namely showing that there is $m \in \{1, \ldots, n\}$ for which $m/\mathsf{c}_{\mathbf{X}}(\ell_1^m) \gtrsim T_2(\mathbf{X})^2$.

Remark 70. It is impossible to improve (1.7) for all the values of the relevant parameters, as seen by considering $\mathbf{X} = \ell_2^{n-m} \oplus_2 \ell_1^m$ for each $m \in \{1, \ldots, n\}$. Indeed, since in this case $T_2(\mathbf{X}) \asymp \sqrt{m}$,

$$\begin{aligned}\mathsf{SEP}(\mathbf{X}) &\overset{(1.100)}{\leqslant} \mathsf{SEP}\big(\ell_2^{n-m}\big) + \mathsf{SEP}\big(\ell_1^m\big)\\ &\asymp \sqrt{n-m} + m \asymp \sqrt{n} + k \asymp \max\Big\{\sqrt{\dim(\mathbf{X})}, T_2(\mathbf{X})^2\Big\}.\end{aligned}$$

Thanks to (1.71), the following theorem is a restatement of the lower bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3.

Theorem 71. *For every $n \in \mathbb{N}$, any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ satisfies*

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \mathrm{evr}(\mathbf{X})\sqrt{n}.$$

As $\mathrm{evr}(\mathbf{X}) \geqslant 1$ (by definition), Theorem 71 implies (1.104), via a proof that differs from the above reasoning. Also, Theorem 71 is stronger than the first inequality in (1.103) because $\mathrm{evr}(\ell_1^n) \asymp \sqrt{n}$, and hence

$$\mathrm{evr}(\mathbf{X})\sqrt{n} \geqslant \frac{\mathrm{evr}(\ell_1^n)}{d_{\mathrm{BM}}(\ell_1^n, \mathbf{X})}\sqrt{n} \asymp \frac{n}{d_{\mathrm{BM}}(\ell_1^n, \mathbf{X})}.$$

We will prove Theorem 71 in Section 2.5 by adapting to the setting of general normed spaces the strategy that was used in [76] to treat ℓ_1^n. The volumetric lower bound on

$\mathsf{SEP}(\mathbf{X})$ of Theorem 71 is typically quite easy to use and it often leads to estimates that are better than the first inequality in (1.103).

For example, by [285, Proposition 2.2] the Schatten–von Neumann trace class S_p^n satisfies

$$\forall p \geqslant 1, \quad \mathrm{evr}\big(\mathsf{S}_p^n\big) \asymp n^{\max\{\frac{1}{p}-\frac{1}{2},0\}}. \tag{1.106}$$

By substituting (1.106) into Theorem 71 we get that

$$\forall 1 \leqslant p \leqslant 2, \quad \mathsf{SEP}(\mathsf{S}_p^n) \gtrsim n^{\frac{1}{p}-\frac{1}{2}}\sqrt{\dim\big(\mathsf{S}_p^n\big)} \asymp n^{\frac{1}{p}+\frac{1}{2}}. \tag{1.107}$$

An upper bound that matches (1.107) is a consequence of the second inequality in (1.103) as follows

$$\mathsf{SEP}\big(\mathsf{S}_p^n\big) \lesssim d_{\mathrm{BM}}\big(\mathsf{S}_p^n, \ell_2^{n^2}\big)\sqrt{\dim\big(\mathsf{S}_p^n\big)} = d_{\mathrm{BM}}\big(\mathsf{S}_p^n, \mathsf{S}_2^n\big)n = n^{\frac{1}{p}+\frac{1}{2}}.$$

We therefore have

$$\forall 1 \leqslant p \leqslant 2, \quad \mathsf{SEP}\big(\mathsf{S}_p^n\big) \asymp n^{\frac{1}{p}+\frac{1}{2}}.$$

At the same time, the first inequality in (1.103) does not imply (1.107) since by a theorem of Davis (which was published only in the monograph [305]; see Theorem 41.10 there), for every $1 \leqslant p \leqslant 2$ we have

$$d_{\mathrm{BM}}\big(\ell_1^{n^2}, \mathsf{S}_p^n\big) \asymp n. \tag{1.108}$$

So, the first inequality in (1.103) only implies the weaker bound $\mathsf{SEP}(\mathsf{S}_p^n) \gtrsim n$. Of course, this rules out a "vanilla" use of (1.103) and a hereditary application of (1.103) as we did above could conceivably lead to (1.107), i.e., there could be $m \in \{1, \dots, n\}$ such that $m/\mathsf{c}_{\mathsf{S}_p^n}(\ell_1^m)$ is at least the right-hand side of (1.107). However, this possibility seems to be unlikely, as it would mean that the following conjecture has a negative answer, which would entail finding a remarkable (and likely valuable elsewhere) subspace of S_p^n.

Conjecture 72. Fix $1 \leqslant p \leqslant 2$ and $0 < \delta \leqslant 1$. If $n, m \in \mathbb{N}$ satisfy $m \geqslant \delta n^2$, then

$$d_{\mathrm{BM}}(\ell_1^m, \mathbf{X}) \gtrsim_{p,\delta} n$$

for every m-dimensional subspace $\mathbf{X}$ of S_p^n.

Thus, (1.108) is the case $\delta = 1$ of Conjecture 72, which asserts that the same asymptotic lower bound persists if we consider subspaces of S_p^n of proportional dimension rather than S_p^n itself. Conjecture 72 is attractive in its own right, but it also implies that (1.107) does not follow from a hereditary application of the first inequality in (1.103). To see this, suppose for contradiction that there were $m \in \{1, \dots, n\}$ such that

$$\frac{m}{\mathsf{c}_{\mathsf{S}_p^n}(\ell_1^m)} \gtrsim_p n^{\frac{1}{p}+\frac{1}{2}}. \tag{1.109}$$

By Rademacher's differentiation theorem [267] there is an m-dimensional subspace $\mathbf{X}$ of S_p^n satisfying

$$\mathsf{c}_{\mathsf{S}_p^n}(\ell_1^m) = d_{\mathrm{BM}}(\ell_1^m, \mathbf{X}) \gtrsim \frac{d_{\mathrm{BM}}(\ell_1^m, \ell_2^m)}{d_{\mathrm{BM}}(\mathsf{S}_p^n, \mathsf{S}_2^n)} = \frac{\sqrt{m}}{n^{\frac{1}{p}-\frac{1}{2}}}. \tag{1.110}$$

By contrasting (1.110) with (1.109) we deduce that necessarily $m \gtrsim_p n^2$, so an application of Conjecture 72 gives $m/\mathsf{c}_{\mathsf{S}_p^n}(\ell_1^m) \lesssim_p n$, which contradicts (1.109) since $p<2$.

Remark 73. The Löwner ellipsoid of $\ell_\infty^n(\ell_1^n)$ is $\sqrt{n}B_{\ell_2^n(\ell_2^n)}$, and $B_{\ell_\infty^n(\ell_1^n)} = (B_{\ell_1^n})^n$. Consequently,

$$\mathrm{evr}\big(\ell_\infty^n(\ell_1^n)\big)n = n\left(\frac{(\pi n)^{\frac{n^2}{2}}/\Gamma\big(\frac{n^2}{2}+1\big)}{2^{n^2}/(n!)^n}\right)^{\frac{1}{n^2}} \asymp n^{\frac{3}{2}}.$$

Therefore, Theorem 71 gives

$$\mathsf{SEP}\big(\ell_\infty^n(\ell_1^n)\big) \gtrsim n^{\frac{3}{2}}. \tag{1.111}$$

We will soon see that (1.111) is optimal, though unlike the above discussion for S_p^n when $1 \leqslant p \leqslant 2$, this does not follow from the second inequality in (1.103) because by [163],

$$d_{\mathrm{BM}}\big(\ell_2^{n^2}, \ell_\infty^n(\ell_1^n)\big) \asymp d_{\mathrm{BM}}\big(\ell_1^{n^2}, \ell_\infty^n(\ell_1^n)\big) \asymp n. \tag{1.112}$$

(1.112) also shows that (1.111) does not follow from the first inequality in (1.103). It seems that the method used in [163] to prove (1.112) is insufficient for proving that (1.111) does not follow from a hereditary application of the first inequality in (1.103). Analogously to Conjecture 72, we conjecture that this is impossible, which is a classical-sounding question about Banach–Mazur distances of independent interest.

Before passing to a description of our upper bounds on the separation modulus, we formulate the following corollary of Theorem 71 on the separation modulus of norms whose unit ball is a polytope; it restates the lower bound (1.6) and establishes its optimality.

Theorem 74. *Fix $n \in \mathbb{N}$ and a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. Suppose that $B_{\mathbf{X}}$ is a polytope that has exactly ρn vertices (note that necessarily $\rho \geqslant 2$, since $B_{\mathbf{X}}$ is origin-symmetric). Then*

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \frac{n}{\sqrt{\log \rho}}. \tag{1.113}$$

Moreover, this bound cannot be improved in general.

As an example of a consequence of Theorem 74, let

$$\mathbf{G} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{G}})$$

be a Gluskin space [111], i.e., it is a certain *random* norm on $\mathbb{R}^n$ whose unit ball has $O(n)$ vertices; see the survey [196] for extensive information about this important construction and its variants. The expected Banach–Mazur distance between two independent copies of $\mathbf{G}$ is at least cn for some universal constant $c > 0$, so the expected Banach–Mazur distance between $\mathbf{G}$ and ℓ_1^n is at least $\sqrt{cn}$. Thus, the first inequality in (1.103) only shows that $\mathsf{SEP}(\mathbf{G}) \gtrsim \sqrt{n}$ in expectation, while Theorem 74 shows that $\mathsf{SEP}(\mathbf{G}) \gtrsim n/\sqrt{\log n}$. It would be interesting to determine the growth rate of $\mathbb{E}[\mathsf{SEP}(\mathbf{G})]$. In particular, can it be that $\mathbb{E}[\mathsf{SEP}(\mathbf{G})] \gtrsim n$?

Proof of Theorem 74. By applying a linear isometry of $\mathbf{X}$ we may assume that $B_{\ell_2^n}$ is the Löwner ellipsoid of $B_{\mathbf{X}}$. Since $B_{\mathbf{X}}$ is a polytope with ρn vertices that is contained in $B_{\ell_2^n}$, we have

$$\sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})} \lesssim \frac{\sqrt{\log \rho}}{n}$$

by a result of Maurey [258] (see also [25, 28, 48, 72, 73, 112, 164] and the expository treatments in [24, 55]). Hence, $\mathrm{evr}(\mathbf{X}) \gtrsim \sqrt{n/\log \rho}$, so (1.113) follows from Theorem 71.

Consider the following (dual of an) example of Figiel and Johnson [98]. Fix $m \in \mathbb{N}$. Let $\mathbf{Z} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{Z}})$ be a normed space with $d_{\mathrm{BM}}(\ell_2^m, \mathbf{Z}) \lesssim 1$ such that $B_{\mathbf{Z}}$ is a polytope of $e^{O(m)}$ vertices; e.g., $B_{\mathbf{Z}}$ can be taken to be the convex hull of a net of S^{m-1}. For $k \in \mathbb{N}$, let $\mathbf{X} = \ell_1^k(\mathbf{Z})$. So, $\dim(\mathbf{X}) = km$ and $B_{\mathbf{X}}$ is a polytope of $2ke^{O(m)}$ vertices. Thus (1.113) becomes $\mathsf{SEP}(\mathbf{X}) \gtrsim k\sqrt{m}$. At the same time, since $d_{\mathrm{BM}}(\ell_2^m, \mathbf{Z}) \lesssim 1$ we have $d_{\mathrm{BM}}(\ell_2^{km}, \mathbf{X}) \lesssim \sqrt{k}$, so by (1.103) in fact $\mathsf{SEP}(\mathbf{X}) \lesssim \sqrt{k} \cdot \sqrt{km} = k\sqrt{m}$, i.e., (1.113) is sharp in this case. ■

Theorem 29 follows from Theorem 66 thanks to the following randomized partitioning theorem.

Theorem 75. *For every $n \in \mathbb{N}$ and every normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$, the metric $\mathfrak{d}$ that is defined by*

$$\forall x, y \in \mathbb{R}^n, \quad \mathfrak{d}(x,y) = \frac{4\|x-y\|_{\Pi^*\mathbf{X}}}{\mathrm{vol}_n(B_{\mathbf{X}})}$$

is a separation profile for $\mathbf{X}$.

To illustrate Theorem 75, fix $1 \leqslant p \leqslant \infty$ and apply it when $\mathbf{X}$ is the space $\mathbf{Y}_p^n$ of Theorem 24. By using Theorem 75 we see that for every $\Delta > 0$ there is a random partition $\mathcal{P}$ of $\mathbb{R}^n$ with the following properties.

(1) For every $x \in \mathbb{R}^n$ we have $\mathrm{diam}_{\ell_p^n}(\mathcal{P}(x)) \leqslant \Delta$.

(2) For every $x, y \in \mathbb{R}^n$ we have

$$\mathbf{Prob}\big[\mathcal{P}(x) \neq \mathcal{P}(y)\big] \lesssim \frac{\|x-y\|_{\Pi^*\mathbf{Y}_p^n}}{\mathrm{vol}_n(B_{\mathbf{Y}_p^n})} \overset{(1.30)\wedge(1.39)}{\lesssim} \frac{n^{\frac{1}{p}}}{\Delta}\|x-y\|_{\ell_2^n}. \quad (1.114)$$

In comparison to the $O(\sqrt{n})$-separating partition of ℓ_2^n from [76], when $p < 2$ the above random partition has smaller clusters in the sense that their diameter in the ℓ_p^n metric is at most Δ, which is more stringent than the requirement that their Euclidean diameter is at most Δ. This improved control on the size of the clusters comes at the cost that in the probabilistic separation requirement (1.114) the quantity that multiplies the Euclidean distance increases from $O(\sqrt{n})$ to $O(n^{1/p})$. When $p > 2$ this tradeoff is reversed, i.e., we get an asymptotic improvement in the separation guarantee (1.114) at the cost of requiring less from the cluster size, namely the diameter of each cluster is now guaranteed to be small in the ℓ_p^n metric rather than the more stringent requirement that it is small in the Euclidean metric.

Theorem 76 below follows from Theorem 75 the same way we deduced Theorem 21 from Theorem 29.

Theorem 76. *Fix $n \in \mathbb{N}$ and two normed spaces $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}}), \mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$. Every closed $\mathcal{C} \subseteq \mathbb{R}^n$ satisfies*

$$\begin{aligned}&\mathsf{SEP}(\mathcal{C}_{\mathbf{X}})\\ &\leqslant 4\Bigg(\sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \frac{\|x-y\|_{\mathbf{X}}}{\|x-y\|_{\mathbf{Y}}}\Bigg) \sup_{\substack{x,y\in\mathcal{C}\\ x\neq y}} \Bigg(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp}(B_{\mathbf{Y}})\big)}{\mathrm{vol}_n(B_{\mathbf{Y}})} \cdot \frac{\|x-y\|_{\ell_2^n}}{\|x-y\|_{\mathbf{X}}}\Bigg). \quad (1.115)\end{aligned}$$

Proof of Theorem 76 assuming Theorem 75. Let M, M' be as in (1.51) and (1.52). By Theorem 75 applied to $\mathbf{Y}$, for every $\Delta > 0$ there is a random partition $\mathcal{P}$ of $\mathbb{R}^n$ that is (Δ/M)-bounded with respect to $\mathbf{Y}$, i.e.,

$$\frac{\mathrm{diam}_{\mathbf{X}}\big(\mathcal{P}(x)\big)}{M} \overset{(1.51)}{\leqslant} \mathrm{diam}_{\mathbf{Y}}\big(\mathcal{P}(x)\big) \leqslant \frac{\Delta}{M}$$

for every $x \in \mathbb{R}^n$, and also, recalling Definition 64, for every distinct $x, y \in \mathbb{R}^n$ we have

$$\begin{aligned}\frac{\Delta}{M}\mathbf{Prob}\big[\mathcal{P}(x) \neq \mathcal{P}(y)\big] &\leqslant \frac{4\|x-y\|_{\Pi^*\mathbf{Y}}}{\mathrm{vol}_n(B_{\mathbf{Y}})}\\ &\overset{(1.30)}{=} \frac{4\,\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{(x-y)^\perp}(B_{\mathbf{Y}})\big)\|x-y\|_{\ell_2^n}}{\mathrm{vol}_n(B_{\mathbf{Y}})}\\ &\overset{(1.52)}{\leqslant} 4M'\|x-y\|_{\mathbf{X}}.\end{aligned}$$

■

The special case $\mathcal{C} = \mathbb{R}^n$ of Theorem 76 coincides (with an explicitly stated constant factor) with the upper bound on $\mathsf{SEP}(\mathbf{X})$ in Theorem 3, since under the normalization $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ we have

$$\begin{aligned}\mathsf{SEP}(\mathbf{X}) &\overset{(1.30)\wedge(1.115)}{\leqslant} 4\frac{\sup_{z\in\partial B_{\mathbf{X}}}\|z\|_{\Pi^*\mathbf{Y}}}{\mathrm{vol}_n(B_{\mathbf{Y}})}\\ &= 4\frac{\|\mathsf{Id}_n\|_{\mathbf{X}\to\Pi^*\mathbf{Y}}}{\mathrm{vol}_n(B_{\mathbf{Y}})} = 4\frac{\|\mathsf{Id}_n\|_{\Pi\mathbf{Y}\to\mathbf{X}^*}}{\mathrm{vol}_n(B_{\mathbf{Y}})} = 2\frac{\mathrm{diam}_{\mathbf{X}^*}(\Pi B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})}.\end{aligned}$$

Also, Theorem 76 is stronger than the second inequality in (1.103) because by applying a linear isometry of $\mathbf{X}$ we may assume without loss of generality that $\|x\|_{\mathbf{X}} \leqslant \|x\|_{\ell_2^n} \leqslant d_{\mathrm{BM}}(\ell_2^n,\mathbf{X})\|x\|_{\mathbf{X}}$ for all $x\in\mathbb{R}^n$, in which case the special case $\mathcal{C}=\mathbb{R}^n$ and $\mathbf{Y}=\ell_2^n$ of (1.115) implies that

$$\begin{aligned}\mathsf{SEP}(\mathbf{X}) &\leqslant \frac{4\,\mathrm{vol}_{n-1}\big(B_{\ell_2^{n-1}}\big)}{\mathrm{vol}_n\big(B_{\ell_2^n}\big)} d_{\mathrm{BM}}(\ell_2^n,\mathbf{X}) = \frac{4\pi^{\frac{n-1}{2}}\Gamma\big(\frac{n}{2}+1\big)}{\pi^{\frac{n}{2}}\Gamma\big(\frac{n-1}{2}+1\big)} d_{\mathrm{BM}}(\ell_2^n,\mathbf{X})\\ &= \frac{2^{\frac{3}{2}}+o(1)}{\sqrt{\pi}} d_{\mathrm{BM}}(\ell_2^n,\mathbf{X})\sqrt{n}.\end{aligned}$$

The right-hand side of (1.115) coincides (up to a universal constant factor) with the right-hand side of (1.28), so all of the upper bounds for the Lipschitz extension modulus that we derived in the previous sections from Theorem 21 hold for the separation modulus, by Theorem 76. For the separation modulus, we get several lower bounds from Theorem 71 that either provably match our upper bounds up to lower order factors, or match them assuming our conjectural isomorphic reverse isoperimetry. We will next spell out some of those consequences on randomized clustering of high dimensional norms.

Theorem 77. *For every $p \geqslant 1$, $n\in\mathbb{N}$ and $k,r\in\{1,\ldots,n\}$ we have*

$$\mathsf{SEP}\big((\ell_p^n)_{\leqslant k}\big) \asymp k^{\max\{\frac{1}{p},\frac{1}{2}\}} \tag{1.116}$$

and

$$\begin{aligned} r^{\max\{\frac{1}{p},\frac{1}{2}\}}\sqrt{n} &\lesssim \mathsf{SEP}\big((\mathsf{S}_p^n)_{\leqslant r}\big)\\ &\lesssim r^{\max\{\frac{1}{p},\frac{1}{2}\}}\sqrt{n}\cdot\begin{cases}\sqrt{\max\{\log(\frac{n}{r}),p\}} & \text{if } p\leqslant \log r,\\ \sqrt{\log n} & \text{if } p\geqslant \log r.\end{cases}\end{aligned} \tag{1.117}$$

Moreover, if Conjecture 49 holds for $\mathbf{X}=\mathsf{S}_p^n$, then in fact

$$\mathsf{SEP}\big((\mathsf{S}_p^n)_{\leqslant r}\big) \asymp r^{\max\{\frac{1}{p},\frac{1}{2}\}}\sqrt{n}.$$

Proof. The deduction of the upper bounds on the separation modulus that appear in (1.116) and (1.117) from Theorem 76 are identical, respectively, to the ways we deduced Theorem 20 and (1.45) from Theorem 21.

For the first inequality in (1.116), since $(\ell_p^n)_{\leqslant k}$ contains an isometric copy of ℓ_p^k, we have

$$\mathsf{SEP}\big((\ell_p^n)_{\leqslant k}\big) \geqslant \mathsf{SEP}\big(\ell_p^k\big) \gtrsim \frac{k}{d_{\mathrm{BM}}\big(\ell_p^k,\ell_1^k\big)} \overset{(1.103)}{\asymp} \frac{k}{k^{\max\{1-\frac{1}{p},\frac{1}{2}\}}} = k^{\min\{\frac{1}{p},\frac{1}{2}\}},$$

where the asymptotic evaluation of $d_{\mathrm{BM}}(\ell_p^k,\ell_q^k)$ for all $p,q\geqslant 1$ is due Gurariĭ, Kadec' and Macaev [125].

For the first inequality in (1.117), use the fact that $(\mathsf{S}_p^n)_{\leqslant r}$ contains an isometric copy of $\mathsf{S}_p^{r\times n}$, which is the Schatten–von Neumann trace class on the r-by-n real matrices $\mathsf{M}_{r\times n}(\mathbb{R})$, whose norm is given by

$$\forall A\in M_{r\times n}(\mathbb{R}),\qquad \|A\|_{\mathsf{S}_p^{r\times n}} = \Big(\mathbf{Tr}\big((AA^*)^{\frac{p}{2}}\big)\Big)^{\frac{1}{p}}. \tag{1.118}$$

We then have the following rectangular version of (1.106) whose derivation is explained in Remark 171:

$$\mathrm{evr}\big(\mathsf{S}_p^{r\times n}\big) \asymp r^{\max\{\frac{1}{p}-\frac{1}{2},0\}}. \tag{1.119}$$

The desired lower bound on $\mathsf{SEP}((\mathsf{S}_p^n)_{\leqslant r})$ is now an application of Theorem 71. ■

Remark 78. Theorem 3.3 in [76] asserts that $\mathsf{SEP}(\ell_p^n)\asymp n^{\max\{1/p,1-1/p\}}$ for every $p\geqslant 1$. Therefore, when $p>2$ it was previously thought that $\mathsf{SEP}(\ell_p^n)\asymp n^{1-1/p}$, which contradicts the case $k=n$ of (1.116). While [76] provides a complete and correct proof that $\mathsf{SEP}(\ell_p^n)\asymp n^{1/p}$ when $1\leqslant p\leqslant 2$, in the range $p>2$ the assertion $\mathsf{SEP}(\ell_p^n)\asymp n^{1-1/p}$ in [76] is justified through the use of a result from reference [14] in [76], which is cited there as a "personal communication" with P. Indyk (dated April 1998). This reference was never published. After discovering Theorem 77, we confirmed with Indyk that his aforementioned personal communication with the authors of [76] contained a gap.

Corollary 79. *Conjecture 49 implies Conjecture 6. Namely, if Conjecture 49 holds for a canonically positioned normed space $\mathbf{X}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{X}})$, then*

$$\mathsf{SEP}(\mathbf{X}) \asymp \mathrm{evr}(\mathbf{X})\sqrt{n} \asymp \mathrm{vr}(\mathbf{X}^*)\sqrt{n}. \tag{1.120}$$

In particular, if $\mathbf{X}$ satisfies the assumptions of Lemma 53 (e.g., if $\mathbf{X}$ is symmetric), then (1.120) holds. Furthermore, if $\mathbf{E}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{E}})$ is a symmetric normed space, then $\mathsf{SEP}(\mathsf{S}_{\mathbf{E}})=\mathrm{evr}(\mathbf{E})n^{1+o(1)}$. More precisely,

$$\mathrm{evr}(\mathbf{E})n \lesssim \mathsf{SEP}(\mathsf{S}_{\mathbf{E}}) \lesssim \mathrm{evr}(\mathbf{E})n\sqrt{\log n}.$$

Proof. The lower bound on $\mathsf{SEP}(\mathbf{X})$ in (1.120) is Theorem 71 (thus, it requires neither Conjecture 49 nor $\mathbf{X}$ being canonically positioned). The matching upper bound on

$\mathsf{SEP}(\mathbf{X})$ in (1.120) follows from Corollary 51 and the fact that by Theorem 76 the separation modulus of any (not necessarily canonically positioned) normed space

$$\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$$

is bounded from above by the right-hand side of (1.54). The rest of the assertions of Corollary 79 follow from Lemma 53 and Lemma 54. ■

By incorporating Proposition 61 into the same reasoning as in the justification of Corollary 79, we also deduce the following stronger version of Theorem 12.

Theorem 80. *If* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ *is a canonically positioned normed space, then*

$$\operatorname{evr}(\mathbf{X})\sqrt{n} \lesssim \mathsf{SEP}(\mathbf{X}) \lesssim K(\mathbf{X}) \operatorname{evr}(\mathbf{X})\sqrt{n} \overset{(1.87)}{\lesssim} \operatorname{evr}(\mathbf{X})\sqrt{n}\log n.$$

Section 6.3 contains volume ratio computations that show how Corollary 79 and Theorem 80 imply Corollary 4, as well as the conjectural (i.e., conditional on the validity of Conjecture 49 for the respective spaces) asymptotic evaluations (1.14) and (1.15), and several further results of this type. Most of the volume ratio computations in Section 6.3 rely on the available literature (notably Schütt's work [285]), with a few new twists that are perhaps of independent geometric/probabilisitic interest (e.g., Lemma 173).

1.7.6 Dimension reduction

Fix $n \in \mathbb{N}$ and a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$. Recall that in Definition 63 we denoted by $\mathsf{SEP}^n(\mathfrak{M}, d_{\mathfrak{M}})$ the supremum over all the separation moduli of subsets of $\mathfrak{M}$ of size at most n. In [76] it was shown that $\mathsf{SEP}^n(\ell_2) \lesssim \sqrt{\log n}$. Indeed, this follows from the Johnson–Lindenstrauss dimension reduction lemma [138], which asserts that any n-point subset of ℓ_2 can be embedded with $O(1)$ distortion into ℓ_2^m with $m \lesssim \log n$, combined with the proof in [76] that $\mathsf{SEP}(\ell_2^m) \lesssim \sqrt{m}$.

One might expect that the optimal bounds that we know for $\mathsf{SEP}(\ell_p^n)$ in the entire range $p \in (1, \infty)$ also translate to improved bounds on $\mathsf{SEP}^n(\ell_p)$. The term "improved" is used here to mean any upper bound of the form $o_p(\log n)$ as $n \to \infty$, since the benchmark general result is the aforementioned upper bound

$$\mathsf{SEP}^n(\mathfrak{M}, d_{\mathfrak{M}}) \lesssim \log n$$

from [29], which holds for any n-point metric space $(\mathfrak{M}, d_{\mathfrak{M}})$. This bound is sharp in general [29], so (because every n-point metric space embeds isometrically into ℓ_∞^n) we cannot hope to get a better bound on $\mathsf{SEP}^n(\ell_\infty)$ despite the fact that we obtained here an improved upper bound on $\mathsf{SEP}(\ell_\infty^n)$.

The obstacle is that when $p \in [1, \infty] \smallsetminus \{2\}$ no bi-Lipschitz dimension reduction result is known for finite subsets of ℓ_p, and poly-logarithmic bi-Lipschitz dimension reduction is impossible if $p \in \{1, \infty\}$; the case $p = \infty$ is due to Matoušek [200] (see also [228, 230]) and the case $p = 1$ is due to Brinkman and Charikar [58] (see also [172, 232, 240, 270]). When $p \in [1, \infty] \smallsetminus \{1, 2, \infty\}$ remarkably nothing is known, i.e., neither positive results nor impossibility results are available for bi-Lipschitz dimension reduction, and it is a major open problem to make any progress in this setting; see [229] for more on this area. Despite this obstacle, we have the following theorem that treats the range $p \in [1, 2]$.

Theorem 81. *For every $p \in (1, 2]$ and $n \in \mathbb{N}$ we have*

$$(\log n)^{\frac{1}{p}} \lesssim \mathsf{SEP}^n(\ell_p) \lesssim \frac{(\log n)^{\frac{1}{p}}}{p-1}.$$

The lower bound on $\mathsf{SEP}^n(\ell_p)$ of Theorem 81 can be deduced from [76]; see Section 4.2 for the details. An upper bound of $\mathsf{SEP}^n(\ell_p) \lesssim_p (\log n)^{1/p}$ was obtained when $p \in (1, 2]$ in the manuscript [170]. As [170] is not intended for publication, a proof of the upper bound on $\mathsf{SEP}^n(\ell_p)$ that is stated in Theorem 81 is included in Section 4.2, where we perform the argument with more care than the way we initially did it in [170], so as to obtain the best dependence on p that is achievable by this approach. Nevertheless, we conjecture that the dependence on p in Theorem 81 could be removed altogether, though this would likely require a substantially new idea.

Conjecture 82. The dependence on p in Theorem 81 can be improved to

$$\mathsf{SEP}^n(\ell_p) \lesssim (\log n)^{\frac{1}{p}}.$$

So, if $p \leqslant 1 + c(\log\log\log n)/\log\log n$ for some universal constant $c > 0$, then Theorem 81 does not improve asymptotically over $\mathsf{SEP}^n(\ell_p) \lesssim \log n$, while Conjecture 82 would imply that $\mathsf{SEP}^n(\ell_p) = o(\log n)$ if and only if

$$\lim_{n\to\infty} (p-1)\log\log n = \infty.$$

For fixed $p \in (2, \infty)$, at present we do not see how to obtain an upper bound on $\mathsf{SEP}^n(\ell_p)$ of the form $o_p(\log n)$ as $n \to \infty$. We state this separately as an interesting and challenging open question.

Question 83. Is it true that for every $n \in \mathbb{N}$ and $p \in (2, \infty)$ we have

$$\lim_{n\to\infty} \frac{\mathsf{SEP}^n(\ell_p)}{\log n} = 0?$$

More ambitiously, is it true that $\mathsf{SEP}^n(\ell_p) \lesssim_p \sqrt{\log n}$?

Note that $\mathsf{SEP}^n(\mathbf{X}) \gtrsim \sqrt{\log n}$ for any infinite-dimensional normed space $\mathbf{X}$, because by Dvoretzky's theorem [90] we have $\mathsf{c}_{\mathbf{X}}(\ell_2^m) = 1$ for every $m \in \mathbb{N}$, and therefore $\mathsf{SEP}^n(\mathbf{X}) \geqslant \mathsf{SEP}^n(\ell_2) \asymp \sqrt{\log n}$.

1.8 Consequences in the linear theory

Even though the purpose of the present article was to investigate the nonlinear invariants $\mathsf{e}(\cdot)$ and $\mathsf{SEP}(\cdot)$, by relating them to volumetric quantities and other linear invariants of Banach spaces (such as type and cotype), we arrive at consequences that have nothing to do with nonlinear issues. In this section, we will give a flavor of such consequences, though we will not be exhaustive since it would be more natural to pursue them separately for their own right in future work.

Denote the Minkowski functional of an origin-symmetric convex body $K \subseteq \mathbb{R}^n$ by $\|\cdot\|_K$, i.e., it is the norm on $\mathbb{R}^n$ whose unit ball is equal to K. The following theorem coincides with the second inequality in (1.1) upon a straightforward application of duality as we did in (1.31); this formulation is intended to highlight how we are bounding a convex-geometric quantity by a bi-Lipschitz invariant.

Theorem 84 (Nonsandwiching between a convex body and its polar projection body). *Fix $n \in \mathbb{N}$ and $\alpha, \beta \in (0,\infty)$. Let $K, L \subseteq \mathbb{R}^n$ be origin-symmetric convex bodies with* $\mathrm{vol}_n(L) = 1$. *Suppose that*

$$\alpha L \subseteq K \subseteq \beta \Pi^* L. \tag{1.121}$$

Then,

$$\frac{\beta}{\alpha} \gtrsim \mathsf{SEP}\big(\mathbb{R}^n, \|\cdot\|_K\big). \tag{1.122}$$

Since the separation modulus of a metric space is at least the separation modulus of any of its subsets, by combining (1.122) with the first inequality in (1.1) we see that the sandwiching hypothesis (1.121) implies the following purely volumetric consequence for every linear subspace $\mathbf{V} \subseteq \mathbb{R}^n$:

$$\frac{\beta}{\alpha} \gtrsim \mathrm{evr}\big(K \cap \mathbf{V}\big)\sqrt{n} \asymp \mathrm{vr}\big(\mathsf{Proj}_{\mathbf{V}} K^\circ\big)\sqrt{n}. \tag{1.123}$$

In particular, using $\mathrm{evr}(\ell_1^n) \asymp \sqrt{n}$, we record separately the following special case of (1.123).

Corollary 85 (Nonsandwiching of the cross-polytope). *Fix $n \in \mathbb{N}$ and $\alpha, \beta \in (0,\infty)$. If $L \subseteq \mathbb{R}^n$ is a convex body of volume 1 that satisfies $\alpha L \subseteq B_{\ell_1^n} \subseteq \beta \Pi^* L$, then necessarily $\beta/\alpha \gtrsim n$.*

The geometric meaning of Theorem 84 when $L = K$ is spelled out in the following corollary.

Corollary 86 (Every origin-symmetric convex body admits a large cone). *For every $n \in \mathbb{N}$, every origin-symmetric convex body $K \subseteq \mathbb{R}^n$ has a boundary point $z \in \partial K$ that satisfies*

$$\frac{\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big)}{\mathrm{vol}_n(K)} \gtrsim \frac{1}{n}\mathsf{SEP}\big(\mathbb{R}^n, \|\cdot\|_K\big). \tag{1.124}$$

To see that Corollary 86 coincides with the case $L = K$ of Theorem 84, simply recall the definition of the polar projection body $\Pi^* K$ in (1.30), while also recalling that for $z \in \mathbb{R}^n \smallsetminus \{0\}$ we denote the cone whose base is $\mathsf{Proj}_{z^\perp}(K) \subseteq z^\perp$ and whose apex is z by $\mathsf{Cone}_z(K)$, and the volume of $\mathsf{Cone}_z(K)$ is given in (1.35).

A substitution of (1.104) into Corollary 86 shows that any origin-symmetric convex body $K \subseteq \mathbb{R}^n$ has a boundary point $z \in \partial K$ that satisfies

$$\frac{\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big)}{\mathrm{vol}_n(K)} \gtrsim \frac{1}{\sqrt{n}}. \tag{1.125}$$

It seems (based on inquiring with experts in convex geometry) that the classical-looking geometric statement (1.125) did not previously appear in the literature. However, in response to our inquiry Lutwak found a different proof of (1.125) which in addition shows that the best possible constant in (1.125) is $1/\sqrt{2\pi}$. More precisely, we have the following proposition, whose proof (which relies on classical Brunn–Minkowski theory, unlike the indirect way by which we found (1.125)), is included in Section 2.7 (this proof is a restructuring of the proof that Lutwak found; we thank him for allowing us to include it here).

Proposition 87 (Lutwak). *For every $n \in \mathbb{N}$, any origin symmetric convex body $K \subseteq \mathbb{R}^n$ satisfies*

$$\max_{z\in\partial K} \frac{\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big)}{\mathrm{vol}_n(K)} \geqslant \frac{\Gamma\big(\frac{n}{2}\big)}{2\sqrt{\pi}\,\Gamma\big(\frac{n+1}{2}\big)} \geqslant \frac{1+\frac{1}{4n}}{\sqrt{2\pi n}}. \tag{1.126}$$

Moreover, the first inequality in (1.126) holds as equality if and only if K is an ellipsoid.

A substitution of (1.105) into Corollary 86 yields the following geometric inequality.

Corollary 88. *Fix $n \in \mathbb{N}$ and suppose that $K \subseteq \mathbb{R}^n$ is an origin-symmetric convex body. There is a boundary point $z \in \partial K$ such that the following inequality holds for every $x_1, \ldots, x_n \in K$:*

$$\frac{\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big)}{\mathrm{vol}_n(K)} \gtrsim \frac{1}{n} \fint_{S^{n-1}} \Big\| \sum_{i=1}^n \theta_i x_i \Big\|_K^2 \,\mathrm{d}\theta. \tag{1.127}$$

By combining [303] with Lemma 102 below, the maximum of the right-hand side of (1.127) over all possible $x_1, \ldots, x_n \in K$ is bounded above and below by universal constant multiples of $T_2(\mathbb{R}^n, \|\cdot\|_K)^2/n$ (recall the definition (1.77) of the type-2 constant), so Corollary 88 is indeed a substitution of (1.105) into (1.124).

Returning to Corollary 86, recall that both the cross-polytope $B_{\ell_1^n}$ and the hypercube $[-1,1]^n$ are examples of extremal symmetric convex bodies $K \subseteq \mathbb{R}^n$ that have a boundary point $z \in \partial K$ for which the volume of $\mathsf{Cone}_z(K)$ is a universal constant proportion of the volume of K (the Euclidean ball is an example of a convex body that is not extremal in this regard). But, there is a difference between the cross-polytope and the hypercube in terms of the stability of this property. Specifically, there is an origin-symmetric convex body $K \subseteq [-1,1]^n \subseteq O(1)K$ such that for every $z \in \partial K$ the left-hand side of (1.124) is at most a universal constant multiple of $1/\sqrt{n}$. In contrast, the following proposition shows that the extremality of $\max_{z\in\partial B_{\ell_1^n}} \mathrm{vol}_n(\mathsf{Cone}_z(B_{\ell_1^n}))/\mathrm{vol}_n(B_{\ell_1^n})$ (up to constant factors) persists under $O(1)$ perturbations.

Proposition 89. *Fix $n \in \mathbb{N}$ and $\alpha, \beta \in (0,\infty)$. Suppose that $K \subseteq \mathbb{R}^n$ is an origin-symmetric convex body that satisfies $\alpha K \subseteq B_{\ell_1^n} \subseteq \beta K$. Then there exists a boundary point $z \in \partial K$ such that*

$$\frac{\mathrm{vol}_n\big(\mathsf{Cone}_z(K)\big)}{\mathrm{vol}_n(K)} \gtrsim \frac{\alpha}{\beta}.$$

Proposition 89 is a direct consequence of Corollary 86, the bi-Lipschitz invariance of the modulus of separated decomposability, and the lower bound $\mathsf{SEP}(\ell_1^n) \gtrsim n$ of [76].

The following proposition is an application in a different direction of the results that we described in the preceding sections.

Proposition 90. *If $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ is a finite dimensional normed space with a 1-symmetric basis, then every subspace $\mathbf{X}$ of $\mathbf{E}$ satisfies*

$$\mathrm{evr}(\mathbf{X})\sqrt{\dim(\mathbf{X})} \lesssim \mathrm{evr}(\mathbf{E})\sqrt{\dim(\mathbf{E})}. \tag{1.128}$$

Proposition 90 holds because $\mathsf{SEP}(\mathbf{E}) \lesssim \mathrm{evr}(\mathbf{E})\sqrt{\dim(\mathbf{E})}$ by Corollary 79, while

$$\mathsf{SEP}(\mathbf{X}) \gtrsim \mathrm{evr}(\mathbf{X})\sqrt{\dim(\mathbf{X})}$$

by Theorem 71, so (1.128) follows from $\mathsf{SEP}(\mathbf{X}) \leqslant \mathsf{SEP}(\mathbf{E})$. This justification shows that Proposition 90 holds for a class of spaces that is larger than those that have a 1-symmetric basis, and Conjecture 6 would imply that Proposition 90 holds when $\mathbf{E}$ is any canonically positioned normed space.

Nevertheless, Proposition 90 fails to hold true without any further assumption on the normed space $\mathbf{E}$. For example, the computation in Remark 52 shows that for any $n, m \in \mathbb{N}$ with $n \geqslant 2$ and $m \asymp n\log n$, the space $\mathbf{E} = \ell_1^n \oplus \ell_2^m$ satisfies

$$\mathrm{evr}(\mathbf{E})\sqrt{\dim(\mathbf{E})} \lesssim \sqrt{n\log n}$$

while its subspace $\mathbf{X} = \ell_1^n$ satisfies $\mathrm{evr}(\mathbf{X})\sqrt{\dim(\mathbf{X})} \asymp n$.

Proposition 90 shows that if $\mathbf{E}$ has a 1-symmetric basis, then among the linear subspaces $\mathbf{X}$ of $\mathbf{E}$ the invariant $\mathrm{evr}(\mathbf{X})\sqrt{\dim(\mathbf{X})}$ is maximized up to universal constant factors at $\mathbf{X} = \mathbf{E}$. The fact we are multiplying here the external volume ratio of $\mathbf{X}$ by the square root of its dimension is an artifact of our proof and it would be interesting to understand what correction factors allow for such a result to hold.

Question 91. Characterize (up to universal constant factors) those $A : [1, \infty) \to [1, \infty)$ with the property that for any $n \geqslant 1$ we have $\mathrm{evr}(\mathbf{X})A(k) \leqslant \mathrm{evr}(\mathbf{E})A(n)$ for every normed space $(\mathbf{E}, \|\cdot\|_{\mathbf{E}})$ of dimension at most n that has a 1-symmetric basis, every $k \in \{1, \ldots, n\}$, and every k-dimensional subspace $\mathbf{X}$ of $\mathbf{E}$.

Proposition 90 shows that if $A(n) \asymp \sqrt{n}$, then $A : [1, \infty) \to [1, \infty)$ has the properties that are described in Question 91. At the same time, no $A : [1, \infty) \to [1, \infty)$ with $A(n) = O(1)$ can be as in Question 91. Indeed, for any such A consider the symmetric normed space $\mathbf{E} = \ell_\infty^n$. There is a universal constant $\eta > 0$ such that any normed space $\mathbf{X}$ with $\dim(\mathbf{X}) \leqslant \eta \log n$ is at Banach–Mazur distance at most 2 from a subspace of ℓ_∞^n.[11] In particular, this holds for $\mathbf{X} = \ell_1^m$ when $m \in \mathbb{N}$ satisfies $m \leqslant \eta \log n$, so we get that

$$A(\eta \log n)\sqrt{\log n} \asymp \mathrm{evr}\big(\ell_1^m\big)A(\eta \log n) \leqslant 2\,\mathrm{evr}\big(\ell_\infty^n\big)A(n) \asymp A(n). \tag{1.129}$$

So, $A(n) \gtrsim \sqrt{\log n}$ and by iterating (1.129) one gets the slightly better lower bound $A(n) \gtrsim \sqrt{(\log n)\log\log n}$, as well as $A(n) \gtrsim \sqrt{(\log n)(\log\log n)\log\log\log n}$ and so forth, yielding in the end the estimate

$$A(n) \geqslant \frac{\left(\prod_{k=1}^{\log^* n} \log^{[k]} n\right)^{\frac{1}{2}}}{e^{O(\log^* n)}}, \tag{1.130}$$

where for $k \in \mathbb{N} \cup \{0\}$ we denote the kth iterant of the logarithm by $\log^{[k]}$, i.e., $\log^{[0]} x = x$ for $x > 0$, and

$$\log^{[k]} x > 0 \implies \log^{[k+1]} x = \log\big(\log^{[k]} x\big). \tag{1.131}$$

There is no reason to expect that the lower bound (1.130) is close to being optimal, but in combination with Proposition 90 it does show that the answer to Question 91 is likely nontrivial.

These considerations lead to the following open-ended question. The literature contains multiple results showing that ℓ_p^n maximizes certain geometric invariants (for examples, Banach–Mazur distance to ℓ_2^n [176], or volume ratio [22]) among all

[11] This assertion is standard, here is a quick sketch. Take a δ-net $\mathfrak{N}$ of the unit sphere of $\mathbf{X}^*$ for a sufficiently small universal constant $\delta > 0$ and consider the embedding $x \mapsto (x^*(x))_{x^* \in \mathfrak{N}}$ from $\mathbf{X}$ to $\ell_\infty(\mathfrak{N})$. Since $\log|\mathfrak{N}| \asymp \dim(\mathbf{X})$, this gives a distortion 2-embedding (say, for $\delta = 1/10$) of $\mathbf{X}$ into ℓ_∞^n provided $\log n$ is at least a sufficiently large universal constant multiple of $\dim(\mathbf{X})$.

the n-dimensional subspaces or quotients of L_p. Is there an analogous theory in the spirit of (1.128) in the much more general setting of spaces that have a 1-symmetric basis? This could be viewed as a symmetric space variant of the classical work of Lewis [176, 177]. An interesting step in this direction can be found in [304]; specifically, see [304, Theorem 1.2], which could be relevant to Question 91 through the approach of [22, Section 2].

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 $x_1,\ldots,x_m \in \mathbf{X}$:

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

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*

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

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*

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

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:

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

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

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

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 $T_2(\mathbf{X})$:

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

■

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

$$T_2(\mathbf{X}) \lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac{3}{2}}. \tag{2.15}$$

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

$$C_2(\mathbf{X}) \lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac{3}{2}}. \tag{2.16}$$

Indeed,

$$\begin{aligned} C_2(\mathbf{X}) \leqslant T_2(\mathbf{X}^*) &\lesssim \mathsf{LT}(\mathbf{X}^*)(\log n)^{\frac{3}{2}} \\ &\lesssim \mathsf{LT}(\mathbf{X})(\log n)^{\frac{3}{2}}, \end{aligned} \tag{2.17}$$

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

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

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

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

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

$$t\sqrt{s} \gtrsim \frac{\delta}{\left(\log\left(\frac{2}{\delta}\right)\right)^{\frac{3}{2}}}, \tag{2.23}$$

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,

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

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

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

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

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

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

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

Consequently,

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

Therefore,

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

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 $x_1, \ldots, x_k$ 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 $x_1, \ldots, x_k \in \partial B_{\mathbf{X}}$ 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

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

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

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

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., $d_\theta(x,y) = \|x-y\|_{\ell_\infty^n(\mathbb{C})}^{\theta}$ 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.,

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

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

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

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

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

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*

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

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

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

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,

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

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

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

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

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

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

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

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

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

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 $i \in \{1, \ldots, n\}$, 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 $d_1(x;\Gamma), \ldots, d_n(x;\Gamma)$. Thus*

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

and

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

Then,

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

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,

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

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 $\mathfrak{x}(w)\in\mathfrak{N}_\varepsilon$ such that $w\in V_{\mathfrak{x}(w)}$.

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

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

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

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

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:

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

Now,

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

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

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

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

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 $w \in \mathcal{P}^*(\mathfrak{x}(w))$ 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 $\mathcal{P}^*(\mathfrak{x}(w)) = \mathcal{P}^*(y)$. 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:

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

Indeed, suppose that $w \in \mathbf{X}$ satisfies $(\mathfrak{x}(w) + (1/\mathfrak{p})B_{\mathbf{X}}) \cap \mathfrak{N}_\varepsilon \subseteq \mathcal{P}(\mathfrak{x}(w))$ 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 $\mathfrak{x}(z) \in \mathfrak{N}_\varepsilon$ is well defined. Now,

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

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

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

Finally,

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

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*

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

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,

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

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

$$\begin{aligned}&\mathrm{vol}_n(K)\\ &= \frac{1}{n}\int_{S^{n-1}}\left(\frac{1}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^{\frac{n}{n+1}}}\right)\left(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^{\frac{n}{n+1}}}{\|u\|_K^n}\right)\mathrm{d}u \qquad (2.83)\\ &\leqslant \frac{1}{n}\left(\int_{S^{n-1}} \frac{\mathrm{d}u}{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)^n}\right)^{\frac{1}{n+1}}\left(\int_{S^{n-1}} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,\mathrm{d}u\right)^{\frac{n}{n+1}}\\ &\qquad (2.84)\\ &\leqslant \frac{1}{n}\left(\frac{2\sqrt{\pi}\,\Gamma\big(\frac{n+1}{2}\big)}{n\Gamma\big(\frac{n}{2}\big)}\right)^{\frac{n}{n+1}} \frac{n^{\frac{1}{n+1}}}{\mathrm{vol}_n(K)^{\frac{n-1}{n+1}}}\left(\int_{S^{n-1}} \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{u^\perp}(K)\big)}{\|u\|_K^{n+1}}\,\mathrm{d}u\right)^{\frac{n}{n+1}},\\ &\qquad (2.85)\end{aligned}$$

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

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

Chapter 3

Preliminaries on random partitions

This section treats basic properties of random partitions, including measurability issues that we need for subsequent applications. As such, it is of a technical/foundational nature and it can be skipped on first reading if one is willing to accept the measurability requirements that are used in the proofs that appear in Section 4 and Section 5.

Recall that a random partition $\mathcal{P}$ of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ was defined in the Introduction as follows. One is given a probability space $(\Omega, \mathbf{Prob})$ and a sequence of set-valued mappings $\{\Gamma^k : \Omega \to 2^{\mathfrak{M}}\}_{k=1}^{\infty}$ such that for each fixed $k \in \mathbb{N}$ the mapping $\Gamma^k : \Omega \to 2^{\mathfrak{M}}$ is strongly measurable relative to the σ-algebra of **Prob**-measurable subsets of Ω, i.e., the set $(\Gamma^k)^-(E) = \{\omega \in \Omega : E \cap \Gamma^k(\omega) \neq \varnothing\}$ is **Prob**-measurable for every closed $E \subseteq \mathfrak{M}$. We require that $\mathcal{P}^{\omega} = \{\Gamma^k(\omega)\}_{k=1}^{\infty}$ is a partition of $\mathfrak{M}$ for every $\omega \in \Omega$.

Definition 63 and Definition 65 (of separating and padded random partitions, respectively) assumed implicitly that the quantities that appear in the left-hand sides of equations (1.92) and (1.95) are well defined, i.e., that the events $\{\mathcal{P}(x) \neq \mathcal{P}(y)\}$ and $\{B_{\mathfrak{M}}(x, r) \subseteq \mathcal{P}(x)\}$ are **Prob**-measurable for every $x, y \in \mathfrak{M}$ and $r > 0$. This follows from the above definition, because for every closed subset $E \subseteq \mathfrak{M}$ we have

$$\begin{aligned}&\big\{\omega \in \Omega : \mathcal{P}^{\omega}(x) \neq \mathcal{P}^{\omega}(y)\big\}\\&\quad= \bigcup_{\substack{k,\ell \in \mathbb{N}\\ k \neq \ell}} \Big(\big\{\omega \in \Omega : \{x\} \cap \Gamma^k(\omega) \neq \varnothing\big\} \cap \big\{\omega \in \Omega : \{y\} \cap \Gamma^{\ell}(\omega) \neq \varnothing\big\}\Big)\end{aligned}$$

and

$$\begin{aligned}&\big\{\omega \in \Omega : E \not\subseteq \mathcal{P}^{\omega}(x)\big\}\\&\quad= \bigcup_{\substack{k,\ell \in \mathbb{N}\\ k \neq \ell}} \Big(\big\{\omega \in \Omega : \{x\} \cap \Gamma^k(\omega) \neq \varnothing\big\} \cap \big\{\omega \in \Omega : E \cap \Gamma^{\ell}(\omega) \neq \varnothing\big\}\Big).\end{aligned}$$

Another "leftover" from the Introduction is the proof of Lemma 67, which asserts that the moduli of Definition 63 and Definition 65 are bi-Lipschitz invariants. The proof of this simple but needed statement is the following direct use of the definition of a Δ-bounded random partition.

Proof of Lemma 67. Fix $D > \mathsf{c}_{(\mathfrak{N}, d_{\mathfrak{N}})}(\mathfrak{M}, d_{\mathfrak{M}})$. There is an embedding $\phi : \mathfrak{M} \to \mathfrak{N}$ and a scaling factor $\lambda > 0$ such that (1.16) holds. Fix $\Delta > 0$ and let $\mathcal{P}$ be a $\lambda\Delta$-bounded random partition of $\mathfrak{N}$. Suppose that $\mathcal{P}$ is induced by the probability space $(\Omega, \mathbf{Prob})$, i.e., there are strongly measurable mappings $\{\Gamma^k : \Omega \to 2^{\mathfrak{N}}\}_{k=1}^{\infty}$ such that $\mathcal{P}^\omega = \{\Gamma^k(\omega)\}_{k=1}^{\infty}$ for every $\omega \in \Omega$. For every $k \in \mathbb{N}$ the mapping $\omega \mapsto \phi^{-1}(\Gamma^k(\omega)) \in 2^{\mathfrak{M}}$ is strongly measurable. Indeed, if $E \subseteq \mathfrak{M}$ is closed then, because $\mathfrak{M}$ is complete and ϕ is a homeomorphism, also $\phi(E) \subseteq \mathfrak{N}$ is closed. So,

$$\{\omega \in \Omega : \phi(E) \cap \Gamma^k(\omega) \neq \varnothing\} = \{\omega \in \Omega : E \cap \phi^{-1}(\Gamma^k(\omega)) \neq \varnothing\}$$

is **Prob**-measurable, as required. Therefore, if we define $\mathcal{Q}^\omega = \{\phi^{-1}(\Gamma^k(\omega))\}_{k=1}^{\infty}$ for $\omega \in \Omega$, then $\mathcal{Q}$ is a random partition of $\mathfrak{M}$.

$\mathcal{Q}$ is Δ-bounded since for $x \in \mathfrak{M}$ and $u, v \in \mathcal{Q}(x)$ we have $\phi(u), \phi(v) \in \mathcal{P}(\phi(x))$, hence $d_{\mathfrak{M}}(u, v) \leqslant d_{\mathfrak{N}}(\phi(u), \phi(v))/\lambda \leqslant \operatorname{diam}_{\mathfrak{N}}(\mathcal{P}(\phi(x)))/\lambda \leqslant \Delta$, using (1.16) and that $\mathcal{P}$ is $\lambda\Delta$-bounded. For every $x, y \in \mathfrak{M}$ the events $\{\mathcal{Q}(x) \neq \mathcal{Q}(y)\}$ and $\{\mathcal{P}(\phi(x)) \neq \mathcal{P}(\phi(y))\}$ coincide. So, if $\mathcal{P}$ is σ-separating for some $\sigma > 0$,

$$\begin{aligned}\textbf{Prob}\big[\mathcal{Q}(x) \neq \mathcal{Q}(y)\big] &= \textbf{Prob}\big[\mathcal{P}\big(\phi(x)\big) \neq \mathcal{P}\big(\phi(y)\big)\big]\\ &\leqslant \frac{\sigma}{\lambda\Delta} d_{\mathfrak{N}}\big(\phi(x), \phi(y)\big) \overset{(1.16)}{\leqslant} \frac{D\sigma}{\Delta} d_{\mathfrak{M}}(x, y).\end{aligned}$$

This shows that $\mathcal{Q}$ is $(D\sigma)$-separating, thus establishing the first assertion (1.97) of Lemma 67.

Suppose that $\mathcal{P}$ is $(\mathfrak{p}, \delta)$-padded for some $\mathfrak{p} > 0$ and $0 < \delta < 1$. Fix $x \in \mathfrak{M}$. Assuming that the event $\{B_{\mathfrak{N}}(\phi(x), \lambda\Delta/\mathfrak{p}) \subseteq \mathcal{P}(\phi(x))\}$ occurs, if $z \in B_{\mathfrak{M}}(x, \Delta/(D\mathfrak{p}))$, then $d_{\mathfrak{N}}(\phi(z), \phi(x)) \leqslant \lambda D d_{\mathfrak{M}}(z, x) \leqslant \lambda\Delta/\mathfrak{p}$ by (1.16). Thus,

$$\phi(z) \in B_{\mathfrak{N}}\Big(\phi(x), \frac{\lambda\Delta}{\mathfrak{p}}\Big)$$

and therefore $\phi(z) \in \mathcal{P}(\phi(x))$, i.e., $z \in \mathcal{Q}(x)$. This shows the inclusion of events

$$\Big\{B_{\mathfrak{N}}\Big(\phi(x), \frac{\lambda\Delta}{\mathfrak{p}}\Big) \subseteq \mathcal{P}\big(\phi(x)\big)\Big\} \subseteq \Big\{B_{\mathfrak{M}}\Big(x, \frac{\Delta}{D\mathfrak{p}}\Big) \subseteq \mathcal{Q}(x)\Big\}.$$

Since $\mathcal{P}$ is $(\mathfrak{p}, \delta)$-padded, it follows from this that also $\mathcal{Q}$ is $(D\mathfrak{p}, \delta)$-padded, thus establishing the second assertion (1.98) of Lemma 67. ∎

The final basic "leftover" from the Introduction is the following simple proof of Lemma 68.

