



# $\varepsilon$ -Isometric Dimension Reduction for Incompressible Subsets of $\ell_p$

Alexandros Eskenazis<sup>1,2</sup>

Received: 13 July 2022 / Revised: 22 August 2023 / Accepted: 23 August 2023 /  
Published online: 21 October 2023  
© The Author(s) 2023

## Abstract

Fix  $p \in [1, \infty)$ ,  $K \in (0, \infty)$ , and a probability measure  $\mu$ . We prove that for every  $n \in \mathbb{N}$ ,  $\varepsilon \in (0, 1)$ , and  $x_1, \dots, x_n \in L_p(\mu)$  with  $\|\max_{i \in \{1, \dots, n\}} |x