Proof of Lemma 68. Fix $\Delta > 0$ and suppose that $\sigma_1 > \mathsf{SEP}(\mathfrak{M}_1)$ and $\sigma_2 > \mathsf{SEP}(\mathfrak{M}_2)$. Define

$$\Delta_1 = \Delta\Big(\frac{\sigma_1}{\sigma_1 + \sigma_2}\Big)^{\frac{1}{s}} \quad \text{and} \quad \Delta_2 = \Delta\Big(\frac{\sigma_2}{\sigma_1 + \sigma_2}\Big)^{\frac{1}{s}}. \tag{3.1}$$

Let $\mathcal{P}_{\Delta_1}$ be a σ_1-separating Δ_1-bounded random partition of $\mathfrak{M}_1$. Similarly, let $\mathcal{P}_{\Delta_2}$ be a σ_2-separating Δ_2-bounded random partition of $\mathfrak{M}_2$. Assume that $\mathcal{P}_{\Delta_1}$ and $\mathcal{P}_{\Delta_2}$

are independent random variables. Let $\mathcal{P}_\Delta$ be the corresponding product random partition of $\mathfrak{M}_1 \times \mathfrak{M}_2$, i.e., its clusters are give by

$$\forall (x_1, x_2) \in \mathfrak{M}_1 \times \mathfrak{M}_2, \quad \mathcal{P}_\Delta(x_1, x_2) = \mathcal{P}_{\Delta_1}(x_1) \times \mathcal{P}_{\Delta_2}(x_2). \tag{3.2}$$

By (3.1) we have $\Delta_1^s + \Delta_2^s = \Delta^s$, so $\mathcal{P}_\Delta$ is a Δ-bounded random partition of the metric space $\mathfrak{M}_1 \oplus_s \mathfrak{M}_2$ (the required measurability is immediate). It therefore remains to observe that every $(x_1, x_2), (y_1, y_2) \in \mathfrak{M}_1 \times \mathfrak{M}_2$ satisfy

$$\begin{aligned}
&\mathbf{Prob}\big[\mathcal{P}_\Delta(x_1, x_2) \neq \mathcal{P}_\Delta(y_1, y_2)\big] \\
&= 1 - \mathbf{Prob}\big[\mathcal{P}_{\Delta_1}(x_1) = \mathcal{P}_{\Delta_1}(y_1)\big]\mathbf{Prob}\big[\mathcal{P}_{\Delta_2}(x_2) = \mathcal{P}_{\Delta_2}(y_2)\big] && (3.3)\\
&\leqslant 1 - \left(1 - \frac{\sigma_1 d_{\mathfrak{M}_1}(x_1, y_1)}{\Delta_1}\right)\left(1 - \frac{\sigma_2 d_{\mathfrak{M}_2}(x_2, y_2)}{\Delta_2}\right) && (3.4)\\
&= \frac{\sigma_1 d_{\mathfrak{M}_1}(x_1, y_1)}{\Delta_1} + \frac{\sigma_2 d_{\mathfrak{M}_2}(x_2, y_2)}{\Delta_2} - \frac{\sigma_1\sigma_2 d_{\mathfrak{M}_1}(x_1, y_1) d_{\mathfrak{M}_2}(x_2, y_2)}{\Delta_1\Delta_2} && (3.5)\\
&\leqslant \left(\left(\frac{\sigma_1}{\Delta_1}\right)^{\frac{s}{s-1}} + \left(\frac{\sigma_2}{\Delta_2}\right)^{\frac{s}{s-1}}\right)^{\frac{s-1}{s}} \left(d_{\mathfrak{M}_1}(x_1, y_1)^s + d_{\mathfrak{M}_2}(x_2, y_2)^s\right)^{\frac{1}{s}} && (3.6)\\
&= \frac{\sigma_1 + \sigma_2}{\Delta} d_{\mathfrak{M}_1 \oplus_s \mathfrak{M}_2}\big((x_1, x_2), (y_1, y_2)\big), && (3.7)
\end{aligned}$$

where (3.3) uses (3.2) and the independence of $\mathcal{P}_{\Delta_1}$ and $\mathcal{P}_{\Delta_2}$, the bound (3.4) is an application of the assumption that $\mathcal{P}_{\Delta_1}$ is σ_1-separating and $\mathcal{P}_{\Delta_2}$ is σ_2-separating, (3.6) is an application of Hölder's inequality, and (3.7) follows from (1.99) and (3.1). This proves (1.100). Note that even though we dropped the quadratic additive improvement in (3.5), this does not change the final bound in (1.100) due to the need to work with all possible scales $\Delta > 0$ and all possible values of $d_{\mathfrak{M}_1}(x_1, y_1)$ and $d_{\mathfrak{M}_2}(x_2, y_2)$.

To prove (1.101), fix $\mathfrak{p}_1 > \mathsf{PAD}_{\delta_1}(\mathfrak{M}_1)$ and $\mathfrak{p}_2 > \mathsf{PAD}_{\delta_2}(\mathfrak{M}_2)$ and replace (3.1) by

$$\Delta_1 = \frac{\Delta \mathfrak{p}_1}{\left(\mathfrak{p}_1^s + \mathfrak{p}_2^s\right)^{\frac{1}{s}}} \quad \text{and} \quad \Delta_2 = \frac{\Delta \mathfrak{p}_2}{\left(\mathfrak{p}_1^s + \mathfrak{p}_2^s\right)^{\frac{1}{s}}}.$$

This time, we choose $\mathcal{P}_{\Delta_1}$ to be a $(\mathfrak{p}_1, \delta_1)$-padded Δ_1-bounded random partition of $\mathfrak{M}_1$. Similarly, let $\mathcal{P}_{\Delta_2}$ be a $(\mathfrak{p}_2, \delta_2)$-padded Δ_2-bounded random partition of $\mathfrak{M}_2$, with $\mathcal{P}_{\Delta_1}$ and $\mathcal{P}_{\Delta_2}$ independent, and we again combine them as in (3.2) to give the product partition $\mathcal{P}_\Delta$ of $\mathfrak{M}_1 \times \mathfrak{M}_2$. By reasoning analogously, $\mathcal{P}_\Delta$ is a $((\mathfrak{p}_1^s + \mathfrak{p}_2^s)^{1/s}, \delta_1\delta_2)$-padded Δ-bounded random partition of $\mathfrak{M}_1 \oplus_s \mathfrak{M}_2$. ∎

3.1 Standard set-valued mappings

Recall that a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ is said to be Polish if it is separable and complete. Polish metric spaces are the appropriate setting for Lipschitz extension theorems that are based on the assumption that for every $\Delta > 0$ there is a probability distribution over Δ-bounded partitions of $\mathfrak{M}$ with certain properties. Indeed, a Banach space-valued Lipschitz function can always be extended to the completion of $\mathfrak{M}$ while preserving the Lipschitz constant, and the mere existence of countably many sets of diameter at most Δ that cover $\mathfrak{M}$ for every $\Delta > 0$ implies that $\mathfrak{M}$ is separable.

Theorem 66 assumes local compactness. Even though this assumption is more restrictive than being Polish, it suffices for the applications that we obtain herein because they deal with finite dimensional normed spaces. It is, however, possible to treat general Polish metric spaces by working with a notion of measurability of set-valued mappings that differs from the strong measurability that was assumed in Section 1.7. We call this notion *standard set-valued mappings*; see Definition 113.

The requirements for a set-valued mapping to be standard are quite innocuous and easy to check. In particular, the clusters of the specific random partitions that we will study are easily seen to be standard set-valued mappings. It is also simple to verify that the clusters of the random partitions that we construct are strongly measurable. So, we have two approaches, which are both easy to work with. We chose to work in the Introduction with the requirement that the clusters are strongly measurable because this directly makes the quantity $\mathsf{SEP}(\cdot)$ be bi-Lipschitz invariant, and it is also slightly simpler to describe. Nevertheless, in practice it is straightforward to check that the clusters are standard, and even though we do not know that this leads to a bi-Lipschitz invariant (we suspect that it *does not*), it does lead to an easily implementable Lipschitz extension criterion that holds in the maximal generality of Polish spaces.

Definition 113 (Standard set-valued mapping). Suppose that $(\mathcal{Z}, d_{\mathcal{Z}})$ is a Polish metric space and that $\Omega \subseteq \mathcal{Z}$ is a Borel subset of $\mathcal{Z}$. Given a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, a set-valued mapping $\Gamma : \Omega \to 2^{\mathfrak{M}}$ is said to be *standard* if the following three conditions hold.

- For every $x \in \mathfrak{M}$ the set $\{\omega \in \Omega : x \in \Gamma(\omega)\}$ is Borel.
- The set $\mathcal{G}_\Gamma = \Gamma^-(\mathfrak{M}) = \{\omega \in \Omega : \Gamma(\omega) \neq \varnothing\}$ is Borel.
- For every $x \in \mathfrak{M}$ the mapping $(\omega \in \mathcal{G}_\Gamma) \mapsto d_{\mathfrak{M}}(x, \Gamma(\omega))$ is Borel measurable on $\mathcal{G}_\Gamma$.

The following extension criterion is a counterpart to Theorem 66 that works in the maximal generality of Polish metric spaces; its proof, which is an adaptation of ideas of [173], appears in Section 5.

Theorem 114. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a Polish metric space and fix another metric $\mathfrak{d}$ on $\mathfrak{M}$. Suppose that for every $\Delta > 0$ there is a Polish metric space $\mathcal{Z}_\Delta$, a Borel subset $\Omega_\Delta \subseteq \mathcal{Z}_\Delta$, a Borel probability measure* $\mathbf{Prob}_\Delta$ *on Ω_Δ and a sequence of standard set-valued mappings $\{\Gamma_\Delta^k : \Omega_\Delta \to 2^{\mathfrak{M}}\}_{k=1}^\infty$ such that $\mathcal{P}_\Delta^\omega = \{\Gamma_\Delta^k(\omega)\}_{k=1}^\infty$ is a partition* *of* $\mathfrak{M}$ *for every* $\omega \in \Omega_\Delta$*, for every* $x \in \mathfrak{M}$ *and* $\omega \in \Omega_\Delta$ *we have* $\mathrm{diam}_{\mathfrak{M}}(\mathcal{P}^\omega_\Delta(x)) \leqslant \Delta$*, and*

$$\forall x, y \in \mathfrak{M}, \quad \Delta \mathbf{Prob}_\Delta\big[\omega \in \Omega_\Delta : \mathcal{P}^\omega_\Delta(x) \neq \mathcal{P}^\omega_\Delta(y)\big] \leqslant \mathfrak{d}(x, y).$$

Then, for every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$*, every subset* $\mathcal{C} \subseteq \mathfrak{M}$ *and every* 1*-Lipschitz mapping* $f : \mathcal{C} \to \mathbf{Z}$*, there exists a mapping* $F : \mathfrak{M} \to \mathbf{Z}$ *that extends* f *and satisfies* $\|F(x) - F(y)\|_{\mathbf{Z}} \lesssim \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$ *(namely,* F *is Lipschitz on* $\mathfrak{M}$ *with respect to the metric* $\mathfrak{d}$*). Moreover,* F *depends linearly on* f*.*

3.2 Proximal selectors

For later applications we need to know that set-valued mappings that are either strongly measurable or standard admit certain auxiliary measurable mappings that are (perhaps approximately) the closest point to a given (but arbitrary) nonempty closed subset of the metric space in question. We will justify this now using classical descriptive set theory.

Lemma 115. *Fix a measurable space* $(\Omega, \mathcal{F})$*. Suppose that* $(\mathfrak{M}, d_{\mathfrak{M}})$ *is a metric space and that* $S \subseteq \mathfrak{M}$ *is nonempty and locally compact. Let* $\Gamma : \Omega \to 2^{\mathfrak{M}}$ *be a strongly measurable set-valued mapping such that* $\Gamma(\omega)$ *is a bounded subset of* $\mathfrak{M}$ *for every* $\omega \in \Omega$*. Then there exists an* $\mathcal{F}$*-to-Borel measurable mapping* $\gamma : \Omega \to S$ *that satisfies* $d_{\mathfrak{M}}(\gamma(\omega), \Gamma(\omega)) = d_{\mathfrak{M}}(S, \Gamma(\omega))$ *for every* $\omega \in \Omega$ *for which* $\Gamma(\omega) \neq \varnothing$*.*

Proof. For every $\omega \in \Omega$ define a subset $\Phi(\omega) \subseteq S$ as follows:

$$\Phi(\omega) \stackrel{\mathrm{def}}{=} \begin{cases} \{s \in S : d_{\mathfrak{M}}(s, \Gamma(\omega)) = d_{\mathfrak{M}}(S, \Gamma(\omega))\} & \text{if } \Gamma(\omega) \neq \varnothing, \\ S & \text{if } \Gamma(\omega) = \varnothing. \end{cases}$$

The goal of Lemma 115 is to show the existence of an $\mathcal{F}$-to-Borel measurable mapping $\gamma : \Omega \to S$ that satisfies $\gamma(\omega) \in \Phi(\omega)$ for every $\omega \in \Omega$. Since $(S, d_{\mathfrak{M}})$ is locally compact, it is in particular Polish, so by the measurable selection theorem of Kuratowski and Ryll-Nardzewski [161] (see also [309] or [291, Chapter 5.2]) it suffices to check that $\Phi(\omega)$ is nonempty and closed for every $\omega \in \Omega$, and that we have $\{\omega \in \Omega : E \cap \Phi(\omega) = \varnothing\} \in \mathcal{F}$ for every closed $E \subseteq S$. Since S is locally compact, every closed subset of S is a countable union of compact subsets, so it suffices to check the latter requirement for compact subsets of S, i.e., to show that $\{\omega \in \Omega : K \cap \Phi(\omega) = \varnothing\} \in \mathcal{F}$ for every compact $K \subseteq S$.

Fix $\omega \in \Omega$. If $\Gamma(\omega) = \varnothing$ then $\Phi(\omega) = S$ is closed (since S is locally compact) and nonempty by assumption. If $\Gamma(\omega) \neq \varnothing$ then the continuity of $s \mapsto d_{\mathfrak{M}}(s, \Gamma(\omega))$ on S implies that $\Phi(\omega)$ is closed. Moreover, in this case since $\Gamma(\omega)$ is bounded and S is locally compact, the continuous mapping $s \mapsto d_{\mathfrak{M}}(s, \Gamma(\omega))$ attains its minimum on S, so that $\Phi(\omega) \neq \varnothing$.

It therefore remains to check that $\{\omega \in \Omega : K \cap \Phi(\omega) = \varnothing\} \in \mathcal{F}$ for every nonempty compact $K \subsetneq S$. Fixing such a K, since S is locally compact and hence separable, there exist $\{\kappa_i\}_{i=1}^{\infty} \subseteq K$ and $\{\sigma_j\}_{j=1}^{\infty} \subseteq S$ that are dense in K and S, respectively. Denote $\mathcal{G}_\Gamma = \{\omega \in \Omega : \Gamma(\omega) \neq \varnothing\}$. Then $\mathcal{G}_\Gamma \in \mathcal{F}$, because Γ is strongly measurable. Observe that the following identity holds:

$$\begin{aligned}
&\big\{\omega \in \Omega : K \cap \Phi(\omega) = \varnothing\big\} \\
&\quad = \big\{\omega \in \mathcal{G}_\Gamma : \forall \kappa \in K,\ d_{\mathfrak{M}}(\kappa, \Gamma(\omega)) > d_{\mathfrak{M}}(S, \Gamma(\omega))\big\} \\
&\quad = \bigcup_{m=1}^{\infty} \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \left\{\omega \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(\kappa_i, \Gamma(\omega)) > d_{\mathfrak{M}}(\sigma_j, \Gamma(\omega)) + \frac{1}{m}\right\}.
\end{aligned} \tag{3.8}$$

The verification of (3.8) proceeds as follows. Since $\Phi(\omega) \neq \varnothing$ for every $\omega \in \Omega$ and $K \neq \varnothing$, if $K \cap \Phi(\omega) = \varnothing$ then $\omega \in \mathcal{G}_\Gamma$ (otherwise $\Phi(\omega) = S$). This explains the first equality (3.8). For the second equality in (3.8), note that since $\Gamma(\omega)$ is bounded and K is compact, $\inf_{\kappa \in K} d_{\mathfrak{M}}(\kappa, \Gamma(\omega))$ is attained. Therefore, the second set in (3.8) is equal to $A = \{\omega \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(K, \Gamma(\omega)) > d_{\mathfrak{M}}(S, \Gamma(\omega))\}$. If $\omega \in A$, then there is $m \in \mathbb{N}$ such that $d_{\mathfrak{M}}(K, \Gamma(\omega)) > d_{\mathfrak{M}}(S, \Gamma(\omega)) + 2/m$, implying in particular that $d_{\mathfrak{M}}(\kappa_i, \Gamma(\omega)) > d_{\mathfrak{M}}(S, \Gamma(\omega)) + 2/m$ for every $i \in \mathbb{N}$. As $\{\sigma_j\}_{j=1}^{\infty}$ is dense in S, for every $i \in \mathbb{N}$ there is $j \in \mathbb{N}$ such that $d_{\mathfrak{M}}(\kappa_i, \Gamma(\omega)) > d_{\mathfrak{M}}(\sigma_j, \Gamma(\omega)) + 1/m$. Hence, the second set in (3.8) is contained in the third set in (3.8). For the reverse inclusion, if ω is in third set in (3.8) then

$$d_{\mathfrak{M}}(K, \Gamma(\omega)) = \inf_{i \in \mathbb{N}} d_{\mathfrak{M}}(\kappa_i, \Gamma(\omega)) > \inf_{j \in \mathbb{N}} d_{\mathfrak{M}}(\sigma_j, \Gamma(\omega)) = d_{\mathfrak{M}}(S, \Gamma(\omega)).$$

By (3.8), it suffices to show that $\{\omega \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(x, \Gamma(\omega)) > d_{\mathfrak{M}}(y, \Gamma(\omega)) + r\} \in \mathcal{F}$ for every fixed $x, y \in S$ and $r > 0$. For this, it suffices to show that for every $z \in \mathfrak{M}$ the mapping $\omega \mapsto d_{\mathfrak{M}}(z, \Gamma(\omega))$ is $\mathcal{F}$-to-Borel measurable on $\mathcal{G}_\Gamma$. Since $\mathcal{G}_\Gamma \in \mathcal{F}$, this is a consequence of the strong measurability of Γ, because for every $t \geqslant 0$ we have

$$\begin{aligned}
&\{\omega \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(z, \Gamma(\omega)) > t\} \\
&\quad = \bigcup_{k=1}^{\infty} \mathcal{G}_\Gamma \cap \left\{\omega \in \Omega : B_{\mathfrak{M}}\left(z, t + \frac{1}{k}\right) \cap \Gamma(\omega) = \varnothing\right\}.
\end{aligned}$$

■

Lemma 115 is a satisfactory treatment of measurable nearest point selectors for strongly measurable set-valued mappings, though under an assumption of local compactness. We did not investigate the minimal assumptions that are required for the conclusion of Lemma 115 to hold. We will next treat the setting of standard set-valued mappings without assuming local compactness.

Let $(\mathcal{Z}, d_{\mathcal{Z}})$ be a Polish metric space. Recall that a subset A of $\mathcal{Z}$ is said to be *universally measurable* if it is measurable with respect to *every* complete σ-finite Borel measure μ on $\mathcal{Z}$ (see, e.g., [154, p. 155]). If $(\mathfrak{M}, d_{\mathfrak{M}})$ is another metric space and

$\Omega \subseteq \mathcal{Z}$ is Borel, then a mapping $\psi : \Omega \to \mathfrak{M}$ is said to be universally measurable if $\psi^{-1}(E)$ is a universally measurable subset of Ω for every Borel subset E of $\mathfrak{M}$. Finally, recall that $A \subseteq \mathfrak{M}$ is said to be *analytic* if it is an image under a continuous mapping of a Borel subset of a Polish metric space (see, e.g., [154, Chapter 14] or [136, Chapter 11]). By Lusin's theorem [189, 192] (see also, e.g., [154, Theorem 21.10]), analytic subsets of Polish metric spaces are universally measurable.

Lemma 116. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ and $(\mathcal{Z}, d_{\mathcal{Z}})$ be Polish metric spaces and fix a Borel subset $\Omega \subseteq \mathcal{Z}$. Fix also $\Delta > 0$ such that* $\mathrm{diam}(\mathfrak{M}) \geqslant \Delta$. *Suppose that $\Gamma : \Omega \to 2^{\mathfrak{M}}$ satisfies the following two properties.*

(1) *For every $\omega \in \Omega$ such that $\Gamma(\omega) \neq \varnothing$ we have* $\mathrm{diam}_{\mathfrak{M}}(\Gamma(\omega)) < \Delta$.

(2) *For every $x \in \mathfrak{M}$ and $t \in \mathbb{R}$ the set $\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing \wedge d_{\mathfrak{M}}(x, \Gamma(\omega)) > t\}$ is analytic.*

Then, for every closed subset $\varnothing \neq S \subseteq \mathfrak{M}$ there exists a universally measurable mapping $\gamma : \Omega \to S$ such that

$$\forall (\omega, x) \in \Omega \times \mathfrak{M}, \quad x \in \Gamma(\omega) \implies d_{\mathfrak{M}}(x, \gamma(\omega)) \leqslant d_{\mathfrak{M}}(x, S) + \Delta.$$

Proof. For every $\omega \in \Omega$, define a subset $\Psi(\omega) \subseteq S$ as follows:

$$\Psi(\omega) \stackrel{\mathrm{def}}{=} \begin{cases} \bigcap_{x \in \mathfrak{M}} \{s \in S : d_{\mathfrak{M}}(x, s) \leqslant 2 d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x, S) + \Delta\} & \text{if } \Gamma(\omega) \neq \varnothing, \\ S & \text{if } \Gamma(\omega) = \varnothing. \end{cases} \tag{3.9}$$

We will show that there exists a universally measurable mapping $\gamma : \Omega \to S$ such that $\gamma(\omega) \in \Psi(\omega)$ for every $\omega \in \Omega$. Since S is a closed subset of $\mathfrak{M}$, it is Polish. Hence, by the Kuratowski–Ryll-Nardzewski measurable selection theorem [161], it suffices to prove that $\Psi(\omega)$ is nonempty and closed for every $\omega \in \Omega$, and that $\Psi^{-}(E) = \{\omega \in \Omega : E \cap \Psi(\omega) \neq \varnothing\}$ is universally measurable for every closed $E \subseteq S$.

By design, $\Psi(\omega) = S$ is nonempty and closed if $\Gamma(\omega) = \varnothing$. So, fix $\omega \in \Omega$ such that $\Gamma(\omega) \neq \varnothing$. Then $\Psi(\omega)$ is closed because if $\{s_k\}_{k=1}^{\infty} \subseteq \Psi(\omega)$ and $s \in \mathfrak{M}$ satisfy $\lim_{k \to \infty} d_{\mathfrak{M}}(s_k, s) = 0$, then for every $k \in \mathbb{N}$ and $x \in \mathfrak{M}$, since $s_k \in \Psi(\omega)$ we have $d_{\mathfrak{M}}(s_k, x) \leqslant 2 d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x, S) + \Delta$. Hence, by continuity also

$$d_{\mathfrak{M}}(s, x) \leqslant 2 d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x, S) + \Delta$$

for every $x \in \mathfrak{M}$, i.e., $s \in \Psi(\omega)$.

We will next check that $\Psi(\omega) \neq \varnothing$ for every $\omega \in \Omega$ such that $\Gamma(\omega) \neq \varnothing$. Denote $\varepsilon_\omega = \Delta - \mathrm{diam}_{\mathfrak{M}}(\Gamma(\omega))$. By assumption (1) of Lemma 116 we have $\varepsilon_\omega > 0$, so we may choose $s_\omega \in S$ and $y_\omega \in \Gamma(\omega)$ that satisfy $d_{\mathfrak{M}}(y_\omega, s_\omega) \leqslant d_{\mathfrak{M}}(\Gamma(\omega), S) + \varepsilon_\omega$.

We claim that $s_\omega \in \Psi(\omega)$. Indeed, for every $x \in \mathfrak{M}$ and $z \in \Gamma(\omega)$ we have

$$\begin{aligned} d_{\mathfrak{M}}(x, s_\omega) &\leqslant d_{\mathfrak{M}}(x,z) + d_{\mathfrak{M}}(z, y_\omega) + d_{\mathfrak{M}}(y_\omega, s_\omega) \\ &\leqslant d_{\mathfrak{M}}(x,z) + \operatorname{diam}_{\mathfrak{M}}(\Gamma(\omega)) + d_{\mathfrak{M}}(\Gamma(\omega), S) + \varepsilon_\omega \\ &\leqslant d_{\mathfrak{M}}(x,z) + d_{\mathfrak{M}}(z,S) + \Delta \\ &\leqslant d_{\mathfrak{M}}(x,z) + d_{\mathfrak{M}}(x,S) + d_{\mathfrak{M}}(x,z) + \Delta, \end{aligned} \tag{3.10}$$

where in the penultimate step of (3.10) we used the fact that $d_{\mathfrak{M}}(\Gamma(\omega), S) \leqslant d_{\mathfrak{M}}(z,S)$, since $z \in \Gamma(\omega)$, and in the final step of (3.10) we used the fact that $p \mapsto d_{\mathfrak{M}}(p,S)$ is 1-Lipschitz on $\mathfrak{M}$. Since (3.10) holds for every $z \in \Gamma(\omega)$, it follows that

$$d_{\mathfrak{M}}(x, s_\omega) \leqslant 2d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x,S) + \Delta.$$

Because this holds for every $x \in \mathfrak{M}$, it follows that $s_\omega \in \Psi(\omega)$.

Having checked that Ψ takes values in closed and nonempty subsets of S, it remains to show that $\Psi^-(E)$ is universally measurable for every closed $E \subseteq S$. To this end, since $\mathfrak{M}$ is separable, we may fix from now on a sequence $\{x_j\}_{j=1}^\infty$ that is dense in $\mathfrak{M}$. Note that by the case $t = 0$ of assumption (2) of Lemma 116, for every $j \in \mathbb{N}$ the following set is analytic:

$$\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing \wedge d_{\mathfrak{M}}(x_j, \Gamma(\omega)) > 0\} = \big\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing \wedge x_j \notin \overline{\Gamma(\omega)}\big\}.$$

Countable unions and intersections of analytic sets are analytic (see, e.g., [154, Proposition 14.4]), so we deduce that the following set is analytic:

$$\begin{aligned} &\bigcup_{j=1}^\infty \big\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing \wedge x_j \notin \overline{\Gamma(\omega)}\big\} \\ &\qquad = \big\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing \wedge \{x_j\}_{j=1}^\infty \not\subseteq \overline{\Gamma(\omega)}\big\} = \big\{\omega \in \Omega : \Gamma(\omega) \neq \varnothing\big\}, \end{aligned} \tag{3.11}$$

where for the final step of (3.11) observe that, since $\{x_j\}_{j=1}^\infty$ is dense in $\mathfrak{M}$, if $\{x_j\}_{j=1}^\infty$ were a subset of $\overline{\Gamma(\omega)}$ then it would follow that $\Gamma(\omega)$ is dense in $\mathfrak{M}$. This would imply that $\operatorname{diam}_{\mathfrak{M}}(\Gamma(\omega)) = \operatorname{diam}(\mathfrak{M}) \geqslant \Delta$, in contradiction to assumption (1) of Lemma 116. We have thus checked that the set $\mathcal{G}_\Gamma = \{\omega \in \Omega : \Gamma(\omega) \neq \varnothing\}$ is analytic, and hence by Lusin's theorem [189, 192] it is universally measurable. Now,

$$\begin{aligned} \Psi^-(E) &\overset{(3.9)}{=} (\Omega \smallsetminus \mathcal{G}_\Gamma) \\ &\quad \cup \{\omega \in \mathcal{G}_\Gamma : \exists s \in E\ \forall x \in \mathfrak{M},\ d_{\mathfrak{M}}(x,s) \leqslant 2d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x,S) + \Delta\}. \end{aligned}$$

Hence, it remains to prove that the following set is universally measurable:

$$\begin{aligned} &\big\{\omega \in \mathcal{G}_\Gamma : \exists s \in E\ \forall x \in \mathfrak{M},\ d_{\mathfrak{M}}(x,s) \leqslant 2d_{\mathfrak{M}}(x, \Gamma(\omega)) + d_{\mathfrak{M}}(x,S) + \Delta\big\} \\ &\quad = \big\{\omega \in \mathcal{G}_\Gamma : \exists s \in E\ \forall j \in \mathbb{N},\ d_{\mathfrak{M}}(x_j,s) \leqslant 2d_{\mathfrak{M}}(x_j, \Gamma(\omega)) + d_{\mathfrak{M}}(x_j,S) + \Delta\big\}, \end{aligned} \tag{3.12}$$

where we used the fact that $\{x_j\}_{j=1}^\infty$ is dense in $\mathfrak{M}$.

Consider the following subset $\mathcal{C}$ of $\Omega \times E$:

$$\mathcal{C} \stackrel{\mathrm{def}}{=} \big\{(\omega, s) \in \mathcal{G}_\Gamma \times E : \forall j \in \mathbb{N},\ d_{\mathfrak{M}}(x_j, s) \leqslant 2d_{\mathfrak{M}}(x_j, \Gamma(\omega)) + d_{\mathfrak{M}}(x_j, S) + \Delta\big\}.$$

The set in (3.12) is $\pi_1(\mathcal{C})$, where $\pi_1 : \Omega \times E \to \Omega$ is the projection to the first coordinate, i.e., $\pi_1(\omega, s) = \omega$ for every $(\omega, s) \in \Omega \times E$. Since continuous images and preimages of analytic sets are analytic (see, e.g., [154, Proposition 14.4]), by another application of Lusin's theorem it suffices to show that $\mathcal{C}$ is analytic. We already proved that $\mathcal{G}_\Gamma \subseteq \Omega$ is analytic, so there is a Borel subset L of a Polish space $\mathcal{Y}$ and a continuous mapping $\phi : L \to \Omega$ such that $\phi(L) = \mathcal{G}_\Gamma$. Denoting the identity mapping on E by $\mathsf{Id}_E : E \to E$, since ϕ maps L onto $\mathcal{G}_\Gamma$, the set $\mathcal{C}$ is the image under the continuous mapping $\phi \times \mathsf{Id}_E$ of the following subset of $\mathcal{Y} \times E$:

$$\begin{aligned}
&\big\{(y, s) \in L \times E : \forall j \in \mathbb{N},\ d_{\mathfrak{M}}(x_j, s) \leqslant 2d_{\mathfrak{M}}\big(x_j, \Gamma(\phi(y))\big) + d_{\mathfrak{M}}(x_j, S) + \Delta\big\} \\
&= \bigcap_{j=1}^{\infty} \big\{(y, s) \in L \times E : d_{\mathfrak{M}}(x_j, s) \leqslant 2d_{\mathfrak{M}}\big(x_j, \Gamma(\phi(y))\big) + d_{\mathfrak{M}}(x_j, S) + \Delta\big\}.
\end{aligned}$$

Hence, since continuous images and countable intersections of analytic sets are analytic, by yet another application of Lusin's theorem we see that it suffices to show that for every fixed $x \in \mathfrak{M}$ the following set is analytic, where for every $q \in \mathbb{Q}$ we denote $A_q = \{(y, s) \in L \times E : q < d_{\mathfrak{M}}(x, s)\} = L \times \{s \in E : q < d_{\mathfrak{M}}(x, s)\}$:

$$\begin{aligned}
&\big\{(y, s) \in L \times E : d_{\mathfrak{M}}(x, s) \leqslant 2d_{\mathfrak{M}}\big(x, \Gamma(\phi(y))\big) + d_{\mathfrak{M}}(x, S) + \Delta\big\} \\
&= \bigcap_{q \in \mathbb{Q}} \Big(\big((L \times E) \smallsetminus A_q\big) \\
&\qquad \cup \Big(A_q \cap \big\{(y, s) \in L \times E : 2d_{\mathfrak{M}}\big(x, \Gamma(\phi(y))\big) > q - d_{\mathfrak{M}}(x, S) - \Delta\big\}\Big)\Big),
\end{aligned}$$

Since A_q is Borel for all $q \in \mathbb{Q}$, it suffices to show that the following set is analytic for every $t \in \mathbb{R}$:

$$\big\{(y, s) \in L \times E : d_{\mathfrak{M}}\big(x, \Gamma(\phi(y))\big) > t\big\} = \phi^{-1}\Big(\big\{\omega \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(x, \Gamma(\omega)) > t\big\}\Big) \times E.$$

Since a preimage under a continuous mapping of an analytic set is analytic, the above set is indeed analytic due to assumption (2) of Lemma 116 and the fact that E is closed. ∎

Remark 117. The proof of Lemma 116 used the assumption $\mathrm{diam}(\mathfrak{M}) \geqslant \Delta$ only to deduce that the set

$$\mathcal{G}_\Gamma = \{\omega \in \Omega : \Gamma(\omega) \neq \varnothing\}$$

is analytic from (the case $t = 0$ of) assumption (2) of Lemma 116. Hence, if we add the assumption that $\mathcal{G}_\Gamma$ is analytic to Lemma 116, then we can drop the restriction

$\operatorname{diam}(\mathfrak{M}) \geqslant \Delta$ altogether. Alternatively, recalling equation (3.11) and the paragraph immediately after it, for the above proof of Lemma 116 to go through it suffices to assume that $\Gamma(\omega)$ is not dense in $\mathfrak{M}$ for any $\omega \in \Omega$.

Recalling Definition 113, Lemma 116 and Remark 117 imply the following corollary. Indeed, by Remark 117 we know that we can drop the assumption $\operatorname{diam}(\mathfrak{M}) \geqslant \Delta$ of Lemma 116, and when Γ is a standard set-valued mapping the sets that appears in assumption (2) of Lemma 116 are Borel.

Corollary 118. *Fix $\Delta > 0$. Let $(\mathfrak{M}, d_{\mathfrak{M}})$ and $(\mathcal{Z}, d_{\mathcal{Z}})$ be Polish metric spaces and fix a Borel subset $\Omega \subseteq \mathcal{Z}$. Suppose that $\Gamma : \Omega \to 2^{\mathfrak{M}}$ is a standard set-valued mapping such that $\operatorname{diam}_{\mathfrak{M}}(\Gamma(\omega)) < \Delta$ for every $\omega \in \mathcal{G}_\Gamma$. Then for every closed $\varnothing \neq S \subseteq \mathfrak{M}$ there exists a universally measurable mapping $\gamma : \Omega \to S$ that satisfies*

$$\forall (\omega, x) \in \Omega \times \mathfrak{M}, \quad x \in \Gamma(\omega) \implies d_{\mathfrak{M}}(x, \gamma(\omega)) \leqslant d_{\mathfrak{M}}(x, S) + \Delta.$$

3.3 Measurability of iterative ball partitioning

The following set-valued mapping is a building block of much of the literature on random partitions, including the present investigation. Fix a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ and $k \in \mathbb{N}$. Define a set-valued mapping $\Gamma : \mathfrak{M}^k \times [0, \infty)^k \to 2^{\mathfrak{M}}$ by setting

$$\Gamma\big(\vec{x}, \vec{r}\big) \stackrel{\mathrm{def}}{=} B_{\mathfrak{M}}(x_k, r_k) \smallsetminus \bigcup_{j=1}^{k-1} B_{\mathfrak{M}}(x_j, r_j) \tag{3.13}$$

for $(\vec{x}, \vec{r}) = (x_1, \ldots, x_k, r_1, \ldots, r_k) \in \mathfrak{M}^k \times [0, \infty)^k$. We can think of Γ as a random subset of $\mathfrak{M}$ if we are given a probability measure $\mathbf{Prob}$ on $\mathfrak{M}^k \times [0, \infty)^k$. The measure $\mathbf{Prob}$ can encode the geometry of $(\mathfrak{M}, d_{\mathfrak{M}})$; for example, if $(\mathfrak{M}, d_{\mathfrak{M}})$ is a complete doubling metric space, then in [173] this measure arises from a doubling measure on $\mathfrak{M}$ (see [191, 308]). The measure $\mathbf{Prob}$ can also have a "smoothing effect" through the randomness of the radii (see, e.g., [1, 30, 71, 96, 173, 208, 238, 239]; choosing a suitable distribution over the random radii is sometimes an important and quite delicate matter, but this intricacy will not arise in the present work. For finite dimensional normed spaces, a random subset as in (3.13) was used in [76, 152]. Note that given $\Delta > 0$, if the measure $\mathbf{Prob}$ is supported on the set of those $(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times [0, \infty)^k$ for which $r_k \leqslant \Delta/2$, then the mapping Γ takes values in subsets of $\mathfrak{M}$ of diameter at most Δ.

While the definition (3.13) is very simple and natural, in order to use it in the ensuing reasoning we need to know that it satisfies certain measurability requirements. Note first that the set-valued mapping Γ in (3.13) has the following basic measurability property: for every fixed $y \in \mathfrak{M}$ the set $\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times [0, \infty)^k : y \in \Gamma(\vec{x}, \vec{r})\}$ is

Borel. Indeed, by definition we have

$$\begin{aligned}&\big\{(\vec{x},\vec{r}) \in \mathfrak{M}^k \times [0,\infty)^k : y \in \Gamma(\vec{x},\vec{r})\big\}\\&= \bigcap_{j=1}^{k-1} \big\{(\vec{x},\vec{r}) \in \mathfrak{M}^k \times [0,\infty)^k : d_{\mathfrak{M}}(y,x_j) > r_j\big\}\\&\quad\cap \big\{(\vec{x},\vec{r}) \in \mathfrak{M}^k \times [0,\infty)^k : d_{\mathfrak{M}}(y,x_k) \leqslant r_k\big\}.\end{aligned}$$

In other words, the indicator mapping $(\vec{x},\vec{r}) \mapsto \mathbf{1}_{\Gamma(\vec{x},\vec{r})}(y)$ is Borel measurable for every fixed $y \in \mathfrak{M}$.

Lemma 119. *Fix $k \in \mathbb{N}$. Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a Polish metric space and suppose that $\Gamma : \mathfrak{M}^k \times [0,\infty)^k \to 2^{\mathfrak{M}}$ be given in (3.13). Then*

$$\Gamma^-(S) = \big\{(\vec{x},\vec{r}) \in \mathfrak{M}^k \times [0,\infty)^k : S \cap \Gamma(\vec{x},\vec{r}) \neq \varnothing\big\}$$

is analytic for every analytic subset $S \subseteq \mathfrak{M}$. Consequently, for every complete σ-finite Borel measure μ on $\mathfrak{M}^k \times [0,\infty)^k$, if $\mathcal{F}_\mu$ denotes the σ-algebra of μ-measurable subsets of $\mathfrak{M}^k \times [0,\infty)^k$, then Γ is a strongly measurable set-valued mapping from the measurable space $(\mathfrak{M}^k \times [0,\infty)^k, \mathcal{F}_\mu)$ to $2^{\mathfrak{M}}$.

Proof. Since S is analytic, there exists a Borel subset T of a Polish metric space $\mathcal{Z}$ and a continuous mapping $\psi : T \to \mathfrak{M}$ such that $\psi(T) = S$. Consider the following Borel subset $\mathfrak{B}$ of the Polish space $\mathfrak{M}^k \times [0,\infty)^k \times \mathcal{Z}$ ($\mathfrak{B}$ is Borel because it is defined using finitely many continuous inequalities)

$$\begin{aligned}\mathfrak{B} \stackrel{\mathrm{def}}{=} \big\{&(\vec{x},\vec{r},t) \in \mathfrak{M}^k \times [0,\infty)^k \times T : d_{\mathfrak{M}}(\psi(t),x_k) \leqslant r_k\\&\wedge \forall j \in \{1,\ldots,k-1\},\ d_{\mathfrak{M}}(\psi(t),x_j) > r_j\big\}.\end{aligned}$$

Then $\Gamma^-(S) = \pi(\mathfrak{B})$, where

$$\pi : \mathfrak{M}^k \times [0,\infty)^k \times \mathcal{Z} \to \mathfrak{M}^k \times [0,\infty)^k$$

is the projection onto the first two coordinates, i.e., $\pi(\vec{x},\vec{r},z) = (\vec{x},\vec{r})$ for $(\vec{x},\vec{r},z) \in \mathfrak{M}^k \times [0,\infty)^k \times \mathcal{Z}$. Since π is continuous, it follows that $\Gamma^-(S)$ is analytic. By Lusin's theorem [189, 192], it follows that $\Gamma^-(S)$ is universally measurable. In particular, if μ is a complete σ-finite Borel measure on $\mathfrak{M}^k \times [0,\infty)^k$ and $\mathcal{F}_\mu$ is the σ-algebra of μ-measurable subsets of $\mathfrak{M}^k \times [0,\infty)^k$, then $\Gamma^-(E) \in \mathcal{F}_\mu$ for every closed subset $E \subseteq \mathfrak{M}$. Recalling (1.91), this means that Γ is a strongly measurable set-valued mapping from the measurable space $(\mathfrak{M}^k \times [0,\infty)^k, \mathcal{F}_\mu)$ to $2^{\mathfrak{M}}$. ∎

Lemma 120 below contains additional Borel measurability assertions that will be used later. Its assumptions are satisfied, for example, when $\mathfrak{M}$ is a separable normed space, which is the case of interest here. We did not investigate the maximal generality under which the conclusion of Lemma 120 holds.

In what follows, given a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, for every $x \in \mathfrak{M}$ and $r > 0$ the open ball of radius r centered at x is denoted $B^{\circ}_{\mathfrak{M}}(x, r) = \{y \in \mathfrak{M} : d_{\mathfrak{M}}(x, y) < r\}$.

Lemma 120. *Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a separable metric space such that*

$$\forall (x, r) \in \mathfrak{M} \times (0, \infty), \quad B_{\mathfrak{M}}(x, r) = \overline{B^{\circ}_{\mathfrak{M}}(x, r)}. \tag{3.14}$$

Fix $k \in \mathbb{N}$ and let $\Gamma : \mathfrak{M}^k \times (0, \infty)^k \to 2^{\mathfrak{M}}$ be given in (3.13). Then the following set is Borel measurable:

$$\mathcal{G}_{\Gamma} = \left\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0, \infty)^k : \Gamma(\vec{x}, \vec{r}) \neq \varnothing\right\}.$$

Also, for each $y \in \mathfrak{M}$ the mapping from $\mathcal{G}_{\Gamma}$ to $\mathbb{R}$ that is given by

$$(\vec{x}, \vec{r}) \mapsto d_{\mathfrak{M}}(y, \Gamma(x, r))$$

is Borel measurable.

Proof. Let $\mathfrak{D} \subseteq \mathfrak{M}$ be a countable dense subset of $\mathfrak{M}$. The assumption (3.14) implies that $\mathfrak{D} \cap \Gamma(\vec{x}, \vec{r})$ is dense in $\Gamma(\vec{x}, \vec{r})$ for every $(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0, \infty)^k$. This is straightforward to check as follows. Fix $y \in \Gamma(\vec{x}, \vec{r})$ and $\delta > 0$. We need to find $q \in \mathfrak{D} \cap \Gamma(\vec{x}, \vec{r})$ with $d_{\mathfrak{M}}(q, y) < \delta$. Recalling (3.13), since $y \in \Gamma(\vec{x}, \vec{r})$ we know that $d_{\mathfrak{M}}(y, x_k) \leqslant r_k$, and also $d_{\mathfrak{M}}(y, x_j) > r_j$ for every $j \in \{1, \ldots, k-1\}$, i.e., $\eta > 0$ where

$$\eta \overset{\text{def}}{=} \min\left\{\delta, d_{\mathfrak{M}}(y, x_1) - r_1, \ldots, d_{\mathfrak{M}}(y, x_{k-1}) - r_{k-1}\right\}.$$

By (3.14) there is $z \in B^{\circ}_{\mathfrak{M}}(x_k, r_k)$ with $d_{\mathfrak{M}}(z, y) < \eta/2$. Denote

$$\rho \overset{\text{def}}{=} \min\left\{r_k - d_{\mathfrak{M}}(z, x_k), \frac{1}{2}\eta\right\}.$$

Then $\rho > 0$, so the density of $\mathfrak{D}$ in $\mathfrak{M}$ implies that there is $q \in \mathfrak{D}$ with $d_{\mathfrak{M}}(q, z) < \rho$. Consequently,

$$d_{\mathfrak{M}}(q, y) \leqslant d_{\mathfrak{M}}(q, z) + d_{\mathfrak{M}}(z, y) < \rho + \frac{\eta}{2} \leqslant \delta.$$

It remains to observe that $q \in \Gamma(\vec{x}, \vec{r})$, because

$$d_{\mathfrak{M}}(q, x_k) \leqslant d_{\mathfrak{M}}(q, z) + d_{\mathfrak{M}}(z, x_k) < \rho + d_{\mathfrak{M}}(z, x_k) \leqslant r_k,$$

and also for every $j \in \{1, \ldots, k-1\}$ we have

$$\begin{aligned} d_{\mathfrak{M}}(q, x_j) &\geqslant d_{\mathfrak{M}}(y, x_j) - d_{\mathfrak{M}}(y, z) - d_{\mathfrak{M}}(z, q) \\ &> d_{\mathfrak{M}}(y, x_j) - \frac{\eta}{2} - \rho \geqslant d_{\mathfrak{M}}(y, x_j) - \eta \geqslant r_j. \end{aligned}$$

For every $(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0, \infty)^k$, we have $\Gamma(\vec{x}, \vec{r}) \neq \varnothing$ if and only if $\mathfrak{D} \cap \Gamma(\vec{x}, \vec{r}) \neq \varnothing$. Consequently,

$$\begin{aligned}\mathcal{G}_\Gamma &= \big\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0,\infty)^k : \Gamma(\vec{x}, \vec{r}) \neq \varnothing\big\} \\ &= \bigcup_{q \in \mathfrak{D}} \big\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0,\infty)^k : q \in \Gamma(\vec{x}, \vec{r})\big\}.\end{aligned}$$

Since $\mathfrak{D}$ is countable and we already checked in the paragraph immediately preceding Lemma 119 that $\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0, \infty)^k : y \in \Gamma(\vec{x}, \vec{r})\}$ is Borel measurable for every $y \in \mathfrak{M}$, we get that $\mathcal{G}_\Gamma$ is Borel measurable.

Next, $d_{\mathfrak{M}}(y, \Gamma(\vec{x}, \vec{r})) = d_{\mathfrak{M}}(y, \mathfrak{D} \cap \Gamma(\vec{x}, \vec{r}))$ for every $(\vec{x}, \vec{r}) \in \mathcal{G}_\Gamma$ and $y \in \mathfrak{M}$. So, for every $t > 0$ we have

$$\begin{aligned}&\big\{(\vec{x}, \vec{r}) \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}\big(y, \Gamma(\vec{x}, \vec{r})\big) < t\big\} \\ &\quad= \bigcup_{q \in \mathfrak{D} \cap B^\circ_{\mathfrak{M}}(y,t)} \big\{(\vec{x}, \vec{r}) \in \mathfrak{M}^k \times (0,\infty)^k : q \in \Gamma(\vec{x}, \vec{r})\big\}.\end{aligned}$$

It follows that $\{(\vec{x}, \vec{r}) \in \mathcal{G}_\Gamma : d_{\mathfrak{M}}(y, \Gamma(\vec{x}, \vec{r})) < t\}$ is Borel measurable for every $t \in \mathbb{R}$. ∎

Corollary 121 below follows directly from the definition of a standard set-valued mapping due to Lemma 120 and the discussion in the paragraph immediately preceding Lemma 119.

Corollary 121. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a Polish metric space satisfying (3.14). Then, for every $k \in \mathbb{N}$ the set-valued mapping $\Gamma : \mathfrak{M}^k \times (0, \infty)^k \to 2^{\mathfrak{M}}$ in (3.13) is standard.*

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

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

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

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

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

$$\chi(z) = \gamma_k \implies \Gamma^{k,z}(\vec{x}, \vec{\gamma}) \stackrel{\mathrm{def}}{=} (z + x_k + B_\mathbf{X}) \setminus \bigcup_{j=1}^{k-1} \bigcup_{\substack{w \in \Lambda \\ \chi(w) = \gamma_j}} (w + x_j + B_\mathbf{X}),$$

$$\chi(z) \neq \gamma_k \implies \Gamma^{k,z}(\vec{x}, \vec{\gamma}) \stackrel{\mathrm{def}}{=} \varnothing. \tag{4.3}$$

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

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

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 $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w + \mathcal{V} + B_{\mathbf{X}}) \neq \varnothing$. 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 $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w + \mathcal{V} + B_{\mathbf{X}}) \neq \varnothing$ then let $w(\gamma, z) \in \Lambda$ be an arbitrary (but fixed) lattice point such that $(z + \mathcal{V} + B_{\mathbf{X}}) \cap (w(\gamma, z) + \mathcal{V} + B_{\mathbf{X}}) = \varnothing$. 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 $\vec{\mathsf{y}} = (\mathsf{y}_1, \dots, \mathsf{y}_k) \in (\mathbb{R}^n)^k$,

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

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 $w(\vec{\chi},z) = (w(\chi_1,z),\dots,w(\chi_k,z)) \in (\mathbb{R}^n)^k$. 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 $\Omega \subseteq \mathcal{Z} = \mathcal{V}^{\mathbb{N}} \times \{1,\dots,5^{2n}\}^{\mathbb{N}}$ 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 $\mathrm{diam}_\mathbf{X}(\Gamma^{k,z}(\vec{x},\vec{\gamma})) \leqslant \mathrm{diam}_\mathbf{X}(z + x_k + B_\mathbf{X}) \leqslant 2$. Note also that by (4.3) if $\Gamma^{k,z}(\vec{x},\vec{\gamma}) \neq \varnothing$, then

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

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* $\operatorname{diam}_{\mathbf{X}}(\mathcal{T})\leqslant 4$. *Then the sets*

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

are pairwise disjoint and

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

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 $\operatorname{diam}_{\mathbf{X}}(\mathcal{M})\leqslant 4$, 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

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

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

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

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 $j\in\{1,\ldots,k-1\}$ 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

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

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

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

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 $(\vec{x},\vec{\gamma}) \in \mathcal{E}_{u,1+\rho}^{j,w}$ 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)*:*

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

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,

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

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

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

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,

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

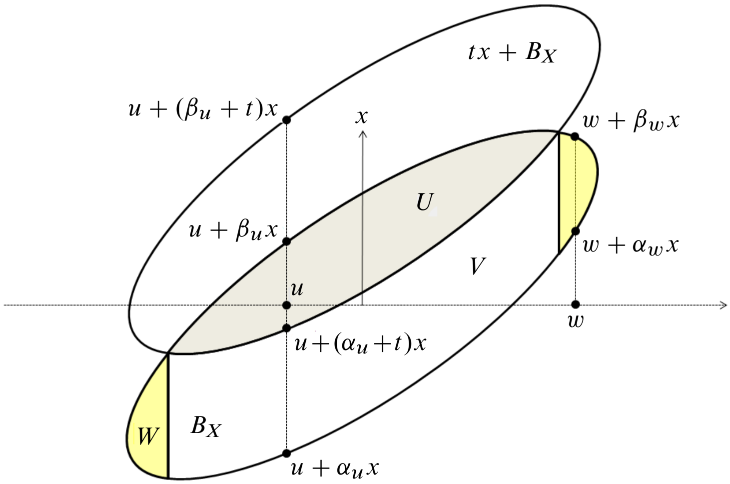

Figure 4.1. A schematic depiction of the partition of $B_{\mathbf{X}}$ into the sets U, V, W (with the sets U, W shaded), as well as the line segments parallel to x that are used in the justification of the estimate (4.23).

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

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

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:*

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

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

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

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:

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

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

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

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 $\mathbf{Prob}$-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

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

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

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

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

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

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 $i, j \in \{1, \dots, n\}$ satisfy

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

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. ∎

Chapter 5

Barycentric-valued Lipschitz extension

In this section, we will explain how separation profiles relate to Lipschitz extension. We cannot invoke [173] as a "black box" because we need a more general result and our definition of random partitions differs from that of [173]. But, the modifications that are required in order to apply the ideas of [173] in the present setting are of a secondary nature, and the main geometric content of the phenomenon that is explained below is the same as in [173].

In addition to making the present article self-contained, there are more advantages to including here complete proofs of Theorem 66 and Theorem 114. Firstly, the reasoning of [173] was designed to deal with a more general setting (treating multiple notions of random partitions at once), and it is illuminating to present a proof for separating decompositions in isolation, which leads to simplifications. Secondly, since [173] appeared, alternative viewpoints have been developed that relate it to optimal transport, as carried out by Kozdoba [158], Brudnyi and Brudnyi [62], Ohta [243], and culminating more recently with a comprehensive treatment by Ambrosio and Puglisi [11]. Here we will frame the construction using the optimal transport methodology, which has conceptual advantages that go beyond yielding a clearer restructuring of the argument. The optimal transport viewpoint had an important role in quantitative improvements that were obtained in [231, 233], as well as results that will appear in forthcoming works. As a byproduct, we will use this viewpoint to easily derive a stability statement for convex hull-valued Lipschitz extension under metric transforms.

5.1 Notational preliminaries

We will start by quickly setting notation and terminology for basic concepts in measure theory and optimal transport. Everything that we describe in this subsection is standard and is included here only in order to avoid any ambiguities in the subsequent discussions.

Given a signed measure μ on a measurable space $(\Omega, \mathcal{F})$, its Hahn–Jordan decomposition is denoted $\mu = \mu^+ - \mu^-$, i.e., μ^+, μ^- are disjointly supported nonnegative measures. The total variation measure of μ is $|\mu| = \mu^+ + \mu^-$. For $A \in \mathcal{F}$, the restriction of μ to A is denoted $\mu\lfloor_A$, i.e., $\mu\lfloor_A(E) = \mu(A \cap E)$ for $E \in \mathcal{F}$. If $(\Omega', \mathcal{F}')$ is another measurable space and $f : \Omega \to \Omega'$ is a measurable mapping, then the push-forward of μ under f is denoted $f_{\#}\mu$. Thus $f_{\#}\mu(E) = \mu(f^{-1}(E))$ for $E \in \mathcal{F}'$, or equivalently

$$\forall h \in L_1(f_\# \mu), \qquad \int_{\Omega'} h(\omega')\, \mathrm{d} f_\#\mu(\omega') = \int_{\Omega} h(f(\omega))\, \mathrm{d}\mu(\omega).$$

Suppose from now on that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space. A signed Borel measure μ on $\mathfrak{M}$ has finite first moment if $\int_{\mathfrak{M}} d_{\mathfrak{M}}(x,y)\, \mathrm{d}|\mu|(y) < \infty$ for all $x \in \mathfrak{M}$. Note that this implies in particular that $|\mu|(\mathfrak{M}) < \infty$, because if $x, x' \in \mathfrak{M}$ are distinct points, then the mapping $(y \in \mathfrak{M}) \mapsto [d_{\mathfrak{M}}(x,y) + d_{\mathfrak{M}}(x',y)]/d_{\mathfrak{M}}(x,x')$ belongs to $L_1(|\mu|)$ and takes values in $[1,\infty)$ by the triangle inequality.

The set of all of the signed Borel measures on $\mathfrak{M}$ of finite first moment is denoted $\mathsf{M}_1(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{M}_1(\mathfrak{M})$ if the metric is clear from the context. The set of all nonnegative measures in $\mathsf{M}_1(\mathfrak{M})$ is denoted $\mathsf{M}_1^+(\mathfrak{M})$, the set of all $\mu \in \mathsf{M}_1(\mathfrak{M})$ with total mass 0, i.e., $\mu^+(\mathfrak{M}) = \mu^-(\mathfrak{M})$, is denoted $\mathsf{M}_1^0(\mathfrak{M})$, and the set of all probability measures in $\mathsf{M}_1(\mathfrak{M})$ is denoted $\mathsf{P}_1(\mathfrak{M})$.

Given $\mu, \nu \in \mathsf{M}_1^+(\mathfrak{M})$ with $\mu(\mathfrak{M}) = \nu(\mathfrak{M})$, a Borel measure π on $\mathfrak{M} \times \mathfrak{M}$ is a coupling of μ and ν if

$$\pi(E \times \mathfrak{M}) = \mu(A) \quad \text{and} \quad \pi(\mathfrak{M} \times E) = \nu(A)$$

for every Borel subset $E \subseteq \mathfrak{M}$. The set of couplings of μ and ν is denoted $\Pi(\mu,\nu) \subseteq \mathsf{M}_1^+(\mathfrak{M} \times \mathfrak{M})$. Note that $(\mu \times \nu)/\mu(\mathfrak{M}) = (\mu \times \nu)/\nu(\mathfrak{M}) \in \Pi(\mu,\nu)$, so $\Pi(\mu,\nu) \neq \varnothing$. The Wasserstein-1 distance between μ and ν that is induced by the metric $d_{\mathfrak{M}}$, denoted $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu,\nu)$ or simply $\mathsf{W}_1(\mu,\nu)$ if the metric is clear from the context, is the infimum of $\int_{\mathfrak{M}\times\mathfrak{M}} d_{\mathfrak{M}}(x,y)\, \mathrm{d}\pi(x,y)$ over all possible couplings $\pi \in \Pi(x,y)$. Since $(\mathfrak{M}, d_{\mathfrak{M}})$ is Polish, the metric space $(\mathsf{P}_1(\mathfrak{M}), \mathsf{W}_1)$ is also Polish; see, e.g., [42] or [10, Proposition 7.1.5]. Throughout what follows, $\mathsf{P}_1(\mathfrak{M})$ will be assumed to be equipped with the metric W_1. The Kantorovich–Rubinstein duality theorem (see, e.g., [307, Theorem 5.10]) asserts that

$$\mathsf{W}_1(\mu,\nu) = \sup_{\substack{\psi:\mathfrak{M}\to\mathbb{R} \\ \|\psi\|_{\mathrm{Lip}(\mathfrak{M})}=1}} \left(\int_{\mathfrak{M}} \psi\, \mathrm{d}\mu - \int_{\mathfrak{M}} \psi\, \mathrm{d}\nu \right). \tag{5.1}$$

Note that (5.1) implies that $\mathsf{W}_1(\mu+\tau, \nu+\tau) = \mathsf{W}_1(\mu,\nu)$ for every $\tau \in \mathsf{M}_1^+(\mathfrak{M})$.

For $\mu \in \mathsf{M}_1^0(\mathfrak{M})$ we have $\mu^+(\mathfrak{M}) = \mu^-(\mathfrak{M})$, so we can define[1]:

$$\|\mu\|_{\mathsf{W}_1(\mathfrak{M})} = \mathsf{W}_1(\mu^+, \mu^-).$$

[1]Note for later use that if $\mu, \nu \in \mathsf{M}_1^+(\mathfrak{M})$ satisfy $\mu(\mathfrak{M}) = \nu(\mathfrak{M})$, then $\mu - \nu \in \mathsf{M}_1^0(\mathfrak{M})$ and $\|\mu - \nu\|_{\mathsf{W}_1(\mathfrak{M})} = \mathsf{W}_1(\mu,\nu)$. For a standard justification of the latter assertion, see, e.g., the simple deduction of [236, equation (2.2)].

This turns $\mathsf{M}_1^0(\mathfrak{M})$ into a normed space whose completion is called the *free space* over $\mathfrak{M}$ (also known as the Arens–Eells space over $\mathfrak{M}$), and is denoted $\mathfrak{F}(\mathfrak{M})$; see [16, 113,310] for more on this topic, and note that while $\mathfrak{F}(\mathfrak{M})$ is commonly defined as the closure of the *finitely supported* measures in $\mathsf{M}_1^0(\mathfrak{M})$ with respect to the Wasserstein-1 norm, since the finitely supported measures are dense in $\mathsf{M}_1^0(\mathfrak{M})$ (see, e.g., [307, Theorem 6.18]), the definitions coincide. It follows from (5.1) that the dual of $\mathfrak{F}(\mathfrak{M})$ is canonically isometric to the space of all the real-valued Lipschitz functions on $\mathfrak{M}$ that vanish at some (arbitrary but fixed) point $x_0 \in \mathfrak{M}$, equipped with the norm $\|\cdot\|_{\mathrm{Lip}(\mathfrak{M})}$.

Suppose that $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ is a separable Banach space and fix $\mu \in \mathsf{M}_1(\mathfrak{M})$. By the Pettis measurability criterion [249] (see also [36, Proposition 5.1]), any $f \in \mathrm{Lip}(\mathfrak{M}; \mathbf{Z})$ is $|\mu|$-measurable. Moreover, we have $\|f\|_{\mathbf{Z}} \in L_1(|\mu|)$ because if we fix $x \in \mathfrak{M}$, then for every $y \in \mathfrak{M}$,

$$\begin{aligned}\|f(y)\|_{\mathbf{Z}} &\leqslant \|f(y) - f(x)\|_{\mathbf{Z}} + \|f(x)\|_{\mathbf{X}}\\ &\leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} d_{\mathfrak{M}}(y, x) + \|f(x)\|_{\mathbf{X}} \in L_1(|\mu|),\end{aligned}$$

where the last step holds by the definition of $\mathsf{M}_1(\mathfrak{M})$ and the fact that it implies that $|\mu|(\mathfrak{M}) < \infty$. By Bochner's integrability criterion [40] (see also [36, Proposition 5.2]), it follows that the Bochner integrals $\int_{\mathfrak{M}} f \,\mathrm{d}\mu^+$ and $\int_{\mathfrak{M}} f \,\mathrm{d}\mu^-$ are well-defined elements of $\mathbf{Z}$, so we can consider the vector

$$\mathfrak{I}_f(\mu) \stackrel{\mathrm{def}}{=} \int_{\mathfrak{M}} f \,\mathrm{d}\mu = \int_{\mathfrak{M}} f \,\mathrm{d}\mu^+ - \int_{\mathfrak{M}} f \,\mathrm{d}\mu^- \in \mathbf{Z}. \tag{5.2}$$

If $\mu \in \mathsf{M}_1^0(\mathfrak{M})$, then $\mathfrak{I}_f(\mu) = \int_{\mathfrak{M}\times\mathfrak{M}} (f(x) - f(y)) \,\mathrm{d}\pi(x, y)$ for every coupling $\pi \in \Pi(\mu^+, \mu^-)$. Consequently, $\|\mathfrak{I}_f(\mu)\|_{\mathbf{Z}} \leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} \int_{\mathfrak{M}\times\mathfrak{M}} d_{\mathfrak{M}}(x, y) \,\mathrm{d}\pi(x, y)$, so by taking the infimum over all $\pi \in \Pi(\mu^+, \mu^-)$ we see that the norm of the linear operator $\mathfrak{I}_f$ from $(\mathsf{M}_1^0(\mathfrak{M}), \|\cdot\|_{\mathsf{W}_1})$ to $\mathbf{Z}$ satisfies

$$\|\mathfrak{I}_f\|_{(\mathsf{M}_1^0(\mathfrak{M}),\|\cdot\|_{\mathsf{W}_1})\to\mathbf{Z}} \leqslant \|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})}. \tag{5.3}$$

Since $\mathsf{M}_1^0(\mathfrak{M})$ is dense in $\mathfrak{F}(\mathfrak{M})$, it follows that $\mathfrak{I}_f$ extends uniquely to a linear operator $\mathfrak{I}_f : \mathfrak{F}(\mathfrak{M}) \to \mathbf{Z}$ of norm at most $\|f\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})}$. So, even though elements of $\mathfrak{F}(\mathfrak{M})$ need not be measures, one can consider the "integral" $\mathfrak{I}_f(\phi) \in \mathbf{Z}$ of $f \in \mathrm{Lip}(\mathfrak{M}; \mathbf{Z})$ with respect to $\phi \in \mathfrak{F}(\mathfrak{M})$; see [114] for more on this topic.

5.2 Refined extension moduli

Continuing with the notation that was introduced by Matoušek [199], we will consider the following parameters related to Lipschitz extension. Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$, $(\mathfrak{N}, d_{\mathfrak{N}})$ are metric spaces and that $\mathfrak{C} \subseteq \mathfrak{M}$. Denote by $\mathsf{e}(\mathfrak{M}, \mathfrak{C}; \mathfrak{N})$ the infimum over those $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathfrak{N}$ with $\|f\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})} < \infty$ there is $F : \mathfrak{M} \to \mathfrak{N}$ that extends f and satisfies

$$\|F\|_{\mathrm{Lip}(\mathfrak{M};\mathfrak{N})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C};\mathfrak{N})}.$$

The supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathfrak{N})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$ will be denoted $\mathsf{e}(\mathfrak{M};\mathfrak{N})$. Note that when $\mathfrak{N}$ is complete, $\mathfrak{N}$-valued Lipschitz functions on $\mathcal{C}$ automatically extend to the closure of $\mathcal{C}$ while preserving the Lipschitz constant, so we may assume here that $\mathcal{C}$ is closed. The supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ over all Banach spaces $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$ will be denoted below by $\mathsf{e}(\mathfrak{M},\mathcal{C})$. Thus, the notation $\mathsf{e}(\mathfrak{M})$ of the Introduction coincides with the supremum of $\mathsf{e}(\mathfrak{M},\mathcal{C})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$.

If $(\mathfrak{M},d_{\mathfrak{M}})$ is a metric space, $\mathcal{C} \subseteq \mathfrak{M}$, and $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$ is a Banach space, then it is natural to consider variants of the above definitions with the additional restrictions that the extended mapping F is required to take values in either the closure of the linear span of $f(\mathcal{C})$ or the closure of the convex hull of $f(\mathcal{C})$. Namely, let $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ be the infimum over those $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathbf{Z}$ there exists

$$F : \mathfrak{M} \to \overline{\mathrm{span}}\big(f(\mathcal{C})\big)$$

that extends f and satisfies

$$\|F\|_{\mathrm{Lip}(\mathfrak{M};\mathbf{Z})} \leqslant K\|f\|_{\mathrm{Lip}(\mathcal{C};\mathbf{Z})}. \tag{5.4}$$

Analogously, let $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ be the infimum over $K \in [1,\infty]$ such that for every $f : \mathcal{C} \to \mathbf{Z}$ there exists

$$F : \mathfrak{M} \to \overline{\mathrm{conv}}\big(f(\mathcal{C})\big)$$

that extends f and satisfies (5.4). We then define $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C})$ to be the supremum of $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C};\mathbf{Z})$ over all possible Banach spaces $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$. Note that while one could attempt to define $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C})$ similarly, there is no point to do so because it would result in the previously defined quantity $\mathsf{e}(\mathfrak{M},\mathcal{C})$. By considering the supremum of $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathcal{C})$ over all subsets $\mathcal{C} \subseteq \mathfrak{M}$, one defines the quantity $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M})$.

Remark 133. By [179] one can have $\mathsf{e}(\mathfrak{M},\mathcal{C};\mathbf{Z}) = \mathsf{e}(\mathfrak{M};\mathbf{Z}) = 1$ yet $\mathsf{e}_{\mathrm{span}}(\mathfrak{M},\mathcal{C},\mathbf{Z}) = \infty$ for some metric space $(\mathfrak{M},d_{\mathfrak{M}})$, some $\mathcal{C} \subseteq \mathfrak{M}$ and some Banach space $(\mathbf{Z},\|\cdot\|_{\mathbf{Z}})$. Indeed, if $\mathbf{X}$ is a closed reflexive subspace of ℓ_∞ and $\mathbf{V} \subseteq \mathbf{X}$ is a closed uncomplemented subspace of $\mathbf{X}$, then by [179] (see also [36, Corollary 7.3]) there is no Lipschitz retraction from $\mathbf{X}$ onto $\mathbf{V}$. Equivalently, the identity mapping from $\mathbf{V}$ to $\mathbf{V}$ cannot be extended to a Lipschitz mapping from $\mathbf{X}$ to $\mathbf{V}$. Hence, since $\mathrm{span}(\mathbf{V}) = \mathbf{V} \subseteq \ell_\infty$, we have $\mathsf{e}_{\mathrm{span}}(\mathbf{X},\mathbf{V};\ell_\infty) = \infty$. In contrast, $\mathsf{e}(\mathbf{X};\ell_\infty) = 1$ by the nonlinear Hahn–Banach theorem (see [206] or, e.g., [36, Lemma 1.1]). By combining [290] with the discretization method of [138] (see also [195]), one can quantify the above example by showing that for arbitrarily large $n \in \mathbb{N}$ there are Banach spaces $(\mathbf{X}, \|\cdot\|_{\mathbf{X}})$ and $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$, and a subset $\mathcal{C} \subseteq \mathbf{X}$ with $|\mathcal{C}| = n$ for which we have

$$\frac{\mathsf{e}_{\mathrm{span}}(\mathbf{X}, \mathcal{C}; \mathbf{Z})}{\mathsf{e}(\mathbf{X}, \mathcal{C}; \mathbf{Z})} \gtrsim \sqrt{\frac{\log n}{\log\log n}}. \tag{5.5}$$

(In fact, in (5.5) one can have $\mathsf{e}(\mathbf{X}, \mathcal{C}; \mathbf{Z}) = \mathsf{e}(\mathbf{X}; \mathbf{Z}) = 1$.) At present, the right-hand side of (5.5) is the largest asymptotic dependence on n that we are able to obtain for this question, and it remains an interesting open problem to determine the best possible asymptotics here.

Most, but not all, of the Lipschitz extension methods in the literature, including Kirszbraun's extension theorem [155], Ball's extension theorem [23] and methods that rely on (variants of) partitions of unity such as in [61, 140, 166, 173], yield convex hull-valued extensions, i.e., they actually provide bounds on the quantity $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z})$. Nevertheless, it seems likely that there is no $\varphi : [1, \infty) \to [1, \infty)$ such that $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}) \leqslant \varphi(\mathsf{e}(\mathfrak{M}))$ for every Polish metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, though if such an estimate were available, then it would be valuable; see, e.g., Remark 141. In fact, we propose the following conjecture.

Conjecture 134. There exists a Polish metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ for which $\mathsf{e}(\mathfrak{M}) < \infty$ yet $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}) = \infty$.

Remark 135. By definition, for every metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and every $\mathcal{C} \subseteq \mathfrak{M}$,

$$\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z}) \geqslant \mathsf{e}_{\mathrm{span}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z}) \geqslant \mathsf{e}(\mathfrak{M}, \mathcal{C}; \mathbf{Z}).$$

We explained in Remark 133 that the second of these inequalities can be strict (in a strong sense). However, as a complement to Conjecture 134, we state that to the best of our knowledge it is unknown whether this is so for the first of these inequalities, i.e., if it could happen that $\mathsf{e}_{\mathrm{span}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z}) < \infty$ yet $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z}) = \infty$. We suspect that this is possible, but if not, then it would be interesting to investigate how one could bound $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z})$ from above by a function of $\mathsf{e}_{\mathrm{span}}(\mathfrak{M}, \mathcal{C}; \mathbf{Z})$. We do know that there are a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, a Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$, a subset $\mathcal{C} \subseteq \mathfrak{M}$ and a Lipschitz mapping $f : \mathcal{C} \to \mathbf{Z}$ that can be extended to a Lipschitz mapping that takes values in $\overline{\mathrm{span}}(f(\mathcal{C}))$ but cannot be extended to a Lipschitz mapping that takes values in $\overline{\mathrm{conv}}(f(\mathcal{C}))$. To see this, let $\{e_j\}_{j=1}^{\infty}$ be the standard basis of ℓ_∞. For $n \in \mathbb{N}$ set $m(n) = n(n-1)/2$ and let $\mathbf{X}_n$ be the span of $\{e_{m(n)+1}, \ldots, e_{m(n+1)}\}$ in ℓ_∞. Thus, $\mathbf{X}_n$ is isometric to ℓ_∞^n and $\ell_\infty = (\bigoplus_{n=1}^{\infty} \mathbf{X}_n)_\infty$. By [290], there is a linear subspace $\mathbf{V}_n$ of $\mathbf{X}_n$ such that every linear projection $\mathsf{Q} : \mathbf{X}_n \to \mathbf{V}_n$ satisfies

$$\|\mathsf{Q}\|_{\mathbf{X}_n \to \mathbf{V}_n} \gtrsim \sqrt{n}.$$

By the method of [138], it follows that there exists[2] $\mathcal{A}_n \subseteq B_{V_n} = V_n \cap B_{\ell_\infty}$ with $|\mathcal{A}_n| \leqslant n^{O(n)}$ such that $\|F_n\|_{\mathrm{Lip}(X_n;V_n)} \gtrsim \sqrt{n}$ for any $F_n : \mathbf{X}_n \to \mathbf{V}_n$ that extends the formal identity $\mathsf{Id}_{\mathcal{A}_n \to \mathbf{V}_n} : \mathcal{A}_n \to \mathbf{V}_n$. By compactness, there exists $\delta_n \in (0,1)$ such that if we define

$$\mathcal{C}_n = \mathcal{A}_n \cup \left\{\delta_n e_{m(n)+1}, \dots, \delta_n e_{m(n+1)}\right\} \cup \{0\},$$

then also $\|\Phi_n\|_{\mathrm{Lip}(\mathbf{X}_n;\mathbf{X}_n)} \gtrsim \sqrt{n}$ for any mapping Φ_n from $\mathbf{X}_n$ to the polytope $\overline{\mathrm{conv}}(\mathcal{C}_n)$ that extends the formal identity $\mathsf{Id}_{\mathcal{C}_n \to \mathbf{X}_n}$. Consider the subset

$$\mathcal{C} = \dot{\bigcup_{n=1}^{\infty}} \mathcal{C}_n \subseteq \ell_\infty.$$

Suppose that $\Phi : \ell_\infty \to \overline{\mathrm{conv}}(\mathcal{C})$ extends $\mathsf{Id}_{\mathcal{C} \to \ell_\infty}$. Then, for each $n \in \mathbb{N}$ the mapping $\mathsf{R}_n \circ (\Phi|_{\mathbf{X}_n}) : \mathbf{X}_n \to \mathbf{X}_n$ extends $\mathsf{Id}_{\mathcal{C}_n \to \ell_\infty}$ and takes values in $\overline{\mathrm{conv}}(\mathcal{C}_n)$, where we denote the canonical restriction operator from ℓ_∞ to $\mathbf{X}_n$ by $\mathsf{R}_n : \ell_\infty \to \mathbf{X}_n$. Hence,

$$\|\Phi\|_{\mathrm{Lip}(\ell_\infty;\mathbf{X}_n)} \geqslant \|\mathsf{R}_n \circ (\Phi|_{\mathbf{X}_n})\|_{\mathrm{Lip}(\mathbf{X}_n;\mathbf{X}_n)} \gtrsim \sqrt{n}.$$

Since this holds for every $n \in \mathbb{N}$, the mapping Φ is not Lipschitz. Consequently, we have $\mathsf{e}_{\mathrm{conv}}(\ell_\infty, \mathcal{C}; \ell_\infty) = \infty$. At the same time, by construction we have $\overline{\mathrm{span}}(\mathcal{C}) = \overline{\mathrm{span}}(\{e_j\}_{j=1}^\infty) = c_0$ (recall that c_0 commonly denotes the subspace of ℓ_∞ consisting of all those sequences that tend to 0). So, any 2-Lipschitz retraction ρ of ℓ_∞ onto c_0 extends $\mathsf{Id}_{\mathcal{C} \to \ell_\infty}$ and takes values in $\overline{\mathrm{span}}(\mathcal{C})$; the existence of such a retraction ρ is due to [179] (see also [36, Example 1.5]). If $\mathsf{e}_{\mathrm{span}}(\ell_\infty, \mathcal{C}; \ell_\infty)$ were finite, then this example would answer the above question,[3] but we suspect that in fact $\mathsf{e}_{\mathrm{span}}(\ell_\infty, \mathcal{C}; \ell_\infty) = \infty$.

Proposition 136 is a convenient characterization of the quantities $\mathsf{e}(\mathfrak{M}, \mathcal{C})$ and $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{C})$; while it was not previously stated explicitly in this form, its proof is based on well-understood ideas.

Proposition 136. *Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space, $\mathcal{C}$ is a Polish subset of $\mathfrak{M}$ and $s_0 \in \mathcal{C}$. Fix two nonnegative functions $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0,\infty)$ and $\varepsilon : \mathcal{C} :\to [0,\infty)$. Then, the following two equivalences hold.*

---

[2]The subset $\mathcal{A}_n$ can be taken to be any ε_n-net of the unit sphere of $\mathbf{V}_n$, for any $\varepsilon_n \lesssim n^{-3/2}$. Note, however, that the bound that follows from [138] (and also [195, Appendix C]) is $\varepsilon_n \lesssim n^{-2}$, and this suffices for the present purposes; see [233, Theorem 23] for the above stated weaker requirement from ε_n.

[3]And, it would show that for arbitrarily large $k \in \mathbb{N}$ there exist a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$, a Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and a subset $\mathcal{S} \subseteq \mathfrak{M}$ with $|\mathcal{S}| = k$ such that $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}, \mathcal{S}; \mathbf{Z})/\mathsf{e}_{\mathrm{span}}(\mathfrak{M}, \mathcal{S}; \mathbf{Z}) \gtrsim \sqrt{(\log k)/\log\log k}$. It would then remain an interesting open question to determine the largest possible asymptotic dependence on k here.

(1) *The following two statements are equivalent.*

- *For every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every mapping* $f : \mathcal{C} \to \mathbf{Z}$ *that is* 1*-Lipschitz with respect to the metric* $d_{\mathfrak{M}}$ *there exists* $F : \mathfrak{M} \to \mathbf{Z}$ *that satisfies the following two conditions.*
  - $\|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.
- *There exists a family* $\{\phi_x\}_{x\in\mathfrak{M}}$ *of elements of the free space* $\mathfrak{F}(\mathcal{C})$ *with the following properties.*
  - $\|\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0}\|_{\mathfrak{F}(\mathcal{C})} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|\phi_x - \phi_y\|_{\mathfrak{F}(\mathcal{C})} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

(2) *The following two statements are equivalent.*

- *For every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every mapping* $f : \mathcal{C} \to \mathbf{Z}$ *that is* 1*-Lipschitz with respect to the metric* $d_{\mathfrak{M}}$ *there exists* $F : \mathfrak{M} \to \overline{\mathrm{conv}}(f(\mathcal{C}))$ *that satisfies the following two conditions.*
  - $\|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.
- *There exists a family* $\{\mu_x\}_{x\in\mathfrak{M}}$ *of probability measures in* $\mathsf{P}_1(\mathcal{C})$ *with the following properties.*
  - $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s)$ *for every* $s \in \mathcal{C}$.
  - $\mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

In the setting of Proposition 136, if $\varepsilon(s) = 0$ for every $s \in \mathcal{C}$ and also $\mathfrak{d} = K d_{\mathfrak{M}}$ for some $K \geqslant 1$, then in [11, Definition 2.7] a family $\{\phi_x\}_{x\in\mathfrak{M}} \subseteq \mathfrak{F}(\mathcal{C})$ as in part (1) of Proposition 136 is called a K-random projection of $\mathfrak{M}$ onto $\mathcal{C}$, and in [243, Definition 3.1] a family $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathcal{C})$ as in part (2) of Proposition 136 is called a stochastic K-Lipschitz retraction of $\mathfrak{M}$ onto $\mathcal{C}$ while in [11, Definition 2.7] it is called a strong K-random projection of $\mathfrak{M}$ onto $\mathcal{C}$.

Proof of Proposition 136. Suppose first that $\{\phi_x\}_{x\in\mathfrak{M}} \subseteq \mathfrak{F}(\mathcal{C})$ and $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathcal{C})$ are as in the two parts of Proposition 136. Let $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ be a Banach space and fix a 1-Lipschitz function $f : \mathcal{C} \to \mathbf{Z}$. Since $\mathcal{C}$ is Polish and hence separable, by replacing $\mathbf{Z}$ with the closure of the linear span of $f(\mathcal{C})$ we may assume that $\mathbf{Z}$ is separable. Recalling the notation (5.2) and the discussion immediately following it for the (integration) operator

$$\mathfrak{I}_f : \mathsf{M}_1(\mathfrak{M}) \cup \mathfrak{F}(\mathfrak{M}) \to \mathbf{Z},$$

define two (linear) mappings

$$\mathsf{Ext}_{\mathcal{C}}^{\phi} f, \mathsf{Ext}_{\mathcal{C}}^{\mu} f : \mathfrak{M} \to \mathbf{Z}$$

by setting for every $x \in \mathfrak{M}$,

$$\mathsf{Ext}_{\mathcal{C}}^{\phi} f(x) \stackrel{\mathrm{def}}{=} f(s_0) + \mathfrak{I}_f(\phi_x) \quad \text{and} \quad \mathsf{Ext}_{\mathcal{C}}^{\mu} f(x) \stackrel{\mathrm{def}}{=} \mathfrak{I}_f(\mu_x) \stackrel{(5.2)}{=} \int_{\mathcal{C}} f \, \mathrm{d}\mu_x. \tag{5.6}$$

Observe that since μ_x is a probability measure, $\mathsf{Ext}_{\mathcal{C}}^{\mu} f(x)$ belongs to the closure of the convex hull of $f(\mathcal{C})$.

For every $x, y \in \mathfrak{M}$ we have

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\phi} f(x) - \mathsf{Ext}_{\mathcal{C}}^{\phi} f(y)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_x - \phi_y)\right\|_{\mathbf{Z}} \stackrel{(5.3)}{\leqslant} \|\phi_x - \phi_y\|_{\mathfrak{F}(\mathcal{C})} \leqslant \mathfrak{d}(x, y),$$

and similarly (using Kantorovich–Rubinstein duality),

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\mu} f(x) - \mathsf{Ext}_{\mu}^{\phi} f(y)\right\|_{\mathbf{Z}} \leqslant \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y).$$

Also, for every $s \in \mathcal{C}$ we have

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\phi} f(s) - f(s)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0})\right\|_{\mathbf{Z}} \leqslant \|\phi_s - \boldsymbol{\delta}_s + \boldsymbol{\delta}_{s_0}\|_{\mathfrak{F}(\mathcal{C})} \leqslant \varepsilon(s),$$

and similarly,

$$\left\|\mathsf{Ext}_{\mathcal{C}}^{\mu} f(s) - f(s)\right\|_{\mathbf{Z}} = \left\|\mathfrak{I}_f(\phi_s - \boldsymbol{\delta}_s)\right\|_{\mathbf{Z}} \leqslant \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s).$$

Conversely, define $f : \mathcal{C} \to \mathfrak{F}(\mathcal{C})$ by setting $f(s) = \boldsymbol{\delta}_s - \boldsymbol{\delta}_{s_0}$ for each $s \in \mathcal{C}$. Then f is 1-Lipschitz. Fix $F : \mathfrak{M} \to \mathfrak{F}(\mathcal{C})$. Writing $F(x) = \phi_x$ for each $x \in \mathfrak{M}$, the assumptions of the first half of part (1) of Proposition 136 coincide with the assertions of its second half. As $\mathcal{C}$ is Polish, $\mathsf{P}_1(\mathcal{C})$ is closed in $\mathfrak{F}(\mathcal{C})$. Therefore,

$$\overline{\mathrm{conv}}\big(f(\mathcal{C})\big) = \mathsf{P}_1(\mathcal{C}) - \boldsymbol{\delta}_{s_0},$$

where the closure is with respect to the topology of $\mathfrak{F}(\mathcal{C})$. Thus, if

$$F(\mathfrak{M}) \subseteq \overline{\mathrm{conv}}\big(f(\mathcal{C})\big),$$

then $\mu_x \stackrel{\mathrm{def}}{=} F(x) + \boldsymbol{\delta}_{s_0} \in \mathsf{P}_1(\mathcal{C})$ and the assumptions of the first half of part (2) of Proposition 136 coincide with the assertions of its second half. ∎

The proof of Proposition 136 shows that even though in the first parts of the two equivalences in Proposition 136 one assumes merely the existence of an F with the desired properties, it follows that such an F can in fact be chosen to depend *linearly* on the input f, per (5.6).

Due to Proposition 136, the following question is closely related to Conjecture 134, though we think that it is also of independent interest.

Question 137. Characterize those Polish metric spaces $(\mathfrak{M}, d_{\mathfrak{M}})$ for which there exists a Lipschitz mapping $\rho : \mathfrak{F}(\mathfrak{M}) \to \mathsf{P}_1(\mathfrak{M})$ (recall that by default $\mathsf{P}_1(\mathfrak{M})$ is equipped with the Wasserstein-1 metric) and $x_0 \in \mathfrak{M}$ such that $\rho(\delta_y - \delta_{x_0}) = \delta_y$ for every $y \in \mathfrak{M}$.

5.3 Barycentric targets

Following [210], say that a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ is W_1-*barycentric* with constant $\beta > 0$ if there is a mapping $\mathfrak{B} : \mathsf{P}_1(\mathfrak{M}) \to \mathfrak{M}$ that satisfies $\mathfrak{B}(\boldsymbol{\delta}_x) = x$ for every $x, \in \mathfrak{M}$, and also

$$\forall \mu, \nu \in \mathsf{P}_1(\mathfrak{M}), \quad d_{\mathfrak{M}}\big(\mathfrak{B}(\mu), \mathfrak{B}(\nu)\big) \leqslant \beta \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu, \nu).$$

The infimal β for which this holds is denoted $\beta_1(\mathfrak{M})$. This notion (and variants thereof) were studied in various contexts; see, e.g., [17,33,94,119,165,173,178,210, 241,243,292]. Any normed space $\mathbf{X}$ is W_1-barycentric with constant 1, as seen by considering $\mathfrak{B}(\mu) = \int_{\mathbf{X}} x \,\mathrm{d}\mu(x)$. Other examples of spaces that are W_1-barycentric with constant 1 include Hadamard spaces and Busemann nonpositively curved spaces [57], or more generally spaces with a conical geodesic bicombing [86].

Thanks to Proposition 136, convex hull-valued (approximate) extension theorems automatically generalize to extension theorems for mappings that take value in W_1-barycentric metric spaces.

Proposition 138. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a metric space and let $\mathcal{C} \subseteq \mathfrak{M}$ be a Polish subset of $\mathfrak{M}$. Fix $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ and $\varepsilon : \mathcal{C} \to [0, \infty)$. Assume that for every Banach space $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ and every $f : \mathcal{C} \to \mathbf{Z}$ that is 1-Lipschitz with respect to $d_{\mathfrak{M}}$ there is $F : \mathfrak{M} \to \overline{\mathrm{conv}}(f(\mathcal{C}))$ that satisfies*

$$\forall s \in \mathcal{C}, \quad \|F(s) - f(s)\|_{\mathbf{Z}} \leqslant \varepsilon(s)$$

and

$$\forall x, y \in \mathfrak{M}, \quad \|F(x) - F(y)\|_{\mathbf{Z}} \leqslant \mathfrak{d}(x, y).$$

Fix $\eta : \mathcal{C} \to (1, \infty)$ and $\tau : \mathfrak{M} \times \mathfrak{M} \to (1, \infty)$, as well as $\beta > 0$ and a concave nondecreasing function $\omega : [0, \infty) \to [0, \infty)$ with $\omega(0) = 0$. If $(\mathfrak{N}, d_{\mathfrak{N}})$ is a W_1-barycentric metric space with constant β and $\phi : \mathcal{C} \to \mathfrak{N}$ has modulus of uniform continuity ω with respect to $d_{\mathfrak{M}}$, namely $d_{\mathfrak{N}}(f(s), f(t)) \leqslant \omega(d_{\mathfrak{M}}(s, t))$ for every $s, t \in \mathcal{C}$, then there is $\Phi : \mathfrak{M} \to \mathfrak{N}$ such that $d_{\mathfrak{N}}(\Phi(s), \phi(s)) \leqslant \omega(\eta(s)\varepsilon(s))$ for every $s \in \mathcal{C}$ and $d_{\mathfrak{N}}(\Phi(x), \Phi(y)) \leqslant \omega(\tau(x, y)\mathfrak{d}(x, y))$ for every $x, y \in \mathfrak{M}$.

Proof. By Proposition 136, there is a collection of measures $\{\mu_x\}_{x \in \mathfrak{M}} \subseteq \mathsf{P}_1(\mathcal{C})$ such that

$$\forall s \in \mathcal{C}, \quad \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_s, \boldsymbol{\delta}_s) \leqslant \varepsilon(s) \quad \text{and} \quad \forall x, y \in \mathfrak{M} \quad \mathsf{W}_1^{d_{\mathfrak{M}}}(\mu_x, \mu_y) \leqslant \mathfrak{d}(x, y).$$

Hence, for every $s \in \mathcal{C}$ and $x, y \in \mathfrak{M}$ there are couplings $\pi_s \in \Pi(\mu_s, \boldsymbol{\delta}_s)$ and $\pi_{x,y} \in \Pi(\mu_x, \mu_y)$ such that

$$\iint_{\mathcal{C} \times \mathcal{C}} d_{\mathfrak{M}}(u, v) \,\mathrm{d}\pi_s(u, v) \leqslant \eta(s)\varepsilon(s)$$

and

$$\iint_{\mathfrak{C}\times\mathfrak{C}} d_{\mathfrak{M}}(u,v)\,\mathrm{d}\pi_{x,y}(u,v) \leqslant \tau(x,y)\mathfrak{d}(x,y).$$

Since $(\phi\times\phi)_\#\pi_s \in \Pi(\phi_\#\mu_s,\phi_\#\boldsymbol{\delta}_s)$ and $(\phi\times\phi)_\#\pi_{x,y} \in \Pi(\phi_\#\mu_x,\phi_\#\mu_y)$, it follows that

$$\begin{aligned}
\mathsf{W}_1^{d_{\mathfrak{N}}}(\phi_\#\mu_s,\phi_\#\boldsymbol{\delta}_s) &\leqslant \iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}(a,b)\,\mathrm{d}(\phi\times\phi)_\#\pi_s(a,b)\\
&= \iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}\big(\phi(u),\phi(v)\big)\,\mathrm{d}\pi_s(u,v)\\
&\leqslant \iint_{\mathfrak{N}\times\mathfrak{N}} \omega\big(d_{\mathfrak{N}}(u,v)\big)\,\mathrm{d}\pi_s(u,v)\\
&\leqslant \omega\bigg(\iint_{\mathfrak{N}\times\mathfrak{N}} d_{\mathfrak{N}}(u,v)\,\mathrm{d}\pi_s(u,v)\bigg)\\
&\leqslant \omega\big(\eta(s)\varepsilon(s)\big),
\end{aligned}$$

where the penultimate step uses the concavity of ω. For the same reason, also

$$\mathsf{W}_1^{d_{\mathfrak{N}}}(\phi_\#\mu_x,\phi_\#\mu_y) \leqslant \omega\big(\tau(x,y)\mathfrak{d}(x,y)\big).$$

Since $(\mathfrak{N},d_{\mathfrak{N}})$ is β-barycentric there is $\mathfrak{B}:\mathsf{P}_1(\mathfrak{N})\to\mathfrak{N}$ satisfying $\mathfrak{B}(\boldsymbol{\delta}_z)=z$ for every $z,\in\mathfrak{N}$, and

$$\forall \nu_1,\nu_2\in\mathsf{P}_1(\mathfrak{N}),\quad d_{\mathfrak{N}}\big(\mathfrak{B}(\nu_1),\mathfrak{B}(\nu_2)\big) \leqslant \beta\mathsf{W}_1^{d_{\mathfrak{N}}}(\nu_1,\nu_2).$$

Define $\Phi:\mathfrak{M}\to\mathfrak{N}$ by

$$\forall x\in\mathfrak{M},\quad \Phi(x) \stackrel{\mathrm{def}}{=} \mathfrak{B}(\phi_\#\mu_x).$$

Then, for every $s\in\mathfrak{C}$ we have

$$d_{\mathfrak{N}}\big(\Phi(s),\phi(s)\big) \leqslant \beta\mathsf{W}_1^{d_{\mathfrak{N}}}\big(\phi_\#\mu_s,\phi_\#\boldsymbol{\delta}_s\big) \leqslant \omega\big(\eta(s)\varepsilon(s)\big).$$

For the same reason also $d_{\mathfrak{N}}(\Phi(x),\phi(y))\leqslant\omega(\tau(x,y)\mathfrak{d}(x,y))$ for every $x,y\in\mathfrak{M}$. ■

Because (as we will soon see) all of our new Lipschitz extension theorems are in fact bounds on $\mathsf{e}_{\mathrm{conv}}(\cdot)$, the following immediate corollary of Proposition 138 (with $\mathfrak{d}$ a multiple of $d_{\mathfrak{M}}$ and ω linear) shows that they apply to barycentric targets and not only to Banach space targets.

Corollary 139. *Fix $\beta>0$. Suppose that $\mathfrak{M}$ is a Polish metric space and that $\mathfrak{N}$ is a complete W_1-barycentric metric space with constant β. Then,*

$$\mathsf{e}_{\mathrm{conv}}(\mathfrak{M},\mathfrak{N}) \leqslant \beta\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}).$$

Another noteworthy special case of Proposition 138 is when $\omega(s) = s^\theta$ for some $0 < \theta \leqslant 1$, i.e., in the setting of Hölder extension that we discussed in Remark 15 and Section 2.3. Analogously to (1.18), we denote the convex hull-valued θ-Hölder extend modulus of a metric space $(\mathfrak{M}, d_{\mathfrak{M}})$ by

$$\mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M}) = \mathsf{e}_{\mathrm{conv}}\big(\mathfrak{M}, d^\theta_{\mathfrak{M}}\big).$$

Corollary 140. *Suppose that $\mathfrak{M}$ is a Polish metric space. Then, for every $0 < \theta \leqslant 1$ we have*

$$\mathsf{e}^\theta(\mathfrak{M}) \leqslant \mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M}) \leqslant \mathsf{e}_{\mathrm{conv}}(\mathfrak{M})^\theta.$$

Because the upper bound on $\mathsf{e}(\ell_\infty^n)$ that we obtain in Theorem 14 is actually an upper bound on $\mathsf{e}_{\mathrm{conv}}(\ell_\infty^n)$, Corollary 140 implies (1.19). More generally, Proposition 138 implies that

$$\mathsf{e}_{\mathrm{conv}}\big(\mathfrak{M}, \omega \circ d_{\mathfrak{M}}\big) \leqslant \sup_{d>0} \frac{\omega\big(\mathsf{e}_{\mathrm{conv}}(\mathfrak{M})d\big)}{\omega(d)} \leqslant \mathsf{e}_{\mathrm{conv}}(\mathfrak{M})$$

for any concave nondecreasing function $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$.

Remark 141. The question of how Lipschitz extension results imply extension results for other moduli of uniform continuity was studied in [224] and treated definitively by Brudnyi and Shvartsman in [65] using an interesting connection to the Brudnyĭ–Krugljak K-divisibility theorem [66] (see also [82]) from the theory of real interpolation of Banach spaces. In particular, by [65] we have $\mathsf{e}^\theta(\mathfrak{M}) \lesssim \mathsf{e}(\mathfrak{M})^2$, which remains the best-known bound on $\mathsf{e}^\theta(\mathfrak{M})$ in terms of $\mathsf{e}(\mathfrak{M})$ and it would be interesting to determine if it could be improved. As Corollary 140 shows that a better bound is available in terms of $\mathsf{e}^\theta_{\mathrm{conv}}(\mathfrak{M})$, Conjecture 134 and Question 137 could be relevant for this purpose.

5.4 Gentle partitions of unity

The following definition describes a numerical parameter that underlies the extension method of [173].

Definition 142 (Modulus of gentle partition of unity). Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. Define the *modulus of gentle partition of unity* of $\mathfrak{M}$ relative to $\mathcal{C}$, denoted $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}}; \mathcal{C})$ or simply $\mathsf{GPU}(\mathfrak{M}; \mathcal{C})$ when the metric is clear from the context, to be the infimum over those $\mathfrak{g} \in (0,\infty]$ such that for every $x \in \mathfrak{M}$ there is a Borel probability measure μ_x supported on $\mathcal{C}$ with the requirements that if $s \in \mathcal{C}$, then $\mu_s = \boldsymbol{\delta}_s$, and also for every $x, y \in \mathfrak{M}$ we have

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s,x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant \mathfrak{g}\, d_{\mathfrak{M}}(x,y).$$

The modulus of gentle partitions of unity of $\mathfrak{M}$, denoted $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}})$ or simply $\mathsf{GPU}(\mathfrak{M})$ when the metric is clear from the context, is defined to be the supremum of $\mathsf{GPU}(\mathfrak{M}, d_{\mathfrak{M}}; \mathcal{C})$ over all nonempty closed subsets $\mathcal{C} \subseteq \mathfrak{M}$.

The nomenclature of Definition 142 is derived from [173], though we warn that Definition 142 considers objects that are not identical to those that were introduced in [173]. In [173] the measures $\{\mu_x\}_{x \in \mathfrak{M} \smallsetminus \mathcal{C}}$ were also required to have a Radon–Nikoým derivative with respect to some reference measure μ. This additional requirement arises automatically from the constructions of [173] but it is not needed for any of the known applications of gentle partitions of unity, so it is beneficial to remove it altogether. The formal connection between [173] and Definition 142 was clarified in [11].

In anticipation of the proof of Theorem 66, one can generalize Definition 142 to the case of general profiles, analogously to what we did in Definition 64.

Definition 143 (Gentle partition of unity profile). Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. A metric $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ is called a *gentle partition of unity profile* for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to $\mathcal{C}$ if for every $x \in \mathfrak{M}$ there is a Borel probability measure μ_x supported on $\mathcal{C}$ with the requirements that if $s \in \mathcal{C}$, then $\mu_s = \boldsymbol{\delta}_s$, and also for every $x, y \in \mathfrak{M}$ we have

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant \mathfrak{d}(x, y).$$

If $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to every closed subset $\varnothing \neq \mathcal{C} \subseteq \mathfrak{M}$, then we say that $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$.

Note in passing that if $\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ relative to $\mathcal{C}$, then for every $x \in \mathfrak{M}$ the probability measure μ_x in Definition 143 has finite first moment. Indeed, for any $s_0 \in \mathcal{C}$,

$$\begin{aligned}\int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\mu_x(s) &= \int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\big(\mu_x - \delta_{s_0}\big)(s) \\ &\leqslant \int_{\mathcal{C}} d_{\mathfrak{M}}(s_0, s)\, \mathrm{d}\big|\mu_x - \mu_{s_0}\big|(s) \leqslant \mathfrak{d}(s_0, x) < \infty, \end{aligned} \tag{5.7}$$

where we used the fact that $\mu_{s_0} = \delta_{s_0}$, since $s_0 \in \mathcal{C}$.

Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space. The following estimate is implicit in [173]:

$$\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}) \leqslant 2\mathsf{GPU}(\mathfrak{M}).$$

In fact, the same reasoning as in [173] leads to the following more general lemma.

Lemma 144. *Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space and that $\mathcal{C} \subseteq \mathfrak{M}$ is nonempty and closed. Assume that $\mathfrak{d} : \mathfrak{M} \times \mathfrak{M} \to [0, \infty)$ is a gentle partition of unity* *profile for* $(\mathfrak{M}, d_{\mathfrak{M}})$ *relative to* $\mathfrak{C}$. *Then, for every Banach space* $(\mathbf{Z}, \|\cdot\|_{\mathbf{Z}})$ *and every* 1*-Lipschitz mapping* $f : \mathfrak{C} \to \mathbf{Z}$ *there exists*

$$F : \mathfrak{M} \to \overline{\mathrm{conv}}\big(f(\mathfrak{C})\big)$$

that extends f *and satisfies* $\|F(x) - F(y)\|_{\mathbf{Z}} \leqslant 2\mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

Proof. Let $\{\mu_x\}_{x\in\mathfrak{M}}$ be probability measures as in Definition 143. Then, $\{\mu_x\}_{x\in\mathfrak{M}} \subseteq \mathsf{P}_1(\mathfrak{C})$ by (5.7). So, by Proposition 136 (with $\varepsilon \equiv 0$) it suffices to check that for every $x, y \in \mathfrak{M}$ we have $\mathsf{W}_1(\mu_x, \mu_y) \leqslant 2\mathfrak{d}(x, y)$. To this end, fix $\eta > 0$ and $s_0 \in \mathfrak{C}$ such that $d_{\mathfrak{M}}(x, s_0) \leqslant d_{\mathfrak{M}}(x, \mathfrak{C}) + \eta$. Then, for every $s \in \mathfrak{C}$ we have

$$d_{\mathfrak{M}}(s, s_0) \leqslant d_{\mathfrak{M}}(s, x) + d_{\mathfrak{M}}(x, s_0) \leqslant d_{\mathfrak{M}}(s, x) + d_{\mathfrak{M}}(x, \mathfrak{C}) + \eta \leqslant 2d_{\mathfrak{M}}(s, x) + \eta.$$

Consequently, every 1-Lipschitz function $\psi : \mathfrak{C} \to \mathbb{R}$ satisfies

$$\begin{aligned}
\int_{\mathfrak{C}} \psi \,\mathrm{d}\mu_x - \int_{\mathfrak{C}} \psi \,\mathrm{d}\mu_y &= \int_{\mathfrak{C}} \big(\psi(s) - \psi(s_0)\big) \,\mathrm{d}(\mu_x - \mu_y)(s) \\
&\leqslant \int_{\mathfrak{C}} |\psi(s) - \psi(s_0)| \,\mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \int_{\mathfrak{C}} d_{\mathfrak{M}}(s, s_0) \,\mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \int_{\mathfrak{C}} (2d_{\mathfrak{M}}(s, x) + \eta) \,\mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant 2\mathfrak{d}(x, y) + 2\eta.
\end{aligned}$$

The desired conclusion follows by letting

$$\eta \to 0$$

and using the Kantorovich–Rubinstein duality (5.1). ∎

5.5 The multi-scale construction

Suppose that $(\mathfrak{M}, d_{\mathfrak{M}})$ is a Polish metric space and fix another metric $\mathfrak{d}$ on $\mathfrak{M}$. In this section we will show that there is a universal constant $\alpha \geqslant 1$ with the following property. Assume that either $(\mathfrak{M}, d_{\mathfrak{M}})$ is locally compact and $\mathfrak{d}$ is a separation modulus for $(\mathfrak{M}, d_{\mathfrak{M}})$ per Definition 64, or the assumptions of Theorem 114 are satisfied. We will prove that either of these assumptions implies that $\alpha\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$. By Lemma 144 this gives Theorems 66 and 114, and will show that in fact these extension results are both convex hull-valued and via a linear extension operator. This also implies that every locally compact metric space $\mathfrak{M}$ satisfies

$$\mathsf{GPU}(\mathfrak{M}) \lesssim \mathsf{SEP}(\mathfrak{M}). \tag{5.8}$$

Remark 145. The bound (5.8) need not be sharp. Indeed, it was proved in [173] that if $\mathfrak{M}$ is finite, then

$$\mathsf{GPU}(\mathfrak{M}) \lesssim \frac{\log|\mathfrak{M}|}{\log\log|\mathfrak{M}|}. \tag{5.9}$$

However, by [29] sometimes $\mathsf{SEP}(\mathfrak{M}) \gtrsim \log|\mathfrak{M}|$ (and always $\mathsf{SEP}(\mathfrak{M}) \lesssim \log|\mathfrak{M}|$). A shorter presentation of the proof of (5.9) can be found in [226], and a different proof of (5.9) will appear in the forthcoming work [207]. Also, in the forthcoming work [212] it is proved that (5.9) is optimal.

The following theorem is a precise formulation of what we will prove in this section.

Theorem 146. *Let $(\mathfrak{M}, d_{\mathfrak{M}})$ be a Polish metric space and fix another metric $\mathfrak{d}$ on $\mathfrak{M}$. Suppose that for every $\Delta > 0$ there is a probability space $(\Omega_\Delta, \mathbf{Prob}_\Delta)$ and a sequence of set-valued mappings $\{\Gamma^k_\Delta : \Omega_\Delta \to 2^{\mathfrak{M}}\}_{k=1}^\infty$ such that one of the following two measurability assumptions hold.*

- *Either $(\mathfrak{M}, d_{\mathfrak{M}})$ is locally compact and Γ^k_Δ is strongly measurable for each fixed $k \in \mathbb{N}$ and $\Delta > 0$,*
- *or Ω_Δ is a Borel subset of some Polish metric space $\mathfrak{Z}_\Delta$ and $\mathbf{Prob}_\Delta$ is a Borel probability measure supported on Ω_Δ, and Γ^k_Δ is a standard set-valued mapping for each fixed $k \in \mathbb{N}$ and $\Delta > 0$.*

Suppose that the following three requirements hold.

(1) $\mathcal{P}^\omega_\Delta = \{\Gamma^k_\Delta(\omega)\}_{k=1}^\infty$ *is a partition of* $\mathfrak{M}$ *for every* $\omega \in \Omega_\Delta$,

(2) $\mathrm{diam}_{\mathfrak{M}}(\mathcal{P}^\omega_\Delta(x)) < \Delta$ *for every* $x \in \mathfrak{M}$ *and* $\omega \in \Omega_\Delta$,

(3) $\Delta \mathbf{Prob}_\Delta\big[\omega \in \Omega_\Delta : \mathcal{P}^\omega_\Delta(x) \neq \mathcal{P}^\omega_\Delta(y)\big] \leqslant \mathfrak{d}(x, y)$ *for every* $x, y \in \mathfrak{M}$.

Then, $\alpha\mathfrak{d}$ is a gentle partition of unity profile for $(\mathfrak{M}, d_{\mathfrak{M}})$ for some universal constant $\alpha \in [1, \infty)$.

Suppose from now on that $\mathcal{C}$ is a nonempty *closed* subset of $\mathfrak{M}$. We will first set notation and record basic properties of a sequence of bump functions that will be used in the proof of Theorem 146; this part of the discussion is entirely standard and has nothing to do with random partitions.

Fix a 1-Lipschitz function $\psi : [0, \infty) \to [0, \infty)$ such that $\mathrm{supp}(\psi) \subseteq [1, 4]$ and $\psi(t) = 1$ for every $t \in [2, 3]$ (these requirements uniquely determine ψ, which is piecewise linear). Define for each $n \in \mathbb{Z}$,

$$\forall x \in \mathfrak{M}, \quad \phi_n(x) = \phi_n^{\mathcal{C}}(x) \stackrel{\mathrm{def}}{=} \psi\big(2^{-n} d_{\mathfrak{M}}(x, \mathcal{C})\big).$$

Then $\|\phi_n\|_{\mathrm{Lip}(\mathfrak{M})} \leqslant 2^{-n}$ and if $\phi_n(x) \neq 0$ then necessarily $2^n \leqslant d_{\mathfrak{M}}(x, \mathcal{C}) \leqslant 2^{n+2}$. We also denote

$$\forall x \in \mathfrak{M}, \quad \Phi(x) = \Phi^{\mathcal{C}}(x) \stackrel{\mathrm{def}}{=} \sum_{m \in \mathbb{Z}} \phi_n(x).$$

For each $x \in \mathfrak{M}$, at most two summands in the sum that defines $\Phi(x)$ do not vanish. If $x \in \mathfrak{M} \smallsetminus \mathcal{C}$, then since $\mathcal{C}$ is closed we have $d_{\mathfrak{M}}(x, \mathcal{C}) > 0$, and therefore there is $n \in \mathbb{Z}$ for which $2^n \leqslant d_{\mathfrak{M}}(x, \mathcal{C}) < 2^{n+1}$. For this value of n we have $\phi_n(x) = 1$, so $\Phi(x) \geqslant 1$ for every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$. Finally, for each $n \in \mathbb{Z}$ define

$$\forall x \in \mathfrak{M}, \qquad \lambda_n(x) = \lambda_n^{\mathcal{C}}(x) \stackrel{\mathrm{def}}{=} \begin{cases} \frac{\phi_n(x)}{\Phi(x)} & \text{if } x \in \mathfrak{M} \smallsetminus \mathcal{C}, \\ 0 & \text{if } x \in \mathcal{C}. \end{cases}$$

By design, $\sum_{n \in \mathbb{Z}} \lambda_n(x) = 1$ for every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$. Further properties of these bump functions are recorded in the following basic lemma, for ease of later reference.

Lemma 147. *Suppose that $x, y \in \mathfrak{M}$ satisfy $d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant d_{\mathfrak{M}}(y, \mathcal{C}) > d_{\mathfrak{M}}(x, y)$. Then for every $n \in \mathbb{Z}$,*

$$\frac{2^n}{d_{\mathfrak{M}}(y, \mathcal{C})} \notin \left(\frac{1}{4}, 2\right) \implies \phi_n(x) = \phi_n(y) = \lambda_n(x) = \lambda_n(y) = 0 \tag{5.10}$$

and

$$2^{n-1} < d_{\mathfrak{M}}(y, \mathcal{C}) < 2^{n+2} \implies \left|\lambda_n(x) - \lambda_n(y)\right| \lesssim \frac{d_{\mathfrak{M}}(x, y)}{d_{\mathfrak{M}}(y, \mathcal{C})}. \tag{5.11}$$

Proof. Our assumption implies that $d_{\mathfrak{M}}(x, \mathcal{C}), d_{\mathfrak{M}}(y, \mathcal{C}) > 0$, so $x, y \in \mathfrak{M} \smallsetminus \mathcal{C}$. To prove (5.10), suppose first that $2^n \geqslant 2 d_{\mathfrak{M}}(y, \mathcal{C})$. Then, since $\mathrm{supp}(\psi) \subseteq [1, 4]$ and $2^{-n} d_{\mathfrak{M}}(y, \mathcal{C}) \leqslant 1$ we have $\phi_n(y) = \lambda_n(y) = 0$. Also,

$$d_{\mathfrak{M}}(x, \mathcal{C}) \leqslant d_{\mathfrak{M}}(x, y) + d_{\mathfrak{M}}(y, \mathcal{C}) < 2 d_{\mathfrak{M}}(y, \mathcal{C}) \leqslant 2^n,$$

so $2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}) \leqslant 1$ and hence $\phi_n(x) = \lambda_n(x) = 0$. The remaining case of (5.10) is when $d_{\mathfrak{M}}(y, \mathcal{C}) \geqslant 2^{n+2}$. When this holds we have $2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant 2^{-n} d_{\mathfrak{M}}(y, \mathcal{C}) \geqslant 4$ and therefore $\{2^{-n} d_{\mathfrak{M}}(x, \mathcal{C}), 2^{-n} d_{\mathfrak{M}}(y, \mathcal{C})\} \cap \mathrm{supp}(\psi) = \emptyset$. Consequently, in this case we have $\phi_n(x) = \phi_n(y) = \lambda_n(x) = \lambda_n(y) = 0$.

To prove (5.11), assume that $2^{n-1} < d_{\mathfrak{M}}(y, \mathcal{C}) < 2^{n+2}$. Recalling that (pointwise) on $\mathfrak{M} \smallsetminus \mathcal{C}$ we have $\lambda_n = \phi_n / \Phi$ for all $n \in \mathbb{Z}$ and $\Phi \geqslant 1$, and moreover $\|\phi_n\|_{\mathrm{Lip}(\mathfrak{M})} \leqslant 2^{-n}$, we conclude as follows:

$$\begin{aligned} \left|\lambda_n(x) - \lambda_n(y)\right| &\leqslant \frac{\left|\phi_n(x) - \phi_n(y)\right|}{\Phi(x)} + \frac{\phi_n(y)}{\Phi(x)\Phi(y)} \left|\Phi(y) - \Phi(x)\right| \\ &\leqslant 2^{-n} d_{\mathfrak{M}}(x, y) + \sum_{n \in \mathbb{Z}} \left|\phi_n(x) - \phi_n(y)\right| \\ &\stackrel{(5.10)}{\leqslant} 2^{-n} d_{\mathfrak{M}}(x, y) + \sum_{\substack{n \in \mathbb{Z} \\ 2^{n-1} < d_{\mathfrak{M}}(y, \mathcal{C}) < 2^{n+2}}} 2^{-n} d_{\mathfrak{M}}(x, y) \\ &\asymp \frac{d_{\mathfrak{M}}(x, y)}{d_{\mathfrak{M}}(y, \mathcal{C})}. \end{aligned}$$

■

The interaction between $\{\lambda_n\}_{n\in\mathbb{Z}}$ and the random partitions of Theorem 146 is the content of the following lemma. Note that by reasoning as in (1.94), the metric $\mathfrak{d}$ in Theorem 146 must satisfy

$$\forall x, y \in \mathfrak{M}, \quad \mathfrak{d}(x,y) \geqslant d_{\mathfrak{M}}(x,y).$$

Lemma 148. *In the setting of Theorem 146, if $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $y \in \mathfrak{M} \smallsetminus \{x\}$ satisfy $d_{\mathfrak{M}}(x,\mathcal{C}) \geqslant d_{\mathfrak{M}}(y,\mathcal{C})$, then*

$$\sum_{n\in\mathbb{Z}}\sum_{k=1}^{\infty}\int_{\Omega_{2^n}}\left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x)-\lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \mathrm{d}\mathbf{Prob}_{2^n}(\omega) \lesssim \frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C})+d_{\mathfrak{M}}(x,y)}. \tag{5.12}$$

Proof. As $\sum_{n\in\mathbb{Z}}\lambda_n(x)=\sum_{n\in\mathbb{Z}}\lambda_n(y)=1$ and

$$\sum_{k=1}^{\infty}\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x)=\sum_{k=1}^{\infty}\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)=1$$

for every $n\in\mathbb{Z}$ and $\omega\in\Omega_{2^n}$, the left-hand side of (5.12) is at most 2. Since $\mathfrak{d}(x,y)\geqslant d_{\mathfrak{M}}(x,y)$, it follows that (5.12) holds if $d_{\mathfrak{M}}(y,\mathcal{C})\leqslant d_{\mathfrak{M}}(x,y)$. So, we will assume in the rest of the proof of Lemma 148 that $d_{\mathfrak{M}}(x,y)<d_{\mathfrak{M}}(y,\mathcal{C})$ (thus, in particular, $y\in\mathfrak{M}\smallsetminus\mathcal{C}$), in which case the right-hand side of (5.12) becomes at least a universal constant multiple of the quantity $\mathfrak{d}(x,y)/d_{\mathfrak{M}}(y,\mathcal{C})$.

We claim that for every $n\in\mathbb{Z}$ the following inequality holds for every $\omega\in\Omega_{2^n}$:

$$\sum_{k=1}^{\infty}\left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x)-\lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \lesssim \left(2^{-n}d_{\mathfrak{M}}(x,y)+\mathbf{1}_{\{\mathcal{P}^{\omega}_{2^m}(x)\neq\mathcal{P}^{\omega}_{2^n}(y)\}}\right)\mathbf{1}_{\{\frac14<\frac{2^n}{d_{\mathfrak{M}}(y,\mathcal{C})}<2\}}. \tag{5.13}$$

Assuming (5.13) for the moment, we will conclude the proof of (5.12) in the remaining case $d_{\mathfrak{M}}(x,y)<d_{\mathfrak{M}}(y,\mathcal{C})$ as follows:

$$\begin{aligned}&\sum_{n\in\mathbb{Z}}\sum_{k=1}^{\infty}\int_{\Omega_{2^n}}\left|\lambda_n(x)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(x)-\lambda_n(y)\mathbf{1}_{\Gamma^k_{2^n}(\omega)}(y)\right| \mathrm{d}\mathbf{Prob}_{2^n}(\omega)\\ &\lesssim \sum_{\substack{n\in\mathbb{Z}\\ 2^{n-1}<d_{\mathfrak{M}}(y,\mathcal{C})<2^{n+2}}}\left(2^{-n}d_{\mathfrak{M}}(x,y)+\mathbf{Prob}_{2^n}\left[\left\{\omega\in\Omega_{2^n}:\mathcal{P}^{\omega}_{2^n}(x)\neq\mathcal{P}^{\omega}_{2^n}(y)\right\}\right]\right)\\ &\lesssim \sum_{\substack{n\in\mathbb{Z}\\ 2^{mn-1}<d_{\mathfrak{M}}(y,\mathcal{C})<2^{n+2}}}2^{-n}\left(d_{\mathfrak{M}}(x,y)+\mathfrak{d}(x,y)\right)\\ &\asymp \frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C})}\asymp\frac{\mathfrak{d}(x,y)}{d_{\mathfrak{M}}(y,\mathcal{C})+d_{\mathfrak{M}}(x,y)},\end{aligned}$$

where the first step uses (5.13), the second step is where we used condition (3) of Theorem 146, the penultimate step uses $\mathfrak{d}(x,y) \geqslant d_{\mathfrak{M}}(x,y)$, and in the final step uses the assumption $d_{\mathfrak{M}}(x,y) < d_{\mathfrak{M}}(y,\mathcal{C})$.

It therefore remains to establish (5.13). By Lemma 147, if it is not the case that $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$, then $\lambda_n(x) = \lambda_n(y) = 0$, so both sides of (5.13) vanish. Thus, we may assume from now that $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$. Under this assumption, if $\mathcal{P}_{2^n}^{\omega}(x) \neq \mathcal{P}_{2^n}^{\omega}(y)$, then the right-hand side of (5.13) is at least 1, while the left-hand side of (5.13) consists of a sum of two numbers, each of which is at most 1. It therefore remains to establish (5.13) when $\mathcal{P}_{2^n}^{\omega}(x) = \mathcal{P}_{2^n}^{\omega}(y)$ (and still $2^{n-1} < d_{\mathfrak{M}}(y,\mathcal{C}) < 2^{n+2}$). In this case, (5.13) becomes the inequality $|\lambda_{2^n}(x) - \lambda_{2^n}(y)| \leqslant d_{\mathfrak{M}}(x,y)/d_{\mathfrak{M}}(y,\mathcal{C})$, which we proved in Lemma 147. ∎

Proof of Theorem 146. By Lemma 115 and Corollary 118, for every $\Delta > 0$ there exists a $\mathbf{Prob}_\Delta$-to-Borel measurable mapping $\gamma_\Delta^k : \Omega_m \to \mathcal{C}$ such that

$$\forall \omega \in \Omega_\Delta, \quad \Gamma_\Delta^k(\omega) \neq \varnothing \implies d_{\mathfrak{M}}\big(\gamma_\Delta^k(\omega), \Gamma_\Delta^k(\omega)\big) \leqslant d_{\mathfrak{M}}\big(\mathcal{C}, \Gamma_\Delta^k(\omega)\big) + \Delta. \quad (5.14)$$

(In fact, in the locally compact setting of Theorem 146, the use of Lemma 115 shows that the additive Δ term in the right-hand side of (5.14) can be removed).

For every $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ define a Borel measure μ_x supported on $\mathcal{C}$ by

$$\mu_x \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \big(\gamma_{2^n}^k\big)_{\#} \big(\mathbf{Prob}_{2^n} \lfloor_{\{\omega\in\Omega_{2^n} : x \in \Gamma_{2^n}^k(\omega)\}}\big). \quad (5.15)$$

In other words, for every Borel-measurable mapping $h : \mathcal{C} \to [0,\infty)$ we have

$$\int_{\mathcal{C}} h(s)\,\mathrm{d}\mu_x(s) = \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \int_{\{\omega\in\Omega_{2^n} : x\in\Gamma_{2^n}^k(\omega)\}} h\big(\gamma_{2^n}^k(\omega)\big)\,\mathrm{d}\mathbf{Prob}_{2^n}(\omega). \quad (5.16)$$

Since $\mathcal{P}_{2^n}^{\omega}$ is a partition of X for every $n \in \mathbb{Z}$ and $\omega \in \Omega_{2^n}$, the special case $h = \mathbf{1}_{\mathcal{C}}$ of (5.16) implies that

$$\begin{aligned}
\mu_x(\mathcal{C}) &= \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \lambda_n(x) \mathbf{Prob}_{2^n}\big[\{\omega \in \Omega_{2^n} : x \in \Gamma_{2^n}^k(\omega)\}\big] \\
&= \sum_{n\in\mathbb{Z}} \lambda_n(x) \mathbf{Prob}_{2^n}\left[\left\{\omega \in \Omega_{2^n} : x \in \bigcup_{k=1}^{\infty} \Gamma_{2^n}^k(\omega)\right\}\right] = \sum_{n\in\mathbb{Z}} \lambda_n(x) = 1.
\end{aligned}$$

Thus μ_x is a probability measure. Consequently, if we also denote $\mu_s = \boldsymbol{\delta}_s$ for every $s \in \mathcal{C}$, then the proof of Theorem 146 will be complete if we show that

$$\forall x, y \in \mathfrak{M}, \qquad \int_{\mathcal{C}} d_{\mathfrak{M}}(s,x)\,\mathrm{d}|\mu_x - \mu_y|(s) \lesssim \mathfrak{d}(x,y). \quad (5.17)$$

It suffices to prove (5.17) when $x, y \in \mathfrak{M}$ are distinct and $\{x, y\} \not\subseteq \mathcal{C}$. Indeed, if $\{x, y\} \subseteq \mathcal{C}$ then $\mu_x = \boldsymbol{\delta}_x$ and $\mu_y = \boldsymbol{\delta}_y$, so the left-hand side of (5.17) is equal to $d_{\mathfrak{M}}(x, y)$, which is at most $\mathfrak{d}(x, y)$. Hence, in the rest of the proof of Theorem 146 we will assume without loss of generality that $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $d_{\mathfrak{M}}(x, \mathcal{C}) \geqslant d_{\mathfrak{M}}(y, \mathcal{C})$.

We claim that the left-hand side of (5.17) can be bounded from above as follows:

$$\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \leqslant d_{\mathfrak{M}}(x, y) + \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\Big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_{2^n}(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\Big|\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega). \tag{5.18}$$

Indeed, if $x, y \in \mathfrak{M} \smallsetminus \mathcal{C}$, then μ_x, μ_y are defined according to (5.15), so that

$$\begin{aligned}
&\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x) \\
&\quad \mathrm{d}\Big(\big(\gamma_{2^n}^k\big)_{\#}\Big|\lambda_n(x)\mathbf{Prob}_{2^n}\lfloor_{\{\omega\in\Omega_{2^n}: x\in\Gamma_{2^n}^k(\omega)\}} - \lambda_n(y)\mathbf{Prob}_{2^n}\lfloor_{\{\omega\in\Omega_{2^n}: y\in\Gamma_{2^n}^k(\omega)\}}\Big|\Big)(s) \\
&= \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\Big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\Big|\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega),
\end{aligned}$$

thus establishing (5.18) in this case. The remaining case is when $x \in \mathfrak{M} \smallsetminus \mathcal{C}$ and $y \in \mathcal{C}$, so that μ_x is given in (5.15) and $\mu_y = \boldsymbol{\delta}_y$. We can then use the following (crude) estimate:

$$\begin{aligned}
&\int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}|\mu_x - \mu_y|(s) \\
&\leqslant \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}\mu_y(s) + \int_{\mathcal{C}} d_{\mathfrak{M}}(s, x)\, \mathrm{d}\mu_x(s) \\
&= d_{\mathfrak{M}}(x, y) + \sum_{n\in\mathbb{Z}} \sum_{k=1}^{\infty} \int_{\Omega_{2^n}} d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x)\, \mathrm{d}\mathbf{Prob}_{2^n}(\omega).
\end{aligned} \tag{5.19}$$

It remains to observe that because $y \in \mathcal{C}$ we have $\lambda_n(y) = 0$ for all $n \in \mathbb{Z}$ and therefore the right-hand side of (5.19) coincides with the right-hand side of (5.18).

Next, we claim that for every $(n, k) \in \mathbb{Z} \times \mathbb{N}$ and every $\omega \in \Omega_{2^n}$ we have

$$\begin{aligned}
&d_{\mathfrak{M}}\big(\gamma_{2^n}^k(\omega), x\big)\Big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\Big| \\
&\qquad \lesssim \big(d_{\mathfrak{M}}(y, \mathcal{C}) + d_{\mathfrak{M}}(x, y)\big)\Big|\lambda_n(x)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(x) - \lambda_n(y)\mathbf{1}_{\Gamma_{2^n}^k(\omega)}(y)\Big|.
\end{aligned} \tag{5.20}$$

By substituting the point-wise estimate (5.20) into (5.18) and using $d_{\mathfrak{m}}(x,y) \leqslant \mathfrak{d}(x,y)$ the desired estimate (5.17) follows from Lemma 148, thus completing the proof of Theorem 146.

To verify (5.20), note first that both sides of (5.20) vanish unless $x \in \Gamma^k_{2^n}(\omega)$ or $y \in \Gamma^k_{2^n}(\omega)$ and also, due to Lemma 147, $2^{n-1} < d_{\mathfrak{m}}(y,\mathcal{C}) < 2^{n+2}$. So, assume from now on that

$$\{x,y\} \cap \Gamma^k_{2^n}(\omega) \neq \varnothing \quad \text{and} \quad 2^{n-1} < d_{\mathfrak{m}}(y,\mathcal{C}) < 2^{n+2}. \tag{5.21}$$

Our goal (5.20) then becomes to deduce that

$$d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) \lesssim d_{\mathfrak{m}}(y,\mathcal{C}) + d_{\mathfrak{m}}(x,y). \tag{5.22}$$

Choose a point $z \in \Gamma^k_m(\omega)$ such that

$$\begin{aligned} d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) &\leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), \Gamma^k_{2^n}(\omega)\big) + 2^n \\ &\overset{(5.14)}{=} d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) + 2^{n+1} \\ &\overset{(5.21)}{\asymp} d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) + d_{\mathfrak{m}}(y,\mathcal{C}). \end{aligned} \tag{5.23}$$

If $x \in \Gamma^k_{2^n}(\omega)$, then

$$d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) \leqslant d_{\mathfrak{m}}(x,\mathcal{C}) \leqslant d_{\mathfrak{m}}(x,y) + d_{\mathfrak{m}}(y,\mathcal{C})$$

and

$$d_{\mathfrak{m}}(x,z) \leqslant \operatorname{diam}_{\mathfrak{m}}\big(\Gamma^k_{2^n}(\omega)\big) \leqslant 2^n \overset{(5.21)}{\asymp} d_{\mathfrak{m}}(y,\mathcal{C}).$$

By combining these two estimates with (5.23) and the triangle inequality, we see that

$$d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) \leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) + d_{\mathfrak{m}}(z,x) \lesssim d_{\mathfrak{m}}(x,y) + d_{\mathfrak{m}}(y,\mathcal{C}).$$

Hence, the desired estimate (5.22) holds when $x \in \Gamma^k_{2^n}(\omega)$.

It remains to check (5.22) when $y \in \Gamma^k_{2^n}(\omega)$, in which case we proceed similarly by noting that now

$$d_{\mathfrak{m}}\big(\mathcal{C}, \Gamma^k_{2^n}(\omega)\big) \leqslant d_{\mathfrak{m}}(y,\mathcal{C}),$$

and

$$d_{\mathfrak{m}}(y,z) \leqslant \operatorname{diam}_{\mathfrak{m}}\big(\Gamma^k_{2^n}(\omega)\big) \leqslant 2^n \overset{(5.21)}{\asymp} d_{\mathfrak{m}}(y,\mathcal{C}).$$

By combining these two estimates with (5.23) and the triangle inequality, we conclude that

$$\begin{aligned} d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), x\big) &\leqslant d_{\mathfrak{m}}\big(\gamma^k_{2^n}(\omega), z\big) + d_{\mathfrak{m}}(z,y) + d_{\mathfrak{m}}(y,x) \\ &\lesssim d_{\mathfrak{m}}(y,\mathcal{C}) + d_{\mathfrak{m}}(x,y). \end{aligned}$$

■

Chapter 6

Volume computations

In this section we will prove volume estimates that occur in our bounds on the separation modulus.

6.1 Direct sums

Fix $n \in \mathbb{N}$ and a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. Throughout what follows, the (normalized) *cone measure* [120] on $\partial B_{\mathbf{X}}$ will be denoted $\kappa_{\mathbf{X}}$. Thus, for every measurable $A \subseteq \partial B_{\mathbf{X}}$,

$$\kappa_{\mathbf{X}}(A) \stackrel{\mathrm{def}}{=} \frac{\mathrm{vol}_n([0,1]A)}{\mathrm{vol}_n(B_{\mathbf{X}})} = \frac{\mathrm{vol}_n(\{sv : (s,v) \in [0,1] \times A\})}{\mathrm{vol}_n(B_{\mathbf{X}})}. \tag{6.1}$$

The probability measure $\kappa_{\mathbf{X}}$ is characterized by the following "generalized polar coordinates" identity, which holds for every $f \in L_1(\mathbb{R}^n)$; see, e.g., [242, Proposition 1]:

$$\int_{\mathbb{R}^n} f(x)\,\mathrm{d}x = n\,\mathrm{vol}_n(B_{\mathbf{X}}) \int_0^\infty r^{n-1}\bigg(\int_{\partial B_{\mathbf{X}}} f(r\theta)\,\mathrm{d}\kappa_{\mathbf{X}}(\theta)\bigg)\,\mathrm{d}r. \tag{6.2}$$

As a quick application of (6.2), we will next record for ease of later reference the following computation of the volume of the unit ball of an ℓ_p direct sum of normed spaces.

Lemma 149. *Fix $n, m_1, \ldots, m_n \in \mathbb{N}$ and normed spaces $\{\mathbf{X}_j = (\mathbb{R}^{m_1}, \|\cdot\|_{\mathbf{X}_{m_j}})\}_{j=1}^n$. Then*

$$\forall\, p \in [1,\infty], \quad \mathrm{vol}_{m_1+\cdots+m_n}(B_{\mathbf{X}_1 \oplus_p \cdots \oplus_p \mathbf{X}_n}) = \frac{\prod_{j=1}^n \Gamma\big(1 + \frac{m_j}{p}\big)\,\mathrm{vol}_{m_j}\big(B_{\mathbf{X}_j}\big)}{\Gamma\big(1 + \frac{m_1+\cdots+m_n}{p}\big)}. \tag{6.3}$$

Proof. This follows by induction on n from the following identity (direct application of Fubini), which holds for every $a, b \in \mathbb{N}$ and any two normed spaces $\mathbf{X} = (\mathbb{R}^a, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^b, \|\cdot\|_{\mathbf{Y}})$:

$$\begin{aligned}\mathrm{vol}_{a+b}(B_{\mathbf{X}\oplus_p \mathbf{Y}}) &= \int_{B_{\mathbf{X}}} \mathrm{vol}_b\Big(\big(1 - \|x\|_{\mathbf{X}}^p\big)^{\frac{1}{p}} B_{\mathbf{Y}}\Big)\,\mathrm{d}x \\ &= \mathrm{vol}_b(B_{\mathbf{Y}}) \int_{B_{\mathbf{X}}} \big(1 - \|x\|_{\mathbf{X}}^p\big)^{\frac{b}{p}}\,\mathrm{d}x\end{aligned}$$

$$\overset{(6.2)}{=} \mathrm{vol}_a(B_{\mathbf{X}})\,\mathrm{vol}_b(B_{\mathbf{Y}}) \int_0^1 a r^{a-1}\big(1-r^p\big)^{\frac{b}{p}}\,\mathrm{d}r$$
$$= \mathrm{vol}_a(B_{\mathbf{X}})\,\mathrm{vol}_b(B_{\mathbf{Y}}) \frac{\Gamma\big(1+\frac{b}{p}\big)\Gamma\big(1+\frac{a}{p}\big)}{\Gamma\big(1+\frac{a+b}{p}\big)}. \qquad \blacksquare$$

By Lemma 149, for every $m \in \mathbb{N}$, every normed space $\mathbf{X} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{X}})$ satisfies

$$\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big) = \frac{\Gamma\big(1+\frac{m}{p}\big)^n}{\Gamma\big(1+\frac{nm}{p}\big)} \mathrm{vol}_m(B_{\mathbf{X}})^n, \tag{6.4}$$

and hence,

$$\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)^{\frac{1}{nm}} \asymp \frac{\mathrm{vol}_m(B_{\mathbf{X}})^{\frac{1}{m}}}{n^{\frac{1}{p}}}. \tag{6.5}$$

In particular, for every $m, n \in \mathbb{N}$ and $1 \leqslant p, q \leqslant \infty$ we have

$$\mathrm{vol}_{nm}\big(B_{\ell_p^n(\ell_q^m)}\big) = \frac{2^{nm}\Gamma\big(1+\frac{1}{q}\big)^{nm}\Gamma\big(1+\frac{m}{p}\big)^n}{\Gamma\big(1+\frac{m}{q}\big)^n\Gamma\big(1+\frac{nm}{p}\big)}, \tag{6.6}$$

and hence,

$$\mathrm{vol}_{nm}\big(B_{\ell_p^n(\ell_q^m)}\big)^{\frac{1}{nm}} \asymp \frac{1}{n^{\frac{1}{p}} m^{\frac{1}{q}}}. \tag{6.7}$$

The following simple lemma records an extension of (6.5) to m-fold iterations of the operation $\mathbf{X} \mapsto \ell_p^n(\mathbf{X})$, i.e., to spaces of the form

$$\ell_{p_m}^{n_m}\Big(\ell_{p_{m-1}}^{n_{m-1}}\big(\cdots \ell_{p_1}^{n_1}(\mathbf{X})\cdots\big)\Big);$$

the main point for us here is that the implicit constants remain bounded as $m \to \infty$.

Lemma 150. *Fix* $\{n_k\}_{k=0}^{\infty} \subseteq \mathbb{N}$ *and* $\{p_k\}_{k=1}^{\infty} \subseteq [1,\infty]$. *Let* $\mathbf{X} = (\mathbb{R}^{n_0}, \|\cdot\|_{\mathbf{X}})$ *be a normed space and define*

$$\forall k \in \mathbb{N}\cup\{0\}, \quad \mathbf{X}_{k+1} = \ell_{p_k}^{n_k}(\mathbf{X}_k), \quad \textit{where } \mathbf{X}_0 = \mathbf{X}.$$

Then, for every $m \in \mathbb{N}$ *we have*

$$\mathrm{vol}_{n_0\cdots n_m}\big(B_{\mathbf{X}_m}\big)^{\frac{1}{n_0\cdots n_k}} \asymp \frac{\mathrm{vol}_{n_0}\big(B_{\mathbf{X}}\big)^{\frac{1}{n_0}}}{\prod_{k=1}^m n_k^{\frac{1}{p_k}}}.$$

Proof. With the convention that an empty product equals 1, by applying (6.4) inductively we see that

$$\mathrm{vol}_{n_0\cdots n_m}\big(B_{\mathbf{X}_m}\big) = \mathrm{vol}_{n_0}\big(B_{\mathbf{X}}\big)^{n_1\cdots n_m} \prod_{k=1}^m \frac{\Gamma\big(1+\frac{n_0\cdots n_{k-1}}{p_k}\big)^{n_k\cdots n_m}}{\Gamma\big(1+\frac{n_0\cdots n_k}{p_k}\big)^{n_{k+1}\cdots n_m}}.$$

Hence,

$$\frac{\operatorname{vol}_{n_0\cdots n_m}\big(B_{\mathbf{X}_m}\big)^{\frac{1}{n_0\cdots n_k}}\prod_{k=1}^m n_k^{\frac{1}{p_k}}}{\operatorname{vol}_{n_0}\big(B_{\mathbf{X}}\big)^{\frac{1}{n_0}}} = \prod_{k=1}^m \frac{\Gamma\big(1+\frac{n_0\cdots n_{k-1}}{p_k}\big)^{\frac{1}{n_0\cdots n_{k-1}}}}{\Gamma\big(1+\frac{n_0\cdots n_k}{p_k}\big)^{\frac{1}{n_0\cdots n_k}}} n_k^{\frac{1}{p_k}}$$
$$= \prod_{k=1}^m f_{n_0\cdots n_{k-1},n_k}\left(\frac{1}{p_k}\right), \tag{6.8}$$

where for $u, v, t > 0$ we denote

$$f_{u,v}(t) \stackrel{\mathrm{def}}{=} \frac{\Gamma(1+ut)^{\frac{1}{u}}}{\Gamma(1+uvt)^{\frac{1}{uv}}} v^t.$$

Since $(\log\Gamma(z))' = \int_0^\infty \frac{se^{-zs}}{1-e^{-s}}\,\mathrm{d}s$ for $z > 0$ (see, e.g., [313, Chapter XII]), if $u, t > 0$ and $v \geqslant 1$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\log f_{u,v}(t) = \log v + \int_0^\infty \big(e^{-uts} - e^{-uvts}\big)\frac{se^{-s}}{1-e^{-s}}\,\mathrm{d}s \geqslant 0.$$

Thus, $f_{u,v}$ is increasing on $[0,\infty)$, and therefore we get from (6.8) that

$$1 = \prod_{k=1}^m f_{n_0\cdots n_{k-1},n_k}(0) \leqslant \frac{\operatorname{vol}_{n_0\cdots n_m}\big(B_{\mathbf{X}_m}\big)^{\frac{1}{n_0\cdots n_k}}\prod_{k=1}^m n_k^{\frac{1}{p_k}}}{\operatorname{vol}_{n_0}\big(B_{\mathbf{X}}\big)^{\frac{1}{n_0}}}$$
$$\leqslant \prod_{k=1}^m f_{n_0\cdots n_{k-1},n_k}(1) = \frac{(n_0!)^{\frac{1}{n_0}} n_1\cdots n_m}{\big((n_0\cdots n_m)!\big)^{\frac{1}{n_0\cdots n_m}}} \leqslant e. \qquad \blacksquare$$

The first part of Lemma 151 below is a restatement of Lemma 37 from the Introduction. Qualitatively, it shows that the class of spaces for which Conjecture 10 holds is closed under unconditional composition, namely, norms of the form (6.9) below. The second part of Lemma 151 is further information that pertains to Conjecture 49, i.e., to the symmetric version of the weak reverse isoperimetric conjecture, for which we want the operator S to be the identity mapping (i.e., weak reverse isoperimetry holds without the need to first change the "position" of the given normed space).

Lemma 151. *Fix $n, m_1, \ldots, m_n \in \mathbb{N}$. Let*

$$\mathbf{X}_1 = (\mathbb{R}^{m_1}, \|\cdot\|_{\mathbf{X}_1}), \ldots, \mathbf{X}_n = (\mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}_n})$$

be normed spaces. Also, let $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ be an unconditional normed space. Define a normed space $\mathbf{X} = (\mathbb{R}^{m_1}\times\cdots\times\mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}})$ by

$$\forall x = (x_1,\ldots,x_n) \in \mathbb{R}^{m_1}\times\cdots\times\mathbb{R}^{m_n}, \quad \|x\|_{\mathbf{X}} \stackrel{\mathrm{def}}{=} \big\|\big(\|x_1\|_{\mathbf{X}_1},\ldots,\|x_n\|_{\mathbf{X}_n}\big)\big\|_{\mathbf{E}}. \tag{6.9}$$

Then, Conjecture 10 (equivalently, Conjecture 35) holds for the space $\mathbf{X}$ *if it holds for all of the spaces* $\mathbf{X}_1, \ldots, \mathbf{X}_n$.

More precisely, suppose that there exist $S_1 \in \mathsf{SL}_{m_1}(\mathbb{R}), \ldots, S_n \in \mathsf{SL}_{m_n}(\mathbb{R})$, *normed spaces* $\mathbf{Y}_1 = (\mathbb{R}^{m_1}, \|\cdot\|_{\mathbf{Y}_1}), \ldots, \mathbf{Y}_n = (\mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{Y}_n})$, *and* $\alpha > 0$ *such that for every* $k \in \{1, \ldots, n\}$ *we have*

$$B_{\mathbf{Y}_k} \subseteq S_k B_{\mathbf{X}_k} \quad \text{and} \quad \frac{\mathrm{iq}(B_{\mathbf{Y}_k})}{\sqrt{m_k}} \left(\frac{\mathrm{vol}_{m_k}(B_{\mathbf{X}_k})}{\mathrm{vol}_{m_k}(B_{\mathbf{Y}_k})} \right)^{\frac{1}{m_k}} \leqslant \alpha. \tag{6.10}$$

Then, there exist a normed space $\mathbf{Y} = (\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}, \|\cdot\|_{\mathbf{X}})$ *and a linear transformation* $S \in \mathsf{SL}(\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n})$ *such that*

$$B_{\mathbf{Y}} \subseteq S B_{\mathbf{X}} \quad \text{and} \quad \frac{\mathrm{iq}(B_{\mathbf{Y}})}{\sqrt{m_1 + \cdots + m_n}} \left(\frac{\mathrm{vol}_{m_1+\cdots+m_n}(B_{\mathbf{X}})}{\mathrm{vol}_{m_1+\cdots+m_n}(B_{\mathbf{Y}})} \right)^{\frac{1}{m_1+\cdots+m_n}} \lesssim \alpha. \tag{6.11}$$

If furthermore $S_1, \ldots, S_n$ *are all identity mappings (of the respective dimensions), then* S *can be taken to be the identity mapping provided the following two conditions hold:*

$$\Big\| \sum_{i=1}^n e_i \Big\|_{\mathbf{E}} \Big\| \sum_{i=1}^n e_i \Big\|_{\mathbf{E}^*} \lesssim n \tag{6.12}$$

and

$$\left(\prod_{k=1}^n m_k^{m_k} \,\mathrm{vol}_{m_k}(B_{\mathbf{X}_k}) \right)^{\frac{1}{m_1+\cdots+m_n}} \lesssim \frac{m_1 + \cdots + m_n}{n} \min_{k \in \{1,\ldots,n\}} \mathrm{vol}_{m_k}(B_{\mathbf{X}_k})^{\frac{1}{m_k}}. \tag{6.13}$$

Note that (6.13) is satisfied in particular if $m_i \asymp m_j$ *and*

$$\mathrm{vol}_{m_i}(B_{\mathbf{X}_i})^{\frac{1}{m_i}} \asymp \mathrm{vol}_{m_j}(B_{\mathbf{X}_j})^{\frac{1}{m_j}}$$

for every $i, j \in \{1, \ldots, n\}$.

Prior to proving Lemma 151 we will make some basic observations. Firstly, (6.9) indeed defines a norm because it is well known that the requirement that $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ is an unconditional normed space is equivalent to (see, e.g., [181, Proposition 1.c.7]) the following "contraction property":

$$\forall a, x \in \mathbb{R}^n, \quad \|(a_1 x_1, \ldots, a_n x_n)\|_{\mathbf{E}} \leqslant \|a\|_{\ell_\infty^n} \|x\|_{\mathbf{E}}. \tag{6.14}$$

Thus, $\|x\|_{\mathbf{E}} \leqslant \|y\|_{\mathbf{E}}$ if $x, y \in \mathbb{R}^n$ satisfy $|x_i| \leqslant |y_i|$ for every $i \in \{1, \ldots, n\}$, so the triangle inequality for (6.9) follows from applying the triangle inequalities entry-wise for each of the norms $\{\|\cdot\|_{\mathbf{X}_i}\}_{i=1}^n$, using this monotonicity property, and then applying the triangle inequality for $\|\cdot\|_{\mathbf{E}}$.

It is well known that condition (6.12) holds (as an equality) when $\mathbf{E}$ is a symmetric normed space (see, e.g., [182, Proposition 3.a.6]). More generally, condition (6.12) holds (also as an equality) in the setting of the following simple averaging lemma, which shows in particular that Lemma 151 implies Lemma 53.

Lemma 152. *Let* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ *be a normed space such that for every two indices* $j, k \in \{1, \dots, n\}$ *there exists a permutation* $\pi = \pi_{jk} \in S_n$ *with* $\pi(j) = k$ *such that* $\|\sum_{i=1}^n a_{\pi(i)} e_i\|_{\mathbf{X}} = \|\sum_{i=1}^n a_i e_i\|_{\mathbf{X}}$ *for every* $a_1, \dots, a_n \in \mathbb{R}$. *Then,*

$$\left\|\sum_{i=1}^n e_i\right\|_{\mathbf{X}} \left\|\sum_{i=1}^n e_i\right\|_{\mathbf{X}^*} = n.$$

Proof. Denote $\mathfrak{S}(\mathbf{X}) = \{\pi \in S_n : T_\pi \in \mathsf{Isom}(\mathbf{X})\}$, where $T_\pi \in \mathsf{GL}_n(\mathbb{R})$ was defined in Example 40 for each $\pi \in S_n$. Then, $\mathfrak{S}(\mathbf{X})$ is a subgroup of S_n that we are assuming acts transitively on $\{1, \dots, n\}$. Consequently,

$$\forall i, j \in \{1, \dots, n\}, \quad |\{\pi \in \mathfrak{S}(\mathbf{X}) : \pi(i) = j\}| = \frac{|\mathfrak{S}(\mathbf{X})|}{n}. \tag{6.15}$$

For every $a_1, \dots, a_n \in \mathbb{R}$ we have

$$\begin{aligned}\frac{1}{|\mathfrak{S}(\mathbf{X})|} \sum_{\pi \in \mathfrak{S}(\mathbf{X})} \sum_{i=1}^n a_{\pi(i)} e_i &= \sum_{i=1}^n \left(\sum_{j=1}^n \frac{|\{\pi \in \mathfrak{S}(\mathbf{X}) : \pi(i) = j\}|}{|\mathfrak{S}(\mathbf{X})|} a_j \right) e_i \\ &\overset{(6.15)}{=} \frac{\sum_{j=1}^n a_j}{n} \sum_{i=1}^n e_i.\end{aligned}$$

Hence,

$$\begin{aligned}\left|\left\langle \sum_{j=1}^n e_j, \sum_{j=1}^n a_j e_j \right\rangle\right| &= \left|\sum_{j=1}^n a_j\right| \\ &= \frac{n \left\| \frac{1}{|\mathfrak{S}(\mathbf{X})|} \sum_{\pi \in \mathfrak{S}(\mathbf{X})} \sum_{i=1}^n a_{\pi(i)} e_i \right\|_{\mathbf{X}}}{\left\|\sum_{i=1}^n e_i\right\|_{\mathbf{X}}} \\ &\leqslant \frac{\frac{n}{|\mathfrak{S}(\mathbf{X})|} \sum_{\pi \in \mathfrak{S}(\mathbf{X})} \left\|\sum_{i=1}^n a_{\pi(i)} e_i\right\|_{\mathbf{X}}}{\left\|\sum_{i=1}^n e_i\right\|_{\mathbf{X}}} \\ &= \frac{n \left\|\sum_{i=1}^n a_i e_i\right\|_{\mathbf{X}}}{\left\|\sum_{i=1}^n e_i\right\|_{\mathbf{X}}},\end{aligned}$$

where the penultimate step uses convexity and the final step uses the assumption that T_π is an isometry of $\mathbf{X}$ for every $\pi \in \mathfrak{S}(\mathbf{X})$. Since this holds for every $a_1, \dots, a_n \in \mathbb{R}$, we have $\|\sum_{i=1}^n e_i\|_{\mathbf{X}^*} \leqslant n / \|\sum_{i=1}^n e_i\|_{\mathbf{X}}$. The reverse inequality holds for any normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ because $\langle \sum_{i=1}^n e_i, \sum_{i=1}^n e_i \rangle = n$. ∎

By combining Lemmas 151 and 152 we obtain the following corollary that establishes Conjecture 49 for the iteratively nested ℓ_p spaces of Lemma 150, provided it holds for the initial space $\mathbf{X}$.

Corollary 153. *Fix* $\{n_k\}_{k=0}^\infty \subseteq \mathbb{N}$ *and* $\{p_k\}_{k=1}^\infty \subseteq [1,\infty]$. *Let* $\mathbf{X} = (\mathbb{R}^{n_0}, \|\cdot\|_{\mathbf{X}})$ *be a normed space and define*

$$\forall k \in \mathbb{N}, \quad \mathbf{X}_{k+1} = \ell_{p_k}^{n_k}(\mathbf{X}_k), \quad \text{where } \mathbf{X}_0 = \mathbf{X}.$$

Suppose that $\alpha > 0$ *and there exists a normed space* $\mathbf{Y} = (\mathbb{R}^{n_0}, \|\cdot\|_{\mathbf{Y}})$ *with* $B_{\mathbf{Y}} \subseteq B_{\mathbf{X}}$ *and that satisfies*

$$\frac{\mathrm{iq}(B_{\mathbf{Y}})}{\sqrt{n_0}}\left(\frac{\mathrm{vol}_{n_0}(B_{\mathbf{X}})}{\mathrm{vol}_{n_0}(B_{\mathbf{Y}})}\right)^{\frac{1}{n_0}} \leqslant \alpha. \tag{6.16}$$

Then, for every $m \in \mathbb{N}$ *there is a normed space* $\mathbf{Y}_m = (\mathbb{R}^{n_0\cdots n_m}, \|\cdot\|_{\mathbf{Y}_m})$ *that satisfies* $B_{\mathbf{Y}_m} \subseteq B_{\mathbf{X}_m}$ *and*

$$\frac{\mathrm{iq}(B_{\mathbf{Y}_m})}{\sqrt{n_0\cdots n_m}}\left(\frac{\mathrm{vol}_{n_0\cdots n_m}(B_{\mathbf{X}_m})}{\mathrm{vol}_{n_0\cdots n_m}(B_{\mathbf{Y}_m})}\right)^{\frac{1}{n_0\cdots n_m}} \lesssim \alpha,$$

To see why Corollary 153 indeed follows from Lemmas 151 and 152, observe that if we start with $\mathbf{E}_0 = \mathbb{R}$ and define inductively $\mathbf{E}_{k+1} = \ell_{p_k}^{n_k}(\mathbf{E}_k)$, then for each $m \in \mathbb{N}$ the space $\mathbf{E}_m$ is unconditional and satisfies the assumptions of Lemma 152. The space $\mathbf{Y}_m$ of Corollary 153 is the same space that is defined in Lemma 151 if we take $\mathbf{E} = \mathbf{E}_m$, and also $\mathbf{X}_1 = \cdots = \mathbf{X}_m = \mathbf{X}$, which ensures that (6.13) holds.

Proof of Lemma 151. Denote

$$M \stackrel{\mathrm{def}}{=} \sum_{k=1}^n m_k = \dim(\mathbf{X}) \quad \text{and} \quad \forall k \in \{1,\ldots,n\}, \quad \rho_k \stackrel{\mathrm{def}}{=} \mathrm{vol}_{m_k}(B_{\mathbf{X}_k})^{\frac{1}{m_k}}. \tag{6.17}$$

Fix numbers $c, C_1, \ldots, C_n, \gamma_1, \ldots, \gamma_n, w_1, \ldots, w_n, w_1^*, \ldots, w_n^*, \beta_1, \ldots, \beta_n > 0$ that satisfy the following conditions (their values will be specified later). Firstly, we require that

$$\Big\|\sum_{i=1}^n w_i e_i\Big\|_{\mathbf{E}} = \Big\|\sum_{i=1}^n w_i^* e_i\Big\|_{\mathbf{E}^*} = 1. \tag{6.18}$$

Secondly, we require that

$$\forall k \in \{1,\ldots,n\}, \quad w_k w_k^* \geqslant \frac{m_k}{\gamma_k M}. \tag{6.19}$$

Finally, we require that

$$\forall k \in \{1,\ldots,n\}, \quad \frac{1}{c w_k \rho_k} \leqslant \beta_k \leqslant \frac{C_k}{w_k \rho_k}, \tag{6.20}$$

Denote

$$D \stackrel{\text{def}}{=} \left(\prod_{k=1}^{n} \beta_k^{m_k}\right)^{\frac{1}{M}}.$$

Consider the block diagonal linear operator $S : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n} \to \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}$ that is given by

$$\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}, \quad Sx \stackrel{\text{def}}{=} \frac{1}{D}\big(\beta_1 S_1 x_1, \ldots, \beta_n S_n x_n\big). \tag{6.21}$$

The normalization by D in (6.21) ensures that $S \in \mathsf{SL}(\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n})$.

Since $\sum_{k=1}^{n} w_k^* e_k$ is a unit functional in $\mathbf{E}^*$, we have

$$\begin{aligned}
\left\|S^{-1}x\right\|_{\mathbf{X}} &\stackrel{(6.9)\wedge(6.21)}{=} D\left\|\sum_{k=1}^{n} \frac{\|S_k^{-1}x_k\|_{\mathbf{X}_k}}{\beta_k} e_k\right\|_{\mathbf{E}} \\
&\stackrel{(6.18)}{\geqslant} D\left\langle \sum_{k=1}^{n} w_k^* e_k, \sum_{k=1}^{n} \frac{\|S_k^{-1}x_k\|_{\mathbf{X}_k}}{\beta_k} e_k \right\rangle \\
&\stackrel{(6.19)}{\geqslant} \frac{D}{M} \sum_{k=1}^{n} \frac{m_k \|S_k^{-1}x_k\|_{\mathbf{X}_k}}{\gamma_k w_k \beta_k},
\end{aligned}$$

for every $x = (x_1, \ldots, x_n) \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n}$. This shows that

$$\begin{aligned}
SB_{\mathbf{X}} &\subseteq \left\{ x \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_n} : \sum_{k=1}^{n} \frac{m_k \|S_k^{-1}x_k\|_{\mathbf{X}_k}}{\gamma_k w_k \beta_k} \leqslant \frac{M}{D} \right\} \\
&= \frac{M}{D} B_{\left(\frac{\gamma_1 w_1 \beta_1}{m_1} S_1 \mathbf{X}_1\right) \oplus_1 \cdots \oplus_1 \left(\frac{\gamma_n w_n \beta_n}{m_n} S_n \mathbf{X}_n\right)}.
\end{aligned}$$

Using Lemma 149, we therefore have

$$\begin{aligned}
\mathrm{vol}_M(B_{\mathbf{X}})^{\frac{1}{M}} &\leqslant \frac{M}{D} \,\mathrm{vol}_M\left(B_{\left(\frac{\gamma_1 w_1 \beta_1}{m_1} S_1 \mathbf{X}_1\right) \oplus_1 \cdots \oplus_1 \left(\frac{\gamma_n w_n \beta_n}{m_n} S_n \mathbf{X}_n\right)}\right)^{\frac{1}{M}} \\
&\stackrel{(6.3)}{=} \frac{1}{D}\left(\frac{M^M}{M!} \prod_{k=1}^{n} m_k! \left(\frac{\gamma_k w_k \beta_k \rho_k}{m_k}\right)^{m_k}\right)^{\frac{1}{M}} \\
&\stackrel{(6.20)}{\leqslant} \frac{1}{D}\left(\frac{M^M}{M!} \prod_{k=1}^{n} \frac{m_k!}{m_k^{m_k}} (\gamma_k C_k)^{m_k}\right)^{\frac{1}{M}} \\
&\leqslant \frac{e}{D}\left(\prod_{k=1}^{n} (\gamma_k C_k)^{m_k}\right)^{\frac{1}{M}}.
\end{aligned} \tag{6.22}$$

Next, for every $x=(x_1,\ldots,x_n)\in\mathbb{R}^{m_1}\times\cdots\times\mathbb{R}^{m_n}$ we have

$$\begin{aligned}\left\|S^{-1}x\right\|_{\mathbf{X}} &\overset{(6.9)\wedge(6.21)}{=} D\left\|\sum_{k=1}^{n}\frac{\|S_k^{-1}x_k\|_{\mathbf{X}_k}}{\beta_k}e_k\right\|_{\mathbf{E}}\\ &\overset{(6.14)}{\leqslant} D\left(\max_{k\in\{1,\ldots,n\}}\frac{\|S_k^{-1}x_k\|_{\mathbf{X}_k}}{w_k\beta_k}\right)\left\|\sum_{k=1}^{n}w_ke_k\right\|_{\mathbf{E}}\\ &\overset{(6.18)}{=} D\max_{k\in\{1,\ldots,n\}}\frac{\|S_k^{-1}x_k\|_{\mathbf{X}_k}}{w_k\beta_k}.\end{aligned}$$

This establishes the following inclusion:

$$SB_{\mathbf{X}}\supseteq\frac{1}{D}\prod_{k=1}^{n}w_k\beta_kS_kB_{\mathbf{X}_k}\overset{\mathrm{def}}{=}\Omega. \tag{6.23}$$

Thanks to (1.62), the assumption (6.10) of Lemma 151 implies that

$$\forall k\in\{1,\ldots,n\},\quad \lambda\big(S_kB_{\mathbf{X}_k}\big)\rho_k^2\overset{(6.17)}{=}\lambda\big(S_kB_{\mathbf{X}_k}\big)\,\mathrm{vol}_{m_k}\big(B_{\mathbf{X}_k}\big)^{\frac{2}{m_k}}\lesssim\alpha^2m_k. \tag{6.24}$$

For each $k\in\{1,\ldots,n\}$ take $f_k:S_kB_{\mathbf{X}_k}\to\mathbb{R}$ that is smooth on the interior of $S_kB_{\mathbf{X}_k}$, vanishes on $\partial S_kB_{\mathbf{X}_k}$, and satisfies $\Delta f_k=-\lambda(S_kB_{\mathbf{X}_k})f_k$ on the interior of $S_kB_{\mathbf{X}_k}$. Define $f:\Omega\to\mathbb{R}$ by

$$\forall x=(x_1,\ldots,x_n)\in\Omega=\frac{1}{D}\prod_{k=1}^{n}w_k\beta_kS_kB_{\mathbf{X}_k},\qquad f(x)\overset{\mathrm{def}}{=}\prod_{k=1}^{n}f_k\left(\frac{D}{w_k\beta_k}x_k\right),$$

Thus $f\equiv0$ on the boundary of Ω and on the interior of Ω it is smooth and satisfies

$$\Delta f=-D^2\left(\sum_{k=1}^{n}\frac{\lambda\big(S_kB_{\mathbf{X}_k}\big)}{(w_k\beta_k)^2}\right)f \tag{6.25}$$

Hence,

$$\begin{aligned}\lambda(S\mathbf{X})=\lambda(SB_{\mathbf{X}})&\overset{(6.23)}{\leqslant}\lambda(\Omega)\overset{(6.25)}{\leqslant}D^2\left(\sum_{k=1}^{n}\frac{\lambda\big(S_kB_{\mathbf{X}_k}\big)}{(w_k\beta_k)^2}\right)\\ &\overset{(6.20)}{\leqslant}(cD)^2\left(\sum_{k=1}^{n}\lambda\big(S_kB_{\mathbf{X}_k}\big)\rho_k^2\right)\overset{(6.24)}{\lesssim}(c\alpha D)^2M.\end{aligned} \tag{6.26}$$

By combining (6.22) and (6.26) we see that

$$\lambda(S\mathbf{X})\,\mathrm{vol}_M(B_{\mathbf{X}})^{\frac{2}{M}}\lesssim c^2\left(\prod_{k=1}^{n}(\gamma_kC_k)^{m_k}\right)^{\frac{2}{M}}\alpha^2M.$$

Another application of (1.62) now shows that the desired conclusion (6.11) holds with $\mathbf{Y} = \mathrm{Ch}\, S\mathbf{X}$ (recall the definition of Cheeger space in Section 1.6.1) provided

$$c\left(\prod_{k=1}^{n}(\gamma_k C_k)^{m_k}\right)^{\frac{1}{M}} \lesssim 1. \tag{6.27}$$

To get (6.11), by the Lozanovskiĭ factorization theorem [186] there exist weights $w_1, \ldots, w_n, w_1^*, \ldots, w_n^* > 0$ such that (6.18) holds and also $w_k w_k^* = m_k/M$ for every $k \in \{1, \ldots, n\}$. Thus (6.19) holds (as equality) if we choose $\gamma_1 = \cdots = \gamma_n = 1$. If we take $c = C_1 = \cdots = C_n = 1$ and $\beta_k = 1/(w_k \rho_k)$ for each $k \in \{1, \ldots, n\}$, then both (6.20) and (6.27) also hold (as equalities). With these choices, (6.11) holds.

Suppose that the additional assumptions (6.12) and (6.13) hold. Denote

$$\eta = \frac{1}{n}\left\|\sum_{i=1}^{n} e_i\right\|_{\mathbf{E}}\left\|\sum_{i=1}^{n} e_i\right\|_{\mathbf{E}^*}.$$

Thus, $\eta = O(1)$ by (6.12). Consider the weights $w_1 = \cdots = w_n = 1/\|\sum_{i=1}^n e_i\|_{\mathbf{E}}$ and $w_1^* = \cdots = w_n^* = 1/\|\sum_{i=1}^n e_i\|_{\mathbf{E}^*}$, so that (6.18) holds by design. This choice also ensures that if we take $\gamma_k = m_k/(\eta M)$ for each $k \in \{1, \ldots, n\}$, then (6.19) holds (as an equality). Next, choose $C_k = \rho_k$ for each $k \in \{1, \ldots, n\}$, as well as $\beta_1 = \cdots = \beta_n = \|\sum_{i=1}^n e_i\|_{\mathbf{E}}$ and $c = 1/\min_{k\in\{1,\ldots,n\}} \rho_k$. This ensures that (6.20) holds, and also that (6.27) coincides with the assumption (6.13), since $\eta = O(1)$. The desired conclusion (6.11) therefore holds with $Sx = (S_1x_1, \ldots, S_nx_n)$ in (6.21). In particular, if $S_k = \mathsf{Id}_{m_k}$ for every $k \in \{1, \ldots, n\}$, then we can take $S = \mathsf{Id}_{\mathbb{R}^{m_1}\times\cdots\times\mathbb{R}^{m_n}}$ in (6.11). ∎

The following lemma provides a formula for the cone measure of Orlicz spaces. Fix a convex increasing function $\psi : [0,\infty) \to [0,\infty]$ that satisfies $\psi(0) = 0$ and $\lim_{x\to\infty}\psi(x) = \infty$ (so, if $\lim_{x\to a^-}\psi(x) = \infty$ for some $a \in (0,\infty)$, then we require that $\psi(x) = \infty$ for every $x \geqslant a$). Henceforth, the associated Orlicz space (see, e.g., [268]) $\ell_\psi^n = (\mathbb{R}^n, \|\cdot\|_{\ell_\psi^n})$ will always be endowed with the Luxemburg norm that is given by

$$\forall x \in \mathbb{R}^n, \quad \|x\|_{\ell_\psi^n} = \inf\left\{s > 0 : \sum_{i=1}^{n}\psi\left(\frac{|x_i|}{s}\right) \leqslant 1\right\}. \tag{6.28}$$

Lemma 154. *Fix $n \in \mathbb{N}$. Suppose that $\psi : [0,\infty) \to [0,\infty]$ is convex, increasing, continuously differentiable on the set $\{x \in (0,\infty) : \psi(x) < \infty\}$, and satisfies $\lim_{x\to\infty}\psi(x) = \infty$ and $\psi(0) = 0$. Define a function $\varphi_\psi^n : \mathbb{R}^n \to [0,\infty)$ by setting*

$$\forall \tau = (\tau_1, \ldots, \tau_n) \in \mathbb{R}^n, \quad \varphi_\psi^n(\tau) \stackrel{\mathrm{def}}{=} \frac{\sum_{i=1}^n \psi^{-1}(|\tau_i|)\psi'\big(\psi^{-1}(|\tau_i|)\big)}{\prod_{i=1}^n \psi'\big(\psi^{-1}(|\tau_i|)\big)}. \tag{6.29}$$

Then, for every $g \in L_1(\kappa_{\ell_\psi^n})$ we have

$$\frac{n!}{2^n}\operatorname{vol}_n\big(B_{\ell_\psi^n}\big)\int_{\partial B_{\ell_\psi^n}} g(\theta)\,\mathrm{d}\kappa_{\ell_\psi^n}(\theta)$$
$$= \int_{\partial B_{\ell_1^n}} g\big(\psi^{-1}(|\tau_1|)\operatorname{sign}(\tau_1),\dots,\psi^{-1}(|\tau_n|)\operatorname{sign}(\tau_n)\big)\varphi_\psi^n(\tau)\,\mathrm{d}\kappa_{\ell_1^n}(\tau). \quad (6.30)$$

For example, when $\psi(t) = t^p$ for some $p \geqslant 1$ and every $t \geqslant 0$, in which case $\ell_\psi^n = \ell_p^n$, Lemma 154 gives

$$\int_{\partial B_{\ell_p^n}} g\,\mathrm{d}\kappa_{\ell_\psi^n} = \frac{\Gamma\big(1+\frac{n}{p}\big)}{n!\Gamma\big(1+\frac{1}{p}\big)^n}\int_{\partial B_{\ell_1^n}} \frac{g\circ M_{1\to p}^n(\tau)}{|\tau_1\cdots\tau_n|^{1-\frac{1}{p}}}\,\mathrm{d}\kappa_{\ell_1^n}(\tau),$$

where $M_{1\to p}:\mathbb{R}^n\to\mathbb{R}^n$ is the Mazur map [205] from ℓ_1^n to ℓ_p^n, i.e.,

$$\forall x\in\mathbb{R}^n,\quad M_{1\to p}^n(x_1,\dots,x_n) = \big(|x_1|^{\frac{1}{p}}\operatorname{sign}(x_1),\dots,|x_n|^{\frac{1}{p}}\operatorname{sign}(x_n)\big).$$

As another special case of Lemma 154, consider the following family of Orlicz spaces $\Omega_\beta^n = (\mathbb{R}^n, \|\cdot\|_{\Omega_\beta^n})$:

$$\forall \beta > 0,\quad \Omega_\beta^n \stackrel{\mathrm{def}}{=} \ell_{\psi_\beta}^n, \quad (6.31)$$

where

$$t\geqslant 0,\quad \psi_\beta(t) \stackrel{\mathrm{def}}{=} \begin{cases} \frac{1}{\beta}\log\big(\frac{1}{1-t}\big) & \text{if } 0\leqslant t<1,\\ \infty & \text{if } t\geqslant 1.\end{cases} \quad (6.32)$$

Observe that by considering the case $g\equiv 1$ of (6.30) we obtain the following identity:

$$\int_{\partial B_{\ell_\psi^n}} g\,\mathrm{d}\kappa_{\ell_\psi^n}$$
$$= \frac{\int_{\partial B_{\ell_1^n}} g\big(\psi^{-1}(|\tau_i|)\operatorname{sign}(\tau_1),\dots,\psi^{-1}(|\tau_n|)\operatorname{sign}(\tau_n)\big)\varphi_\psi^n(\tau)\,\mathrm{d}\kappa_{\ell_1^n}(\tau)}{\int_{\partial B_{\ell_1^n}}\varphi_\psi^n(\tau)\,\mathrm{d}\kappa_{\ell_1^n}(\tau)}, \quad (6.33)$$

where we recall that φ_ψ^n is defined in (6.29). When $\psi = \psi_\beta$ as in (6.32) for some $\beta > 0$ (we will eventually need to work with $\beta \asymp n$), for every $\tau\in\partial B_{\ell_1^n}$ we have

$$\varphi_{\psi_\beta}^n(\tau) = \frac{\sum_{i=1}^n \psi_\beta^{-1}(|\tau_i|)\psi_\beta'\big(\psi_\beta^{-1}(|\tau_i|)\big)}{\prod_{i=1}^n\psi_\beta'\big(\psi_\beta^{-1}(|\tau_i|)\big)} = \frac{\sum_{i=1}^n\big(1-e^{-\beta|\tau_i|}\big)\frac{e^{\beta|\tau_i|}}{\beta}}{\prod_{i=1}^n\frac{e^{\beta|\tau_i|}}{\beta}}$$
$$= \frac{\beta^{n-1}\sum_{i=1}^n\big(e^{\beta|\tau_i|}-1\big)}{e^{\beta\|\tau\|_{\ell_1^n}}} = \frac{\beta^{n-1}}{e^\beta}\sum_{i=1}^n\big(e^{\beta|\tau_i|}-1\big). \quad (6.34)$$

Consequently, (6.33) gives the following identity, which we will need later:

$$\int_{\partial B_{\Omega_\beta^n}} g \,\mathrm{d}\kappa_{\Omega_\beta^n} = \frac{\int_{\partial B_{\ell_1^n}} g\big((e^{\beta|\tau_1|}-1)\operatorname{sign}(\tau_1),\ldots,(e^{\beta|\tau_n|}-1)\operatorname{sign}(\tau_n)\big)\sum_{i=1}^n\big(e^{\beta|\tau_i|}-1\big)\,\mathrm{d}\kappa_{\ell_1^n}(\tau)}{\int_{\partial B_{\ell_1^n}}\sum_{i=1}^n\big(e^{\beta|\tau_i|}-1\big)\,\mathrm{d}\kappa_{\ell_1^n}(\tau)}.$$

Proof of Lemma 154. For each $i\in\{1,\ldots,n\}$ define $f_i:\mathbb{R}^n\to\mathbb{R}$ by setting $f_i(0)=0$ and

$$\forall y\in\mathbb{R}^n\smallsetminus\{0\},\qquad f_i(y)=\|y\|_{\ell_1^n}\psi^{-1}\left(\frac{|y_i|}{\|y\|_{\ell_1^n}}\right)\operatorname{sign}(y_i).$$

Consider $f=(f_1,\ldots,f_n):\mathbb{R}^n\to\mathbb{R}^n$. Then, $\|f(y)\|_{\ell_\psi^n}=\|y\|_{\ell_1^n}$ for every $y\in\mathbb{R}^n$. Hence, $f(B_{\ell_1^n})=B_{\ell_\psi^n}$. Now,

$$\begin{aligned}\int_{\partial B_{\ell_\psi^n}} g(\theta)\,\mathrm{d}\kappa_{\ell_\psi^n}(\theta) &\overset{(6.2)}{=} \frac{1}{\mathrm{vol}_n\big(B_{\ell_\psi^n}\big)}\int_{f(B_{\ell_1^n})} g\left(\frac{1}{\|x\|_{\ell_\psi^n}}x\right)\mathrm{d}x\\ &= \frac{1}{\mathrm{vol}_n\big(B_{\ell_\psi^n}\big)}\int_{B_{\ell_1^n}} g\left(\frac{1}{\|f(y)\|_{\ell_\psi^n}}f(y)\right)|\det f'(y)|\,\mathrm{d}y\\ &\overset{(6.2)}{=} \frac{\mathrm{vol}_n\big(B_{\ell_1^n}\big)}{\mathrm{vol}_n\big(B_{\ell_\psi^n}\big)}\int_{\partial B_{\ell_1^n}} g\big(f(\tau)\big)|\det f'(\tau)|\,\mathrm{d}\kappa_{\ell_1^n}(\tau),\end{aligned}$$

where in the final step we used the fact f is positively homogeneous of order 1, and hence its derivative is homogeneous of order 0 almost everywhere (f is continuously differentiable on $\{y\in\mathbb{R}^n;\ y_1,\ldots,y_n\neq 0\}$). Since the volume of the unit ball of ℓ_1^n equals $2^n/n!$, it remains to check that the Jacobian of f satisfies

$$\det f'(\tau)=\frac{\sum_{i=1}^n\psi^{-1}(|\tau_i|)\psi'\big(\psi^{-1}(|\tau_i|)\big)}{\prod_{i=1}^n\psi'\big(\psi^{-1}(|\tau_i|)\big)}=\varphi_\psi^n(\tau),$$

for every $\tau\in\partial B_{\ell_1^n}$ with $\tau_1,\ldots,\tau_n\neq 0$. This indeed holds because for every such τ and $i,j\in\{1,\ldots,n\}$ we have

$$\partial_j f_i(\tau)=\frac{\delta_{ij}-\tau_i\operatorname{sign}(\tau_j)}{\psi'\big(\psi^{-1}(|\tau_i|)\big)}+\psi^{-1}(|\tau_i|)\operatorname{sign}(\tau_i)\operatorname{sign}(\tau_j).$$

Hence, $f'(\tau)=A(\tau)+u(\tau)\otimes v(\tau)$, where $A(\tau)\in\mathsf{M}_n(\mathbb{R})$ is the diagonal matrix $\mathrm{Diag}((1/\psi'(\psi^{-1}(|\tau_i|)))_{i=1}^n)$ and the vectors $u(\tau),v(\tau)\in\mathbb{R}^n$ are defined by setting

$$u(\tau)=\left(\psi^{-1}(|\tau_i|)\operatorname{sign}(\tau_i)-\frac{\tau_i}{\psi'\big(\psi^{-1}(|\tau_i|)\big)}\right)_{i=1}^n,\qquad v(\tau)=\big(\operatorname{sign}(\tau_i)\big)_{i=1}^n\in\mathbb{R}^n.$$

By the textbook formula for the determinant of a rank-1 perturbation of an invertible matrix (e.g., [214, Section 6.2]), it follows that

$$\begin{aligned}\det f'(\tau) &= \big(1+\langle A(\tau)^{-1}u(\tau), v(\tau)\rangle\big)\det A(\tau)\\ &= \frac{1+\sum_{i=1}^n \psi'\big(\psi^{-1}(|\tau_i|)\big)\Big(\psi^{-1}(|\tau_i|)\operatorname{sign}(\tau_i)-\frac{\tau_i}{\psi'(\psi^{-1}(|\tau_i|))}\Big)\operatorname{sign}(\tau_i)}{\prod_{i=1}^n \psi'\big(\psi^{-1}(|\tau_i|)\big)}\\ &= \frac{\sum_{i=1}^n \psi^{-1}(|\tau_i|)\psi'\big(\psi^{-1}(|\tau_i|)\big)}{\prod_{i=1}^n \psi'\big(\psi^{-1}(|\tau_i|)\big)}. \qquad \blacksquare\end{aligned}$$

Another description of $\kappa_{\mathbf{X}}$ is the fact (see, e.g., [242, Lemma 1]) that the Radon–Nikodým derivative of the $(n-1)$-dimensional Hausdorff (non-normalized surface area) measure on $\partial B_{\mathbf{X}}$ with respect to the (non-normalized cone) measure $\mathrm{vol}_n(B_{\mathbf{X}})\kappa_{\mathbf{X}}$ is equal at almost every $x\in\partial B_{\mathbf{X}}$ to n times the Euclidean length of the gradient at x of the function $u\mapsto\|u\|_{\mathbf{X}}$. In other words, for any $g\in L_1(\partial B_{\mathbf{X}})$,

$$\int_{\partial B_{\mathbf{X}}} g(x)\,\mathrm{d}x = n\,\mathrm{vol}_n(B_{\mathbf{X}})\int_{\partial B_{\mathbf{X}}} g(x)\big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}\,\mathrm{d}\kappa_{\mathbf{X}}(x). \tag{6.35}$$

The special case $g\equiv 1$ of (6.35) gives the following identity:

$$\begin{aligned}\frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})} &= n\int_{\partial B_{\mathbf{X}}}\big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}\,\mathrm{d}\kappa_{\mathbf{X}}(x)\\ &= \fint_{B_{\mathbf{X}}}\frac{\|\nabla\|\cdot\|_{\mathbf{X}}(x)\|_{\ell_2^n}}{\|x\|_{\mathbf{X}}^{n-1}}\,\mathrm{d}x,\end{aligned} \tag{6.36}$$

where the second equality in (6.36) is an application of (6.2) because it is straightforward to check that $\|\nabla\|\cdot\|_{\mathbf{X}}(rx)\|_{\ell_2^n}=\|\nabla\|\cdot\|_{\mathbf{X}}(x)\|_{\ell_2^n}$ for any $r>0$ and $x\in\mathbb{R}^n$ at which the norm $\|\cdot\|_{\mathbf{X}}$ is smooth.

Remark 155. By applying Cauchy–Schwarz to the first equality in (6.36), we see that

$$\begin{aligned}\frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})} &\leqslant n\bigg(\int_{\partial B_{\mathbf{X}}}\|\nabla\|\cdot\|_{\mathbf{X}}(x)\|_{\ell_2^n}^2\,\mathrm{d}\kappa_{\mathbf{X}}(x)\bigg)^{\frac12}\\ &= \bigg(\frac{n}{\mathrm{vol}_n(B_{\mathbf{X}})}\int_{\partial B_{\mathbf{X}}}\big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}\,\mathrm{d}x\bigg)^{\frac12},\end{aligned} \tag{6.37}$$

where the final step of (6.37) is an application of (6.35) with $g(x)=\|\nabla\|\cdot\|_{\mathbf{X}}(x)\|_{\ell_2^n}$. If $\|\cdot\|_{\mathbf{X}}$ is twice continuously differentiable on $\mathbb{R}^n\smallsetminus\{0\}$ and $\varphi:\mathbb{R}\to[0,\infty)$ is twice continuously differentiable with $\varphi'(1)>0$ and $\varphi''(0)=0$, then because for every $x\in\partial B_{\mathbf{X}}$ the vector $\nabla\|\cdot\|_{\mathbf{X}}(x)/\|\nabla\|\cdot\|_{\mathbf{X}}(x)\|_{\ell_2^n}$ is the unit outer normal to $\partial B_{\mathbf{X}}$ at x, by the divergence theorem we have

$$\begin{aligned}\int_{\partial B_{\mathbf{X}}} \Delta\big(\varphi \circ \|\cdot\|_{\mathbf{X}}\big)(x)\,\mathrm{d}x &= \int_{\partial B_{\mathbf{X}}} \operatorname{div}\nabla\big(\varphi \circ \|\cdot\|_{\mathbf{X}}\big)(x)\,\mathrm{d}x \\ &= \int_{\partial B_{\mathbf{X}}} \frac{\langle \nabla(\varphi \circ \|\cdot\|_{\mathbf{X}})(x), \nabla\|\cdot\|_{\mathbf{X}}(x)\rangle}{\big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}}\,\mathrm{d}x \\ &= \int_{\partial B_{\mathbf{X}}} \varphi'\big(\|x\|_{\mathbf{X}}\big)\big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}\,\mathrm{d}x \\ &= \varphi'(1)\int_{\partial B_{\mathbf{X}}} \big\|\nabla\|\cdot\|_{\mathbf{X}}(x)\big\|_{\ell_2^n}\,\mathrm{d}x.\end{aligned}$$

A substitution of this identity into (6.37) give the following bound:

$$\frac{\operatorname{vol}_{n-1}(\partial B_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})} \leqslant \frac{\sqrt{n}}{\sqrt{\varphi'(1)}}\left(\fint_{\partial B_{\mathbf{X}}} \Delta\big(\varphi \circ \|\cdot\|_{\mathbf{X}}\big)(x)\,\mathrm{d}x\right)^{\frac{1}{2}}. \tag{6.38}$$

In particular, for every $p > 2$ we have

$$\frac{\operatorname{vol}_{n-1}(\partial B_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})} \leqslant \sqrt{\frac{n}{p}}\left(\fint_{B_{\mathbf{X}}} \Delta\big(\|\cdot\|_{\mathbf{X}}^p\big)(x)\,\mathrm{d}x\right)^{\frac{1}{2}}. \tag{6.39}$$

It is worthwhile to record (6.38) separately because this estimate is sometimes convenient for getting good bounds on $\operatorname{vol}_{n-1}(\partial B_{\mathbf{X}})$. In particular, by using (6.39) when $\mathbf{X}$ is an ℓ_p direct sum one can obtain an alternative derivation of some of the ensuing estimates. Another noteworthy consequence of (6.37) is when there is a transitive subgroup of permutations $G \leqslant S_n$ such that $\|(x_{\pi(1)}, \dots, x_{\pi(n)})\|_{\mathbf{X}} = \|x\|_{\mathbf{X}}$ for all $x \in \mathbb{R}^n$ and $\pi \in G$. Under this further symmetry assumption, the first inequality of (6.37) becomes

$$\frac{\operatorname{vol}_{n-1}(\partial B_{\mathbf{X}})}{\operatorname{vol}_n(B_{\mathbf{X}})} \leqslant n^{\frac{3}{2}}\left(\int_{\partial B_{\mathbf{X}}} \left(\frac{\partial\|\cdot\|_{\mathbf{X}}}{\partial x_1}(x)\right)^2 \mathrm{d}\kappa_{\mathbf{X}}(x)\right)^{\frac{1}{2}}.$$

The following lemma provides a probabilistic interpretation of the cone measure which generalizes the treatment of the special case $\mathbf{X} = \ell_p^n$ by Schechtman–Zinn [279] and Rachev–Rüschendorf [266].

Lemma 156 (Probabilistic representation of cone measure). *Fix $n \in \mathbb{N}$ and let $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ be a normed space. Suppose that $\varphi : [0, \infty) \to [0, \infty)$ is a continuous function such that $\varphi(0) = 0$, $\varphi(t) > 0$ when $t > 0$ and $\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r < \infty$. Let V be a random vector in $\mathbb{R}^n$ whose density at each $x \in \mathbb{R}^n$ is equal to*

$$\frac{1}{n\operatorname{vol}_n(B_{\mathbf{X}})\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r}\varphi\big(\|x\|_{\mathbf{X}}\big), \tag{6.40}$$

where we note that (6.40) *in indeed a probability density by* (6.2)*. Then, the density of* $\|\mathsf{V}\|_{\mathbf{X}}$ *at* $s \in [0,\infty)$ *is equal to* $s^{n-1}\varphi(s)/\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r$*. Moreover, the following two assertions hold:*

- $\mathsf{V}/\|\mathsf{V}\|_{\mathbf{X}}$ *is distributed according to the cone measure* $\kappa_{\mathbf{X}}$*,*
- $\|\mathsf{V}\|_{\mathbf{X}}$ *and* $\mathsf{V}/\|\mathsf{V}\|_{\mathbf{X}}$ *are (stochastically) independent.*

Proof. The density of $\|\mathsf{V}\|_{\mathbf{X}}$ at $s \in [0,\infty)$ is equal to

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{Prob}\big[\|\mathsf{V}\|_{\mathbf{X}} \leqslant s\big] &\overset{(6.40)}{=} \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{n\,\mathrm{vol}_n(B_{\mathbf{X}})\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r}\int_{sB_{\mathbf{X}}}\varphi\big(\|x\|_{\mathbf{X}}\big)\,\mathrm{d}x\right)\\
&\overset{(6.2)}{=} \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\int_0^s t\,r^{n-1}\varphi(r)\,\mathrm{d}r}{\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r}\right)\\
&= \frac{s^{n-1}\varphi(s)}{\int_0^\infty r^{n-1}\varphi(r)\,\mathrm{d}r}.
\end{aligned}
$$

The rest of Lemma 156 is equivalent to showing that for every measurable $A \subseteq \partial B_{\mathbf{X}}$ and $\rho > 0$,

$$
\mathbf{Prob}\left[\frac{\mathsf{V}}{\|\mathsf{V}\|_{\mathbf{X}}} \in A \mid \|\mathsf{V}\|_{\mathbf{X}} = \rho\right] = \kappa_{\mathbf{X}}(A).
$$

To prove this identity, observe first that for every $a, b \in \mathbb{R}$ with $a < b$ we have

$$
\begin{aligned}
\mathrm{vol}_n([a,b]A) &= \mathrm{vol}_n\Big(b\Big(\big([0,1]A\big) \smallsetminus \Big(\frac{a}{b}[0,1]A\Big)\Big)\Big)\\
&= (b^n - a^n)\,\mathrm{vol}_n([0,1]A).
\end{aligned}
$$

Hence, it follows from the definition (6.1) that

$$
\kappa_{\mathbf{X}}(A) = \frac{\mathrm{vol}_n([a,b]A)}{\mathrm{vol}_n([a,b]\partial B_{\mathbf{X}})}. \tag{6.41}
$$

Consequently,

$$
\begin{aligned}
\mathbf{Prob}\left[\frac{\mathsf{V}}{\|\mathsf{V}\|_{\mathbf{X}}} \in A \mid \|\mathsf{V}\|_{\mathbf{X}} = \rho\right] &= \lim_{\varepsilon \to 0}\frac{\mathbf{Prob}[\mathsf{V} \in \|\mathsf{V}\|_{\mathbf{X}}A \text{ and } \rho-\varepsilon \leqslant \|\mathsf{V}\|_{\mathbf{X}} \leqslant \rho+\varepsilon]}{\mathbf{Prob}[\rho-\varepsilon \leqslant \|\mathsf{V}\|_{\mathbf{X}} \leqslant \rho+\varepsilon]}\\
&= \lim_{\varepsilon \to 0}\frac{\int_{([0,\infty)A)\cap([\rho-\varepsilon,\rho+\varepsilon]\partial B_{\mathbf{X}})}\varphi(\|x\|_{\mathbf{X}})\,\mathrm{d}x}{\int_{[\rho-\varepsilon,\rho+\varepsilon]\partial B_{\mathbf{X}}}\varphi(\|x\|_{\mathbf{X}})\,\mathrm{d}x}\\
&= \lim_{\varepsilon \to 0}\frac{\mathrm{vol}_n([\rho-\varepsilon,\rho+\varepsilon]A)}{\mathrm{vol}_n([\rho-\varepsilon,\rho+\varepsilon]\partial B_{\mathbf{X}})}\\
&= \kappa_{\mathbf{X}}(A),
\end{aligned}
$$

where the penultimate step holds as φ is continuous at ρ and $\varphi(\rho) > 0$, and the final step uses (6.41). ∎

Lemma 157. *Fix $m, n \in \mathbb{N}$ and $p \in (1,\infty)$. Suppose that $\mathbf{X} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{X}})$ is a normed space. Let $\mathsf{R}_1, \ldots, \mathsf{R}_n$ be i.i.d. random variables taking values in $[0,\infty)$ whose density at each $t \in (0,\infty)$ is equal to*

$$\frac{p}{2(p-1)\Gamma\left(\frac{m}{p}\right)} t^{\frac{m}{2p-2}-1} e^{-t^{\frac{p}{2p-2}}}. \tag{6.42}$$

Then,

$$\frac{\mathrm{vol}_{nm-1}\left(\partial B_{\ell_p^n(\mathbf{X})}\right)}{\mathrm{vol}_{nm}\left(B_{\ell_p^n(\mathbf{X})}\right)} = \frac{p\Gamma\left(1+\frac{nm}{p}\right)}{\Gamma\left(1+\frac{nm-1}{p}\right)} \int_{(\partial B_{\mathbf{X}})^n} \mathbb{E}\left[\left(\sum_{i=1}^n \mathsf{R}_i \|\nabla\|\cdot\|_{\mathbf{X}}(x_i)\|_{\ell_2^m}^2\right)^{\frac12}\right] \mathrm{d}\kappa_{\mathbf{X}}^{\otimes n}(x_1,\ldots,x_n). \tag{6.43}$$

Furthermore,

$$\int_{\partial B_{\ell_p^n(\mathbf{X})}} \|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\|_{\ell_2^n(\ell_2^m)}^2 \,\mathrm{d}\kappa_{\ell_p^n(\mathbf{X})} = \frac{n\Gamma\left(\frac{nm}{p}\right)\Gamma\left(\frac{m+2p-2}{p}\right)}{\Gamma\left(\frac{m}{p}\right)\Gamma\left(\frac{nm+2p-2}{p}\right)} \int_{\partial B_{\mathbf{X}}} \left\|\nabla\|\cdot\|_{\mathbf{X}}\right\|_{\ell_2^m}^2 \mathrm{d}\kappa_{\mathbf{X}}. \tag{6.44}$$

Proof. For almost every $x = (x_1,\ldots,x_n) \in \ell_p^n(\mathbf{X})$ we have

$$\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}(x) = \frac{1}{\|x\|_{\ell_p^n(\mathbf{X})}^{p-1}}\left(\|x_1\|_{\mathbf{X}}^{p-1}\nabla\|\cdot\|_{\mathbf{X}}(x_1),\ldots,\|x_n\|_{\mathbf{X}}^{p-1}\nabla\|\cdot\|_{\mathbf{X}}(x_n)\right).$$

Consequently,

$$\begin{aligned}
&\|x\|_{\ell_p^n(\mathbf{X})}^{p-1}\left\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\left(\frac{x}{\|x\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^n(\ell_2^m)}\\
&= \left(\sum_{i=1}^n \|x_i\|_{\mathbf{X}}^{2p-2}\left\|\nabla\|\cdot\|_{\mathbf{X}}\left(\frac{x_i}{\|x\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^m}^2\right)^{\frac12}\\
&= \left(\sum_{i=1}^n \|x_i\|_{\mathbf{X}}^{2p-2}\left\|\nabla\|\cdot\|_{\mathbf{X}}\left(\frac{x_i}{\|x_i\|_{\mathbf{X}}}\right)\right\|_{\ell_2^m}^2\right)^{\frac12},
\end{aligned} \tag{6.45}$$

where we used the straightforward fact that the gradient of any (finite dimensional) norm is homogeneous of order 0 (on its domain of definition, which is almost everywhere).

Let

$$\mathsf{V} = (\mathsf{V}_1,\ldots,\mathsf{V}_n)$$

be a random vector on $\ell_p^n(\mathbf{X})$ whose density at $x=(x_1,\ldots,x_n)\in\ell_p^n(\mathbf{X})$ is

$$\frac{1}{\Gamma\big(1+\frac{nm}{p}\big)\,\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)}e^{-\|x\|_{\ell_p^n(\mathbf{X})}^p}=\frac{1}{\Gamma\big(1+\frac{nm}{p}\big)\,\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)}\prod_{i=1}^n e^{-\|x_i\|_{\mathbf{X}}^p}. \tag{6.46}$$

By combining Lemma 156 with the first equality in (6.36), we see that

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\mathbf{X})}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)}=nm\mathbb{E}\left[\left\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\left(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^n(\ell_2^m)}\right]. \tag{6.47}$$

Also, using the formula from Lemma 156 for the density of $\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}$, for $q>-nm$ we have

$$\mathbb{E}\big[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^q\big]=\frac{\int_0^\infty s^{nm+q-1}e^{-s^p}\,\mathrm{d}s}{\int_0^\infty r^{nm-1}e^{-r^p}\,\mathrm{d}r}=\frac{\Gamma\big(\frac{nm+q}{p}\big)}{\Gamma\big(\frac{nm}{p}\big)}. \tag{6.48}$$

Consequently,

$$\begin{aligned}&\mathbb{E}\left[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^{p-1}\left\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\left(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^n(\ell_2^m)}\right]\\&\quad=\mathbb{E}\big[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^{p-1}\big]\mathbb{E}\left[\left\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\left(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^n(\ell_2^m)}\right]\\&\quad=\frac{\Gamma\big(\frac{nm+p-1}{p}\big)}{nm\Gamma\big(\frac{nm}{p}\big)}\cdot\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\mathbf{X})}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)},\end{aligned} \tag{6.49}$$

where the first step of (6.49) uses the independence of $\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}$ and $\mathsf{V}/\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}$, by Lemma 156, and the final step of (6.49) is a substitution of (6.47) and the case $q=p-1$ of (6.48). Hence,

$$\begin{aligned}&\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\mathbf{X})}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)}\\&\quad=\frac{nm\Gamma\big(1+\frac{nm}{p}\big)}{\Gamma\big(1+\frac{nm-1}{p}\big)}\mathbb{E}\left[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^{p-1}\left\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\left(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\right)\right\|_{\ell_2^n(\ell_2^m)}\right]\\&\quad=\frac{p\Gamma\big(1+\frac{nm}{p}\big)}{\Gamma\big(1+\frac{nm-1}{p}\big)}\mathbb{E}\left[\left(\sum_{i=1}^n\|\mathsf{V}_i\|_{\mathbf{X}}^{2p-2}\left\|\nabla\|\cdot\|_{\mathbf{X}}\left(\frac{\mathsf{V}_i}{\|\mathsf{V}_i\|_{\mathbf{X}}}\right)\right\|_{\ell_2^m}^2\right)^{\frac12}\right],\end{aligned} \tag{6.50}$$

where in the last step we used the identity (6.45).

The product structure of the density of V in (6.46) means that $\mathsf{V}_1,\ldots,\mathsf{V}_n$ are (stochastically) independent. By Lemma 156, for each $i\in\{1,\ldots,n\}$ the random vector $\mathsf{V}_i/\|\mathsf{V}_i\|_{\mathbf{X}}$ is distributed on $\partial B_{\mathbf{X}}$ according to the cone measure $\kappa_{\mathbf{X}}$, and it is independent of the random variable

$$\mathsf{R}_i\stackrel{\mathrm{def}}{=}\|\mathsf{V}_i\|_{\mathbf{X}}^{2p-2}, \tag{6.51}$$

whose density at $t \in (0,\infty)$ is equal (using Lemma 156 once more) to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{Prob}\big[\|\mathsf{V}_i\|_{\mathbf{X}} \leqslant t^{\frac{1}{2p-2}}\big] &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^{t^{\frac{1}{2p-2}}} \frac{s^{m-1}e^{-s^p}}{\int_0^\infty r^{m-1}e^{-r^p}\,\mathrm{d}r}\,\mathrm{d}s\\ &= \frac{p}{2(p-1)\Gamma\big(\frac{m}{p}\big)}t^{\frac{m}{2p-2}-1}e^{-t^{\frac{p}{2p-2}}}.\end{aligned}$$

Hence, the identity (6.50) which we established above coincides with the desired identity (6.43).

To prove the identity (6.44), let R be a random variable whose density at each $t \in (0,\infty)$ is given by (6.42), i.e., $\mathsf{R}_1,\ldots,\mathsf{R}_n$ are independent copies of R. Then, for every $\alpha > -m/(2p-2)$ we have

$$\mathbb{E}\big[\mathsf{R}^\alpha\big] = \frac{p}{2(p-1)\Gamma\big(\frac{m}{p}\big)}\int_0^\infty t^{\frac{m}{2p-2}+\alpha-1}e^{-t^{\frac{p}{2p-2}}}\,\mathrm{d}t = \frac{\Gamma\big(2\alpha+\frac{m-2\alpha}{p}\big)}{\Gamma\big(\frac{m}{p}\big)}. \tag{6.52}$$

Using Lemma 156 (including the independence of $\mathsf{V}_i/\|\mathsf{V}_i\|_{\mathbf{X}}$ and $\|\mathsf{V}_i\|_{\mathbf{X}}$), we have

$$\begin{aligned}&\mathbb{E}\Bigg[\sum_{i=1}^n \|\mathsf{V}_i\|_{\mathbf{X}}^{2p-2}\Big\|\nabla\|\cdot\|_{\mathbf{X}}\Big(\frac{\mathsf{V}_i}{\|\mathsf{V}_i\|_{\mathbf{X}}}\Big)\Big\|_{\ell_2^m}^2\Bigg]\\ &= n\mathbb{E}[\mathsf{R}]\int_{\partial B_{\mathbf{X}}}\big\|\nabla\|\cdot\|_{\mathbf{X}}\big\|_{\ell_2^m}^2\,\mathrm{d}\kappa_{\mathbf{X}}\\ &= \frac{n\Gamma\big(\frac{m+2p-2}{p}\big)}{\Gamma\big(\frac{m}{p}\big)}\int_{\partial B_{\mathbf{X}}}\big\|\nabla\|\cdot\|_{\mathbf{X}}\big\|_{\ell_2^m}^2\,\mathrm{d}\kappa_{\mathbf{X}},\end{aligned} \tag{6.53}$$

where we recall (6.51) and the last step of (6.53) is the case $\alpha = 1$ of (6.52). At the same time,

$$\begin{aligned}&\mathbb{E}\Bigg[\sum_{i=1}^n \|\mathsf{V}_i\|_{\mathbf{X}}^{2p-2}\Big\|\nabla\|\cdot\|_{\mathbf{X}}\Big(\frac{\mathsf{V}_i}{\|\mathsf{V}_i\|_{\mathbf{X}}}\Big)\Big\|_{\ell_2^m}^2\Bigg]\\ &= \mathbb{E}\Bigg[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^{2p-2}\Big\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\Big(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\Big)\Big\|_{\ell_2^n(\ell_2^m)}^2\Bigg]\\ &= \mathbb{E}\big[\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}^{2p-2}\big]\mathbb{E}\Bigg[\Big\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\Big(\frac{\mathsf{V}}{\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}}\Big)\Big\|_{\ell_2^n(\ell_2^m)}^2\Bigg]\\ &= \frac{\Gamma\big(\frac{nm+2p-2}{p}\big)}{\Gamma\big(\frac{nm}{p}\big)}\int_{\partial B_{\ell_p^n(\mathbf{X})}}\big\|\nabla\|\cdot\|_{\ell_p^n(\mathbf{X})}\big\|_{\ell_2^n(\ell_2^m)}^2\,\mathrm{d}\kappa_{\ell_p^n(\mathbf{X})},\end{aligned} \tag{6.54}$$

where the first step of (6.54) uses the identity (6.45), the second step of (6.54) uses the independence of $\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}$ and $\mathsf{V}/\|\mathsf{V}\|_{\ell_p^n(\mathbf{X})}$ per Lemma 156, and the final step of uses the case $q = 2p - 2$ of (6.48) and Lemma 156. The desired identity (6.44) now follows by substituting (6.54) into (6.53). ■

The following lemma will have a central role in the proof of Theorems 24 and 48.

Lemma 158. *Suppose that $n, m \in \mathbb{N}$ and $\beta > 0$ satisfy $\beta \leqslant \frac{m-1}{2}$. Then,*

$$\forall 1 \leqslant p \leqslant m, \quad \mathrm{iq}\big(B_{\ell_p^n(\Omega_\beta^m)}\big) \asymp \sqrt{nm} = \sqrt{\dim\big(\ell_p^n(\Omega_\beta^m)\big)}.$$

Recall that the normed space $\Omega_\beta^m = (\mathbb{R}^m, \|\cdot\|_{\Omega_\beta^m})$ was defined in (6.31) and (6.32).

Prior to proving Lemma 158, we will show how it implies Theorem 48, and then deduce Theorem 24.

Proof of Theorem 48 assuming Lemma 158. By the assumption (1.73) of Theorem 48, write $n = km$ for some $k, m \in \mathbb{N}$ with $\max\{2, p\} \leqslant m \leqslant e^p$. Then $(m-1)/2 > 0$ and $m \geqslant p$, so we may apply Lemma 158 with n replaced by k and $\beta = (m-1)/2$. Denoting $\mathbf{Y} = \ell_p^k(\Omega_\beta^m)$, the conclusion of Lemma 158 is that $\mathrm{iq}(B_{\mathbf{Y}}) \asymp \sqrt{n}$.

$\mathbf{Y}$ is canonically positioned (it is a space from Example 40). To prove Theorem 48, it remains to check that $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\ell_p^n}$, where, since $n = km$, we identify $\mathbb{R}^n$ with $\mathsf{M}_{k\times n}(\mathbb{R})$, namely we identify ℓ_p^n with $\ell_p^k(\ell_p^m)$.

In fact, for any $\beta > 0$ (not only our choice $\beta = (m-1)/2$ above) we will check that

$$\forall x \in \mathbb{R}^m, \quad \big(1 - e^{-\frac{\beta}{m}}\big)\|x\|_{\Omega_\beta^m} \leqslant \|x\|_{\ell_\infty^m} \leqslant \|x\|_{\Omega_\beta^m}. \tag{6.55}$$

It follows from (6.55) that $\|\cdot\|_{\Omega_\beta^m} \asymp \|\cdot\|_{\ell_\infty^m}$ when $\beta \asymp m$. But, $\|\cdot\|_{\ell_p^m} \asymp \|\cdot\|_{\ell_\infty^m}$ by the assumption $e^p \geqslant m$. So,

$$\beta \asymp n \implies \|\cdot\|_{\mathbf{Y}} = \|\cdot\|_{\ell_p^k(\Omega_\beta^m)} \asymp \|\cdot\|_{\ell_p^k(\ell_\infty^m)} \asymp \|\cdot\|_{\ell_p^k(\ell_p^m)} = \|\cdot\|_{\ell_p^n}.$$

Fix $x \in \mathbb{R}^m$. To verify the second inequality in (6.55), the definition (6.32) gives $\sum_{i=1}^m \psi_\beta(|x_i|/s) = \infty$ when $0 < s \leqslant \|x\|_{\ell_\infty^m}$, so $\|x\|_{\Omega_\beta^m} \geqslant \|x\|_{\ell_\infty^m}$ by (6.28) and (6.31). For the first inequality in (6.55), by direct differentiation it is elementary to verify that the function $u \mapsto \log(1/(1-u))/u$ is increasing on the interval $[0, 1)$. Thus,

$$0 \leqslant t \leqslant \alpha < 1 \implies \psi_\beta(t) = \frac{1}{\beta}\log\left(\frac{1}{1-t}\right) \leqslant \frac{\log\left(\frac{1}{1-\alpha}\right)}{\alpha\beta} t.$$

Hence, for every fixed $0 < \alpha < 1$,

$$s \geqslant \frac{1}{\alpha}\|x\|_{\ell_\infty^m} \implies \sum_{i=1}^m \psi_\beta\left(\frac{|x_i|}{s}\right) \leqslant \sum_{i=1}^m \frac{\log\left(\frac{1}{1-\alpha}\right)}{\alpha\beta s}|x_i| \leqslant \frac{m\log\left(\frac{1}{1-\alpha}\right)}{\alpha\beta s}\|x\|_{\ell_\infty^m}. \tag{6.56}$$

Provided $\alpha \geqslant 1 - e^{-\beta/m}$, the choice $s = m\log(1/(1-\alpha))\|x\|_{\ell_\infty^m}/(\alpha\beta)$ satisfies the requirement $s \geqslant \|x\|_{\ell_\infty^m}/\alpha$, so we get from (6.28) and (6.56) that

$$\|x\|_{\Omega_\beta^m} \leqslant \frac{m\log\left(\frac{1}{1-\alpha}\right)}{\alpha\beta}\|x\|_{\ell_\infty^m}. \tag{6.57}$$

The optimal choice of α in (6.57) is $\alpha = 1 - e^{-\beta/m}$, giving the first inequality in (6.55). ∎

Having proved Theorem 48 (assuming Lemma 158, which we will soon prove), we have also already established Theorem 24 provided $n \in \mathbb{N}$ and $p \geqslant 1$ satisfy the divisor condition (1.73). Indeed, the space $\mathbf{Y}$ that Theorem 48 provides is canonically positioned and hence by the discussion in Section 1.6.2 it is also in its minimum surface area position, so by [104, Proposition 3.1] we have

$$\frac{\mathrm{MaxProj}(B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \asymp \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})\sqrt{n}} = \left(\frac{\mathrm{iq}(B_{\mathbf{Y}})}{\sqrt{n}}\right)\frac{1}{\mathrm{vol}_n(B_{\mathbf{Y}})^{\frac{1}{n}}} \asymp \frac{1}{\mathrm{vol}_n(B_{\ell_p^n})^{\frac{1}{n}}} \overset{(6.4)}{\asymp} n^{\frac{1}{p}},$$

where the penultimate step uses the fact that $\mathrm{iq}(B_{\mathbf{Y}}) \asymp \sqrt{n}$ by Theorem 48, and also that by Theorem 48 we have $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\ell_p^n}$, which implies that the nth root of the volume of the unit ball of $\mathbf{Y}$ is proportional to the nth root of the volume of the unit ball of ℓ_p^n.

The deduction of Theorem 24 for the remaining values of $p \geqslant 1$ and $n \in \mathbb{N}$ uses the following identity, which we will also use in the proof of Proposition 164 below.

Lemma 159. *Fix $n, m \in \mathbb{N}$. Suppose that $K \subseteq \mathbb{R}^n$ and $L \subseteq \mathbb{R}^m$ are convex bodies. Then,*

$$\frac{\mathrm{MaxProj}(K\times L)}{\mathrm{vol}_{n+m}(K\times L)} = \left(\frac{\mathrm{MaxProj}(K)^2}{\mathrm{vol}_n(K)^2} + \frac{\mathrm{MaxProj}(L)^2}{\mathrm{vol}_m(L)^2}\right)^{\frac{1}{2}}.$$

Proof. Fix $z \in S^{n+m-1}$. By Cauchy's projection formula [102] that we recalled in (1.30), we have

$$\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big) = \frac{1}{2}\int_{\partial(K\times L)} \big|\langle z, N_{K\times L}(w)\rangle\big|\,\mathrm{d}w,$$

where $N_{K\times L}(w)$ is the (almost-everywhere defined) unit outer normal to $\partial(K\times L)$ at $w \in \partial(K\times L)$. Now,

$$\partial(K\times L) = (\partial K\times L)\cup(K\times\partial L) \quad \text{and} \quad \mathrm{vol}_{n+m-1}\big((\partial K\times L)\cap(K\times\partial L)\big) = 0.$$

Consequently,

$$\begin{aligned}&\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big)\\ &\quad= \frac{1}{2}\int_{\partial K\times L}\big|\langle z, N_{K\times L}(w)\rangle\big|\,\mathrm{d}w + \frac{1}{2}\int_{K\times\partial L}\big|\langle z, N_{K\times L}(w)\rangle\big|\,\mathrm{d}w.\end{aligned}$$

If we write each $w \in \mathbb{R}^n$ as $w = (w_1, w_2)$ where $w_1 \in \mathbb{R}^n$ and $w_2 \in \mathbb{R}^m$, then for almost every (with respect to the $(n+m-1)$-dimensional Hausdorff measure) $w \in \partial K \times L$ we have $N_{K\times L}(w) = (N_K(w_1), 0)$. Also, $N_{K\times L}(w) = (0, N_L(w_2))$ for almost every $w \in K \times \partial L$. We therefore have

$$\begin{aligned}
&\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big)\\
&\quad= \frac{\mathrm{vol}_m(L)}{2}\int_{\partial K}\big|\langle z_1, N_K(x)\rangle\big|\,\mathrm{d}x + \frac{\mathrm{vol}_n(K)}{2}\int_{\partial L}\big|\langle z_2, N_L(y)\rangle\big|\,\mathrm{d}y\\
&\quad= \mathrm{vol}_m(L)\,\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z_1^\perp}K\big)\|z_1\|_{\ell_2^n} + \mathrm{vol}_n(K)\,\mathrm{vol}_{m-1}\big(\mathsf{Proj}_{z_2^\perp}L\big)\|z_2\|_{\ell_2^m},
\end{aligned}$$

where the last step is two applications of the Cauchy projection formula (in $\mathbb{R}^n$ and $\mathbb{R}^m$). Hence,

$$\begin{aligned}
&\frac{\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big)}{\mathrm{vol}_{n+m}(K\times L)}\\
&\quad= \frac{\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big)}{\mathrm{vol}_n(K)\,\mathrm{vol}_m(L)}\\
&\quad= \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{z_1^\perp}K\big)}{\mathrm{vol}_n(K)}\|z_1\|_{\ell_2^n} + \frac{\mathrm{vol}_{m-1}\big(\mathsf{Proj}_{z_2^\perp}L\big)}{\mathrm{vol}_m(L)}\|z_2\|_{\ell_2^m}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\frac{\mathrm{MaxProj}(K\times L)}{\mathrm{vol}_{n+m}(K\times L)}\\
&\quad= \max_{z\in S^{n+m-1}} \frac{\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{z^\perp}(K\times L)\big)}{\mathrm{vol}_{n+m}(K\times L)}\\
&\quad= \max_{(u,v)\in S^1}\max_{x\in S^{n-1}}\max_{y\in S^{m-1}} \frac{\mathrm{vol}_{n+m-1}\big(\mathsf{Proj}_{(ux+vy)^\perp}(K\times L)\big)}{\mathrm{vol}_{n+m}(K\times L)}\\
&\quad= \max_{(u,v)\in S^1}\max_{x\in S^{n-1}}\max_{y\in S^{m-1}} \left(\frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}K\big)}{\mathrm{vol}_n(K)}|u| + \frac{\mathrm{vol}_{m-1}\big(\mathsf{Proj}_{y^\perp}L\big)}{\mathrm{vol}_m(L)}|v|\right)\\
&\quad= \max_{(u,v)\in S^1}\left(\frac{\mathrm{MaxProj}(K)}{\mathrm{vol}_n(K)}|u| + \frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_m(L)}|v|\right)\\
&\quad= \left(\frac{\mathrm{MaxProj}(K)^2}{\mathrm{vol}_n(K)^2} + \frac{\mathrm{MaxProj}(L)^2}{\mathrm{vol}_m(L)^2}\right)^{\frac{1}{2}}. \qquad \blacksquare
\end{aligned}$$

We can now prove Theorem 24 in its full generality using the fact that we proved Theorem 48.

Proof of Theorem 24. Let $m \in \mathbb{N}$ satisfy $\max\{2, p\} \leqslant m \leqslant e^p$ (if $1 \leqslant p \leqslant 2$, then take $m = 2$, and if $p \geqslant 2$, then such an m exists because $e^p - p \geqslant e^2 - 2 > 5$). Write $n = km + r$ for some $k \in \mathbb{N}\cup\{0\}$ and $r \in \{0, \dots, m-1\}$. If $r = 0$, then m divides n

and we can conclude by applying Theorem 48 as we did above (recall the paragraph immediately before Lemma 159). So, assume from now that $r \geqslant 1$.

By Theorem 48 there is a canonically positioned normed space $\mathbf{Y} = (\mathbb{R}^{km}, \|\cdot\|_{\mathbf{Y}})$ such that $\mathrm{iq}(B_{\mathbf{Y}}) \asymp \sqrt{km}$ and $\|\cdot\|_{\mathbf{Y}} \asymp \|\cdot\|_{\ell_p^{km}}$. Define $\mathbf{Y}_p^n = \mathbf{Y} \oplus_\infty \Omega_\beta^r$, where $\beta \asymp r$ and $\mathrm{iq}(\Omega_\beta^r) \asymp \sqrt{r}$; such β exists trivially if $r = 1$, and if $r \geqslant 2$, then its existence follows from an application of Lemma 158 (with the choices $n = 1$ and $p = m = r$).

Since $\beta \asymp r$, by (6.55) we have $\|\cdot\|_{\Omega_\beta^r} \asymp \|\cdot\|_{\ell_\infty^r}$. Also, $\|\cdot\|_{\ell_\infty^r} \asymp \|\cdot\|_{\ell_p^r}$ since $e^p \geqslant m > r$. Consequently, for every $(x, y) \in \mathbb{R}^{km} \times \mathbb{R}^r$ we have

$$\max\{\|x\|_{\mathbf{Y}}, \|y\|_{\Omega_\beta^r}\} \asymp \max\{\|x\|_{\ell_p^{km}}, \|y\|_{\ell_p^r}\} \asymp \left(\|x\|_{\ell_p^{km}}^p + \|y\|_{\ell_p^r}^p\right)^{\frac{1}{p}}.$$

Recalling the definition of $\mathbf{Y}_p^n$, this means that $\|\cdot\|_{\mathbf{Y}_p^n} \asymp \|\cdot\|_{\ell_p^n}$.

Since both $\mathbf{Y}$ and Ω_β^r are canonically positioned and hence in their minimum surface area positions,

$$\frac{\mathrm{MaxProj}(B_{\mathbf{Y}})}{\mathrm{vol}_{km}(B_{\mathbf{Y}})} \asymp \Big(\frac{\mathrm{iq}(B_{\mathbf{Y}})}{\sqrt{km}}\Big)\frac{1}{\mathrm{vol}_{km}(B_{\mathbf{Y}})^{\frac{1}{km}}} \asymp \frac{1}{\mathrm{vol}_{km}\big(B_{\ell_p^{km}}\big)^{\frac{1}{km}}} \asymp (km)^{\frac{1}{p}}$$

and

$$\frac{\mathrm{MaxProj}\big(B_{\Omega_\beta^r}\big)}{\mathrm{vol}_r\big(B_{\Omega_\beta^r}\big)} \asymp \left(\frac{\mathrm{iq}\big(\Omega_\beta^r\big)}{\sqrt{r}}\right)\frac{1}{\mathrm{vol}\big(\Omega_\beta^r\big)^{\frac{1}{r}}} \asymp \frac{1}{\mathrm{vol}\big(\ell_\infty^r\big)^{\frac{1}{r}}} \asymp 1 \asymp r^{\frac{1}{p}}.$$

Consequently, since $B_{\mathbf{Y}_p^n} = B_{\mathbf{Y}} \times B_{\Omega_\beta^r}$, by Lemma 159 we conclude that

$$\begin{aligned}\frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}_p^n}\big)}{\mathrm{vol}_n\big(B_{\mathbf{Y}_p^n}\big)} &= \left(\frac{\mathrm{MaxProj}(B_{\mathbf{Y}})^2}{\mathrm{vol}_{km}(B_{\mathbf{Y}})^2} + \frac{\mathrm{MaxProj}\big(B_{\Omega_\beta^r}\big)^2}{\mathrm{vol}_r\big(B_{\Omega_\beta^r}\big)^2}\right)^{\frac{1}{2}} \\ &\asymp \big((km)^{\frac{2}{p}} + r^{\frac{2}{p}}\big)^{\frac{1}{2}} \asymp (km + r)^{\frac{1}{p}} = n^{\frac{1}{p}}. \qquad \blacksquare\end{aligned}$$

The following lemma will be used in the proof of Lemma 158.

Lemma 160. *Suppose that $m \in \mathbb{N}$, $r \in \mathbb{N} \cup \{0\}$ and $\beta > 0$ satisfy $\beta \leqslant \frac{m+r-2}{2}$. Then*

$$\begin{aligned}\int_{\partial B_{\ell_1^m}} \left(e^{\beta|\tau_1|} - \sum_{k=r-1}^{\infty} \frac{\beta^k|\tau_1|^k}{k!}\right) \mathrm{d}\kappa_{\ell_1^m}(\tau) &= \int_{\partial B_{\ell_1^m}} \left(\sum_{k=r}^{\infty} \frac{\beta^k|\tau_1|^k}{k!}\right) \mathrm{d}\kappa_{\ell_1^m}(\tau) \\ &\asymp \frac{\beta^r(m-1)!}{(m+r-1)!}. \end{aligned} \tag{6.58}$$

Proof. Let $\mathsf{H}_1, \ldots, \mathsf{H}_m$ be independent random variables whose density at each $s \in \mathbb{R}$ is equal to $e^{-|s|}/2$. Then, $|\mathsf{H}_1|, \ldots, |H_m|$ are exponential random variables of rate 1, and therefore if we denote

$$\Gamma \stackrel{\text{def}}{=} \sum_{i=1}^{m} |\mathsf{H}_i|,$$

then Γ has $\Gamma(m,1)$ distribution, i.e., its density at $s \geqslant 0$ equals $s^{m-1}e^{-s}/(m-1)!$; the proof of this standard probabilistic fact can be found in, e.g., [89]. By [266, 279] (or Lemma 156), the random vector $(\mathsf{H}_1, \ldots, \mathsf{H}_m)/\Gamma$ is distributed according to $\kappa_{\ell_1^m}$ and is independent of Γ. Thus, for every $k \in \mathbb{N}$,

$$\begin{aligned}\int_{\partial B_{\ell_1^m}} |\tau_1|^k \,\mathrm{d}\kappa_{\ell_1^m}(\tau) &= \mathbb{E}\left[\frac{|\mathsf{H}_1|^k}{\Gamma^k}\right] = \frac{\mathbb{E}\left[|\mathsf{H}_1|^k\right]}{\mathbb{E}\left[\Gamma^k\right]} \\ &= \frac{\int_0^\infty s^k e^{-s}\,\mathrm{d}s}{\frac{1}{(m-1)!}\int_0^\infty s^{k+m-1}e^{-s}\,\mathrm{d}s} \\ &= \frac{k!(m-1)!}{(k+m-1)!}.\end{aligned}$$

Consequently,

$$\begin{aligned}\int_{\partial B_{\ell_1^m}} \left(\sum_{k=r}^{\infty} \frac{\beta^k |\tau_1|^k}{k!}\right) \mathrm{d}\kappa_{\ell_1^m}(\tau) &= \frac{(m-1)!}{\beta^{m-1}} \sum_{k=r}^{\infty} \frac{\beta^{k+m-1}}{(k+m-1)!} \\ &= \frac{\beta^r (m-1)!}{(m+r-2)!} \int_0^1 e^{\beta t}(1-t)^{m+r-2}\,\mathrm{d}t, \end{aligned} \tag{6.59}$$

where the last step is the integral form of the remainder of the Taylor series of the exponential function.

It is mechanical to check that (6.58) holds for $m \in \{1,2\}$, so assume for the rest of the proof of Lemma 160 that $m \geqslant 3$. We then see from (6.59) that our goal (6.58) is equivalent to showing that

$$\int_0^1 e^{\beta t}(1-t)^{m+r-2}\,\mathrm{d}t \asymp \frac{1}{m+r}. \tag{6.60}$$

For the upper bound in (6.60), estimate the integrand using

$$(1-t)^{m+r-2} \leqslant e^{-(m+r-2)t}$$

to get

$$\begin{aligned}\int_0^1 e^{\beta t}(1-t)^{m+r-2}\,\mathrm{d}t &\leqslant \int_0^1 e^{-(m+r-2-\beta)t}\,\mathrm{d}t \\ &= \frac{1-e^{-(m+r-2-\beta)}}{m+r-2-\beta} \asymp \frac{1}{m+r},\end{aligned}$$

where we used $\beta < \frac{m+r-2}{2}$. For the lower bound in (6.60), since $(1-t)^{m+r-2} \gtrsim 1$ when $0 \leqslant t \leqslant \frac{1}{m+r-2}$,

$$\int_0^1 e^{\beta t}(1-t)^{m+r-2}\,\mathrm{d}t \geqslant \int_0^{\frac{1}{m+r-2}} e^{\beta t}(1-t)^{m+r-2}\,\mathrm{d}t$$
$$\gtrsim \int_0^{\frac{1}{m+r-2}} e^{\beta t}\,\mathrm{d}t = \frac{e^{\frac{\beta}{m+r-2}}-1}{\beta} \asymp \frac{1}{m+r},$$

where in the last step we used the assumption $\beta < \frac{m+r-2}{2}$ once more. ■

Proof of Lemma 158. By combining the case $g \equiv 1$ of (6.30) with (6.34), we see that

$$\mathrm{vol}_m\big(B_{\Omega_\beta^m}\big) = \frac{\beta^{m-1}2^m}{e^\beta m!} m \int_{\partial B_{\ell_1^m}} \big(e^{\beta|\tau_1|}-1\big)\,\mathrm{d}\kappa_{\ell_1^m}(\tau) \overset{(6.58)}{\asymp} \frac{(2\beta)^m}{e^\beta m!}. \tag{6.61}$$

Since we are assuming in Lemma 158 that $\beta \lesssim m$, in combination with (6.4) we get from (6.61) that

$$\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^m)}\big)^{\frac{1}{nm}} \asymp \frac{\beta}{n^{\frac{1}{p}}m}, \tag{6.62}$$

At the same time, by applying Cauchy–Schwarz to the identity (6.43) of Lemma 157 we have

$$\begin{aligned}&\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\Omega_\beta^m)}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^m)}\big)}\\
&\leqslant \frac{p\Gamma\big(1+\frac{nm}{p}\big)}{\Gamma\big(1+\frac{nm-1}{p}\big)}\bigg(n\big(\mathbb{E}\big[\mathsf{R}_1\big]\big)\int_{\partial B_{\Omega_\beta^m}} \|\nabla\|\cdot\|_{\Omega_\beta^m}(\theta)\|_{\ell_2^m}^2\,\mathrm{d}\kappa_{\Omega_\beta^m}(\theta)\bigg)^{\frac{1}{2}}\\
&\asymp n^{\frac{1}{p}+\frac{1}{2}} m\bigg(\int_{\partial B_{\Omega_\beta^m}} \|\nabla\|\cdot\|_{\Omega_\beta^m}(\theta)\|_{\ell_2^m}^2\,\mathrm{d}\kappa_{\Omega_\beta^m}(\theta)\bigg)^{\frac{1}{2}},\end{aligned} \tag{6.63}$$

where the random variable R_1 is as in Lemma 157, i.e., its density is in (6.42), and the last step is an application the evaluation (6.52) of its moments and Stirling's formula, using the assumption $1 \leqslant p \lesssim m$.

Recalling (6.31) and (6.32), even though $\|\cdot\|_{\Omega_\beta^m}$ is defined implicitly by (6.28), we can compute $\nabla\|\cdot\|_{\Omega_\beta^m}(\theta)$ for almost every $\theta \in \partial B_{\Omega_\beta^m}$ as the unique vector $v \in \mathbb{R}^m$ that is normal to $\partial B_{\Omega_\beta^m}$ and satisfies $\langle v, \theta\rangle = 1$. Indeed, since $\partial\Omega_\beta^m$ is parameterized as the zero set of the function $\Psi_\beta : \mathbb{R}^n \to \mathbb{R}^n$ that is given by

$$\forall x \in \mathbb{R}^n, \quad \Psi_\beta(x) \overset{\mathrm{def}}{=} 1 - \sum_{i=1}^m \psi_\beta(|x_i|),$$

the following vector is normal to $\partial B_{\Omega_\beta^m}$ for almost every $\theta \in \partial B_{\Omega_\beta^m}$:

$$v_\beta(\theta) \stackrel{\text{def}}{=} \nabla\Psi_\beta(\theta) = -\big(\psi_\beta'(|\theta_1|)\,\mathrm{sign}(\theta_1),\ldots,\psi_\beta'(|\theta_m|)\,\mathrm{sign}(\theta_m)\big).$$

So, $\nabla\|\cdot\|_{\Omega_\beta^m}(\theta) = \lambda_\beta(\theta)v_\beta(\theta)$ for almost every $\theta \in \partial B_{\Omega_\beta^m}$, where $\lambda_\beta(\theta) \in \mathbb{R}$ is such that $\langle\lambda_\beta(\theta)v_\beta(\theta),\theta\rangle = 1$, i.e., $\lambda_\beta(\theta) = -1/\langle v_\beta(\theta),\theta\rangle$. This shows that for almost every $\theta \in \partial B_{\Omega_\beta^m}$,

$$\begin{aligned}\nabla\|\cdot\|_{\Omega_\beta^m}(\theta) &= \frac{1}{\sum_{i=1}^m |\theta_i|\psi_\beta'(|\theta_i|)}\big(\psi_\beta'(|\theta_1|)\,\mathrm{sign}(\theta_1),\ldots,\psi_\beta'(|\theta_m|)\,\mathrm{sign}(\theta_m)\big)\\ &= \frac{1}{\sum_{i=1}^m \frac{|\theta_i|}{1-|\theta_i|}}\left(\frac{\mathrm{sign}(\theta_1)}{1-|\theta_1|},\ldots,\frac{\mathrm{sign}(\theta_m)}{1-|\theta_m|}\right),\end{aligned} \tag{6.64}$$

where the first equality in (6.64) holds for any ψ_β that satisfies the conditions of Lemma 154, and for the second equality in (6.64) recall the definition (6.32) of the specific ψ_β that we are using here. Therefore,

$$\begin{aligned}\int_{\partial B_{\Omega_\beta^m}} \big\|\nabla\|\cdot\|_{\Omega_\beta^m}(\theta)\big\|_{\ell_2^m}^2 \,\mathrm{d}\kappa_{\Omega_\beta^m}(\theta) &= \frac{\int_{\partial B_{\ell_1^m}} \frac{\sum_{i=1}^m e^{2\beta|\tau_i|}}{\sum_{i=1}^m (e^{\beta|\tau_i|}-1)}\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}{m\int_{\partial B_{\ell_1^m}}\big(e^{\beta|\tau_1|}-1\big)\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}\\ &\leqslant \frac{\int_{\partial B_{\ell_1^m}} \frac{\sum_{i=1}^m e^{2\beta|\tau_i|}}{\beta\sum_{i=1}^m |\tau_i|}\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}{m\int_{\partial B_{\ell_1^m}}\big(e^{\beta|\tau_1|}-1\big)\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}\\ &= \frac{\int_{\partial B_{\ell_1^m}} e^{2\beta|\tau_1|}\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}{\beta\int_{\partial B_{\ell_1^m}}\big(e^{\beta|\tau_1|}-1\big)\,\mathrm{d}\kappa_{\ell_1^m}(\tau)}\\ &\asymp \frac{m}{\beta^2},\end{aligned} \tag{6.65}$$

where the first step of (6.65) is a substitution of (6.64) into (6.33) while using (6.34) and that $\psi_\beta^{-1}(t) = 1-e^{-\beta t}$ for every $t \geqslant 0$, the second step of (6.65) uses the inequality $e^t \geqslant t+1$ which holds for any $t \in \mathbb{R}$, and the final step of (6.65) is an application of Lemma 160. Now, a combination of (6.63) and (6.65) gives

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\Omega_\beta^m)}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^m)}\big)} \lesssim \frac{n^{\frac{1}{p}+\frac{1}{2}}m^{\frac{3}{2}}}{\beta}. \tag{6.66}$$

By combining (6.62) and (6.66) we conclude that

$$\mathrm{iq}\big(B_{\ell_p^n(\Omega_\beta^m)}\big) = \frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\Omega_\beta^m)}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^m)}\big)}\,\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^n)}\big)^{\frac{1}{nm}} \lesssim \sqrt{nm}.$$

The reverse inequality, namely $\mathrm{iq}(B_{\ell_p^n(\Omega_\beta^m)}) \gtrsim \sqrt{nm}$, follows from the isoperimetric theorem (1.12), so the proof of Lemma 158 is complete. Note that this also shows that all of the inequalities that we derived in the above proof of Lemma 158 are in fact asymptotic equivalences. This holds in particular for (6.66), i.e.,

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\Omega_\beta^m)}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\Omega_\beta^m)}\big)} \asymp \frac{n^{\frac{1}{p}+\frac{1}{2}}m^{\frac{3}{2}}}{\beta}. \qquad \blacksquare$$

The following asymptotic evaluation of the surface area of the sphere of $\ell_p^n(\ell_q^m)$ in the entire range of possible values of $p,q \geqslant 1$ and $m,n \in \mathbb{N}$ is an application of Lemma 157; by (6.7) it is equivalent to (1.82).

Theorem 161. *For every $n,m \in \mathbb{N}$ and $p,q \in [1,\infty]$ we have*

$$\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\ell_q^m)}\big)$$
$$\asymp \frac{2^{nm}\Gamma\big(1+\frac{1}{q}\big)^{nm}\Gamma\big(1+\frac{m}{p}\big)^n}{\Gamma\big(1+\frac{m}{q}\big)^n\Gamma\big(1+\frac{nm}{p}\big)} \cdot \begin{cases} n^{1+\frac{1}{p}}m^{1+\frac{1}{q}} & m \leqslant \min\{\frac{p}{n},q\}, \\ \sqrt{q}n^{1+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & q \leqslant m \leqslant \frac{p}{n}, \\ \sqrt{p}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & \frac{p}{n} \leqslant m \leqslant \min\{p,q\}, \\ \sqrt{pq}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{q}} & \max\{\frac{p}{n},q\} \leqslant m \leqslant p, \\ n^{\frac{1}{2}+\frac{1}{p}}m^{1+\frac{1}{q}} & p \leqslant m \leqslant q, \\ \sqrt{q}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & m \geqslant \max\{p,q\}. \end{cases}$$

Proof. By continuity we may assume that $p,q \in (1,\infty)$. Suppose that G is a symmetric real-valued random variable whose density at each $s \in \mathbb{R}$ is equal to

$$\frac{1}{2\Gamma\big(1+\frac{1}{q}\big)}e^{-|s|^q}. \tag{6.67}$$

Let $\mathsf{G}_1,\ldots,\mathsf{G}_m$ be independent copies of G. Set $\mathsf{U} \stackrel{\mathrm{def}}{=} (\mathsf{G}_1,\ldots,\mathsf{G}_m) \in \mathbb{R}^m$. By the probabilistic representation of the cone measure on $\partial B_{\ell_q^m}$ in [266, 279] (or Lemma 156), the random vector $\mathsf{U}/\|\mathsf{U}\|_{\ell_q^m}$ is distributed according to the cone measure on $\partial B_{\ell_q^m}$, and moreover it is independent of $\|\mathsf{U}\|_{\ell_q^m}$.

Consider the following random variable:

$$\mathsf{N} \stackrel{\mathrm{def}}{=} \left\|\nabla\|\cdot\|_{\ell_q^m}\left(\frac{\mathsf{U}}{\|\mathsf{U}\|_{\ell_q^m}}\right)\right\|_{\ell_2^m}^2 = \frac{1}{\|\mathsf{U}\|_{\ell_q^m}^{2q-2}}\sum_{j=1}^m |\mathsf{G}_j|^{2q-2} = \frac{\|\mathsf{U}\|_{\ell_{2q-2}^m}^{2q-2}}{\|\mathsf{U}\|_{\ell_q^m}^{2q-2}}. \tag{6.68}$$

If we let $\mathsf{N}_1,\ldots,\mathsf{N}_n,\mathsf{R}_1,\ldots,\mathsf{R}_n$ be independent random variables such that $\mathsf{N}_1,\ldots,\mathsf{N}_n$ have the same distribution as N, and $\mathsf{R}_1,\ldots,\mathsf{R}_n$ are as in Lemma 157, then Lemma 157 gives that

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\ell_q^m)}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\ell_q^m)}\big)} = \frac{p\Gamma\big(1+\frac{nm}{p}\big)}{\Gamma\big(1+\frac{nm-1}{p}\big)}\mathbb{E}[\mathsf{Z}] \asymp pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}[\mathsf{Z}], \tag{6.69}$$

where for (6.69) we introduce the following notation:

$$\mathsf{Z} \stackrel{\mathrm{def}}{=} \left(\sum_{i=1}^{n} \mathsf{R}_i \mathsf{N}_i\right)^{\frac{1}{2}}. \tag{6.70}$$

Let R be a random variable that takes values in $[0,\infty)$ whose density at each $t \in (0,\infty)$ is given by (6.42), i.e., $\mathsf{R}_1,\ldots,\mathsf{R}_n$ are independent copies of R. We computed the moments of R in (6.52) and by Stirling's formula this gives the following asymptotic evaluations:

$$\mathbb{E}\big[\mathsf{R}^{\frac{1}{2}}\big] \asymp \frac{m^{1-\frac{1}{p}}}{p}, \tag{6.71}$$

$$\mathbb{E}[\mathsf{R}] \asymp \max\left\{\frac{m}{p}, 1\right\} \frac{m^{1-\frac{2}{p}}}{p}, \tag{6.72}$$

$$\mathbb{E}\big[\mathsf{R}^2\big] \asymp \max\left\{\frac{m^3}{p^3}, 1\right\} \frac{m^{1-\frac{4}{p}}}{p}. \tag{6.73}$$

We also need an analogous asymptotic evaluation of moments of the random variable N in (6.68). Observe that the random variables N and $\|\mathsf{U}\|_{\ell_q^m}$ are independent, since $\mathsf{U}/\|\mathsf{U}\|_{\ell_q^m}$ and $\|\mathsf{U}\|_{\ell_q^m}$ are independent and N is a function $\mathsf{U}/\|\mathsf{U}\|_{\ell_q^m}$. Consequently, for every $\beta > 0$ we have

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_q^m}^{(2q-2)\beta}\big]\mathbb{E}\big[\mathsf{N}^\beta\big] = \mathbb{E}\big[\|\mathsf{U}\|_{\ell_q^m}^{(2q-2)\beta}\mathsf{N}^\beta\big] \stackrel{(6.68)}{=} \mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{(2q-2)\beta}\big]. \tag{6.74}$$

Since (e.g., by Lemma 156) the density of $\|\mathsf{U}\|_{\ell_q^m}$ at $s \in (0,\infty)$ is proportional to $s^{m-1}e^{-s^q}$, we can compute analogously to (6.48) that

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_q^m}^{(2q-2)\beta}\big] = \frac{\int_0^\infty s^{m-1+(2q-2)\beta} e^{-s^q}\,\mathrm{d}s}{\int_0^\infty r^{m-1}e^{-r^q}\,\mathrm{d}r} = \frac{\Gamma\big(2\beta + \frac{m-2\beta}{q}\big)}{\Gamma\big(\frac{m}{q}\big)}.$$

Therefore, (6.74) implies that

$$\mathbb{E}\big[\mathsf{N}^\beta\big] = \frac{\Gamma\big(\frac{m}{q}\big)}{\Gamma\big(2\beta + \frac{m-2\beta}{q}\big)}\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{(2q-2)\beta}\big].$$

By considering each of the values $\beta \in \{\frac{1}{2}, 1, 2\}$ in this identity and using Stirling's formula, we get the following asymptotic evaluations of moments of N in terms of moments of $\|\mathsf{U}\|_{\ell_{2q-2}^m}$:

$$\mathbb{E}\big[\mathsf{N}^{\frac{1}{2}}\big] \asymp \frac{q}{m^{1-\frac{1}{q}}}\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big], \tag{6.75}$$

$$\mathbb{E}[\mathsf{N}] \asymp \min\left\{\frac{q}{m}, 1\right\} \frac{q}{m^{1-\frac{2}{q}}} \mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{2q-2}\big], \tag{6.76}$$

$$\mathbb{E}\big[\mathsf{N}^2\big] \asymp \min\left\{\frac{q^3}{m^3}, 1\right\} \frac{q}{m^{1-\frac{4}{q}}} \mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{4q-4}\big]. \tag{6.77}$$

Due to (6.75), (6.76), (6.77), we will next evaluate the corresponding moments of $\|\mathsf{U}\|_{\ell_{2q-2}^m}$. Recalling the density (6.67) of G, for every $\beta > -1/(2q-2)$ we have

$$\mathbb{E}\big[|\mathsf{G}|^{(2q-2)\beta}\big] = \frac{1}{\Gamma\big(1+\frac{1}{q}\big)} \int_0^\infty s^{(2q-2)\beta} e^{-s^q}\, ds = \frac{\Gamma\big(\frac{2q-2}{q}\beta + \frac{1}{q}\big)}{q\Gamma\big(1+\frac{1}{q}\big)}.$$

Hence,

$$\mathbb{E}\big[|\mathsf{G}|^{q-1}\big] \asymp \mathbb{E}\big[|\mathsf{G}|^{2q-2}\big] \asymp \mathbb{E}\big[|\mathsf{G}|^{4q-4}\big] \asymp \frac{1}{q}. \tag{6.78}$$

We therefore have

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{2q-2}\big] = m\mathbb{E}\big[|\mathsf{G}|^{2q-2}\big] \overset{(6.78)}{\asymp} \frac{m}{q} \tag{6.79}$$

and

$$\begin{aligned} \mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{4q-4}\big] &= \mathbb{E}\left[\left(\sum_{j=1}^m |\mathsf{G}_j|^{2q-2}\right)^2\right] \\ &= m\mathbb{E}\big[|\mathsf{G}|^{4q-4}\big] + m(m-1)\big(\mathbb{E}\big[|\mathsf{G}|^{2q-2}\big]\big)^2 \\ &\overset{(6.78)}{\asymp} \max\left\{\frac{m}{q}, 1\right\}\frac{m}{q}. \end{aligned} \tag{6.80}$$

Consequently, using Hölder's inequality we get the following estimate:

$$\begin{aligned} \frac{m}{q} &\overset{(6.79)}{\asymp} \mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{2q-2}\big] \\ &= \mathbb{E}\Big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{\frac{2}{3}(q-1)} \|\mathsf{U}\|_{\ell_{2q-2}^m}^{\frac{1}{3}(4q-4)}\Big] \\ &\leqslant \Big(\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big]\Big)^{\frac{2}{3}} \Big(\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{4q-4}\big]\Big)^{\frac{1}{3}} \\ &\overset{(6.80)}{\asymp} \Big(\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big]\Big)^{\frac{2}{3}} \left(\max\left\{\frac{m}{q}, 1\right\}\frac{m}{q}\right)^{\frac{1}{3}}. \end{aligned} \tag{6.81}$$

This simplifies to give

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big] \gtrsim \min\left\{\sqrt{\frac{m}{q}}, \frac{m}{q}\right\}. \tag{6.82}$$

At the same time, by Cauchy–Schwarz,

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big] \leqslant \Big(\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{2q-2}\big]\Big)^{\frac{1}{2}} \overset{(6.79)}{\asymp} \sqrt{\frac{m}{q}}. \tag{6.83}$$

Also, by the subadditivity of the square root on $[0,\infty)$,

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big] = \mathbb{E}\left[\left(\sum_{j=1}^m |\mathsf{G}_j|^{2q-2}\right)^{\frac{1}{2}}\right] \leqslant \mathbb{E}\left[\sum_{j=1}^m |\mathsf{G}_j|^{q-1}\right]$$
$$= m\mathbb{E}\big[|\mathsf{G}|^{q-1}\big] \overset{(6.78)}{\asymp} \frac{m}{q}. \tag{6.84}$$

By combining (6.83) and (6.84) we see that (6.82) is in fact sharp, i.e.,

$$\mathbb{E}\big[\|\mathsf{U}\|_{\ell_{2q-2}^m}^{q-1}\big] \asymp \min\left\{\sqrt{\frac{m}{q}}, \frac{m}{q}\right\}. \tag{6.85}$$

By substituting (6.85) into (6.75), and correspondingly (6.79) into (6.76) and (6.80) into (6.77), we get the following asymptotic identities:

$$\mathbb{E}\big[\mathsf{N}^{\frac{1}{2}}\big] \asymp \min\left\{\sqrt{\frac{q}{m}}, 1\right\} m^{\frac{1}{q}}, \tag{6.86}$$
$$\mathbb{E}[\mathsf{N}] \asymp \min\left\{\frac{q}{m}, 1\right\} m^{\frac{2}{q}}, \tag{6.87}$$
$$\mathbb{E}\big[\mathsf{N}^2\big] \asymp \min\left\{\frac{q^2}{m^2}, 1\right\} m^{\frac{4}{q}}. \tag{6.88}$$

By combining (6.72) and (6.87) we see that

$$\mathbb{E}\big[\mathsf{Z}^2\big] = n\big(\mathbb{E}[\mathsf{R}]\big)\big(\mathbb{E}[\mathsf{N}]\big) \asymp \frac{\max\{m,p\}\min\{q,m\}}{p^2} n m^{\frac{2}{q}-\frac{2}{p}}.$$

Using Cauchy–Schwarz, this implies the following upper bound on the final term in (6.69):

$$pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}[\mathsf{Z}] \leqslant pn^{\frac{1}{p}}m^{\frac{1}{p}}\big(\mathbb{E}\big[\mathsf{Z}^2\big]\big)^{\frac{1}{2}}$$
$$\asymp n^{\frac{1}{2}+\frac{1}{p}} m^{\frac{1}{q}} \sqrt{\max\{m,p\}\min\{m,q\}}. \tag{6.89}$$

Also, recalling (6.70) and using the subadditivity of the square root on $[0,\infty)$ in combination with (6.71) and (6.86), we have the following additional upper bound on the final term in (6.69):

$$pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}[\mathsf{Z}] \leqslant pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}\left[\sum_{i=1}^n \mathsf{R}_i^{\frac{1}{2}}\mathsf{N}_i^{\frac{1}{2}}\right]$$
$$= pn^{1+\frac{1}{p}}m^{\frac{1}{p}}\big(\mathbb{E}\big[\mathsf{R}^{\frac{1}{2}}\big]\big)\big(\mathbb{E}\big[\mathsf{N}^{\frac{1}{2}}\big]\big)$$
$$\asymp n^{1+\frac{1}{p}} m^{\frac{1}{2}+\frac{1}{q}} \sqrt{\min\{m,q\}}. \tag{6.90}$$

It follows from (6.89) and (6.90) that

$$
\begin{aligned}
pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}[\mathsf{Z}] &\lesssim n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{q}}\sqrt{\min\{m,q\}}\min\left\{\sqrt{nm},\sqrt{\max\{m,p\}}\right\}\\
&= \begin{cases} n^{1+\frac{1}{p}}m^{1+\frac{1}{q}} & m \leqslant \min\{\frac{p}{n},q\},\\ \sqrt{q}n^{1+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & q \leqslant m \leqslant \frac{p}{n},\\ \sqrt{p}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & \frac{p}{n} \leqslant m \leqslant \min\{p,q\},\\ \sqrt{pq}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{q}} & \max\{\frac{p}{n},q\} \leqslant m \leqslant p,\\ n^{\frac{1}{2}+\frac{1}{p}}m^{1+\frac{1}{q}} & p \leqslant m \leqslant q,\\ \sqrt{q}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & m \geqslant \max\{p,q\}. \end{cases}
\end{aligned} \tag{6.91}
$$

We will next prove that (6.91) is optimal in all of the six ranges that appear in (6.91); by (6.69) and (6.6), this will complete the proof of Theorem 161. Recalling (6.70) and using (6.72), (6.73), (6.87), (6.88), the fourth moment of Z can be evaluated (up to universal constant factors) as follows:

$$
\begin{aligned}
\mathbb{E}\big[\mathsf{Z}^4\big] &= \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mathsf{R}_i\mathsf{R}_j\mathsf{N}_i\mathsf{N}_j\right]\\
&= n\big(\mathbb{E}\big[\mathsf{R}^2\big]\big)\big(\mathbb{E}\big[\mathsf{N}^2\big]\big) + n(n-1)\big(\mathbb{E}[\mathsf{R}]\big)^2\big(\mathbb{E}[\mathsf{N}]\big)^2\\
&\asymp \frac{(\max\{m,p\})^3(\min\{m,q\})^2}{p^4}nm^{\frac{4}{q}-\frac{4}{p}-1}\\
&\quad+\frac{(\max\{m,p\}\min\{m,q\})^2}{p^4}n^2m^{\frac{4}{q}-\frac{4}{p}}\\
&\asymp \frac{(\max\{m,p\}\min\{m,q\})^2\max\{nm,p\}}{p^4}nm^{\frac{4}{q}-\frac{4}{p}-1}.
\end{aligned} \tag{6.92}
$$

By using Hölder's inequality similarly to (6.81), we conclude that

$$
\begin{aligned}
pn^{\frac{1}{p}}m^{\frac{1}{p}}\mathbb{E}[\mathsf{Z}] &\geqslant pn^{\frac{1}{p}}m^{\frac{1}{p}}\frac{\big(\mathbb{E}\big[\mathsf{Z}^2\big]\big)^{\frac{3}{2}}}{\big(\mathbb{E}\big[\mathsf{Z}^4\big]\big)^{\frac{1}{2}}}\\
&\overset{(6.89)\wedge(6.92)}{\asymp} n^{1+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}}\frac{\sqrt{\max\{m,p\}\min\{m,q\}}}{\sqrt{\max\{nm,p\}}}\\
&= \begin{cases} n^{1+\frac{1}{p}}m^{1+\frac{1}{q}} & m \leqslant \min\{\frac{p}{n},q\},\\ \sqrt{q}n^{1+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & q \leqslant m \leqslant \frac{p}{n},\\ \sqrt{p}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & \frac{p}{n} \leqslant m \leqslant \min\{p,q\},\\ \sqrt{pq}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{q}} & \max\{\frac{p}{n},q\} \leqslant m \leqslant p,\\ n^{\frac{1}{2}+\frac{1}{p}}m^{1+\frac{1}{q}} & p \leqslant m \leqslant q,\\ \sqrt{q}n^{\frac{1}{2}+\frac{1}{p}}m^{\frac{1}{2}+\frac{1}{q}} & m \geqslant \max\{p,q\}. \end{cases}
\end{aligned}
$$

■

Lemma 162 below applies Theorem 161 iteratively to obtain an upper bound on the surface area of the unit sphere of nested ℓ_p norms on k-tensors (the case $k=2$ corresponds to n by m matrices equipped with the $\ell_p^n(\ell_q^m)$ norm). The second part of Lemma 162, namely the conclusion (6.94) below, is an implementation of the approach towards Conjecture 9 for the hypercube that we described in Remark 56.

Lemma 162. *Suppose that $k, n_1, \ldots, n_k \in \mathbb{N}$ and $p_1, \ldots, p_k \in [1,\infty]$ are such that $n_1 \geqslant \max\{3, p_1-2\}$ and $n_1n_2\cdots n_{j-1} \geqslant p_j - 2$ for every $j \in \{2,\ldots,k\}$. Define normed spaces $\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_k$ by setting $\mathbf{Y}_0 = \mathbb{R}$ and inductively $\mathbf{Y}_j = \ell_{p_j}^{n_j}(\mathbf{Y}_{j-1})$ for $j \in \{1,\ldots,k\}$. Then,*

$$\frac{\mathrm{vol}_{n_1\cdots n_k - 1}\big(\partial B_{\mathbf{Y}_k}\big)}{\mathrm{vol}_{n_1\cdots n_k}\big(B_{\mathbf{Y}_k}\big)} \leqslant e^{O(k)}\sqrt{p_1}\prod_{j=1}^{k} n_j^{\frac12+\frac{1}{p_j}}. \tag{6.93}$$

Hence, using the natural identification of the vector space that underlies $\mathbf{Y}_k$ with $\mathbb{R}^{\dim(\mathbf{Y}_k)} = \mathbb{R}^{n_1n_2\cdots n_k}$, if in addition we have $n_1 = O(1)$ and $p_j = \log n_j$ for every $j \in \{1,\ldots,k\}$, then

$$B_{\mathbf{Y}_k} \subseteq B_{\ell_\infty^{\dim(\mathbf{Y}_k)}} \subseteq e^{O(k)} B_{\mathbf{Y}_k} \quad \textit{and} \quad \frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}_k}\big)}{\mathrm{vol}_{\dim(\mathbf{Y}_k)}\big(B_{\mathbf{Y}_k}\big)} \leqslant e^{O(k)}, \tag{6.94}$$

where we recall the notation (1.53).

Proof. Suppose that $n, m \in \mathbb{N}$ and $p \in (1,\infty)$. By applying Cauchy–Schwarz to the right-hand side of (6.43) while using the case $\alpha = 1$ of (6.52), we see that for every normed space $\mathbf{X} = (\mathbb{R}^m, \|\cdot\|_{\mathbf{X}})$ we have

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\mathbf{X})}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)} \leqslant \frac{p\Gamma\big(1+\frac{nm}{p}\big)}{\Gamma\big(1+\frac{nm-1}{p}\big)}\left(\frac{n\Gamma\big(\frac{m+2p-2}{p}\big)}{\Gamma\big(\frac{m}{p}\big)}\int_{\partial B_{\mathbf{X}}} \big\|\nabla\|\cdot\|_{\mathbf{X}}\big\|_{\ell_2^m}^2 \,\mathrm{d}\kappa_{\mathbf{X}}\right)^{\frac12}. \tag{6.95}$$

If also $m \geqslant \max\{3, p-2\}$, then by Stirling's formula (6.95) gives the following estimate:

$$\frac{\mathrm{vol}_{nm-1}\big(\partial B_{\ell_p^n(\mathbf{X})}\big)}{\mathrm{vol}_{nm}\big(B_{\ell_p^n(\mathbf{X})}\big)} \lesssim n^{\frac12+\frac1p} m\left(\int_{\partial B_{\mathbf{X}}} \big\|\nabla\|\cdot\|_{\mathbf{X}}\big\|_{\ell_2^m}^2 \,\mathrm{d}\kappa_{\mathbf{X}}\right)^{\frac12}. \tag{6.96}$$

By continuity we may assume that $p_1, \ldots, p_k \in (1,\infty)$. Denote $d_0 = 1$ and for $j \in \{1,\ldots,k\}$ denote $d_j = \dim(\mathbf{Y}_j) = n_1n_2\cdots n_j$. We will naturally identify $\mathbf{Y}_j$ with $(\mathbb{R}^{d_j}, \|\cdot\|_{\mathbf{Y}_j})$. As $\mathbf{Y}_k = \ell_{p_k}^{n_k}(\mathbf{Y}_{k-1})$, we deduce from (6.96) that

$$\frac{\mathrm{vol}_{d_k-1}\big(\partial B_{\mathbf{Y}_k}\big)}{\mathrm{vol}_{d_k}\big(B_{\mathbf{Y}_k}\big)} \lesssim n_k^{\frac12+\frac{1}{p_k}}\left(\prod_{j=1}^{k-1} n_j\right)\left(\int_{\partial B_{\mathbf{Y}_{k-1}}} \big\|\nabla\|\cdot\|_{\mathbf{Y}_{k-1}}\big\|_{\ell_2^{d_{k-1}}}^2 \,\mathrm{d}\kappa_{\mathbf{Y}_{k-1}}\right)^{\frac12}. \tag{6.97}$$

At the same time, by (6.44) for every $j \in \{1,\ldots,k\}$ we have

$$\int_{\partial B_{\mathbf{Y}_j}} \|\nabla\|\cdot\|_{\mathbf{Y}_j}\|^2_{\ell_2^{d_j}}\, \mathrm{d}\kappa_{\mathbf{Y}_j} = \frac{n_j\Gamma\big(\frac{d_j}{p_j}\big)\Gamma\big(\frac{d_{j-1}+2p_j-2}{p_j}\big)}{\Gamma\big(\frac{d_{j-1}}{p_j}\big)\Gamma\big(\frac{d_j+2p_j-2}{p_j}\big)} \int_{\partial B_{\mathbf{Y}_{j-1}}} \|\nabla\|\cdot\|_{\mathbf{Y}_{j-1}}\|^2_{\ell_2^{d_{j-1}}}\, \mathrm{d}\kappa_{\mathbf{Y}_{j-1}}. \tag{6.98}$$

If also $j \geqslant 2$, then $d_{j-1} \geqslant n_1 \geqslant 3$ and by assumption $d_{j-1} \geqslant p_j - 2$, so by Stirling's formula (6.98) gives that for every $j \in \{2,\ldots,k\}$ we have

$$\int_{\partial B_{\mathbf{Y}_j}} \|\nabla\|\cdot\|_{\mathbf{Y}_j}\|^2_{\ell_2^{d_j}}\, \mathrm{d}\kappa_{\mathbf{Y}_j} \asymp n_j^{\frac{2}{p_j}-1} \int_{\partial B_{\mathbf{Y}_{j-1}}} \|\nabla\|\cdot\|_{\mathbf{Y}_{j-1}}\|^2_{\ell_2^{d_{j-1}}}\, \mathrm{d}\kappa_{\mathbf{Y}_{j-1}}. \tag{6.99}$$

When $j = 1$ we have $d_0 = 1$ and $n_1 \geqslant \max\{3, p_1 - 2\}$, and therefore by Stirling's formula (6.98) gives

$$\int_{\partial B_{\mathbf{Y}_1}} \|\nabla\|\cdot\|_{\mathbf{Y}_1}\|^2_{\ell_2^{d_1}}\, \mathrm{d}\kappa_{\mathbf{Y}_1} \asymp p_1 n_1^{\frac{2}{p_1}-1}. \tag{6.100}$$

Hence, by applying (6.99) iteratively in combination with the base case (6.100), we conclude that

$$\int_{\partial B_{\mathbf{Y}_{k-1}}} \|\nabla\|\cdot\|_{\mathbf{Y}_{k-1}}\|^2_{\ell_2^{d_{k-1}}}\, \mathrm{d}\kappa_{\mathbf{Y}_{k-1}} \leqslant e^{O(k)} p_1 \prod_{j=1}^{k-1} n_j^{\frac{2}{p_j}-1}. \tag{6.101}$$

A substitution of (6.101) into (6.97) yields the desired estimate (6.93).

To deduce the conclusion (6.94), note that for every $j \in \{1,\ldots,k\}$ we have the point-wise bounds

$$\|\cdot\|_{\ell_\infty^{n_j}(\mathbf{Y}_{j-1})} \leqslant \|\cdot\|_{\mathbf{Y}_j} = \|\cdot\|_{\ell_{p_j}^{n_j}(\mathbf{Y}_{j-1})} \leqslant n_j^{\frac{1}{p_j}} \|\cdot\|_{\ell_\infty^{n_j}(\mathbf{Y}_{j-1})}.$$

It follows by induction that

$$\|\cdot\|_{\ell_\infty^{d_k}} \leqslant \|\cdot\|_{\mathbf{Y}_k} \leqslant \Bigg(\prod_{j=1}^{k} n_j^{\frac{1}{p_j}}\Bigg)\|\cdot\|_{\ell_\infty^{d_k}} = e^{O(k)}\|\cdot\|_{\ell_\infty^{d_k}},$$

where the final step holds if $p_j = \log n_j$ for every $j \in \{1,\ldots,k\}$. This implies the inclusions in (6.94). Furthermore, $\mathbf{Y}_k$ belongs to the class of spaces from Example 40. Hence $\mathbf{Y}_k$ is canonically positioned and by the discussion in Section 1.6.2 know that $B_{\mathbf{Y}'}$ is in its minimum surface area position. Therefore,

$$\frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}_k}\big)}{\mathrm{vol}_{d_k}\big(B_{\mathbf{Y}_k}\big)} \asymp \frac{\mathrm{vol}_{d_k-1}\big(\partial B_{\mathbf{Y}_k}\big)}{\mathrm{vol}_{d_k}\big(B_{\mathbf{Y}_k}\big)\sqrt{d_k}} \leqslant e^{O(k)}\sqrt{p_1}\prod_{j=1}^{k} n_j^{\frac{1}{p_j}} \asymp e^{O(k)},$$

where the first step uses [104, Proposition 3.1], the second step is (6.93), and the final step holds because $p_1 = O(1)$ and $p_j = \log n_j$. This completes the proof of (6.94). ■

The following technical lemma replaces a more ad-hoc argument that we previously had to deduce Proposition 164 below from Lemma 162; it is due to Noga Alon and we thank him for allowing us to include it here. This lemma shows that the set of super-lacunary products $n_1 n_2 \cdots n_k$ that can serve as dimensions of the space $\mathbf{Y}_k$ in Lemma 162 for which (6.94) holds is quite dense in $\mathbb{N}$.

Lemma 163. *For every integer $n \geqslant 3$ there are $k, m \in \mathbb{N} \cup \{0\}$ and integers $n_1 < n_2 < \cdots < n_k$ that satisfy*

- $n = n_1 n_2 \cdots n_k + m$,
- $n_1 \in \{6, 7\}$ *and* $n_{i+1} \lesssim 2^{n_i} \lesssim n_{i+1}^3$ *for every* $i \in \{1, \ldots, k-1\}$,
- $m \lesssim (\log n)^{1+o(1)}$.

Prior to proving Lemma 163, we will make some preparatory (mechanical) observations for ease of later reference. Note first that the conclusion $n_{i+1} \lesssim 2^{n_i} \lesssim n_{i+1}^3$ of Lemma 163 can be rewritten as

$$\forall i \in \{1, \ldots, k-1\}, \quad \log_2 n_{i+1} \lesssim n_i \lesssim \log_{\sqrt[3]{2}} n_{i+1}.$$

It follows by induction that

$$\forall i \in \{1, \ldots, k\}, \quad \log_2^{[k-i]} n_k \lesssim n_i \lesssim \log_{\sqrt[3]{2}}^{[k-i]} n_k, \tag{6.102}$$

where, as we also did in (1.131), we denote the iterates of a function $\varphi : (0, \infty) \to \mathbb{R}$ by $\varphi^{[j]} = \varphi \circ \varphi^{[j-1]} : (\varphi^{[j-1]})^{-1}(0, \infty) \to \mathbb{R}$ for each $j \in \mathbb{N}$, with the convention $\varphi^{[0]}(x) = x$ for every $x \in (0, \infty)$. Since $n_1 \in \{6, 7\}$, it follows from (6.102) that

$$k \asymp \log^* n_k \lesssim \log^* n. \tag{6.103}$$

Consequently,

$$\begin{aligned} n_k \log n_k \asymp n_k n_{k-1} \leqslant \prod_{i=1}^{k} n_k \leqslant n = m + \prod_{i=1}^{k} n_k \lesssim (\log n)^{1+o(1)} + \prod_{i=1}^{k} \log_{\sqrt[3]{2}}^{[k-i]} n_k \\ \lesssim (\log n)^2 + n_k (\log n_k)(\log\log n_k)^{O(\log^* n_k)} \lesssim (\log n)^2 + n_k (\log n_k)^2. \end{aligned}$$

This implies the following (quite crude) bounds on n_k:

$$\frac{n}{(\log n)^2} \lesssim n_k \lesssim \frac{n}{\log n}. \tag{6.104}$$

Note in particular that thanks to (6.104) we know that (6.103) can be improved to $k \asymp \log^* n$.

Proof of Lemma 163. Let $\mathbb{M} \subseteq \mathbb{N}$ be the set of all those $x \in \mathbb{N}$ that can be written as $x = n_1 n_2 \cdots n_k$ for some $k, n_1, \ldots, n_k \in \mathbb{N}$ that satisfy $n_k > n_{k-1} > \cdots > n_1 \in \{6, 7\}$ and

$$\forall i \in \{1, \ldots, k-1\}, \quad n_{i+1} \leqslant 2^{n_i} \leqslant n_{i+1}^3. \tag{6.105}$$

The goal of Lemma 163 is to show that there exists $x \in \mathbb{M}$ such that

$$n - (\log n)^{1+o(1)} \leqslant x \leqslant n. \tag{6.106}$$

By adjusting the $o(1)$ term, we may assume that n is sufficiently large, say, $n \geqslant n(0)$ for some fixed $n(0) \in \mathbb{N}$ that will be determined later. We will then find $x \in \mathbb{M}$ with a representation $x = n_1 n_2 \cdots n_k$ as above and

$$n - n_1 n_2 \cdots n_{k-1} \leqslant x \leqslant n. \tag{6.107}$$

This would imply the desired bound (6.106) because

$$\prod_{i=1}^{k-1} n_i \overset{(6.102)}{\leqslant} \prod_{i=1}^{k-1} \log_{\sqrt[3]{2}}^{[k-i]} n_k \overset{(6.103)}{\lesssim} (\log n_k)^{1+o(1)} \overset{(6.104)}{\lesssim} (\log n)^{1+o(1)}.$$

We will first construct $\{y_i\}_{i=1}^{\infty} \subseteq \mathbb{M}$ such that $y_1 = 7$ and $y_i < y_{i+1} < 12 y_i$ for every $i \in \mathbb{N}$. Furthermore, for each $i \in \mathbb{N}$ there are $k, n_1, \ldots, n_k \in \mathbb{N}$ with $y_i = n_1 n_2 \cdots n_k$ such that $n_k > n_{k-1} > \cdots > n_1 \in \{6, 7\}$ and

$$\forall j \in \{1, \ldots, k-1\}, \quad n_{j+1}^2 \leqslant 2^{n_j} \leqslant 2 n_{j+1}^2, \tag{6.108}$$

which is a more stringent requirement than (6.105). Note in passing that (6.108) implies the (crude) bound

$$\prod_{j=1}^{k} \left(1 + \frac{1}{n_j}\right) \leqslant 2. \tag{6.109}$$

To verify (6.109), note that since $\{n_j\}_{j=1}^k$ is strictly increasing and the second inequality in (6.108) holds, it is mechanical to check that $n_1 \geqslant 6$, $n_2 \geqslant 7$, $n_3 \geqslant 8$, $n_4 \geqslant 12$ and $n_{j+1} \geqslant 3 n_j$ for every $j \in \{4, 5, \ldots, k-1\}$. So,

$$\begin{aligned} \prod_{j=1}^{k} \left(1 + \frac{1}{n_j}\right) &\leqslant \left(1 + \frac{1}{6}\right)\left(1 + \frac{1}{7}\right)\left(1 + \frac{1}{8}\right) e^{\sum_{s=0}^{\infty} \frac{1}{12 \cdot 3^s}} \\ &= \left(1 + \frac{1}{6}\right)\left(1 + \frac{1}{7}\right)\left(1 + \frac{1}{8}\right) e^{\frac{1}{8}} \leqslant 2. \end{aligned}$$

Suppose that y_i has been defined with a representation $y_i = n_1 n_2 \cdots n_k$ that fulfils the above requirements. Define $m_0, m_1, \ldots, m_k \in \mathbb{N}$ with $m_0 = 6$, $m_k = n_k + 1$ and $m_j \in \{n_j, n_j + 1\}$ for all $j \in \{1, \ldots, k-1\}$ by induction as follows. Assuming that

m_{j+1} has already been constructed for some $j \in \{1, \ldots, k-1\}$, let

$$m_j \stackrel{\text{def}}{=} \begin{cases} n_j & \text{if } m_{j+1}^2 \leqslant 2^{n_j}, \\ n_j + 1 & \text{if } m_{j+1}^2 > 2^{n_j}. \end{cases} \tag{6.110}$$

Definition (6.110) implies that $m_j < m_{j+1}$. Indeed, $n_j < n_{j+1}$ so if $m_j = n_j$, then $n_j < n_{j+1} \leqslant m_{j+1}$ since $m_{j+1} \geqslant n_{j+1}$ by the induction hypothesis. On the other hand, if $m_j = n_j + 1$, then since the first inequality in (6.108) holds, the definition (6.110) necessitates that $m_{j+1} = n_j + 1$, so $m_j < m_{j+1}$ in this case as well.

Next, Definition (6.110) also ensures that the requirement (6.108) is inherited by $\{m_j\}_{j=1}^k$, i.e.,

$$\forall j \in \{1, \ldots, k-1\}, \quad m_{j+1}^2 \leqslant 2^{m_j} \leqslant 2m_{j+1}^2. \tag{6.111}$$

Indeed, if $m_j = n_j$, then $m_{j+1}^2 \leqslant 2^{n_j} = 2^{m_j}$ by (6.110), i.e., the first inequality in (6.111) holds, and the second inequality in (6.111) holds because $m_{j+1} \geqslant n_{j+1}$ and (6.108) holds. On the other hand, if $m_j = n_j + 1$, then by (6.110) we have $m_{j+1} = n_j + 1$ and $m_{j+1}^2 > 2^{n_j}$, which directly gives the second inequality in (6.111), and in combination with (6.108) we also get the first inequality in (6.111) because

$$\frac{m_{j+1}}{2^{m_j}} = \frac{(n_j+1)^2}{2^{n_j+1}} \stackrel{(6.108)}{\leqslant} \frac{(n_j+1)^2}{2n_j^2} \leqslant 1,$$

where the final step uses $n_j \geqslant 6$, though $n_j \geqslant 1/(\sqrt{2}-1) = 2.414\ldots$ is all that is needed for this purpose.

If the above construction produces $m_1 \in \{6, 7\}$, then define $y_{i+1} = m_1 m_2 \cdots m_k$. Otherwise necessarily $m_1 = n_1 + 1 = 8$, so (6.111) holds also when $j = 0$ (recall that $m_0 = 6$, hence $m_1^2 = 2^6 = 2^{m_0}$), so we can define $y_{i+1} = m_0 m_1 \cdots m_k$ and thanks to (6.111) in both cases y_{i+1} has the desired form. Moreover,

$$\frac{y_{i+1}}{y_i} \leqslant 6 \prod_{j=1}^{k} \left(1 + \frac{1}{n_j}\right) \stackrel{(6.109)}{\leqslant} 12.$$

This completes the inductive construction of the desired sequence $\{y_i\}_{i=1}^\infty \subseteq \mathbb{M}$.

With the sequence $\{y_i\}_{i=1}^\infty \subseteq \mathbb{M}$ at hand, will next explain how to obtain for each integer $n \geqslant n(0)$, where $n(0) \in \mathbb{N}$ is a sufficiently large universal constant that is yet to be determined, an element $x \in \mathbb{M}$ that approximates n as in (6.107). Let $i \in \mathbb{N}$ be such that $y_i \leqslant n \leqslant y_{i+1}$ and denote $y = y_i$. Thus, there are $k, n_1, \ldots, n_k \in \mathbb{N}$ for which $y = n_1 n_2 \cdots n_k$ such that $n_k > n_{k-1} > \cdots > n_1 \in \{6, 7\}$ and (6.108) holds.

If $y \geqslant n - n_1 n_2 \cdots n_{k-1}$, then $x = y$ has the desired approximation property, so suppose from now that $y < n - n_1 n_2 \cdots n_{k-1}$, or equivalently

$$\frac{n}{n_1 n_2 \cdots n_{k-1}} > \frac{y}{n_1 n_2 \cdots n_{k-1}} + 1 = n_k + 1.$$

Hence, if we define

$$n_k' \stackrel{\mathrm{def}}{=} \left\lfloor \frac{n}{n_1 n_2, \ldots, n_{k-1}} \right\rfloor \quad \text{and} \quad x = n_1 n_2 \cdots n_{k-1} n_k',$$

then $n_k' \geqslant n_k + 1 \gtrsim n/(\log n)^2$, where we used (6.104). Consequently, recalling (6.102), there is a universal constant $n(0) \in \mathbb{N}$ such that if $n \geqslant n(0)$, then $n_k' > \max\{144, n_{k-1}\}$. Thus, the sequence $n_1, n_2, \ldots, n_{k-1}, n_k'$ is still increasing. Since by design x satisfies (6.107), it remains to check that $x \in \mathbb{M}$, i.e., that (6.105) holds. Since $n_1, \ldots, n_k$ are assumed to satisfy the more stringent requirement (6.108), we only need to check that

$$n_k' \leqslant 2^{n_{k-1}} \leqslant (n_k')^3. \tag{6.112}$$

The second inequality in (6.112) is valid since (6.108) holds and $n_k' > n_k$. To justify the first inequality in (6.112), observe that $y \leqslant n \leqslant 12y$, as $y_{i+1} \leqslant 12y_i$. Consequently,

$$n_k' \leqslant n/(n_1 n_2 \cdots n_{k-1}) \leqslant 12y/(n_1 n_2 \cdots n_{k-1}) = 12 n_k.$$

Therefore,

$$2^{n_{k-1}} \stackrel{(6.108)}{\geqslant} n_k^2 \geqslant \left(\frac{n_k'}{12}\right)^2 > n_k',$$

where the last step uses the fact that $n_k' > 144$. ∎

We are now ready to extend the conclusion (6.94) of Lemma 162 to all dimensions $n \in \mathbb{N}$. Namely, we will prove the following proposition, which comes very close to proving Conjecture 9 for the hypercube $[-1,1]^n$ via a route that differs from the way by which we proved Theorem 24.

Proposition 164. *For any $n \in \mathbb{N}$ there is a normed space $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ that for every $x \in \mathbb{R}^n \smallsetminus \{0\}$ we have*

$$\|x\|_{\ell_\infty^n} \leqslant \|x\|_{\mathbf{Y}} \leqslant e^{O(\log^* n)} \|x\|_{\ell_\infty^n} \quad \textit{and} \quad \frac{\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp} B_{\mathbf{Y}}\big)}{\mathrm{vol}_n(B_{\mathbf{Y}})} \leqslant e^{O(\log^* n)}.$$

Furthermore, $\mathbf{Y}$ can be taken to be an ℓ_∞ direct sum of nested ℓ_p spaces as in Lemma 162.

Proof. Let $\mathbb{M} \subseteq \mathbb{N}$ be the set of integers from the proof of Lemma 163, namely $m \in \mathbb{M}$ if and only if there are integers $n_k > n_{k-1} > \cdots > n_1 \in \{6,7\}$ that satisfy (6.105) such that $m = n_1 n_2 \cdots n_k$. By Lemma 162, there exists $C > 1$ such that for every $m \in \mathbb{M}$ there is a normed space $\mathbf{Y}^m = (\mathbb{R}^m, \|\cdot\|_{\mathbf{Y}^m})$ that satisfies

$$\|\cdot\|_{\ell_\infty^m} \leqslant \|\cdot\|_{\mathbf{Y}^m} \leqslant e^{C\log^* m} \|\cdot\|_{\ell_\infty^m} \quad \text{and} \quad \frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}^m}\big)}{\mathrm{vol}_n\big(B_{\mathbf{Y}^m}\big)} \leqslant e^{C\log^* m}.$$

By applying Lemma 163 iteratively write $n = m_1 + \cdots + m_{s+1}$ for $m_1, \ldots, m_s \in \mathbb{M}$ and $m_{s+1} \in \{1, 2\}$ that satisfy $m_{i+1} \leqslant (\log m_i)^c$ for every $i \in \{1, \ldots, s\}$, where $c > 1$ is a universal constant. Denote $\mathbf{Y}^{m_{s+1}} = \ell_\infty^{m_{s+1}}$ and consider the ℓ_∞ direct sum

$$\mathbf{Y} \stackrel{\mathrm{def}}{=} \mathbf{Y}^{m_1} \oplus_\infty \mathbf{Y}^{m_2} \oplus_\infty \cdots \oplus_\infty \mathbf{Y}^{m_{s+1}} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}}).$$

Then $\|\cdot\|_{\ell_\infty^n} \leqslant \|\cdot\|_{\mathbf{Y}} \leqslant \max_{i\in\{1,\ldots,s+1\}} e^{C\log^* m_i} \|\cdot\|_{\ell_\infty^{m_i}} \leqslant e^{C\log^* n}\|\cdot\|_{\ell_\infty^n}$. We claim that

$$\frac{\mathrm{MaxProj}(B_{\mathbf{Y}})}{\mathrm{vol}_n(B_{\mathbf{Y}})} \leqslant e^{O(\log^* n)}.$$

Since $B_{\mathbf{Y}} = B_{\mathbf{Y}^{m_1}} \times B_{\mathbf{Y}^{m_2}} \times \cdots \times B_{\mathbf{Y}^{m_{s+1}}}$, by an inductive application of Lemma 159 we have

$$\frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}}\big)}{\mathrm{vol}_n\big(B_{\mathbf{Y}}\big)} \leqslant \left(\sum_{i=1}^{s+1} \frac{\mathrm{MaxProj}\big(B_{\mathbf{Y}^{m_i}}\big)^2}{\mathrm{vol}_{m_i}\big(B_{\mathbf{Y}^{m_i}}\big)^2}\right)^{\frac{1}{2}} \leqslant \left(\sum_{i=1}^{s+1} e^{2C\log^* m_i}\right)^{\frac{1}{2}} \lesssim e^{C\log^* n},$$

where the first step uses Lemma 159, the penultimate step is our assumption on $\mathbf{Y}^{m_i}$, and the final step has the following justification. Recall that $m_{i+1} \leqslant (\log m_i)^c$ for every $i \in \{1, \ldots, s\}$, where $c > 1$ is a universal constant. So, $m_{i+2} \leqslant c^c (\log\log m_i)^c$ for every $i \in \{1, \ldots, s-1\}$. Fix $n_0 \in \mathbb{N}$ such that $c^c(\log\log n)^c \leqslant \log n$ for every $n \geqslant n_0$. Then, $m_{i+2} \leqslant \log m_i$ if $m_i \geqslant n_0$, hence $\log^* m_{i+2} \leqslant \log^* m_i - 1$. Let i_0 be the largest $i \in \{1, \ldots, s+1\}$ for which $m_i < n_0$. Then,

$$\log^* m_{2i} \leqslant \log^* m_2 - i \leqslant \log^* n - i$$

and $\log^* m_{2j+1} \leqslant \log^* m_1 - j \leqslant \log^* n - j$ if $2i, 2j+1 \in \{1, \ldots, i_0 - 1\}$. We also have $|\{i_0, \ldots, s+1\}| = O(1)$. Consequently,

$$\sum_{i=1}^{s+1} e^{2C\log^* m_i} \leqslant e^{2C\log^* n} \sum_{k=0}^{\infty} e^{-2Ck} + O(1) \lesssim e^{2C\log^* n}. \qquad \blacksquare$$

Remark 165. A straightforward way to attempt to compute the surface area of the unit sphere of a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is to fix a direction $z \in S^{n-1}$ and consider $\partial B_{\mathbf{X}}$ as the union of the two graphs of the functions $\Psi_z^{\mathbf{X}}, \psi_z^{\mathbf{X}} : \mathsf{Proj}_{z^\perp}(B_{\mathbf{X}}) \to \mathbb{R}$ that are defined by setting $\Psi_z^{\mathbf{X}}(x)$ and $\psi_z^{\mathbf{X}}(x)$ for each $x \in \mathsf{Proj}_{z^\perp}(B_{\mathbf{X}})$ to be, respectively, the largest and smallest $s \in \mathbb{R}$ for which $x + sz \in \partial B_{\mathbf{X}}$. We then have

$$\begin{aligned}\mathrm{vol}_{n-1}(\partial B_{\mathbf{X}}) = &\int_{\mathsf{Proj}_{z^\perp}(B_{\mathbf{X}})} \sqrt{1 + \|\nabla \Psi_z^{\mathbf{X}}(x)\|_{\ell_2^n}^2}\, \mathrm{d}x \\ &+ \int_{\mathsf{Proj}_{z^\perp}(B_{\mathbf{X}})} \sqrt{1 + \|\nabla \psi_z^{\mathbf{X}}(x)\|_{\ell_2^n}^2}\, \mathrm{d}x. \end{aligned} \qquad (6.113)$$

When $\mathbf{X} = \ell_p^n$ for some $p \in (1, \infty)$ and $z = e_n$,

$$\forall x \in \mathsf{Proj}_{e_n^\perp}\big(B_{\ell_p^n}\big) = B_{\ell_p^{n-1}}, \quad \Psi_{e_n}^{\ell_p^n}(x) = -\psi_{e_n}^{\ell_p^n}(x) = \big(1 - \|x\|_{\ell_p^{n-1}}^p\big)^{\frac{1}{p}}.$$

Therefore, (6.113) becomes

$$\frac{\mathrm{vol}_{n-1}\big(\partial B_{\ell_p^n}\big)}{\mathrm{vol}_{n-1}\big(B_{\ell_p^{n-1}}\big)} = 2\fint_{B_{\ell_p^{n-1}}} \bigg(1 + (1 - \|x\|_{\ell_p^{n-1}}^p)^{-\frac{2(p-1)}{p}} \sum_{i=1}^{n-1} |x_i|^{2(p-1)}\bigg)^{\frac{1}{2}} \,\mathrm{d}x.$$

By [31], a point chosen from the normalized volume measure on $B_{\ell_p^{n-1}}$ is equidistributed with

$$\big(|\mathsf{G}_1|^p + \cdots + |\mathsf{G}_{n-1}|^p + \mathsf{Z}\big)^{-\frac{1}{p}}(\mathsf{G}_1, \ldots, \mathsf{G}_{n-1}) \in \mathbb{R}^{n-1},$$

where $\mathsf{G}_1, \ldots, \mathsf{G}_{n-1}, \mathsf{Z}$ are independent random variables, the density of $\mathsf{G}_1, \ldots, \mathsf{G}_{n-1}$ at $s \in \mathbb{R}$ is equal to $2\Gamma(1 + 1/p)^{-1} e^{-|s|^p}$ and the density of Z at $t \in [0, \infty)$ is equal to e^{-t}. Consequently,

$$\frac{\mathrm{vol}_{n-1}\big(\partial B_{\ell_p^n}\big)}{\mathrm{vol}_{n-1}\big(B_{\ell_p^{n-1}}\big)} = 2\mathbb{E}\Bigg[\bigg(1 + \mathsf{Z}^{-\frac{2(p-1)}{p}} \sum_{i=1}^{n-1} |\mathsf{G}_i|^{2(p-1)}\bigg)^{\frac{1}{2}}\Bigg]. \tag{6.114}$$

Optimal estimates on moments such as the right-hand side of (6.114) were derived (in greater generality) in [225], using which one can quickly get asymptotically sharp bounds on the left-hand side of (6.114). It is possible to implement this approach to get an alternative treatment of $\ell_p^n(\ell_q^m)$, though it is significantly more involved than the different way by which we proceeded above, and it becomes much more tedious and technically intricate when one aims to treat hierarchically nested ℓ_p norms as we did in Lemma 162. Nevertheless, an advantage of (6.113) is that it applies to normed spaces that do not have a product structure as in Lemma 157, which is helpful in other settings that we will study elsewhere.

6.2 Negatively correlated normed spaces

Our goal here is to further elucidate the role of symmetries in the context of the discussion in Section 1.6.2. Fix $n \in \mathbb{N}$ and $\gamma \geqslant 1$. Say that a normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is *γ-negatively correlated* if the standard scalar product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^n$ is invariant under its isometry group $\mathsf{Isom}(\mathbf{X})$, i.e., $\mathsf{Isom}(\mathbf{X}) \leqslant \mathsf{O}_n$, and there exists a Borel probability measure μ on $\mathsf{Isom}(\mathbf{X})$ such that

$$\forall x, y \in \mathbb{R}^n, \quad \int_{\mathsf{Isom}(\mathbf{X})} |\langle Ux, y\rangle| \,\mathrm{d}\mu(U) \leqslant \frac{\gamma}{\sqrt{n}} \|x\|_{\ell_2^n} \|y\|_{\ell_2^n}. \tag{6.115}$$

We were inspired to formulate this notion by the proof of [286, Theorem 1.1]. It is tailored for the purpose of bounding volumes of hyperplane projections of $B_{\mathbf{X}}$ from above in terms of the surface area of $\partial B_{\mathbf{X}}$, as exhibited by the following lemma which generalizes the reasoning in [286].

Lemma 166. *Fix $n \in \mathbb{N}$ and $\gamma \geqslant 1$. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is γ-negatively correlated, then*

$$\mathrm{MaxProj}(B_{\mathbf{X}}) \leqslant \frac{\gamma}{2\sqrt{n}} \mathrm{vol}_{n-1}(\partial B_{\mathbf{X}}).$$

Proof. Recall that for every $y \in \partial B_{\mathbf{X}}$ at which $\partial B_{\mathbf{X}}$ is smooth we denote the unit outer normal to $\partial B_{\mathbf{X}}$ at y by $N_{\mathbf{X}}(y) \in S^{n-1}$. By the Cauchy projection formula (1.30) for every $x \in S^{n-1}$ we have

$$\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})\big) = \frac{1}{2}\int_{\partial B_{\mathbf{X}}} |\langle x, N_{\mathbf{X}}(y)\rangle| \,\mathrm{d}y.$$

Since every $U \in \mathsf{Isom}(\mathbf{X})$ is an orthogonal transformation and $N_{\mathbf{X}} \circ U^* = U^* \circ N_{\mathbf{X}}$ almost surely on $\partial B_{\mathbf{X}}$,

$$\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})\big) = \frac{1}{2}\int_{\partial B_{\mathbf{X}}} |\langle Ux, N_{\mathbf{X}}(y)\rangle| \,\mathrm{d}y.$$

By integrating this identity with respect to μ, we therefore conclude that

$$\begin{aligned}\mathrm{vol}_{n-1}\big(\mathsf{Proj}_{x^\perp}(B_{\mathbf{X}})\big) &= \frac{1}{2}\int_{\partial B_{\mathbf{X}}} \bigg(\int_{\mathsf{Isom}(\mathbf{X})} |\langle Ux, N_{\mathbf{X}}(y)\rangle| \,\mathrm{d}\mu(U)\bigg) \mathrm{d}y\\ &\leqslant \frac{\gamma}{2\sqrt{n}} \mathrm{vol}_{n-1}(\partial B_{\mathbf{X}}),\end{aligned}$$

where we used (6.115) and the fact that $\|x\|_{\ell_2^n} = 1$ and $\|N_{\mathbf{X}}(y)\|_{\ell_2^n} = 1$ for almost every $y \in \partial B_{\mathbf{X}}$. ■

By substituting Lemma 166 into Theorem 76 and using (1.96), we get the following corollary.

Corollary 167. *Fix $n \in \mathbb{N}$ and $\gamma \geqslant 1$. If $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ is γ-negatively correlated, then*

$$\mathsf{e}(\mathbf{X}) \lesssim \mathsf{SEP}(\mathbf{X}) \leqslant 2\gamma \frac{\mathrm{vol}_{n-1}(\partial B_{\mathbf{X}})\, \mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})\sqrt{n}}.$$

Corollary 167 generalizes Corollary 45 since any canonically positioned normed space is 1-negatively correlated. Indeed, suppose that

$$\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$$

is canonically positioned. Recall that in Section 1.6.2 we denoted the Haar probability measure on $\mathsf{Isom}(\mathbf{X})$ by $\hslash_{\mathbf{X}}$. Fix $x, y \in \mathbb{R}^n$. The distribution of the random vector

Ux when U is distributed according to $\hbar_{\mathbf{X}}$ is $\mathsf{Isom}(\mathbf{X})$-invariant, and therefore it is isotropic. Hence,

$$\begin{aligned}\int_{\mathsf{Isom}(\mathbf{X})} |\langle Ux, y\rangle| \,\mathrm{d}\hbar_{\mathbf{X}}(U) &\leqslant \left(\int_{\mathsf{Isom}(\mathbf{X})} \langle Ux, y\rangle^2 \,\mathrm{d}\hbar_{\mathbf{X}}(U)\right)^{\frac12}\\ &\overset{(1.69)}{=} \frac{\|y\|_{\ell_2^n}}{\sqrt{n}}\left(\int_{\mathsf{Isom}(\mathbf{X})} \|Ux\|_{\ell_2^n}^2 \,\mathrm{d}\hbar_{\mathbf{X}}(U)\right)^{\frac12}\\ &= \frac{1}{\sqrt{n}}\|x\|_{\ell_2^n}\|y\|_{\ell_2^n},\end{aligned}$$

where the final step uses the fact that each $U \in \mathsf{Isom}(\mathbf{X})$ is an orthogonal transformation.

One way to achieve (6.115), which is close in spirit to the considerations in [286], is when there are $\Gamma \subseteq \{-1,1\}^n$ and $\Pi \subseteq S_n$ such that $U_{\varepsilon,\pi} \in \mathsf{Isom}(\mathbf{X})$ for every $(\varepsilon,\pi) \in \Gamma \times \Pi$, where $U_{\varepsilon,\pi} \in \mathsf{GL}_n(\mathbb{R})$ is given by

$$\forall x = (x_1,\ldots,x_n) \in \mathbb{R}^n, \quad U_{\varepsilon,\pi}x \overset{\mathrm{def}}{=} \big(\varepsilon_1 x_{\pi(1)},\ldots \varepsilon_n x_{\pi(n)}\big),$$

and also there are $\alpha, \beta > 0$ such that

$$\forall w \in \mathbb{R}^n, \quad \frac{1}{|\Gamma|}\sum_{\varepsilon\in\Gamma} |\langle \varepsilon, w\rangle| \leqslant \alpha\|w\|_{\ell_2^n} \tag{6.116}$$

and

$$\forall i,j \in \{1,\ldots,n\}, \quad |\{\pi \in \Pi : \pi(i) = j\}| \leqslant \beta\frac{|\Pi|}{n}. \tag{6.117}$$

Under these assumptions, $\mathbf{X}$ is γ-negatively correlated with $\gamma = \alpha\sqrt{\beta}$. Indeed, we can take μ in (6.115) to be the uniform distribution over the finite set

$$\{U_{\varepsilon,\pi} : (\varepsilon,\pi) \in \Gamma\times\Pi\} \subseteq \mathsf{Isom}(\mathbf{X}),$$

since every $x, y \in \mathbb{R}^n$ satisfy

$$\begin{aligned}\frac{1}{|\Gamma\times\Pi|}\sum_{(\varepsilon,\pi)\in\Gamma\times\Pi} |\langle U_{\varepsilon,\pi}x, y\rangle| &\overset{(6.116)}{\leqslant} \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\alpha\left(\sum_{i=1}^n (x_{\pi(i)}y_i)^2\right)^{\frac12}\\ &\leqslant \alpha\left(\sum_{i=1}^n\left(\frac{1}{|\Pi|}\sum_{\pi\in\Pi} x_{\pi(i)}^2\right)y_i^2\right)^{\frac12}\\ &= \alpha\left(\sum_{i=1}^n\left(\frac{1}{|\Pi|}\sum_{j=1}^n |\{\pi\in\Pi : \pi(i)=j\}|x_j^2\right)y_i^2\right)^{\frac12}\\ &\overset{(6.117)}{\leqslant} \frac{\alpha\sqrt{\beta}}{\sqrt{n}}\|x\|_{\ell_2^n}\|y\|_{\ell_2^n}.\end{aligned}$$

Condition (6.116) can be viewed as a negative correlation property of the coordinates of sign vectors that are chosen uniformly from Γ. Condition (6.117) roughly means that for each $i \in \{1,\ldots,n\}$ the sets $\{\pi \in \Pi : \pi(i) = 1\},\ldots,\{\pi \in \Pi : \pi(i) = n\}$ form an approximately equitable partition of Π. This holds with $\beta = 1$ if Π is a transitive subgroup of S_n. One could formulate weaker conditions that ensure the validity of the conclusion of Lemma 166 (e.g., considering bi-Lipschitz automorphisms of $\mathbf{X}$ rather than isometries of $\mathbf{X}$), and hence also the conclusion of Corollary 167, though we will not pursue this here as we expect that in concrete cases such issues should be easy to handle.

6.3 Volume ratio computations

Here we will present asymptotic evaluations of volume ratios of some normed spaces, for the purpose of plugging them into results that we stated in the Introdcution. Due to the large amount of knowledge on this topic that is available in the literature, we will only give a flavor of such applications. The main reference for the contents of this section is the valuable work [285].

We will start by examining the iteratively nested ℓ_p products $\{\mathbf{X}_k\}_{k=0}^{\infty}$ that we considered in Corollary 153, in the special case when the initial space $\mathbf{X} = \mathbf{X}_0$ is a canonically positioned normed space for which Conjecture 49 holds. Thus, we are fixing $\{n_k\}_{k=0}^{\infty} \subseteq \mathbb{N}$ and $\{p_k\}_{k=1}^{\infty} \subseteq [1,\infty]$, and assuming that

$$\mathbf{X} = (\mathbb{R}^{n_0}, \|\cdot\|_{\mathbf{X}})$$

is a canonically positioned normed space satisfying Conjecture 49, i.e., (6.16) holds with $\alpha = O(1)$; the case $\mathbf{X} = \mathbb{R}$ is sufficiently rich for our present illustrative purposes, but one can also take $\mathbf{X} = \mathbf{E}$ to be any symmetric space, per Lemma 54. By Corollary 153 and Corollary 79, if we define inductively

$$\forall k \in \mathbb{N}, \quad \mathbf{X}_{k+1} = \ell_{p_k}^{n_k}(\mathbf{X}_k), \quad \text{where } \mathbf{X}_0 = \mathbf{X},$$

then, because $\{\mathbf{X}_k\}_{k=1}^{\infty}$ are canonically positioned (they belong to the class of spaces in Example 40),

$$\forall m \in \mathbb{N}, \quad \mathsf{SEP}(\mathbf{X}_m) \asymp \mathrm{evr}(\mathbf{X}_m)\sqrt{\dim(\mathbf{X}_m)} = \mathrm{evr}(\mathbf{X}_m)\sqrt{n_0\cdots n_m}. \qquad (6.118)$$

Let $\{\mathbf{H}_k\}_{k=0}^{\infty}$ be the sequence of Euclidean spaces that arise from the above construction with the same $\{n_k\}_{k=0}^{\infty} \subseteq \mathbb{N}$ but with $p_k = 2$ for all $k \in \mathbb{N}$ and $\mathbf{X} = \ell_2^{n_0}$. Thus, for each $m\in\mathbb{N}$ the Euclidean space $\mathbf{H}_m$ can be identified naturally with $\ell_2^{n_0\cdots n_m}$. Under this identification, by a straightforward inductive application of Hölder's inequality and the fact that the ℓ_p norm deceases with p, the Löwner ellipsoid of

$\mathbf{X}_m$ satisfies[1]

$$\mathcal{L}_{\mathbf{X}_m} \subseteq \left(\prod_{k=1}^{m} n_k^{\max\{\frac{1}{2}-\frac{1}{p_k},0\}}\right)(\mathcal{L}_{\mathbf{X}})^{n_1\cdots n_m}.$$

Also, by Lemma 150 we have

$$\mathrm{vol}_{n_0\cdots n_m}(B_{\mathbf{X}_m})^{\frac{1}{n_0\cdots n_k}} \asymp \frac{\mathrm{vol}_{n_0}(B_{\mathbf{X}})^{\frac{1}{n_0}}}{\prod_{k=1}^{m} n_k^{\frac{1}{p_k}}}.$$

These facts combine to give the following consequence of (6.118):

$$\mathsf{SEP}(\mathbf{X}_m) \asymp \mathrm{evr}(\mathbf{X}) \prod_{k=1}^{m} n_k^{\max\{\frac{1}{2},\frac{1}{p_k}\}}.$$

In particular, when we take $\mathbf{X} = \mathbb{R}$ and consider only two steps of the above iteration, we get the following asymptotic evaluation of the separation modulus of the $\ell_p^n(\ell_q^m)$ norm the space of n-by-m matrices $M_{n\times m}(\mathbb{R})$ for any $n, m \in \mathbb{N}$ and $p, q \geqslant 1$; the case of square matrices was stated in the Introduction as (1.5):

$$\mathsf{SEP}\big(\ell_p^n(\ell_q^m)\big) \asymp n^{\max\{\frac{1}{p},\frac{1}{2}\}} m^{\max\{\frac{1}{q},\frac{1}{2}\}} = \max\Big\{\sqrt{nm}, m^{\frac{1}{q}}\sqrt{n}, n^{\frac{1}{p}}\sqrt{m}, n^{\frac{1}{p}} m^{\frac{1}{q}}\Big\}.$$

Next, fix an integer $n \geqslant 2$ and let $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ be an unconditional normed space. Given $q \in [2,\infty]$ and $\Lambda \geqslant 1$, one says (see, e.g., [182, Definition 1.f.4]) that $\mathbf{E}$ satisfies a lower q-estimate with constant Λ if for every $\{u_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^n$ with pairwise disjoint supports we have

$$\left(\sum_{k=1}^{\infty} \|u_k\|_{\mathbf{E}}^{q}\right)^{\frac{1}{q}} \leqslant \Lambda \left\|\sum_{k=1}^{\infty} u_k\right\|_{\mathbf{E}}. \tag{6.119}$$

Note that by (6.14) this always holds with $\Lambda = 1$ if $q = \infty$.

In concrete cases it is often mechanical to evaluate up to universal constant factors the minimum radius of a Euclidean ball that circumscribes $B_{\mathbf{X}}$, but it is always within a $O(\sqrt{\log n})$ factor of the expression

$$R_{\mathbf{E}} \stackrel{\mathrm{def}}{=} \max_{\varnothing\neq S\subseteq\{1,\ldots,n\}} \left(\frac{\sqrt{|S|}}{\|\sum_{i\in S} e_i\|_{\mathbf{E}}}\right). \tag{6.120}$$

More precisely, if $\mathbf{E}$ satisfies a lower q-estimate with constant Λ, then

$$R_{\mathbf{E}} \leqslant \mathrm{outradius}_{\ell_2^n}(B_{\mathbf{X}}) \lesssim \Lambda(\log n)^{\frac{1}{2}-\frac{1}{q}} R_{\mathbf{E}}. \tag{6.121}$$

[1] As $\mathbf{X}_m$ is canonically positioned, this holds as an equality, but for the present purposes we just need the stated inclusion.

The first inequality in (6.121) is immediate because $\|\sum_{i\in S} e_i\|_{\mathbf{E}}^{-1}\sum_{i\in S} e_i \in B_{\mathbf{E}}$ if $\varnothing \neq S \subseteq \{1,\ldots,n\}$. For a quick justification of the second inequality in (6.121), note that by homogeneity we may assume without loss of generality that $\|e_i\|_{\mathbf{E}} \geqslant 1$ for every $i \in \mathbb{N}$. Therefore, using (6.14) we see that if $x = (x_1,\ldots,x_n) \in B_{\mathbf{E}}$, then $\max_{i\in\{1,\ldots,n\}} |x_i| \leqslant 1$. Consequently, if we fix $x \in B_{\mathbf{E}}$ and denote for each $k \in \mathbb{N}$,

$$S_k = S_k(x) \stackrel{\mathrm{def}}{=} \left\{ i \in \{1,\ldots,n\} : \frac{1}{2^k} < |x_i| \leqslant \frac{1}{2^{k-1}} \right\}, \tag{6.122}$$

then the sets $\{S_k\}_{k=1}^{\infty}$ are a partition of $\{1,\ldots,n\}$ and in particular $\sum_{k=1}^{\infty} |S_k| = n$. Next,

$$\begin{aligned}\Lambda R_{\mathbf{E}} \geqslant \Lambda R_{\mathbf{E}}\|x\|_{\mathbf{E}} &\geqslant R_{\mathbf{E}}\left(\sum_{k=1}^{\infty}\left\|\sum_{i\in S_k} x_i e_i\right\|_{\mathbf{E}}^q\right)^{\frac{1}{q}}\\ &\geqslant \left(\sum_{k=1}^{\infty} R_{\mathbf{E}}^q \left\|\sum_{i\in S_k}\frac{1}{2^k} e_i\right\|_{\mathbf{E}}^q\right)^{\frac{1}{q}} \geqslant \left(\sum_{k=1}^{\infty}\frac{|S_k|^{\frac{q}{2}}}{2^{kq}}\right)^{\frac{1}{q}}.\end{aligned} \tag{6.123}$$

The second step of (6.123) uses (6.119), the penultimate step of (6.123) uses (6.14) and (6.122), and the final step of (6.123) uses (6.120). Now, for every $0 < \theta < 1$ we have

$$\begin{aligned}\|x\|_{\ell_2^n} &= \left(\sum_{k=1}^{\infty}\sum_{i\in S_k} x_i^2\right)^{\frac{1}{2}}\\ &\leqslant \left(\sum_{k=1}^{\infty}\frac{|S_k|}{2^{2(k-1)}}\right)^{\frac{1}{2}}\\ &= 2\left(\sum_{k=1}^{\infty}\frac{|S_k|^{1-\theta}}{2^{2k(1-\theta)}}|S_k|^{\theta}2^{-2k\theta}\right)^{\frac{1}{2}}\\ &\leqslant 2\left(\sum_{k=1}^{\infty}\frac{|S_k|^{\frac{q}{2}}}{2^{2kq}}\right)^{\frac{1-\theta}{q}}\left(\sum_{k=1}^{\infty}|S_k|\right)^{\frac{\theta}{2}}\left(\sum_{k=1}^{\infty}2^{-\frac{2kq\theta}{(q-2)(1-\theta)}}\right)^{(\frac{1}{2}-\frac{1}{q})(1-\theta)}\\ &\lesssim (\Lambda R_{\mathbf{E}})^{1-\theta} n^{\frac{\theta}{2}}\theta^{-(\frac{1}{2}-\frac{1}{q})},\end{aligned} \tag{6.124}$$

where the second step of (6.124) uses (6.122), the penultimate step of (6.124) uses trilinear Hölder with exponents $1/\theta, q/(2(1-\theta))$ and $1/((1-2/q)(1-\theta))$, and the final step of (6.124) uses (6.123), the fact that

$$\sum_{k=1}^{\infty}|S_k| = n,$$

and elementary calculus. By choosing $\theta = 1/\log n$ in (6.124), we get (6.121).

By the Lozanovskiĭ factorization theorem [186] there exist $w_1, \ldots, w_n > 0$ such that

$$\left\| \sum_{i=1}^{n} w_i e_i \right\|_{\mathbf{E}} = \left\| \sum_{i=1}^{n} \frac{1}{w_i} e_i \right\|_{\mathbf{E}^*} = \sqrt{n}. \tag{6.125}$$

We will call any $w_1, \ldots, w_n > 0$ that satisfy (6.125) Lozanovskiĭ weights for $\mathbf{E}$. They can be found by maximizing the concave function $w \mapsto \sum_{i=1}^{n} \log w_i$ over $w \in B_{\mathbf{E}}$ (see also, e.g., [263, Chapter 3]), which can be done efficiently if $\mathbf{E}$ is given by an efficient oracle; their existence can also be established non-constructively using the Brouwer fixed point theorem [135]. By [285, Lemma 1.2] (note that we are using a different normalization of the weights than in [285]),

$$\mathrm{vol}_n(B_{\mathbf{E}})^{\frac{1}{n}} \asymp \frac{(w_1 \cdots w_n)^{\frac{1}{n}}}{\sqrt{n}}. \tag{6.126}$$

By combining (6.121) and (6.126), we get the following lemma.

Lemma 168. *Fix an integer $n \geqslant 2$ and let $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$ be an unconditional normed space. Suppose that $\mathbf{E}$ satisfies a lower q-estimate with constant Λ for some $q \geqslant 2$ and $\Lambda \geqslant 1$. Then,*

$$\mathrm{evr}(\mathbf{E}) \lesssim \frac{\max_{\varnothing \neq S \subseteq \{1,\ldots,n\}} \left(\frac{\sqrt{|S|}}{\|\sum_{i \in S} e_i\|_{\mathbf{E}}} \right)}{\sqrt[n]{w_1 \cdots w_n}} \Lambda (\log n)^{\frac{1}{2} - \frac{1}{q}},$$

for any Lozanovskiĭ weights $w_1, \ldots, w_n > 0$ for $\mathbf{E}$. If the Löwner ellipsoid of $\mathbf{E}$ is a multiple of $B_{\ell_2^n}$, then

$$\begin{aligned} \frac{\max_{\varnothing \neq S \subseteq \{1,\ldots,n\}} \left(\frac{\sqrt{|S|}}{\|\sum_{i \in S} e_i\|_{\mathbf{E}}} \right)}{\sqrt[n]{w_1 \cdots w_n}} &\lesssim \mathrm{evr}(\mathbf{E}) \\ &\lesssim \frac{\max_{\varnothing \neq S \subseteq \{1,\ldots,n\}} \left(\frac{\sqrt{|S|}}{\|\sum_{i \in S} e_i\|_{\mathbf{E}}} \right)}{\sqrt[n]{w_1 \cdots w_n}} \Lambda (\log n)^{\frac{1}{2} - \frac{1}{q}}. \end{aligned}$$

The following corollary is a consequence of Lemma 168 because if

$$\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$$

is a normed space that satisfies the assumptions of Lemma 53 (in particular, $\mathbf{E}$ is unconditional), then by Lemma 152

$$w_1 = w_2 = \cdots = w_n = \frac{\sqrt{n}}{\|e_1 + \cdots + e_n\|_{\mathbf{E}}}$$

are Lozanovskiĭ weights for $\mathbf{E}$.

Corollary 169. *If* $\mathbf{E}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{E}})$ *a normed space that satisfies the assumptions of Lemma 53, then*

$$\frac{\|e_1+\cdots+e_n\|_{\mathbf{E}}}{\sqrt{n}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)$$
$$\lesssim \operatorname{evr}(\mathbf{E})\lesssim \frac{\|e_1+\cdots+e_n\|_{\mathbf{E}}}{\sqrt{n}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)\sqrt{\log n}.$$

Hence, by Corollary 79 we have

$$\|e_1+\cdots+e_n\|_{\mathbf{E}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)$$
$$\lesssim \mathsf{SEP}(\mathbf{E})\lesssim \|e_1+\cdots+e_n\|_{\mathbf{E}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)\sqrt{\log n},$$

More succinctly, this can be written in the following form, which we already stated in Corollary 4:

$$\mathsf{SEP}(\mathbf{E})=\|e_1+\cdots+e_n\|_{\mathbf{E}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)n^{o(1)}.$$

By [285, Proposition 2.2], the unitary ideal of any symmetric normed space $\mathbf{E}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{E}})$ satisfies

$$\operatorname{vr}(\mathsf{S}_{\mathbf{E}})\asymp\operatorname{vr}(\mathbf{E}). \tag{6.127}$$

This implies that

$$\operatorname{evr}(\mathsf{S}_{\mathbf{E}})\asymp\operatorname{evr}(\mathbf{E}), \tag{6.128}$$

by (1.71) combined with $\mathsf{S}_{\mathbf{E}}^*=\mathsf{S}_{\mathbf{E}^*}$, though a straightforward adjustment of the proof of (6.127) in [285] yields (6.128) directly, without using the much deeper result (1.71). We therefore have the following corollary.

Corollary 170. *If* $\mathbf{E}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{E}})$ *is a symmetric normed space, then*

$$\frac{\|e_1+\cdots+e_n\|_{\mathbf{E}}}{\sqrt{n}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)$$
$$\lesssim \operatorname{evr}(\mathsf{S}_{\mathbf{E}})\lesssim \frac{\|e_1+\cdots+e_n\|_{\mathbf{E}}}{\sqrt{n}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)\sqrt{\log n}.$$

Hence, by Corollary 79 we have

$$\|e_1+\cdots+e_n\|_{\mathbf{E}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)\sqrt{n}$$
$$\lesssim \mathsf{SEP}(\mathsf{S}_{\mathbf{E}})\lesssim \|e_1+\cdots+e_n\|_{\mathbf{E}}\left(\max_{k\in\{1,\ldots,n\}}\frac{\sqrt{k}}{\|e_1+\cdots+e_k\|_{\mathbf{E}}}\right)\sqrt{n}\log n,$$

More succinctly, this can be written in the following form, which we already stated in Corollary 4:

$$\mathsf{SEP}(\mathsf{S}_{\mathbf{E}}) = \|e_1 + \cdots + e_n\|_{\mathbf{E}} \left(\max_{k\in\{1,\ldots,n\}} \frac{\sqrt{k}}{\|e_1 + \cdots + e_k\|_{\mathbf{E}}} \right) n^{\frac{1}{2}+o(1)}.$$

Remark 171. In the above discussion, as well as in the ensuing treatment of tensor products, we prefer to consider square matrices rather than rectangular matrices because the setting of square matrices exhibits all of the key issues while being notationally simpler. Nevertheless, there are two places in which we do need to work with rectangular matrices, namely the above proof of Proposition 164 and the proof of the first inequality in (1.117). For the latter, fix $p \geqslant 1$ and $n, m \in \mathbb{N}$. As in the proof of Theorem 77, denote the Schatten–von Neumann trace class on the n-by-m real matrices $\mathsf{M}_{n\times m}(\mathbb{R})$ by $\mathsf{S}_p^{n\times m}$; recall (1.118). The following asymptotic identity implies (1.119) (recall that in the setting of (1.119) we have $r \in \{1, \ldots, n\}$)

$$\mathrm{evr}\big(\mathsf{S}_p^{n\times m}\big) \asymp \big(\min\{n,m\}\big)^{\max\{\frac{1}{p}-\frac{1}{2},0\}}. \tag{6.129}$$

Volumes of unit balls of Schatten–von Neumann trace classes have been satisfactorily estimated in the literature, starting with [293] and the comprehensive work [285], through the more precise asymptotics in [146, 277]. Unfortunately, all of these works dealt only with square matrices. Nevertheless, these references could be mechanically adjusted to treat rectangular matrices as well. Since (6.129) does not seem to have been stated in the literature, we will next sketch its derivation by mimicking the reasoning of [285], though the more precise statements of [146, 277] could be derived as well via similarly straightforward modifications of the known proofs for square matrices. We claim that

$$\mathrm{vol}_{nm}\big(B_{\mathsf{S}_p^{n\times m}}\big)^{\frac{1}{nm}} \asymp \frac{1}{\big(\min\{n,m\}\big)^{\frac{1}{p}}\sqrt{\max\{n,m\}}}. \tag{6.130}$$

(6.130) gives (6.129) since $\mathsf{S}_p^{n\times m}$ is canonically positioned, so by Hölder's inequality its Löwner ellipsoid is

$$\mathcal{L}_{\mathsf{S}_p^{n\times m}} = \big(\min\{n,m\}\big)^{\max\left\{\frac{1}{2}-\frac{1}{p},0\right\}} B_{\mathsf{S}_2^{n\times m}}.$$

To prove (6.130), note first that it follows from its special case $p = \infty$. Indeed, as $\mathsf{S}_1^{n\times m} = (\mathsf{S}_\infty^{n\times m})^*$, by the Blaschke–Santaló inequality [39, 278] and the Bourgain–Milman inequality [50] the case $p = 1$ of (6.130) follows from its case $p = \infty$. Now, (6.130) follows in full generality since by Hölder's inequality:

$$\frac{1}{\big(\min\{n,m\}\big)^{\frac{1}{p}}} B_{\mathsf{S}_\infty^{n\times m}} \subseteq B_{\mathsf{S}_p^{n\times m}} \subseteq \big(\min\{n,m\}\big)^{1-\frac{1}{p}} B_{\mathsf{S}_1^{n\times m}}.$$

The upper bound $\mathrm{vol}_{nm}(B_{\mathsf{S}_\infty^{n\times m}})^{1/(nm)} \lesssim 1/\sqrt{\max\{n,m\}}$ follows from the inclusion $B_{\mathsf{S}_\infty^{n\times m}} \subseteq \sqrt{\min\{n,m\}}B_{\mathsf{S}_2^{n\times m}}$. To justify the matching lower bound, if $\{\varepsilon_{ij}\}_{i,j\in\mathbb{N}}$ are i.i.d. Bernoulli random variables, then by [35, Theorem 1],

$$\mathbb{E}\left[\left\|\sum_{i=1}^{n}\sum_{j=1}^{m}\varepsilon_{ij}e_i\otimes e_j\right\|_{\mathsf{S}_\infty^{n\times m}}\right] \lesssim \sqrt{\max\{n,m\}},$$

This implies the lower bound $\mathrm{vol}_{nm}(B_{\mathsf{S}_\infty^{n\times m}})^{1/(nm)} \gtrsim 1/\sqrt{\max\{n,m\}}$ by an application of [285, Lemma 1.5].

Proof of Lemma 54. By [285, equation (2.2)] we have

$$\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{E}}}\big)^{\frac{1}{n^2}} \asymp \frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}\sqrt{n}}. \tag{6.131}$$

In particular,

$$\forall q\geqslant 1,\quad \mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)^{\frac{1}{n^2}} \asymp \frac{1}{n^{\frac12+\frac1q}}. \tag{6.132}$$

Because S_q^n is canonically positioned (it belongs to the class of spaces in Example 40), and hence it is in its minimum surface area position, by combining [104, Proposition 3.1] and (1.55) we see that

$$\frac{\mathrm{vol}_{n^2-1}\big(\partial B_{\mathsf{S}_q^n}\big)}{\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)} \asymp \frac{n\mathrm{MaxProj}\big(B_{\mathsf{S}_q^n}\big)}{\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)} \overset{(1.43)}{\asymp} n^{\frac32+\frac1q}\sqrt{\min\{q,n\}}. \tag{6.133}$$

Consequently,

$$\begin{aligned}\mathrm{iq}\big(B_{\mathsf{S}_q^n}\big) &= n\frac{\mathrm{vol}_{n^2-1}\big(\partial B_{\mathsf{S}_q^n}\big)}{\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)}\mathrm{vol}_{n^2}\big(B_{\mathsf{S}_q^n}\big)^{\frac{1}{n^2}}\\ &\overset{(6.132)\wedge(6.133)}{\asymp} \frac{n^{\frac32+\frac1q}\sqrt{\min\{q,n\}}}{n^{\frac12+\frac1q}} = n\sqrt{\min\{q,n\}}.\end{aligned} \tag{6.134}$$

Because by (6.14) we have

$$\forall x\in\mathbb{R}^n,\quad \|x\|_{\mathbf{E}} \leqslant \|e_1+\cdots+e_n\|_{\mathbf{E}}\|x\|_{\ell_\infty^n},$$

every matrix $A\in\mathsf{M}_n(\mathbb{R})$ satisfies

$$\|A\|_{\mathsf{S}_{\mathbf{E}}} \leqslant \|e_1+\cdots+e_n\|_{\mathbf{E}}\|A\|_{\mathsf{S}_\infty^n} \leqslant \|e_1+\cdots+e_n\|_{\mathbf{E}}\|A\|_{\mathsf{S}_q^n}.$$

Consequently,

$$\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}B_{\mathsf{S}_q^n} \subseteq B_{\mathsf{S}_{\mathbf{E}}}. \tag{6.135}$$

Moreover,

$$\operatorname{iq}\left(\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}B_{\mathsf{S}_q^n}\right)=\operatorname{iq}\big(B_{\mathsf{S}_q^n}\big)\overset{(6.134)}{\asymp}n\sqrt{\min\{q,n\}}$$

and

$$\operatorname{vol}_{n^2}\left(\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}B_{\mathsf{S}_q^n}\right)^{\frac{1}{n^2}}\overset{(6.132)}{\asymp}\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}n^{\frac12+\frac1q}}\overset{(6.131)}{\asymp}\frac{\operatorname{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{E}}}\big)^{\frac{1}{n^2}}}{n^{\frac1q}}.$$

By choosing $q=\log n$ we get (1.80) for the normed space $\mathbf{Y}$ whose unit ball is the left-hand side of (6.135). ∎

Remark 172. An inspection of the proof of Lemma 54 reveals that if Conjecture 49 holds for S_∞^n, then also Conjecture 49 holds for $\mathsf{S}_{\mathbf{E}}$ for any symmetric normed space $\mathbf{E}=(\mathbb{R}^n,\|\cdot\|_{\mathbf{E}})$. Indeed, we would then take $\mathbf{Y}'=(\mathsf{M}_n(\mathbb{R}),\|\cdot\|_{\mathbf{Y}'})$ to be the normed space whose unit ball is

$$B_{\mathbf{Y}'}=\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}\operatorname{Ch}S_\infty^n=\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}\mathsf{S}_{\chi\ell_\infty^n},$$

where we recall Corollary 43. If Conjecture 49 holds for S_∞^n, then we would have $n\asymp\operatorname{iq}(\operatorname{Ch}S_\infty^n)=\operatorname{iq}(B_{\mathbf{Y}'})$, and also

$$\operatorname{vol}_{n^2}\big(\operatorname{Ch}S_\infty^n\big)^{\frac{1}{n^2}}\asymp\operatorname{vol}_{n^2}\big(S_\infty^n\big)^{\frac{1}{n^2}}\overset{(6.132)}{\asymp}\frac{1}{\sqrt{n}},$$

from which we see that

$$\operatorname{vol}_{n^2}\big(B_{\mathbf{Y}'}\big)^{\frac{1}{n^2}}=\frac{\operatorname{vol}_{n^2}\big(\operatorname{Ch}S_\infty^n\big)^{\frac{1}{n^2}}}{\|e_1+\cdots+e_n\|_{\mathbf{E}}}\asymp\frac{1}{\|e_1+\cdots+e_n\|_{\mathbf{E}}\sqrt{n}}\overset{(6.131)}{\asymp}\operatorname{vol}_{n^2}\big(B_{\mathsf{S}_{\mathbf{E}}}\big)^{\frac{1}{n^2}}.$$

This proves Conjecture 49 for $\mathsf{S}_{\mathbf{E}}$. Note in passing that this also implies that

$$\frac{1}{\sqrt{n}}\asymp\operatorname{vol}_{n^2}\big(S_{\chi\ell_\infty^n}\big)^{\frac{1}{n^2}}\overset{(6.131)}{\asymp}\frac{1}{\|e_1+\cdots+e_n\|_{\chi\ell_\infty^n}\sqrt{n}}.$$

Hence, if Conjecture 49 holds for S_∞^n, then we would have $\|e_1+\cdots+e_n\|_{\chi\ell_\infty^n}\asymp 1$. More generally, by mimicking the above reasoning we deduce that if Conjecture 49 holds for $\mathsf{S}_{\mathbf{E}}$, then $\|e_1+\cdots+e_n\|_{\chi\mathbf{E}}\asymp\|e_1+\cdots+e_n\|_{\mathbf{E}}$, which would be a modest step towards Problem 44.

Fix $n\in\mathbb{N}$ and $p,q\geqslant 1$. We claim that the volume ratio of the projective tensor product $\ell_p^n\widehat{\otimes}\ell_q^n$ satisfies

$$\operatorname{vr}\big(\ell_p^n\widehat{\otimes}\ell_q^n\big)\asymp\Phi_{p,q}(n),\tag{6.136}$$

where

$$\Phi_{p,q}(n) \stackrel{\mathrm{def}}{=} \begin{cases} 1 & \text{if } 1 \leqslant p, q \leqslant 2, \\ n^{\frac{1}{2}-\frac{1}{p}} & \text{if } q \leqslant 2 \leqslant p \leqslant \frac{q}{q-1}, \\ n^{\frac{1}{q}-\frac{1}{2}} & \text{if } q \leqslant 2 \leqslant \frac{q}{q-1} \leqslant p, \\ n^{\frac{1}{2}-\frac{1}{q}} & \text{if } p \leqslant 2 \leqslant q \leqslant \frac{p}{p-1}, \\ n^{\frac{1}{p}-\frac{1}{2}} & \text{if } p \leqslant 2 \leqslant \frac{p}{p-1} \leqslant q, \\ 1 & \text{if } p, q \geqslant 2 \text{ and } \frac{1}{p}+\frac{1}{q} \geqslant \frac{1}{2}, \\ n^{\frac{1}{2}-\frac{1}{p}-\frac{1}{q}} & \text{if } \frac{1}{p}+\frac{1}{q} \leqslant \frac{1}{2}. \end{cases} \tag{6.137}$$

Assuming (6.137) for the moment, by substituting it into Theorem 3 we get that

$$\begin{aligned} \mathsf{SEP}\big(\ell_p^n \check{\otimes} \ell_q^n\big) &\gtrsim n\,\mathrm{vr}\big(\big(\ell_p^n \check{\otimes} \ell_q^n\big)^*\big) = n\,\mathrm{vr}\big(\ell_{p^*}^n \hat{\otimes} \ell_{q^*}^n\big) \\ &\asymp n\Phi_{p^*,q^*}(n) = \begin{cases} n & \text{if } p, q \geqslant 2, \\ n^{\frac{1}{p}+\frac{1}{2}} & \text{if } \frac{q}{q-1} \leqslant p \leqslant 2 \leqslant q, \\ n^{\frac{3}{2}-\frac{1}{q}} & \text{if } p \leqslant \frac{q}{q-1} \leqslant 2 \leqslant q, \\ n^{\frac{1}{q}+\frac{1}{2}} & \text{if } \frac{p}{p-1} \leqslant q \leqslant 2 \leqslant p, \\ n^{\frac{3}{2}-\frac{1}{p}} & \text{if } q \leqslant \frac{p}{p-1} \leqslant 2 \leqslant p, \\ n & \text{if } p, q \leqslant 2 \text{ and } \frac{1}{p}+\frac{1}{q} \leqslant \frac{3}{2}, \\ n^{\frac{1}{p}+\frac{1}{q}-\frac{1}{2}} & \text{if } \frac{1}{p}+\frac{1}{q} \geqslant \frac{3}{2}. \end{cases} \end{aligned}$$

Since for any two normed spaces $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ and $\mathbf{Y} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{Y}})$ the space of operators from $\mathbf{X}^*$ to $\mathbf{Y}$ is isometric to the injective tensor product $\mathbf{X}^* \check{\otimes} \mathbf{Y}$ (see, e.g., [87]), we get from this that

$$\begin{aligned} \mathsf{SEP}\big(\mathsf{M}_n(\mathbb{R}), \|\cdot\|_{\ell_p^n \to \ell_q^n}\big) &= \mathsf{SEP}\big(\ell_{p^*}^n \check{\otimes} \ell_q^n\big) \\ &\gtrsim \begin{cases} n & \text{if } p \leqslant 2 \leqslant q, \\ n^{\frac{3}{2}-\frac{1}{p}} & \text{if } 2 \leqslant p \leqslant q, \\ n^{\frac{3}{2}-\frac{1}{q}} & \text{if } 2 \leqslant q \leqslant p, \\ n^{\frac{1}{q}+\frac{1}{2}} & \text{if } p \leqslant q \leqslant 2, \\ n^{\frac{1}{p}+\frac{1}{2}} & \text{if } q \leqslant p \leqslant 2, \\ n & \text{if } \frac{2p}{p+2} \leqslant q \leqslant 2 \leqslant p, \\ n^{\frac{1}{q}-\frac{1}{p}+\frac{1}{2}} & \text{if } q \leqslant \frac{2p}{p+2}. \end{cases} \end{aligned} \tag{6.138}$$

Note that the rightmost quantity in (6.138) coincides with the right-hand side of (1.14). Since $\ell_p^n \check{\otimes} \ell_q^n$ belongs to the class of spaces in Example 40, a positive answer to Conjecture 11 for $\ell_p^n \check{\otimes} \ell_q^n$ would imply the following asymptotic evaluation of $\mathsf{SEP}(\ell_p^n \check{\otimes} \ell_q^n)$, which is equivalent to (1.14):

$$\mathsf{SEP}\big(\ell_p^n \check{\otimes} \ell_q^n\big) \asymp \begin{cases} n & \text{if } p,q \geqslant 2, \\ n^{\frac{1}{2}+\frac{1}{p}} & \text{if } \frac{q}{q-1} \leqslant p \leqslant 2 \leqslant q, \\ n^{\frac{3}{2}-\frac{1}{q}} & \text{if } p \leqslant \frac{q}{q-1} \leqslant 2 \leqslant q, \\ n^{\frac{1}{2}+\frac{1}{q}} & \text{if } \frac{p}{p-1} \leqslant q \leqslant 2 \leqslant p, \\ n^{\frac{3}{2}-\frac{1}{p}} & \text{if } q \leqslant \frac{p}{p-1} \leqslant 2 \leqslant p, \\ n & \text{if } p,q \leqslant 2 \text{ and } \frac{1}{p}+\frac{1}{q} \leqslant \frac{3}{2}, \\ n^{\frac{1}{p}+\frac{1}{q}-\frac{1}{2}} & \text{if } \frac{1}{p}+\frac{1}{q} \geqslant \frac{3}{2}. \end{cases}$$

Furthermore, by Theorem 80 the leftmost quantity in (6.138) is bounded from above by $O(\log n)$ times the rightmost quantity in (6.138), thus implying the fourth bullet point of Corollary 4.

The asymptotic evaluation (6.136) of $\mathrm{vr}(\ell_p^n \hat{\otimes} \ell_q^n)$ was proved in [285] up to constant factors that depend on p, q, namely [285, Theorem 3.1] states that

$$\forall\, p,q > 1, \quad \mathrm{vr}\big(\ell_p^n \hat{\otimes} \ell_q^n\big) \asymp_{p,q} \Phi_{p,q}(n). \tag{6.139}$$

If $2 \in \{p,q\}$ and also $\min\{p,q\} \leqslant 2$, then (6.139) is due to Szarek and Tomczak-Jaegermann [293]. More recently, Defant and Michels [84] generalized (6.139) to projective tensor products of symmetric normed spaces that are either 2-convex or 2-concave. The proof of (6.139) in [285] yields constants that degenerate as $\min\{p,q\}$ tends to 1. We will therefore next improve the reasoning in [285] to get (6.136).

Lemma 173. *Fix $n \in \mathbb{N}$ and $p,q \geqslant 1$. Let $\{\varepsilon_{ij}\}_{i,j \in \{1,\ldots,n\}}$ be i.i.d. Bernoulli random variables (namely, they are independent and each of them is uniformly distributed over $\{-1,1\}$). Then,*

$$\mathbb{E}\left[\bigg\| \sum_{i=1}^n \sum_{j=1}^n \varepsilon_{ij} e_i \otimes e_j \bigg\|_{\ell_p^n \check{\otimes} \ell_q^n}\right] \asymp n^{\beta(p,q)}$$

$$\stackrel{\mathrm{def}}{=} \begin{cases} n^{\frac{1}{p}+\frac{1}{q}-\frac{1}{2}} & \textit{if } \max\{p,q\} \leqslant 2, \\ n^{\frac{1}{\min\{p,q\}}} & \textit{if } \max\{p,q\} \geqslant 2. \end{cases} \tag{6.140}$$

Citing the work [79] of Chevet, a version of Lemma 173 appears as [285, Lemma 2.3], except that in [285, Lemma 2.3] the implicit constants in (6.140) depend on p, q. An inspection of the proof of (6.139) in [285] reveals that this is the only source of the dependence of the constants on p, q (in fact, for this purpose [285] only needs half of (6.140), namely to bound from above its left-hand side by its right-hand side). Specifically, all of the steps within [285] incur only a loss of a universal constant factor, and the proof of (6.139) in [285] also appeals to inequalities in the earlier work [284] of Schütt, as well a classical inequality of Hardy and Littlewood [127]; all of the constants in these cited inequalities are universal. Therefore, (6.136) will be established after we prove Lemma 173.

Proof of Lemma 173. We will denote the random matrix whose (i, j) entry is ε_{ij} by $\mathcal{E} \in \mathsf{M}_n(\mathbb{R})$. Then, the goal is

$$\mathbb{E}\big[\|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_q^n}\big] \asymp n^{\beta(p,q)}. \tag{6.141}$$

In fact, the lower bound on the expected norm in (6.141) holds always, i.e., for a universal constant $c > 0$,

$$\forall A \in \mathsf{M}_n(\{-1, 1\}), \quad \|A\|_{\ell_{p^*}^n \to \ell_q^n} \geqslant cn^{\beta(p,q)}. \tag{6.142}$$

A justification of (6.142) appears in the *proof of* Proposition 3.2 of Bennett's work [34] (specifically, see the reasoning immediately after [34, inequality (15)]), where it is explained that we can take $c = 1$ if $\min\{p^*, q\} \geqslant 2$ or $\max\{p^*, q\} \leqslant 2$, and that we can take $c = 1/\sqrt{2}$ otherwise.

Next, let $\{\mathsf{g}_{ij}\}_{i,j\in\{1,\ldots,n\}}$ be i.i.d. standard Gaussian random variables. By [79, Lemme 3.1],

$$\mathbb{E}\left[\left\|\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{g}_{ij} e_i \otimes e_j\right\|_{\ell_p^n \check{\otimes} \ell_q^n}\right] \asymp n^{\max\{\frac{1}{p}+\frac{1}{q}-\frac{1}{2},\frac{1}{p}\}}\sqrt{p} + n^{\max\{\frac{1}{p}+\frac{1}{q}-\frac{1}{2},\frac{1}{q}\}}\sqrt{q}. \tag{6.143}$$

Consequently,

$$\begin{aligned}\mathbb{E}\left[\left\|\sum_{i=1}^{n}\sum_{j=1}^{n} \varepsilon_{ij} e_i \otimes e_j\right\|_{\ell_p^n \check{\otimes} \ell_q^n}\right] &\leqslant \sqrt{\frac{\pi}{2}}\,\mathbb{E}\left[\left\|\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{g}_{ij} e_i \otimes e_j\right\|_{\ell_p^n \check{\otimes} \ell_q^n}\right] \\ &\lesssim n^{\beta(p,q)}\sqrt{\max\{p, q\}},\end{aligned} \tag{6.144}$$

where the first step of (6.144) is a standard comparison between Rademacher and Gaussian averages (a quick consequence of Jensen's inequality; e.g., [204]) and the final step of (6.144) uses (6.143). This proves the desired bound (6.140) when

$$\max\{p, q\} \leqslant 2,$$

so suppose from now on that $\max\{p, q\} \geqslant 2$.

It suffices to treat the case $p \geqslant 2$. Indeed, if $p \leqslant 2$, then $q \geqslant 2$ since $\max\{p, q\} \geqslant 2$, so by the duality

$$\|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_q^n} = \|\mathcal{E}^*\|_{\ell_{q^*}^n \to \ell_p^n},$$

and the fact that the transpose $\mathcal{E}^*$ has the same distribution as $\mathcal{E}$, the case $p \leqslant 2$ follows from the case $p \geqslant 2$. It also suffices to treat the case $q \leqslant p$ because if $q \geqslant p$, then $\|\cdot\|_{\ell_q^n} \leqslant \|\cdot\|_{\ell_p^n}$ point-wise, and therefore

$$\|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_q^n} \leqslant \|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_p^n}.$$

Consequently, since $\beta(p,q) = \beta(p,p)$ when $q \geqslant p$, the case $q \geqslant p$ follows from the case $q = p$.

So, suppose from now that $p \geqslant 2$ and $q \leqslant p$. If we denote

$$r \stackrel{\mathrm{def}}{=} \frac{q(p-2)}{p-q},$$

with the convention $r = \infty$ if $q = p$, then $r \geqslant 1$ and

$$\frac{1}{q} = \frac{1-\theta}{r} + \frac{\theta}{2}, \quad \text{where } \theta \stackrel{\mathrm{def}}{=} \frac{2}{p} \in [0,1]. \tag{6.145}$$

Hence, by the Riesz–Thorin interpolation theorem [272, 301] we have

$$\begin{aligned} \|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_q^n} &\leqslant \|\mathcal{E}\|_{\ell_1^n \to \ell_r^n}^{1-\theta} \|\mathcal{E}\|_{\ell_2^n \to \ell_2^n}^{\theta} \\ &= \Big(\max_{i \in \{1,\ldots,n\}} \|\mathcal{E} e_i\|_{\ell_r^n} \Big)^{1-\theta} \|\mathcal{E}\|_{\ell_2^n \to \ell_2^n}^{\theta} = n^{\frac{1-\theta}{r}} \|\mathcal{E}\|_{\ell_2^n \to \ell_2^n}^{\theta}. \end{aligned}$$

By taking expectations of this inequality, we get that

$$\begin{aligned} \mathbb{E}\big[\|\mathcal{E}\|_{\ell_{p^*}^n \to \ell_q^n}\big] &\leqslant n^{\frac{1-\theta}{r}} \mathbb{E}\big[\|\mathcal{E}\|_{\ell_2^n \to \ell_2^n}^{\theta}\big] \leqslant n^{\frac{1-\theta}{r}} \big(\mathbb{E}\big[\|\mathcal{E}\|_{\ell_2^n \to \ell_2^n}\big]\big)^{\theta} \\ &\lesssim n^{\frac{1-\theta}{r} + \frac{\theta}{2}} = n^{\frac{1}{q}} = n^{\beta(p,q)}, \end{aligned} \tag{6.146}$$

where the second step of (6.146) uses Jensen's inequality, the third step of (6.146) uses the classical fact that the expectation of the operator norm from ℓ_2^n to ℓ_2^n of an $n \times n$ matrix whose entries are i.i.d. symmetric Bernoulli random variables is $O(\sqrt{n})$ (this follows from (6.144), though it is older; see, e.g., [35]), the penultimate step of (6.146) uses (6.145), and the last step of (6.146) uses the definition of $\beta(p,q)$ in (6.140) while recalling that we are now treating the case $p \geqslant 2$ and $q \leqslant p$. ■

A substitution of Lemma 173 into the proof of [285, Lemma 3.2] yields the following asymptotic evaluations of the n^2-roots of volumes of the unit balls of injective and projective tensor products; the statement of [285, Lemma 3.2] is identical, except that the constant factors depend on p, q, but that is due only to the dependence of the constants on p, q in [285, Lemma 2.3], which Lemma 173 removes

$$\mathrm{vol}_{n^2}\big(B_{\ell_p^n \check{\otimes} \ell_q^n}\big)^{\frac{1}{n^2}} \asymp n^{-\beta(p,q)} \quad \text{and} \quad \mathrm{vol}_{n^2}\big(B_{\ell_p^n \hat{\otimes} \ell_q^n}\big)^{\frac{1}{n^2}} \asymp n^{\beta(p^*,q^*)-2}. \tag{6.147}$$

Since $\ell_p^n \widehat{\otimes} \ell_q^n$ belongs to the class of spaces in Example 40, its Löwner ellipsoid is the minimal multiple of the standard Euclidean ball $B_{\mathsf{S}_2^n}$ that superscribes the unit ball of $\ell_p^n \widehat{\otimes} \ell_q^n$, namely

$$\mathcal{L}_{\ell_p^n \widehat{\otimes} \ell_q^n} = R(n,p,q) B_{\mathsf{S}_2^n},$$

where, since $B_{\ell_p^n \widehat{\otimes} \ell_q^n}$ is the convex hull of $B_{\ell_p^n} \otimes B_{\ell_q^n}$,

$$\begin{aligned} R(n,p,q) &= \max_{\substack{x \in B_{\ell_p^n} \\ y \in B_{\ell_q^n}}} \|x \otimes y\|_{\mathsf{S}_2^n} \\ &= \Big(\max_{x \in B_{\ell_p^n}} \|x\|_{\ell_2^n}\Big)\Big(\max_{y \in B_{\ell_q^n}} \|y\|_{\ell_2^n}\Big) = n^{\max\{\frac{1}{2}-\frac{1}{p},0\}+\max\{\frac{1}{2}-\frac{1}{q},0\}}. \end{aligned} \tag{6.148}$$

By combining (6.147) and (6.148) we get that

$$\begin{aligned} \mathrm{vr}\big(\ell_{p^*}^n \check{\otimes} \ell_{q^*}^n\big) &\overset{(1.71)}{\asymp} \mathrm{evr}\big(\ell_p^n \widehat{\otimes} \ell_q^n\big) \\ &= R(n,p,q)\left(\frac{\mathrm{vol}_{n^2}(B_{\mathsf{S}_2^n})}{\mathrm{vol}_{n^2}(B_{\ell_p^n \widehat{\otimes} \ell_q^n})}\right)^{\frac{1}{n^2}} \\ &\asymp n^{\max\{\frac{1}{2}-\frac{1}{p},0\}+\max\{\frac{1}{2}-\frac{1}{q},0\}-\beta(p^*,q^*)+1} \\ &\overset{(6.140)}{=} \begin{cases} \sqrt{n} & \text{if } \max\{p,q\} \geqslant 2, \\ n^{\frac{1}{\max\{p,q\}}} & \text{if } \max\{p,q\} \leqslant 2. \end{cases} \end{aligned} \tag{6.149}$$

A substitution of (6.149) into Theorem 3 gives

$$\mathsf{SEP}\big(\ell_p^n \widehat{\otimes} \ell_q^n\big) \gtrsim \begin{cases} n^{\frac{3}{2}} & \text{if } \max\{p,q\} \geqslant 2, \\ n^{1+\frac{1}{\max\{p,q\}}} & \text{if } \max\{p,q\} \leqslant 2. \end{cases} \tag{6.150}$$

Furthermore, if Conjecture 11 holds for $\ell_p^n \widehat{\otimes} \ell_q^n$, then (6.150) is sharp, namely (1.15) holds. Also, by Theorem 80 the left-hand side of (6.150) is bounded from above by $O(\log n)$ times the right-hand side of (6.150), thus implying the fifth bullet point of Corollary 4.

Remark 174. The above results imply clustering statements (and impossibility thereof) for norms that have significance to algorithms and complexity theory. For example, the *cut norm* [101] on $\mathsf{M}_n(\mathbb{R})$ is $O(1)$-equivalent [6] to the operator norm from ℓ_∞^n to ℓ_1^n. So, by (1.13) the separation modulus of the cut norm on $\mathsf{M}_n(\mathbb{R})$ is predicted to be bounded above and below by universal constant multiples of $n^{3/2}$, and by Theorem 80 we know that it is at least a universal constant multiple of $n^{3/2}$ and at most a universal constant multiple of $n^{3/2}\log n$. As another notable example, we proved that

$$\mathsf{SEP}(\ell_\infty^n \widehat{\otimes} \ell_\infty^n) \gtrsim n^{\frac{3}{2}}.$$

Moreover, if Conjecture 11 holds for $\ell_\infty^n \hat{\otimes} \ell_\infty^n$, then $\mathsf{SEP}(\ell_\infty^n \hat{\otimes} \ell_\infty^n) \asymp n^{3/2}$ and by Theorem 80 we have

$$\mathsf{SEP}(\ell_\infty^n \hat{\otimes} \ell_\infty^n) \lesssim n^{\frac{3}{2}} \log n.$$

Grothendieck's inequality [121] implies that

$$\forall A \in \mathsf{M}_n(\mathbb{R}), \quad \|A\|_{\ell_\infty^n \hat{\otimes} \ell_\infty^n} \asymp \gamma_2^{1\to\infty}(A), \tag{6.151}$$

where $\gamma_2^{1\to\infty}(A)$ is the factorization-through-ℓ_2 norm (see [261]) of A as an operator from ℓ_1^n to ℓ_∞^n, i.e.,

$$\begin{aligned}\gamma_2^{1\to\infty}(A) &\stackrel{\text{def}}{=} \min_{\substack{X,Y\in\mathsf{M}_n(\mathbb{R})\\ A=XY}} \|X\|_{\ell_2^n\to\ell_\infty^n}\|Y\|_{\ell_1^n\to\ell_2^n}\\ &= \min_{\substack{X,Y\in\mathsf{M}_n(\mathbb{R})\\ A=XY}} \max_{i,j\in\{1,\ldots,n\}} \|\mathrm{row}_i(X)\|_{\ell_2^n}\|\mathrm{column}_j(Y)\|_{\ell_2^n}.\end{aligned}$$

Above, for $i, j \in \{1,\ldots,n\}$ and $M \in \mathsf{M}_n(\mathbb{R})$ we denote by $\mathrm{row}_i(M)$ and $\mathrm{column}_j(M)$ the ith row and jth column of M, respectively. See [183] for the justification of (6.151), as well as the importance of the factorization norm $\gamma_2^{1\to\infty}$ to complexity theory (see [38, 202] for further algorithmic significance of factorization norms). Thanks to the above discussion, we know that

$$n^{\frac{3}{2}} \lesssim \mathsf{SEP}\big(\mathsf{M}_n(\mathbb{R}), \gamma_2^{1\to\infty}\big) \lesssim n^{\frac{3}{2}} \log n,$$

and that $\mathsf{SEP}(\mathsf{M}_n(\mathbb{R}), \gamma_2^{1\to\infty}) \asymp n^{3/2}$ assuming Conjecture 11. To check that this does not follow from the previously known bounds (1.2), we need to know the asymptotic growth rate of the Banach–Mazur distance between $\ell_\infty^n \hat{\otimes} \ell_\infty^n$ and each of the spaces $\ell_1^{n^2}, \ell_2^{n^2}$. However, these Banach–Mazur distances do not appear in the literature. In response to our inquiry, Carsten Schütt answered this question, by showing that

$$d_{\mathrm{BM}}\big(\ell_2^{n^2}, \ell_\infty^n \hat{\otimes} \ell_\infty^n\big) \asymp d_{\mathrm{BM}}\big(\ell_1^{n^2}, \ell_\infty^n \hat{\otimes} \ell_\infty^n\big) \asymp n. \tag{6.152}$$

More generally, Schütt succeeded to evaluate the asymptotic growth rate of the Banach–Mazur distance between $\ell_p^n \hat{\otimes} \ell_q^n$ and $\ell_p^n \check{\otimes} \ell_q^n$ to each of $\ell_1^{n^2}, \ell_2^{n^2}$ for every $p, q \in [1,\infty]$ (this is a substantial matter that Schütt communicated to us privately and he will publish it elsewhere). Due to (6.152), an application of (1.2) only gives the bounds $n \lesssim \mathsf{SEP}(\ell_\infty^n \hat{\otimes} \ell_\infty^n) \lesssim n^2$, which hold for *every* n^2-dimensional normed space. More generally, Schütt's result shows that (1.13) and (1.15) do not follow from (1.2).

The volume computations of this section are only an indication of the available information. The literature contains many more volume estimates that could be substituted into Theorem 3 and Conjecture 6 to yield new results (and conjectures) on separation moduli of various spaces; examples of further pertinent results appear in [20, 85, 88, 104, 110, 115–117, 145, 146, 285].

Chapter 7

Logarithmic weak isomorphic isoperimetry in minimum dual mean width position

In this section we will prove the results that we stated in Section 1.6.3. We first claim that for every integer $n \geqslant 2$ and every $r > 0$ we have

$$\begin{aligned}\mathrm{iq}\big(B_{\ell_\infty^n} \cap (rB_{\ell_2^n})\big) &= \mathrm{iq}\big([-1,1]^n \cap (rB_{\ell_2^n})\big) \\ &\gtrsim \left(\min\{\sqrt{n}, r\}\left(1 - \frac{1}{\max\{1, r^2\}}\right)^{\frac{n-1}{2}} + 1\right)\sqrt{n}. \end{aligned} \tag{7.1}$$

Observe that (7.1) implies (1.85). Furthermore, (7.1) implies the direction $\gtrsim$ in (1.86) because

$$\begin{aligned}&\min_{r>0} \frac{\mathrm{iq}\big(B_{\ell_\infty^n} \cap (rB_{\ell_2^n})\big)}{\sqrt{n}} \left(\frac{\mathrm{vol}_n(B_{\ell_\infty^n})}{\mathrm{vol}_n\big(B_{\ell_\infty^n} \cap (rB_{\ell_2^n})\big)}\right)^{\frac{1}{n}} \\ &\geqslant \min_{r>0} \frac{\mathrm{iq}\big(B_{\ell_\infty^n} \cap (rB_{\ell_2^n})\big)}{\sqrt{n}} \left(\frac{2^n}{\mathrm{vol}_n\big(rB_{\ell_2^n}\big)}\right)^{\frac{1}{n}} \\ &\gtrsim \min_{r>0}\left(\min\left\{\frac{\sqrt{n}}{r}, 1\right\}\left(1 - \frac{1}{\max\{1, r^2\}}\right)^{\frac{n-1}{2}} + \frac{1}{r}\right)\sqrt{n} \\ &\asymp \sqrt{\log n},\end{aligned}$$

where the penultimate step uses (7.1) and the final step is elementary calculus. Since the K-convexity constant of ℓ_∞^n satisfies $K(\ell_\infty^n) \asymp \sqrt{\log n}$ (see [263, Chapter 2]), the matching upper bound in (1.86) will follow after we will prove (below) Proposition 61. This will also show that Proposition 61 is sharp, though it would be worthwhile to find out if it is sharp even for some normed space $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ for which $K(\mathbf{X}) \asymp \log n$; such a space exists by a remarkable (randomized) construction of Bourgain [44].

To prove (7.1), note first that if $0 < r \leqslant 1$, then $rB_{\ell_2^n} \subseteq [-1,1]^n$ and therefore

$$\forall 0 < r \leqslant 1, \quad \mathrm{iq}\big([-1,1]^n \cap (rB_{\ell_2^n})\big) = \mathrm{iq}(rB_{\ell_2^n}) \asymp \sqrt{n}. \tag{7.2}$$

Similarly, note that if $r \geqslant \sqrt{n}$, then $rB_{\ell_2^n} \subseteq [-1,1]^n$ and therefore

$$\forall r \geqslant \sqrt{n}, \quad \mathrm{iq}\big([-1,1]^n \cap (rB_{\ell_2^n})\big) = \mathrm{iq}\big([-1,1]^n\big) \asymp n. \tag{7.3}$$

Both (7.2) and (7.3) coincide with (7.1) in the respective ranges. The less trivial range of (7.1) is when $1 < r < \sqrt{n}$, in which case the boundary of $[-1,1]^n \cap (rB_{\ell_2^n})$ contains the disjoint union of the intersection of $rB_{\ell_2^n}$ with the $2n$ faces of $[-1,1]^n$, each of which is isometric to the following set:

$$[-1,1]^{n-1} \cap \big(\sqrt{r^2-1}B_{\ell_2^{n-1}}\big).$$

Together with the straightforward inclusion

$$[-1,1]^{n-1} \cap \big(\sqrt{r^2-1}B_{\ell_2^{n-1}}\big) \supseteq \sqrt{1-\frac{1}{r^2}}\big([-1,1]^{n-1} \cap (rB_{\ell_2^{n-1}})\big),$$

the above observation implies that if $1 < r < \sqrt{n}$, then

$$\begin{aligned}
&\mathrm{vol}_{n-1}\big(\partial\big([-1,1]^n \cap (rB_{\ell_2^n})\big)\big)\\
&\geqslant 2n\left(1-\frac{1}{r^2}\right)^{\frac{n-1}{2}} \mathrm{vol}_{n-1}\big([-1,1]^{n-1} \cap (rB_{\ell_2^{n-1}})\big)\\
&= n\left(1-\frac{1}{r^2}\right)^{\frac{n-1}{2}} \mathrm{vol}_{n}\big(\big([-1,1]^{n-1} \cap (rB_{\ell_2^{n-1}})\big) \times [-1,1]\big)\\
&\geqslant n\left(1-\frac{1}{r^2}\right)^{\frac{n-1}{2}} \mathrm{vol}_{n}\big([-1,1]^{n} \cap (rB_{\ell_2^{n}})\big),
\end{aligned} \tag{7.4}$$

where the final step (7.4) is a consequence of the straightforward inclusion

$$\big([-1,1]^{n-1} \cap (rB_{\ell_2^{n-1}})\big) \times [-1,1] \supseteq [-1,1]^n \cap (rB_{\ell_2^n}).$$

By combining (7.4) with the definition (1.11) of the isoperimetric quotient, we see that

$$\begin{aligned}
\mathrm{iq}\big([-1,1]^n \cap (rB_{\ell_2^n})\big) &\geqslant \frac{n\big(1-\frac{1}{r^2}\big)^{\frac{n-1}{2}} \mathrm{vol}_n\big([-1,1]^n \cap (rB_{\ell_2^n})\big)}{\mathrm{vol}_n\big([-1,1]^n \cap (rB_{\ell_2^n})\big)^{\frac{n-1}{n}}}\\
&= n\left(1-\frac{1}{r^2}\right)^{\frac{n-1}{2}} \mathrm{vol}_n\big([-1,1]^n \cap (rB_{\ell_2^n})\big)^{\frac{1}{n}}.
\end{aligned} \tag{7.5}$$

When $r \leqslant \sqrt{n}$ we have

$$[-1,1]^n \cap (rB_{\ell_2^n}) \supseteq \left[-\frac{r}{\sqrt{n}}, \frac{r}{\sqrt{n}}\right]^n.$$

In combination with (7.5), this implies that

$$\forall 1 < r < \sqrt{n}, \quad \mathrm{iq}\big([-1,1]^n \cap (rB_{\ell_2^n})\big) \geqslant 2r\sqrt{n}\left(1-\frac{1}{r^2}\right)^{\frac{n-1}{2}}.$$

As also $\mathrm{iq}([-1,1]^n \cap (rB_{\ell_2^n})) \gtrsim \sqrt{n}$ by the isoperimetric theorem (1.12), this completes the proof of (7.1). ∎

Passing to the proof of Proposition 61, observe first that for every $r > 0$ we have

$$\frac{\mathrm{vol}_n\big(B_{\mathbf{X}} \cap (rB_{\ell_2^n})\big)}{\mathrm{vol}_n(rB_{\ell_2^n})} = \frac{\mathrm{vol}_n\big(\{x \in rB_{\ell_2^n} : \|x\|_{\mathbf{X}} \leqslant 1\}\big)}{\mathrm{vol}_n(rB_{\ell_2^n})}$$
$$\geqslant 1 - \fint_{rB_{\ell_2^n}} \|x\|_{\mathbf{X}}\,\mathrm{d}x = 1 - \frac{nr}{n+1}M(\mathbf{X}), \tag{7.6}$$

where the penultimate step in (7.6) is Markov's inequality and the final step in (7.6) is integration in polar coordinates using the following standard notation for the mean of the norm on the Euclidean sphere:

$$M(\mathbf{X}) \stackrel{\mathrm{def}}{=} \fint_{S^{n-1}} \|z\|_{\mathbf{X}}\,\mathrm{d}z.$$

We will also use the common notation $M^*(\mathbf{X}) \stackrel{\mathrm{def}}{=} M(\mathbf{X}^*)$. By setting $r = 1/(2M(\mathbf{X}))$ in (7.6) we get that

$$\mathrm{vol}_n\left(B_{\mathbf{X}} \cap \left(\frac{1}{2M(\mathbf{X})}B_{\ell_2^n}\right)\right)^{\frac{1}{n}} \geqslant \left(\frac{1}{2}\,\mathrm{vol}_n\left(\frac{1}{2M(\mathbf{X})}B_{\ell_2^n}\right)\right)^{\frac{1}{n}} \asymp \frac{1}{M(\mathbf{X})\sqrt{n}}. \tag{7.7}$$

This simple consideration gives the following general elementary lemma.

Lemma 175. *Let $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$ be a normed space. If we denote $r = 1/(2M(\mathbf{X}))$ and $L = B_{\mathbf{X}} \cap (rB_{\ell_2^n})$, then we have*

$$\mathrm{vol}_n(L)^{\frac{1}{n}} \gtrsim \frac{1}{M(\mathbf{X})\sqrt{n}} \quad \textit{and} \quad \frac{\mathrm{MaxProj}(L)}{\mathrm{vol}_n(L)^{\frac{n-1}{n}}} \lesssim 1. \tag{7.8}$$

Proof. The first inequality in (7.8) follows from (7.7). For the second inequality in (7.8), observe that $\mathsf{Proj}_{z^\perp}(L) \subseteq \mathsf{Proj}_{z^\perp}(rB_{\ell_2^n})$ for every $z \in S^{n-1}$, since $L \subseteq rB_{\ell_2^n}$. Consequently,

$$\mathrm{MaxProj}(L) \leqslant \mathrm{MaxProj}\big(rB_{\ell_2^n}\big) = r^{n-1}\,\mathrm{vol}_{n-1}\big(B_{\ell_2^{n-1}}\big)$$
$$\asymp \mathrm{vol}_n\big(rB_{\ell_2^n}\big)^{\frac{n-1}{n}} \lesssim \mathrm{vol}_n(L)^{\frac{n-1}{n}}, \tag{7.9}$$

where the penultimate step of (7.9) is a standard computation using Stirling's formula and the final step of (7.9) uses the first inequality in (7.7). ∎

By (1.55), the second inequality in (7.8) implies that $\mathrm{iq}(L) \lesssim \sqrt{n}$. Hence, in order to use Lemma 175 in the context of Conjecture 10 it would be beneficial to choose $S \in \mathsf{SL}_n(\mathbb{R})$ for which $M(S\mathbf{X})$ is small. So, fix $\delta > 0$ and suppose that

$$\delta M(S\mathbf{X}) \leqslant \min_{T \in \mathsf{SL}_n(\mathbb{R})} M(T\mathbf{X}). \tag{7.10}$$

By compactness, this holds for some $S \in \mathsf{SL}_n(\mathbb{R})$ with $\delta = 1$, in which case the polar of $SB_{\mathbf{X}}$ is in *minimum mean width position* and we will say that $S\mathbf{X}$ is in *minimum dual mean width position* (the terminology that is used in [108] is that $SB_{\mathbf{X}}$ *has minimal* M). By [107], the matrix in $\mathsf{SL}_n(\mathbb{R})$ at which $\min_{T \in \mathsf{SL}_n(\mathbb{R})} M(T\mathbf{X})$ is attained is unique up to orthogonal transformations. We allow the flexibility of working with some universal constant $0 < \delta < 1$ rather than considering only the minimum dual mean width position since this will encompass other commonly used positions, such as the ℓ-position (see [55, Section 1.11]). By [107], $\mathbf{X}$ is in minimum dual mean width position if and only if the measure $\mathrm{d}\nu_{\mathbf{X}}(z) = \|z\|_{\mathbf{X}}\,\mathrm{d}z$ on S^{n-1} is isotropic. Since $\nu_{\mathbf{X}}$ is evidently $\mathsf{Isom}(\mathbf{X})$-invariant, by (1.69) if $\mathbf{X}$ is canonically positioned, then it is in minimum dual mean width position.

Let γ denote the standard Gaussian measure on $\mathbb{R}$, i.e., its density at $u \in \mathbb{R}$ equals $\exp(-u^2/2)/\sqrt{2\pi}$. The (Gaussian) K-convexity constant $K(\mathbf{X})$ of $\mathbf{X}$ is defined [204] to be the infimum over those $K > 0$ that satisfy

$$\begin{aligned}\bigg(\int_{\mathbb{R}^{\aleph_0}} \bigg\| \sum_{i=1}^{\infty} \mathsf{g}_i' \int_{\mathbb{R}^{\aleph_0}} \mathsf{g}_i f(\mathsf{g})\,\mathrm{d}\gamma^{\otimes \aleph_0}(\mathsf{g}) \bigg\|_{\mathbf{X}}^2 \mathrm{d}\gamma^{\otimes \aleph_0}(\mathsf{g}')\bigg)^{\frac12} \\ \leqslant K\bigg(\int_{\mathbb{R}^{\aleph_0}} \|f(\mathsf{g})\|_{\mathbf{X}}^2\,\mathrm{d}\gamma^{\otimes \aleph_0}(\mathsf{g})\bigg)^{\frac12},\end{aligned}$$

for every measurable $f : \mathbb{R}^{\aleph_0} \to \mathbf{X}$ with $\int_{\mathbb{R}^{\aleph_0}} \|f(\mathsf{g})\|_{\mathbf{X}}^2\,\mathrm{d}\gamma^{\otimes \aleph_0}(\mathsf{g}) < \infty$. By [100] there is $T \in \mathsf{SL}_n(\mathbb{R})$ such that $M(T\mathbf{X})M^*(T\mathbf{X}) \leqslant K(\mathbf{X})$. By the above assumption (7.10) we know that $\delta M(S\mathbf{X}) \leqslant M(T\mathbf{X})$, so $\delta M(S\mathbf{X}) \leqslant K(\mathbf{X})/M^*(T\mathbf{X})$. Next, we always have

$$M(\mathbf{X}) \geqslant \left(\frac{\mathrm{vol}_n(B_{\ell_2^n})}{\mathrm{vol}_n(B_{\mathbf{X}})}\right)^{\frac1n};$$

see, e.g., [218, Section 2] and [132, Lemma 30] for two derivations of this well-known volumetric lower bound on $M(\mathbf{X})$. Applying this lower bound to the dual of $T\mathbf{X}$, we get $M^*(T\mathbf{X}) \geqslant (\mathrm{vol}_n(B_{\ell_2^n})/\,\mathrm{vol}_n(B_{\mathbf{X}^*}))^{1/n}$. The Blaschke–Santaló inequality [39, 278] states that

$$\frac{\mathrm{vol}_n(B_{\ell_2^n})}{\mathrm{vol}_n(B_{\mathbf{X}^*})} \geqslant \frac{\mathrm{vol}_n(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\ell_2^n})},$$

so we conclude that $\delta M(S\mathbf{X})\sqrt{n} \lesssim K(\mathbf{X})/\sqrt[n]{\mathrm{vol}_n(B_{\mathbf{X}})}$. A substitution of this bound into Lemma 175 gives the following proposition:

Proposition 176. *Fix* $0 < \delta \leqslant 1$ *and a normed space* $\mathbf{X} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{X}})$. *Suppose that* $S \in \mathsf{SL}_n(\mathbb{R})$ *satisfies*

$$\delta M(S\mathbf{X}) \leqslant \min_{T \in \mathsf{SL}_n(\mathbb{R})} M(T\mathbf{X}).$$

Then, denoting $r = 1/(2M(S\mathbf{X}))$ we have

$$\mathrm{vol}_n\big((SB_{\mathbf{X}}) \cap (rB_{\ell_2^n})\big)^{\frac{1}{n}} \gtrsim \frac{\delta}{K(\mathbf{X})} \mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}$$

and

$$\mathrm{MaxProj}\big((SB_{\mathbf{X}}) \cap (rB_{\ell_2^n})\big) \lesssim \mathrm{vol}_n\big((SB_{\mathbf{X}}) \cap (rB_{\ell_2^n})\big)^{\frac{n-1}{n}}.$$

Furthermore, if $\mathbf{X}$ is canonically positioned, then this holds when S is the identity matrix and $\delta = 1$.

By (1.55), Proposition 176 implies Proposition 61, with the additional information that the conclusion of Proposition 61 holds with S the identity matrix if $\mathbf{X}$ is in minimum dual mean width position, in which case we obtain an upper bound on $\mathrm{MaxProj}(L)$. Hence, by the reasoning in Section 1.6, if $\mathbf{X}$ is in minimum dual mean width position, then

$$\mathsf{SEP}(\mathbf{X}) \lesssim K(\mathbf{X}) \frac{\mathrm{diam}_{\ell_2^n}(B_{\mathbf{X}})}{\mathrm{vol}_n(B_{\mathbf{X}})^{\frac{1}{n}}}.$$

References

[1] I. Abraham, Y. Bartal, and O. Neiman, Advances in metric embedding theory. *Adv. Math.* **228** (2011), no. 6, 3026–3126

[2] D. Achlioptas, Database-friendly random projections: Johnson–Lindenstrauss with binary coins. *J. Comput. System Sci.* **66** (2003), no. 4, 671–687

[3] F. Albiac and N. J. Kalton, *Topics in Banach space theory*. Grad. Texts in Math. 233, Springer, New York, 2006

[4] N. Alon, R. M. Karp, D. Peleg, and D. B. West, A graph-theoretic game and its application to the k-server problem (extended abstract). In *On-line algorithms, proceedings of a DIMACS workshop, new brunswick, new jersey, usa, february 11-13, 1991*, edited by L. A. McGeoch and D. D. Sleator, pp. 1–10, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 7, DIMACS/AMS, 1991

[5] N. Alon and V. D. Milman, Embedding of l_∞^k in finite-dimensional Banach spaces. *Israel J. Math.* **45** (1983), no. 4, 265–280

[6] N. Alon and A. Naor, Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.* **35** (2006), no. 4, 787–803

[7] F. Alter and V. Caselles, Uniqueness of the Cheeger set of a convex body. *Nonlinear Anal.* **70** (2009), no. 1, 32–44

[8] F. Alter, V. Caselles, and A. Chambolle, A characterization of convex calibrable sets in $\mathbb{R}^N$. *Math. Ann.* **332** (2005), no. 2, 329–366

[9] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*. Oxford Math. Monogr., The Clarendon Press, Oxford University Press, New York, 2000

[10] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. 2nd edn., Lectures Math. ETH Zürich, Birkhäuser, Basel, 2008

[11] L. Ambrosio and D. Puglisi, Linear extension operators between spaces of Lipschitz maps and optimal transport. *J. Reine Angew. Math.* **764** (2020), 1–21

[12] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*. Cambridge Stud. Adv. Math. 118, Cambridge University Press, Cambridge, 2010

[13] A. Andoni and P. Indyk, Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In *47th annual IEEE symposium on foundations of computer science (FOCS 2006), 21-24 october 2006, berkeley, california, usa, proceedings*, pp. 459–468, IEEE Computer Society, 2006

[14] G. E. Andrews, R. Askey, and R. Roy, *Special functions*. Encyclopedia Math. Appl. 71, Cambridge University Press, Cambridge, 1999

[15] R. Anisca, A. Tcaciuc, and N. Tomczak-Jaegermann, Structure of normed spaces with extremal distance to the Euclidean space. *Houston J. Math.* **31** (2005), no. 1, 267–283

[16] R. F. Arens and J. Eells, Jr., On embedding uniform and topological spaces. *Pacific J. Math.* **6** (1956), 397–403

[17] T. Austin, A CAT(0)-valued pointwise ergodic theorem. *J. Topol. Anal.* **3** (2011), no. 2, 145–152

[18] B. Awerbuch and D. Peleg, Sparse partitions (extended abstract). In *31st annual symposium on foundations of computer science, st. louis, missouri, usa, october 22-24, 1990, volume II*, pp. 503–513, IEEE Computer Society, 1990

[19] K. Ball, Volumes of sections of cubes and related problems. In *Geometric aspects of functional analysis (1987–88)*, pp. 251–260, Lecture Notes in Math. 1376, Springer, Berlin, 1989

[20] K. Ball, Normed spaces with a weak-Gordon–Lewis property. In *Functional analysis (Austin, TX, 1987/1989)*, pp. 36–47, Lecture Notes in Math. 1470, Springer, Berlin, 1991

[21] K. Ball, Shadows of convex bodies. *Trans. Amer. Math. Soc.* **327** (1991), no. 2, 891–901

[22] K. Ball, Volume ratios and a reverse isoperimetric inequality. *J. London Math. Soc. (2)* **44** (1991), no. 2, 351–359

[23] K. Ball, Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.* **2** (1992), no. 2, 137–172

[24] K. Ball, Convex geometry and functional analysis. In *Handbook of the geometry of Banach spaces, Vol. I*, pp. 161–194, North-Holland, Amsterdam, 2001

[25] K. Ball and A. Pajor, Convex bodies with few faces. *Proc. Amer. Math. Soc.* **110** (1990), no. 1, 225–231

[26] K. M. Ball, K. J. Böröczky, and A. Naor, Approximate isoperimetry for convex polytopes. Forthcoming manuscript

[27] S. Banach, *Théorie des opérations linéaires*. Éditions Jacques Gabay, Sceaux, 1993

[28] I. Bárány and Z. Füredi, Computing the volume is difficult. *Discrete Comput. Geom.* **2** (1987), no. 4, 319–326

[29] Y. Bartal, Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996)*, pp. 184–193, IEEE Computer Society Press, Los Alamitos, CA, 1996

[30] Y. Bartal, On approximating arbitrary metrices by tree metrics. In *STOC '98 (Dallas, TX)*, pp. 161–168, ACM, New York, 1999

[31] F. Barthe, O. Guédon, S. Mendelson, and A. Naor, A probabilistic approach to the geometry of the l_p^n-ball. *Ann. Probab.* **33** (2005), no. 2, 480–513

[32] F. Barthe and A. Naor, Hyperplane projections of the unit ball of ℓ_p^n. *Discrete Comput. Geom.* **27** (2002), no. 2, 215–226

[33] G. Basso, Fixed point theorems for metric spaces with a conical geodesic bicombing. *Ergodic Theory Dynam. Systems* **38** (2018), no. 5, 1642–1657

[34] G. Bennett, Schur multipliers. *Duke Math. J.* **44** (1977), no. 3, 603–639

[35] G. Bennett, V. Goodman, and C. M. Newman, Norms of random matrices. *Pacific J. Math.* **59** (1975), no. 2, 359–365

[36] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. Vol. 1*. Amer. Math. Soc. Colloq. Publ. 48, American Mathematical Society, Providence, RI, 2000

[37] R. Bhatia, *Matrix analysis*. Grad. Texts in Math. 169, Springer, New York, 1997

[38] V. Bhattiprolu, E. Lee, and A. Naor, A framework for quadratic form maximization over convex sets through nonconvex relaxations. In *STOC '21: 53rd annual ACM SIGACT symposium on theory of computing, virtual event, italy, june 21-25, 2021*, edited by S. Khuller and V. V. Williams, pp. 870–881, ACM, 2021

[39] W. Blaschke, Über affine Geometrie VII: Neue Extremeigenschaften von Ellipse und Ellipsoid. Leipz. Ber. 69, 306–318, 1917

[40] S. Bochner, Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind. *Fundam. Math.* **20** (1933), 262–276

[41] E. D. Bolker, A class of convex bodies. *Trans. Amer. Math. Soc.* **145** (1969), 323–345

[42] F. Bolley, Separability and completeness for the Wasserstein distance. In *Séminaire de probabilités XLI*, pp. 371–377, Lecture Notes in Math. 1934, Springer, Berlin, 2008

[43] J. Bourgain, A remark on finite-dimensional P_λ-spaces. *Studia Math.* **72** (1982), no. 3, 285–289

[44] J. Bourgain, On martingales transforms in finite-dimensional lattices with an appendix on the K-convexity constant. *Math. Nachr.* **119** (1984), 41–53

[45] J. Bourgain, On dimension free maximal inequalities for convex symmetric bodies in $\mathbf{R}^n$. In *Geometrical aspects of functional analysis (1985/86)*, pp. 168–176, Lecture Notes in Math. 1267, Springer, Berlin, 1987

[46] J. Bourgain, Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical aspects of functional analysis (1985/86)*, pp. 157–167, Lecture Notes in Math. 1267, Springer, Berlin, 1987

[47] J. Bourgain and J. Lindenstrauss, Projection bodies. In *Geometric aspects of functional analysis (1986/87)*, pp. 250–270, Lecture Notes in Math. 1317, Springer, Berlin, 1988

[48] J. Bourgain, J. Lindenstrauss, and V. Milman, Approximation of zonoids by zonotopes. *Acta Math.* **162** (1989), no. 1-2, 73–141

[49] J. Bourgain and V. D. Milman, Sections euclidiennes et volume des corps symétriques convexes dans $\mathbf{R}^n$. *C. R. Acad. Sci. Paris Sér. I Math.* **300** (1985), no. 13, 435–438

[50] J. Bourgain and V. D. Milman, New volume ratio properties for convex symmetric bodies in $\mathbf{R}^n$. *Invent. Math.* **88** (1987), no. 2, 319–340

[51] J. Bourgain and S. J. Szarek, The Banach–Mazur distance to the cube and the Dvoretzky-Rogers factorization. *Israel J. Math.* **62** (1988), no. 2, 169–180

[52] J. Bourgain and L. Tzafriri, Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis. *Israel J. Math.* **57** (1987), no. 2, 137–224

[53] L. Brasco, On principal frequencies and isoperimetric ratios in convex sets. *Ann. Fac. Sci. Toulouse Math. (6)* **29** (2020), no. 4, 977–1005

[54] M. Braverman and A. Naor, Quantitative Wasserstein rounding. Forthcoming manuscript

[55] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.-H. Vritsiou, *Geometry of isotropic convex bodies*. Math. Surveys Monogr. 196, American Mathematical Society, Providence, RI, 2014

[56] E. Breuillard, M. W. Liebeck, A. Naor, and A. Rizzoli, On the inverse problem for isometry groups of Banach spaces. Forthcoming manuscript

[57] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Grundlehren Math. Wiss. 319, Springer, Berlin, 1999

[58] B. Brinkman and M. Charikar, On the impossibility of dimension reduction in l_1. *J. ACM* **52** (2005), no. 5, 766–788

[59] R. L. Brooks, On colouring the nodes of a network. *Proc. Cambridge Philos. Soc.* **37** (1941), 194–197

[60] A. Brudnyĭ and Y. Brudnyĭ, Simultaneous extensions of Lipschitz functions. *Uspekhi Mat. Nauk* **60** (2005), no. 6(366), 53–72

[61] A. Brudnyi and Y. Brudnyi, Extension of Lipschitz functions defined on metric subspaces of homogeneous type. *Rev. Mat. Complut.* **19** (2006), no. 2, 347–359

[62] A. Brudnyi and Y. Brudnyi, Linear and nonlinear extensions of Lipschitz functions from subsets of metric spaces. *Algebra i Analiz* **19** (2007), no. 3, 106–118

[63] A. Brudnyi and Y. Brudnyi, Metric spaces with linear extensions preserving Lipschitz condition. *Amer. J. Math.* **129** (2007), no. 1, 217–314

[64] A. Brudnyi and Y. Brudnyi, *Methods of geometric analysis in extension and trace problems. Volume 2*. Monogr. Math. 103, Birkhäuser/Springer, Basel, 2012

[65] Y. Brudnyi and P. Shvartsman, Stability of the Lipschitz extension property under metric transforms. *Geom. Funct. Anal.* **12** (2002), no. 1, 73–79

[66] J. A. Brudnyĭ and N. J. Krugljak, Functors of real interpolation. *Dokl. Akad. Nauk SSSR* **256** (1981), no. 1, 14–17

[67] D. Bucur and I. Fragalà, Blaschke–Santaló and Mahler inequalities for the first eigenvalue of the Dirichlet Laplacian. *Proc. Lond. Math. Soc. (3)* **113** (2016), no. 3, 387–417

[68] D. Bucur and I. Fragalà, Reverse Faber–Krahn and Mahler inequalities for the Cheeger constant. *Proc. Roy. Soc. Edinburgh Sect. A* **148** (2018), no. 5, 913–937

[69] P. Buser, A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup. (4)* **15** (1982), no. 2, 213–230

[70] G. J. Butler, Simultaneous packing and covering in euclidean space. *Proc. London Math. Soc. (3)* **25** (1972), 721–735

[71] G. Calinescu, H. Karloff, and Y. Rabani, Approximation algorithms for the 0-extension problem. *SIAM J. Comput.* **34** (2004/05), no. 2, 358–372

[72] B. Carl, Inequalities of Bernstein–Jackson-type and the degree of compactness of operators in Banach spaces. *Ann. Inst. Fourier (Grenoble)* **35** (1985), no. 3, 79–118

[73] B. Carl and A. Pajor, Gel'fand numbers of operators with values in a Hilbert space. *Invent. Math.* **94** (1988), no. 3, 479–504

[74] V. Caselles, A. Chambolle, and M. Novaga, Uniqueness of the Cheeger set of a convex body. *Pacific J. Math.* **232** (2007), no. 1, 77–90

[75] G. D. Chakerian and P. Filliman, The measures of the projections of a cube. *Studia Sci. Math. Hungar.* **21** (1986), no. 1-2, 103–110

[76] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. A. Plotkin, Approximating a finite metric by a small number of tree metrics. In *39th annual symposium on foundations of computer science, FOCS '98, november 8-11, 1998, palo alto, california, USA*, pp. 379–388, IEEE Computer Society, 1998

[77] I. Chavel, *Eigenvalues in Riemannian geometry*. Pure Appl. Math. 115, Academic Press, Orlando, FL, 1984

[78] J. Cheeger, A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pp. 195–199, Princeton University Press, Princeton, NJ, 1970

[79] S. Chevet, Séries de variables aléatoires gaussiennes à valeurs dans $E\hat{\otimes}_\varepsilon F$. Application aux produits d'espaces de Wiener abstraits. In *Séminaire sur la Géométrie des Espaces de Banach (1977–1978)*, Exp. No. 19, 15 pp., École polytechnique, Palaiseau, 1978

[80] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer, Some intersection theorems for ordered sets and graphs. *J. Combin. Theory Ser. A* **43** (1986), no. 1, 23–37

[81] R. Courant and D. Hilbert, *Methods of mathematical physics. Vol. I*. Interscience Publishers, New York, NY, 1953

[82] M. Cwikel, K-divisibility of the K-functional and Calderón couples. *Ark. Mat.* **22** (1984), no. 1, 39–62

[83] M. M. Day, Polygons circumscribed about closed convex curves. *Trans. Amer. Math. Soc.* **62** (1947), 315–319

[84] A. Defant and C. Michels, Norms of tensor product identities. *Note Mat.* **25** (2005/06), no. 1, 129–166

[85] A. Defant and C. Prengel, Volume estimates in spaces of homogeneous polynomials. *Math. Z.* **261** (2009), no. 4, 909–932

[86] D. Descombes and U. Lang, Convex geodesic bicombings and hyperbolicity. *Geom. Dedicata* **177** (2015), 367–384

[87] J. Diestel, J. H. Fourie, and J. Swart, *The metric theory of tensor products*. American Mathematical Society, Providence, RI, 2008

[88] A. Doležalová and J. Vybíral, On the volume of unit balls of finite-dimensional Lorentz spaces. *J. Approx. Theory* **255** (2020), article no. 105407, 20 pp.

[89] R. Durrett, *Probability—theory and examples*. Camb. Ser. Stat. Probab. Math. 49, Cambridge University Press, Cambridge, 2019

[90] A. Dvoretzky, Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pp. 123–160, Jerusalem Academic Press, Jerusalem; Pergamon, Oxford, 1961

[91] A. Dvoretzky and C. A. Rogers, Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.* **36** (1950), 192–197

[92] J. Elton, Sign-embeddings of l_1^n. *Trans. Amer. Math. Soc.* **279** (1983), no. 1, 113–124

[93] P. Enflo, Uniform structures and square roots in topological groups. Part II. *Israel J. Math. 8 (1970), 230-252; ibid.* **8** (1970), 253–272

[94] A. Es-Sahib and H. Heinich, Barycentre canonique pour un espace métrique à courbure négative. In *Séminaire de Probabilités, XXXIII*, pp. 355–370, Lecture Notes in Math. 1709, Springer, Berlin, 1999

[95] G. Faber, Beweis, daß unter allen homogenen Membranen von gleicher Fläche und gleicher Spannung die kreisförmige den tiefsten Grundton gibt. *Münch. Ber.* **1923** (1923), 169–172

[96] J. Fakcharoenphol, S. Rao, and K. Talwar, A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. System Sci.* **69** (2004), no. 3, 485–497

[97] T. Figiel, Review of [147]. *Math. Rev.* **53** (1977), no. 3649

[98] T. Figiel and W. B. Johnson, Large subspaces of l_∞^n and estimates of the Gordon–Lewis constant. *Israel J. Math.* **37** (1980), no. 1-2, 92–112

[99] T. Figiel, J. Lindenstrauss, and V. D. Milman, The dimension of almost spherical sections of convex bodies. *Acta Math.* **139** (1977), no. 1-2, 53–94

[100] T. Figiel and N. Tomczak-Jaegermann, Projections onto Hilbertian subspaces of Banach spaces. *Israel J. Math.* **33** (1979), no. 2, 155–171

[101] A. Frieze and R. Kannan, Quick approximation to matrices and applications. *Combinatorica* **19** (1999), no. 2, 175–220

[102] R. J. Gardner, *Geometric tomography*. 2nd edn., Encyclopedia Math. Appl. 58, Cambridge University Press, New York, 2006

[103] D. J. H. Garling and Y. Gordon, Relations between some constants associated with finite dimensional Banach spaces. *Israel J. Math.* **9** (1971), 346–361

[104] A. Giannopoulos and M. Papadimitrakis, Isotropic surface area measures. *Mathematika* **46** (1999), no. 1, 1–13

[105] A. A. Giannopoulos, A note on the Banach–Mazur distance to the cube. In *Geometric aspects of functional analysis (Israel, 1992–1994)*, pp. 67–73, Oper. Theory Adv. Appl. 77, Birkhäuser, Basel, 1995

[106] A. A. Giannopoulos, A proportional Dvoretzky–Rogers factorization result. *Proc. Amer. Math. Soc.* **124** (1996), no. 1, 233–241

[107] A. A. Giannopoulos and V. D. Milman, Extremal problems and isotropic positions of convex bodies. *Israel J. Math.* **117** (2000), 29–60

[108] A. A. Giannopoulos, V. D. Milman, and M. Rudelson, Convex bodies with minimal mean width. In *Geometric aspects of functional analysis*, pp. 81–93, Lecture Notes in Math. 1745, Springer, Berlin, 2000

[109] O. Giladi, A. Naor, and G. Schechtman, Bourgain's discretization theorem. *Ann. Fac. Sci. Toulouse Math. (6)* **21** (2012), no. 4, 817–837

[110] O. Giladi, J. Prochno, C. Schütt, N. Tomczak-Jaegermann, and E. Werner, On the geometry of projective tensor products. *J. Funct. Anal.* **273** (2017), no. 2, 471–495

[111] E. D. Gluskin, The diameter of the Minkowski compactum is roughly equal to n. *Funktsional. Anal. i Prilozhen.* **15** (1981), no. 1, 72–73

[112] E. D. Gluskin, Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces. *Mat. Sb. (N.S.)* **136(178)** (1988), no. 1, 85–96

[113] G. Godefroy, A survey on Lipschitz-free Banach spaces. *Comment. Math.* **55** (2015), no. 2, 89–118

[114] G. Godefroy and N. J. Kalton, Lipschitz-free Banach spaces. *Studia Math.* **159** (2003), no. 1, 121–141

[115] Y. Gordon and M. Junge, Volume formulas in L_p-spaces. *Positivity* **1** (1997), no. 1, 7–43

[116] Y. Gordon and M. Junge, Volume ratios in L_p-spaces. *Studia Math.* **136** (1999), no. 2, 147–182

[117] Y. Gordon, M. Junge, and N. J. Nielsen, The relations between volume ratios and new concepts of GL constants. *Positivity* **1** (1997), no. 4, 359–379

[118] Y. Gordon and R. Loewy, Uniqueness of (Δ) bases and isometries of Banach spaces. *Math. Ann.* **241** (1979), no. 2, 159–180

[119] M. Gromov, Random walk in random groups. *Geom. Funct. Anal.* **13** (2003), no. 1, 73–146

[120] M. Gromov and V. D. Milman, Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces. *Compositio Math.* **62** (1987), no. 3, 263–282

[121] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo* **8** (1953), 1–79

[122] B. Grünbaum, Projection constants. *Trans. Amer. Math. Soc.* **95** (1960), 451–465

[123] M. A. Gunes and A. Naor, The separation modulus of unitary ideals. Forthcoming manuscript

[124] A. Gupta, R. Krauthgamer, and J. R. Lee, Bounded geometries, fractals, and low-distortion embeddings. In *44th symposium on foundations of computer science (FOCS 2003), 11-14 october 2003, cambridge, ma, usa, proceedings*, pp. 534–543, IEEE Computer Society, 2003

[125] V. I. Gurariĭ, M. Ĭ. Kadec, and V. I. Macaev, Distances between finite-dimensional analogs of the L_p-spaces. *Mat. Sb. (N.S.)* **70 (112)** (1966), 481–489

[126] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*. Springer, Berlin, 1957

[127] G. H. Hardy and J. E. Littlewood, Bilinear forms bounded in spaces (p, q). *Quart. J. Math. Oxford* **5** (1934), no. 1, 241–254

[128] M. Henk, Löwner–John ellipsoids. In *Optimization Stories: 21st International Symposium on Mathematical Programming. Berlin, Germany, 19–24 August 2012*, pp. 95–106, Doc. Math. Ser., European Mathematical Society, Zürich, 2012

[129] A. Henrot, *Extremum problems for eigenvalues of elliptic operators*. Front. Math., Birkhäuser, Basel, 2006

[130] A. Henrot and M. Pierre, *Shape variation and optimization*. EMS Tracts Math. 28, European Mathematical Society, Zürich, 2018

[131] T. Hytönen, S. Li, and A. Naor, Quantitative affine approximation for UMD targets. *Discrete Anal.* (2016), article no. 6, 37 pp.

[132] T. Hytönen and A. Naor, Heat flow and quantitative differentiation. *J. Eur. Math. Soc. (JEMS)* **21** (2019), no. 11, 3415–3466

[133] M. Q. Jacobs, Measurable multivalued mappings and Lusin's theorem. *Trans. Amer. Math. Soc.* **134** (1968), 471–481

[134] R. C. James, Nonreflexive spaces of type 2. *Israel J. Math.* **30** (1978), no. 1-2, 1–13

[135] R. E. Jamison and W. H. Ruckle, Factoring absolutely convergent series. *Math. Ann.* **224** (1976), no. 2, 143–148

[136] T. Jech, *Set theory*. Springer Monogr. Math., Springer, Berlin, 2003

[137] F. John, Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pp. 187–204, Interscience Publishers, New York, NY, 1948

[138] W. B. Johnson and J. Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, pp. 189–206, Contemp. Math. 26, American Mathematical Society, Providence, RI, 1984

[139] W. B. Johnson and J. Lindenstrauss, Basic concepts in the geometry of Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pp. 1–84, North-Holland, Amsterdam, 2001

[140] W. B. Johnson, J. Lindenstrauss, and G. Schechtman, Extensions of Lipschitz maps into Banach spaces. *Israel J. Math.* **54** (1986), no. 2, 129–138

[141] W. B. Johnson and A. Naor, The Johnson–Lindenstrauss lemma almost characterizes Hilbert space, but not quite. *Discrete Comput. Geom.* **43** (2010), no. 3, 542–553

[142] W. B. Johnson and G. Schechtman, On subspaces of L_1 with maximal distances to Euclidean space. In *Proceedings of research workshop on Banach space theory (Iowa City, Iowa, 1981)*, pp. 83–96, The University of Iowa, Iowa City, IA, 1982

[143] W. B. Johnson and G. Schechtman, Finite dimensional subspaces of L_p. In *Handbook of the geometry of Banach spaces, Vol. I*, pp. 837–870, North-Holland, Amsterdam, 2001

[144] I. Joó and L. L. Stachó, Generalization of an inequality of G. Pólya concerning the eigenfrequences of vibrating bodies. *Publ. Inst. Math. (Beograd) (N.S.)* **31(45)** (1982), 65–72

[145] Z. Kabluchko and J. Prochno, The maximum entropy principle and volumetric properties of Orlicz balls. *J. Math. Anal. Appl.* **495** (2021), no. 1, article no. 124687, 19 pp.

[146] Z. Kabluchko, J. Prochno, and C. Thäle, Exact asymptotic volume and volume ratio of Schatten unit balls. *J. Approx. Theory* **257** (2020), article no. 105457, 13 pp.

[147] M. I. Kadec and B. S. Mitjagin, Complemented subspaces in Banach spaces. *Uspehi Mat. Nauk* **28** (1973), no. 6(174), 77–94

[148] J.-P. Kahane, Sur les sommes vectorielles $\sum \pm u_n$. *C. R. Acad. Sci. Paris* **259** (1964), 2577–2580

[149] N. J. Kalton, Spaces of Lipschitz and Hölder functions and their applications. *Collect. Math.* **55** (2004), no. 2, 171–217

[150] N. J. Kalton, The complemented subspace problem revisited. *Studia Math.* **188** (2008), no. 3, 223–257

[151] N. J. Kalton, The uniform structure of Banach spaces. *Math. Ann.* **354** (2012), no. 4, 1247–1288

[152] D. Karger, R. Motwani, and M. Sudan, Approximate graph coloring by semidefinite programming. *J. ACM* **45** (1998), no. 2, 246–265

[153] B. S. Kašin, The widths of certain finite-dimensional sets and classes of smooth functions. *Izv. Akad. Nauk SSSR Ser. Mat.* **41** (1977), no. 2, 334–351, 478

[154] A. S. Kechris, *Classical descriptive set theory*. Grad. Texts in Math. 156, Springer, New York, 1995

[155] M. D. Kirszbraun, Über die zusammenziehenden und Lipschitzchen Transformationen. *Fundam. Math.* **22** (1934), 77–108

[156] P. N. Klein, S. A. Plotkin, and S. Rao, Excluded minors, network decomposition, and multicommodity flow. In *Proceedings of the twenty-fifth annual ACM symposium on theory of computing, may 16-18, 1993, san diego, ca, USA*, edited by S. R. Kosaraju, D. S. Johnson, and A. Aggarwal, pp. 682–690, ACM, 1993

[157] A. Koldobsky, D. Ryabogin, and A. Zvavitch, Projections of convex bodies and the Fourier transform. *Israel J. Math.* **139** (2004), 361–380

[158] M. Kozdoba, *Extension of Banach space valued Lipschitz functions*. Master's thesis, Technion-Israel Institute of Technology, 2005

[159] E. Krahn, Über Minimaleigenschaften der Kugel in drei und mehr Dimensionen. *Acta Univ.* **9** (1926), 1–44

[160] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor, Measured descent: A new embedding method for finite metrics. *Geom. Funct. Anal.* **15** (2005), no. 4, 839–858

[161] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **13** (1965), 397–403

[162] S. Kwapień, Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. *Studia Math.* **44** (1972), 583–595

[163] S. a. Kwapień and C. Schütt, Some combinatorial and probabilistic inequalities and their application to Banach space theory. II. *Studia Math.* **95** (1989), no. 2, 141–154

[164] I. Kyrezi, On the entropy of the convex hull of finite sets. *Proc. Amer. Math. Soc.* **128** (2000), no. 8, 2393–2403

[165] U. Lang, B. Pavlović, and V. Schroeder, Extensions of Lipschitz maps into Hadamard spaces. *Geom. Funct. Anal.* **10** (2000), no. 6, 1527–1553

[166] U. Lang and T. Schlichenmaier, Nagata dimension, quasisymmetric embeddings, and Lipschitz extensions. *Int. Math. Res. Not.* (2005), no. 58, 3625–3655

[167] R. Latała and K. Oleszkiewicz, Gaussian measures of dilatations of convex symmetric sets. *Ann. Probab.* **27** (1999), no. 4, 1922–1938

[168] M. Ledoux and M. Talagrand, *Probability in Banach spaces*. Ergeb. Math. Grenzgeb. 23, Springer, Berlin, 1991

[169] J. R. Lee, M. Mendel, and A. Naor, Metric structures in L_1: Dimension, snowflakes, and average distortion. *European J. Combin.* **26** (2005), no. 8, 1180–1190

[170] J. R. Lee and A. Naor, Metric decomposition, smooth measures, and clustering. 2003, unpublished manuscript, available on request

[171] J. R. Lee and A. Naor, Absolute Lipschitz extendability. *C. R. Math. Acad. Sci. Paris* **338** (2004), no. 11, 859–862

[172] J. R. Lee and A. Naor, Embedding the diamond graph in L_p and dimension reduction in L_1. *Geom. Funct. Anal.* **14** (2004), no. 4, 745–747

[173] J. R. Lee and A. Naor, Extending Lipschitz functions via random metric partitions. *Invent. Math.* **160** (2005), no. 1, 59–95

[174] L. Lefton and D. Wei, Numerical approximation of the first eigenpair of the p-Laplacian using finite elements and the penalty method. *Numer. Funct. Anal. Optim.* **18** (1997), no. 3-4, 389–399

[175] F. T. Leighton and S. Rao, An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *29th annual symposium on foundations of computer science, white plains, new york, usa, 24-26 october 1988*, pp. 422–431, IEEE Computer Society, 1988

[176] D. R. Lewis, Finite dimensional subspaces of L_p. *Studia Math.* **63** (1978), no. 2, 207–212

[177] D. R. Lewis, Ellipsoids defined by Banach ideal norms. *Mathematika* **26** (1979), no. 1, 18–29

[178] Y. Lim, Contractive barycentric maps and L^1 ergodic theorems on the cone of positive definite matrices. *J. Math. Anal. Appl.* **459** (2018), no. 1, 291–306

[179] J. Lindenstrauss, On nonlinear projections in Banach spaces. *Michigan Math. J.* **11** (1964), 263–287

[180] J. Lindenstrauss and L. Tzafriri, On the complemented subspaces problem. *Israel J. Math.* **9** (1971), 263–269

[181] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. I*. Springer, Berlin, 1977

[182] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. II*. Ergeb. Math. Grenzgeb. 97, Springer, Berlin, 1979

[183] N. Linial, S. Mendelson, G. Schechtman, and A. Shraibman, Complexity measures of sign matrices. *Combinatorica* **27** (2007), no. 4, 439–463

[184] N. Linial and M. E. Saks, Decomposing graphs into regions of small diameter. In *Proceedings of the second annual ACM/SIGACT-SIAM symposium on discrete algorithms, 28-30 january 1991, san francisco, california.*, edited by A. Aggarwal, pp. 320–330, ACM/SIAM, 1991

[185] L. H. Loomis and H. Whitney, An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc* **55** (1949), 961–962

[186] G. J. Lozanovskiĭ, Certain Banach lattices. *Sibirsk. Mat. Ž.* **10** (1969), 584–599

[187] M. Ludwig, Projection bodies and valuations. *Adv. Math.* **172** (2002), no. 2, 158–168

[188] M. Ludwig, Minkowski valuations. *Trans. Amer. Math. Soc.* **357** (2005), no. 10, 4191–4213

[189] N. Lusin, *Leçons sur les ensembles analytiques et leurs applications*. Chelsea, New York, 1972

[190] E. Lutwak, Selected affine isoperimetric inequalities. In *Handbook of convex geometry, Vol. A, B*, pp. 151–176, North-Holland, Amsterdam, 1993

[191] J. Luukkainen and E. Saksman, Every complete doubling metric space carries a doubling measure. *Proc. Amer. Math. Soc.* **126** (1998), no. 2, 531–534

[192] N. N. Luzin, Sur la classification de M. Baire. *C. R. Math. Acad. Sci. Paris* **164** (1917), 91–94

[193] K. Mahler, Ein Übertragungsprinzip für konvexe Körper. *Časopis Pěst. Mat. Fys.* **68** (1939), 93–102

[194] E. Makai, Jr. and H. Martini, The cross-section body, plane sections of convex bodies and approximation of convex bodies. I. *Geom. Dedicata* **63** (1996), no. 3, 267–296

[195] K. Makarychev and Y. Makarychev, Metric extension operators, vertex sparsifiers and Lipschitz extendability. *Israel J. Math.* **212** (2016), no. 2, 913–959

[196] P. Mankiewicz and N. Tomczak-Jaegermann, Quotients of finite-dimensional Banach spaces; random phenomena. In *Handbook of the geometry of Banach spaces, Vol. 2*, pp. 1201–1246, North-Holland, Amsterdam, 2003

[197] M. B. Marcus and G. Pisier, Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta Math.* **152** (1984), no. 3-4, 245–301

[198] E. Markessinis, G. Paouris, and C. Saroglou, Comparing the M-position with some classical positions of convex bodies. *Math. Proc. Cambridge Philos. Soc.* **152** (2012), no. 1, 131–152

[199] J. Matoušek, Extension of Lipschitz mappings on metric trees. *Comment. Math. Univ. Carolin.* **31** (1990), no. 1, 99–104

[200] J. Matoušek, On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.* **93** (1996), 333–344

[201] J. Matoušek, *Lectures on discrete geometry*. Grad. Texts in Math. 212, Springer, New York, 2002

[202] J. Matoušek, A. Nikolov, and K. Talwar, Factorization norms and hereditary discrepancy. *Int. Math. Res. Not. IMRN* **2020** (2020), no. 3, 751–780

[203] B. Maurey, Type, cotype and K-convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, pp. 1299–1332, North-Holland, Amsterdam, 2003

[204] B. Maurey and G. Pisier, Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach. *Studia Math.* **58** (1976), no. 1, 45–90

[205] S. Mazur, Une remarque sur l'homéomorphie des champs fonctionels. *Studia Math.* **1** (1929), 83–85

[206] E. J. McShane, Extension of range of functions. *Bull. Amer. Math. Soc.* **40** (1934), no. 12, 837–842

[207] M. Mendel and A. Naor, Formulae for absolute extendability operators. Forthcoming manuscript

[208] M. Mendel and A. Naor, Ramsey partitions and proximity data structures. *J. Eur. Math. Soc. (JEMS)* **9** (2007), no. 2, 253–275

[209] M. Mendel and A. Naor, Metric cotype. *Ann. of Math. (2)* **168** (2008), no. 1, 247–298

[210] M. Mendel and A. Naor, Spectral calculus and Lipschitz extension for barycentric metric spaces. *Anal. Geom. Metr. Spaces* **1** (2013), 163–199

[211] M. Mendel and A. Naor, Expanders with respect to Hadamard spaces and random graphs. *Duke Math. J.* **164** (2015), no. 8, 1471–1548

[212] M. Mendel, A. Naor, and Y. Rabani, A weighted Sobolev embedding on the discrete torus and nonexistence of gentle partitions of unity. Forthcoming manuscript

[213] S. Mendelson and R. Vershynin, Entropy and the combinatorial dimension. *Invent. Math.* **152** (2003), no. 1, 37–55

[214] C. Meyer, *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000

[215] V. D. Milman, A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkcional. Anal. i Priložen.* **5** (1971), no. 4, 28–37

[216] V. D. Milman, Almost Euclidean quotient spaces of subspaces of a finite-dimensional normed space. *Proc. Amer. Math. Soc.* **94** (1985), no. 3, 445–449

[217] V. D. Milman, Some remarks on Urysohn's inequality and volume ratio of cotype 2-spaces. In *Geometrical aspects of functional analysis (1985/86)*, pp. 75–81, Lecture Notes in Math. 1267, Springer, Berlin, 1987

[218] V. D. Milman and A. Pajor, Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. In *Geometric aspects of functional analysis (1987–88)*, pp. 64–104, Lecture Notes in Math. 1376, Springer, Berlin, 1989

[219] V. D. Milman and G. Pisier, Banach spaces with a weak cotype 2 property. *Israel J. Math.* **54** (1986), no. 2, 139–158

[220] V. D. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces*. Lecture Notes in Math. 1200, Springer, Berlin, 1986

[221] V. D. Mil'man and H. Wolfson, Minkowski spaces with extremal distance from the Euclidean space. *Israel J. Math.* **29** (1978), no. 2-3, 113–131

[222] L. Mirsky, Symmetric gauge functions and unitarily invariant norms. *Quart. J. Math. Oxford Ser. (2)* **11** (1960), 50–59

[223] D. Müller, A geometric bound for maximal functions associated to convex bodies. *Pacific J. Math.* **142** (1990), no. 2, 297–312

[224] A. Naor, A phase transition phenomenon between the isometric and isomorphic extension problems for Hölder functions between L_p spaces. *Mathematika* **48** (2001), no. 1-2, 253–271 (2003)

[225] A. Naor, The surface measure and cone measure on the sphere of l_p^n. *Trans. Amer. Math. Soc.* **359** (2007), no. 3, 1045–1079

[226] A. Naor, Class notes on Lipschitz extension from finite subsets. 2015, https://web.math.princeton.edu/~naor/homepage%20files/extension-from-finite.pdf, visited on 29 March 2024

[227] A. Naor, Probabilistic clustering of high dimensional norms. In *Proceedings of the twenty-eighth annual ACM-SIAM symposium on discrete algorithms, SODA 2017*, pp. 690–709, 2017

[228] A. Naor, A spectral gap precludes low-dimensional embeddings. In *33rd International Symposium on Computational Geometry*, pp. article no. 50, 16 pp., LIPIcs. Leibniz Int. Proc. Inform. 77, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2017

[229] A. Naor, Metric dimension reduction: A snapshot of the Ribe program. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures*, pp. 759–837, World Scientific, Hackensack, NJ, 2018

[230] A. Naor, An average John theorem. *Geom. Topol.* **25** (2021), no. 4, 1631–1717

[231] A. Naor, Impossibility of almost extension. *Adv. Math.* **384** (2021), article no. 107761, 34 pp.

[232] A. Naor, G. Pisier, and G. Schechtman, Impossibility of dimension reduction in the nuclear norm. *Discrete Comput. Geom.* **63** (2020), no. 2, 319–345

[233] A. Naor and Y. Rabani, On Lipschitz extension from finite subsets. *Israel J. Math.* **219** (2017), no. 1, 115–161

[234] A. Naor and G. Schechtman, Lipschitz almost-extension and nonexistence of uniform embeddings of balls in Schatten classes. Forthcoming manuscript

[235] A. Naor and G. Schechtman, Obstructions to metric embeddings of Schatten classes. Forthcoming manuscript

[236] A. Naor and G. Schechtman, Planar earthmover is not in L_1. *SIAM J. Comput.* **37** (2007), no. 3, 804–826

[237] A. Naor and G. Schechtman, Pythagorean powers of hypercubes. *Ann. Inst. Fourier (Grenoble)* **66** (2016), no. 3, 1093–1116

[238] A. Naor and T. Tao, Random martingales and localization of maximal inequalities. *J. Funct. Anal.* **259** (2010), no. 3, 731–779

[239] A. Naor and T. Tao, Scale-oblivious metric fragmentation and the nonlinear Dvoretzky theorem. *Israel J. Math.* **192** (2012), no. 1, 489–504

[240] A. Naor and R. Young, Foliated corona decompositions. *Acta Math.* **229** (2022), no. 1, 55–200

[241] A. Navas, An L^1 ergodic theorem with values in a non-positively curved space via a canonical barycenter map. *Ergodic Theory Dynam. Systems* **33** (2013), no. 2, 609–623

[242] I. Newman and Y. Rabinovich, A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.* **29** (2003), no. 1, 77–81

[243] S.-i. Ohta, Extending Lipschitz and Hölder maps between metric spaces. *Positivity* **13** (2009), no. 2, 407–425

[244] M. I. Ostrovskii, *Metric embeddings. Bilipschitz and coarse embeddings into Banach spaces*. De Gruyter Stud. Math. 49, De Gruyter, Berlin, 2013

[245] A. Pajor, Plongement de l_1^n dans les espaces de Banach complexes. *C. R. Acad. Sci. Paris Sér. I Math.* **296** (1983), no. 17, 741–743

[246] E. Parini, Reverse Cheeger inequality for planar convex sets. *J. Convex Anal.* **24** (2017), no. 1, 107–122

[247] A. Pełczyński, Linear extensions, linear averagings, and their applications to linear topological classification of spaces of continuous functions. *Dissertationes Math. Rozprawy Mat.* **58** (1968), 92 pp.

[248] D. Peleg and E. Reshef, Deterministic polylog approximation for minimum communication spanning trees (extended abstract). In *Automata, languages and programming (Aalborg, 1998)*, pp. 670–681, Lecture Notes in Comput. Sci. 1443, Springer, Berlin, 1998

[249] B. J. Pettis, Linear functionals and completely additive set functions. *Duke Math. J.* **4** (1938), no. 3, 552–565

[250] C. M. Petty, Surface area of a convex body under affine transformations. *Proc. Amer. Math. Soc.* **12** (1961), 824–828

[251] C. M. Petty, Projection bodies. In *Proc. Colloquium on Convexity (Copenhagen, 1965)*, pp. 234–241, Kobenhavns Univ. Mat. Inst., Copenhagen, 1967

[252] C. M. Petty, Isoperimetric problems. In *Proceedings of the Conference on Convexity and Combinatorial Geometry (Univ. Oklahoma, Norman, Okla., 1971)*, pp. 26–41, Department of Mathematics, The University of Oklahoma, Norman, Okla., 1971

[253] A. Pietsch and J. Wenzel, *Orthonormal systems and Banach space geometry*. Encyclopedia Math. Appl. 70, Cambridge University Press, Cambridge, 1998

[254] G. Pisier, Sur les espaces de Banach qui ne contiennent pas uniformément de l_n^1. *C. R. Acad. Sci. Paris Sér. A-B* **277** (1973), A991–A994

[255] G. Pisier, Sur les espaces de Banach de dimension finie à distance extrémale d'un espace euclidien [d'après V. D. Milman et H. Wolfson]. In *Séminaire d'Analyse Fonctionnelle (1978–1979)*, Exp. No. 16, 10 pp., École polytechnique, Palaiseau, 1979

[256] G. Pisier, Sur les espaces de Banach K-convexes. In *Seminar on Functional Analysis, 1979–1980 (French)*, Exp. No. 11, 15 pp., École polytechnique, Palaiseau, 1980

[257] G. Pisier, Un théorème sur les opérateurs linéaires entre espaces de Banach qui se factorisent par un espace de Hilbert. *Ann. Sci. École Norm. Sup. (4)* **13** (1980), no. 1, 23–43

[258] G. Pisier, Remarques sur un résultat non publié de B. Maurey. In *Seminar on Functional Analysis, 1980–1981*, Exp. No. V, 13 pp., École polytechnique, Palaiseau, 1981

[259] G. Pisier, Holomorphic semigroups and the geometry of Banach spaces. *Ann. of Math. (2)* **115** (1982), no. 2, 375–392

[260] G. Pisier, On the dimension of the l_p^n-subspaces of Banach spaces, for $1 \leqslant p < 2$. *Trans. Amer. Math. Soc.* **276** (1983), no. 1, 201–211

[261] G. Pisier, *Factorization of linear operators and geometry of Banach spaces*. CBMS Reg. Conf. Ser. Math. 60, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1986

[262] G. Pisier, Weak Hilbert spaces. *Proc. London Math. Soc. (3)* **56** (1988), no. 3, 547–579

[263] G. Pisier, *The volume of convex bodies and Banach space geometry*. Cambridge Tracts in Math. 94, Cambridge University Press, Cambridge, 1989

[264] G. Pisier, Dvoretzky's theorem for operator spaces. *Houston J. Math.* **22** (1996), no. 2, 399–416

[265] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*. Ann. of Math. Stud. 27, Princeton University Press, Princeton, NJ, 1951

[266] S. T. Rachev and L. Rüschendorf, Approximate independence of distributions on spheres and their stability properties. *Ann. Probab.* **19** (1991), no. 3, 1311–1337

[267] H. Rademacher, Über partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale. *Math. Ann.* **79** (1919), no. 4, 340–359

[268] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*. Monogr. Textbooks Pure Appl. Math. 146, Marcel Dekker, New York, 1991

[269] S. Rao, Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry (Miami Beach, FL, 1999)*, pp. 300–306, ACM, New York, 1999

[270] O. Regev, Entropy-based bounds on dimension reduction in L_1. *Israel J. Math.* **195** (2013), no. 2, 825–832

[271] S. Reisner, Zonoids with minimal volume-product. *Math. Z.* **192** (1986), no. 3, 339–346

[272] M. Riesz, Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires. *Acta Math.* **49** (1927), no. 3-4, 465–497

[273] C. A. Rogers, A note on coverings and packings. *J. London Math. Soc.* **25** (1950), 327–331

[274] M. Rudelson and R. Vershynin, Combinatorics of random processes and sections of convex bodies. *Ann. of Math. (2)* **164** (2006), no. 2, 603–648

[275] D. Rutovitz, Some parameters associated with finite-dimensional Banach spaces. *J. London Math. Soc.* **40** (1965), 241–255

[276] R. A. Ryan, *Introduction to tensor products of Banach spaces*. Springer Monogr. Math., Springer, London, 2002

[277] J. Saint-Raymond, Le volume des idéaux d'opérateurs classiques. *Studia Math.* **80** (1984), no. 1, 63–75

[278] L. A. Santaló, An affine invariant for convex bodies of n-dimensional space. *Portugal. Math.* **8** (1949), 155–161

[279] G. Schechtman and J. Zinn, On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.* **110** (1990), no. 1, 217–224

[280] M. Schmuckenschläger, The distribution function of the convolution square of a convex symmetric body in $\mathbf{R}^n$. *Israel J. Math.* **78** (1992), no. 2-3, 309–334

[281] M. Schmuckenschläger, Petty's projection inequality and Santalo's affine isoperimetric inequality. *Geom. Dedicata* **57** (1995), no. 3, 285–295

[282] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*. expanded edn., Encyclopedia Math. Appl. 151, Cambridge University Press, Cambridge, 2014

[283] R. Schneider and W. Weil, Zonoids and related topics. In *Convexity and its applications*, pp. 296–317, Birkhäuser, Basel, 1983

[284] C. Schütt, Unconditionality in tensor products. *Israel J. Math.* **31** (1978), no. 3-4, 209–216

[285] C. Schütt, On the volume of unit balls in Banach spaces. *Compositio Math.* **47** (1982), no. 3, 393–407

[286] C. Schütt, The isoperimetric quotient and some classical Banach spaces. *Israel J. Math.* **67** (1989), no. 1, 43–61

[287] G. C. Shephard, Shadow systems of convex sets. *Israel J. Math.* **2** (1964), 229–236

[288] E. Silverman, Lower semicontinuity of parametric integrals. *Trans. Amer. Math. Soc.* **175** (1973), 499–508

[289] B. Simon, *Trace ideals and their applications*. London Math. Soc. Lecture Note Ser. 35, Cambridge University Press, Cambridge, 1979

[290] A. Sobczyk, Projections in Minkowski and Banach spaces. *Duke Math. J.* **8** (1941), 78–106

[291] S. M. Srivastava, *A course on Borel sets*. Grad. Texts in Math. 180, Springer, New York, 1998

[292] K.-T. Sturm, Probability measures on metric spaces of nonpositive curvature. In *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002)*, pp. 357–390, Contemp. Math. 338, American Mathematical Society, Providence, RI, 2003

[293] S. a. Szarek and N. Tomczak-Jaegermann, On nearly Euclidean decomposition for some classes of Banach spaces. *Compositio Math.* **40** (1980), no. 3, 367–385

[294] S. a. J. Szarek, On Kashin's almost Euclidean orthogonal decomposition of l_n^1. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **26** (1978), no. 8, 691–694

[295] S. a. J. Szarek, Spaces with large distance to l_∞^n and random matrices. *Amer. J. Math.* **112** (1990), no. 6, 899–942

[296] S. J. Szarek, On the geometry of the Banach–Mazur compactum. In *Functional analysis (Austin, TX, 1987/1989)*, pp. 48–59, Lecture Notes in Math. 1470, Springer, Berlin, 1991

[297] S. J. Szarek and M. Talagrand, An "isomorphic" version of the Sauer–Shelah lemma and the Banach–Mazur distance to the cube. In *Geometric aspects of functional analysis (1987–88)*, pp. 105–112, Lecture Notes in Math. 1376, Springer, Berlin, 1989

[298] M. Talagrand, Type, infratype and the Elton–Pajor theorem. *Invent. Math.* **107** (1992), no. 1, 41–59

[299] M. Talagrand, Embedding of l_k^∞ and a theorem of Alon and Milman. In *Geometric aspects of functional analysis (Israel, 1992–1994)*, pp. 289–293, Oper. Theory Adv. Appl. 77, Birkhäuser, Basel, 1995

[300] A. E. Taylor, A geometric theorem and its application to biorthogonal systems. *Bull. Amer. Math. Soc.* **53** (1947), 614–616

[301] G. O. Thorin, Convexity theorems generalizing those of M. Riesz and Hadamard with some applications. *Comm. Sem. Math. Univ. Lund [Medd. Lunds Univ. Mat. Sem.]* **9** (1948), 1–58

[302] K. Tikhomirov, On the Banach–Mazur distance to cross-polytope. *Adv. Math.* **345** (2019), 598–617

[303] N. Tomczak-Jaegermann, Computing 2-summing norm with few vectors. *Ark. Mat.* **17** (1979), no. 2, 273–277

[304] N. Tomczak-Jaegermann, Finite-dimensional subspaces of uniformly convex and uniformly smooth Banach lattices and trace classes S_p. *Studia Math.* **66** (1979/80), no. 3, 261–281

[305] N. Tomczak-Jaegermann, *Banach–Mazur distances and finite-dimensional operator ideals*. Pitman Monogr. Surveys Pure Appl. Math. 38, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, New York, 1989

[306] F. Tricomi, Sulle funzioni di Bellel di ordine e argomento pressochè uguali. *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* **83** (1949), 3–20

[307] C. Villani, *Optimal transport*. Grundlehren Math. Wiss. 338, Springer, Berlin, 2009

[308] A. L. Vol'berg and S. V. Konyagin, On measures with the doubling condition. *Izv. Akad. Nauk SSSR Ser. Mat.* **51** (1987), no. 3, 666–675

[309] D. H. Wagner, Survey of measurable selection theorems. *SIAM J. Control Optimization* **15** (1977), no. 5, 859–903

[310] N. Weaver, *Lipschitz algebras*. World Scientific, River Edge, NJ, 1999

[311] H. Weyl, *The Classical Groups. Their Invariants and Representations*. Princeton University Press, Princeton, NJ, 1939

[312] H. Whitney, Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36** (1934), no. 1, 63–89

[313] E. T. Whittaker and G. N. Watson, *A course of modern analysis. An introduction to the general theory of infinite processes and of analytic functions: with an account of the principal transcendental functions*. Cambridge University Press, New York, 1962

[314] P. Wojtaszczyk, *Banach spaces for analysts*. Cambridge Stud. Adv. Math. 25, Cambridge University Press, Cambridge, 1991

[315] C. Zong, From deep holes to free planes. *Bull. Amer. Math. Soc. (N.S.)* **39** (2002), no. 4, 533–555

Added in proof. Since the initial posting of this work, the following progress was made on some of the issues that are discussed herein. Conjecture 10 is proved in the forthcoming article [26] in the special case when $K \subseteq \mathbb{R}^n$ is an origin-symmetric convex polytope that has $O(n)$ faces. The forthcoming article [123] proves Conjecture 10 when $K \subseteq \mathsf{M}_n(\mathbb{R})$ is the unit ball of the unitary ideal $\mathsf{S}_{\mathbf{E}}$ of any 1-symmetric normed space $\mathbf{E} = (\mathbb{R}^n, \|\cdot\|_{\mathbf{E}})$; consequently, Lemma 54 and Proposition 55 hold with all of the logarithmic factors that appear in them replaced by universal constants. The forthcoming work [54] builds on the results herein while adding multiple innovations and ideas to obtain several new results. These include the estimate $\mathsf{e}(\mathbf{X}) \gtrsim \sqrt[4]{\dim(\mathbf{X})}$ for every normed space $\mathbf{X}$, which is an improvement over the value of the universal constant c that we obtained in the proof of Theorem 1. Conjecture 134 is resolved (negatively) in [54], where it is proved that $\mathsf{e}_{\mathrm{conv}}(\mathfrak{M}) \lesssim \mathsf{e}(\mathfrak{M})^2$ for every Polish metric space $(\mathfrak{M}, d_{\mathfrak{M}})$. It is also proved in [54] that $\mathsf{e}(\ell_2^n, \mathfrak{N}) \asymp \sqrt[4]{n}$ for any 1-net $\mathfrak{N}$ of ℓ_2^n and $\mathsf{e}(\ell_2^n, \mathbb{Z}^n) \asymp \sqrt[6]{n}$; both of these asymptotic evaluations of Lipschitz extension moduli answer questions that were posed in the precursor [227] of the present work. Finally, the lower order factor in the main result of [231] is removed in [54], thus showing that an old Lipschitz almost-extension result of Bourgain [46] is sharp up to universal constant factors. Beyond the aforementioned examples of statements from [54], multiple other new results on Lipschitz extension and separation moduli are obtained in [54]. The discussion in Remark 41 (more generally, the role that canonically positioned norms play herein), evolved (very substantially) to the forthcoming work [56] which investigates the question of when is it possible to construct a norm with prescribed group of isometries; as demonstrated in [56], it turns out that the answer to this old inverse problem is quite subtle.

MEMOIRS OF THE EUROPEAN MATHEMATICAL SOCIETY

Assaf Naor

Extension, Separation and Isomorphic Reverse Isoperimetry

The Lipschitz extension modulus $\mathsf{e}(\mathcal{M})$ of a metric space $\mathcal{M}$ is the infimum over those $L \in [1,\infty]$ such that for any Banach space Z and any $\mathscr{C} \subseteq \mathcal{M}$, any 1-Lipschitz function $f : \mathscr{C} \to \mathsf{Z}$ can be extended to an L-Lipschitz function $F : \mathcal{M} \to \mathsf{Z}$. Johnson, Lindenstrauss and Schechtman proved (1986) that if X is an n-dimensional normed space, then $\mathsf{e}(\mathsf{X}) \lesssim n$. In the reverse direction, we prove that every n-dimensional normed space X satisfies $\mathsf{e}(\mathsf{X}) \gtrsim n^c$, where $c > 0$ is a universal constant. Our core technical contribution is a geometric structural result on stochastic clustering of finite dimensional normed spaces which implies upper bounds on their Lipschitz extension moduli using an extension method of Lee and the author (2005). The separation modulus of a metric space $(\mathcal{M}, d_{\mathcal{M}})$ is the infimum over those $\sigma \in (0,\infty]$ such that for any $\Delta > 0$ there is a distribution over random partitions of $\mathcal{M}$ into clusters of diameter at most Δ such that for every two points $x, y \in \mathcal{M}$ the probability that they belong to different clusters is at most $\sigma d_{\mathcal{M}}(x,y)/\Delta$. We obtain upper and lower bounds on the separation moduli of finite dimensional normed spaces that relate them to well-studied volumetric invariants (volume ratios and projection bodies). Using these connections, we determine the asymptotic growth rate of the separation moduli of various normed spaces. If X is an n-dimensional normed space with enough symmetries, then our bounds imply that its separation modulus is equal to $\mathrm{vr}(\mathsf{X}^*)\sqrt{n}$ up to factors of lower order, where $\mathrm{vr}(\mathsf{X}^*)$ is the volume ratio of the unit ball of the dual of X. We formulate a conjecture on isomorphic reverse isoperimetric properties of symmetric convex bodies (akin to Ball's reverse isoperimetric theorem (1991), but permitting a non-isometric perturbation in addition to the choice of position) that can be used with our volumetric bounds on the separation modulus to obtain many more exact asymptotic evaluations of the separation moduli of normed spaces. Our estimates on the separation modulus imply asymptotically improved upper bounds on the Lipschitz extension moduli of various classical spaces. In particular, we deduce an improved upper bound on $\mathsf{e}(\ell_p^n)$ when $p > 2$ that resolves a conjecture of Brudnyi and Brudnyi (2005), and we prove that $\mathsf{e}(\ell_\infty^n) \asymp \sqrt{n}$, which is the first time that the growth rate of $\mathsf{e}(\mathsf{X})$ has been evaluated (as $\dim(\mathsf{X}) \to \infty$) for *any* finite dimensional normed space X.

https://ems.press
ISSN 2747-9080
ISBN 978-3-98547-069-3

PRESS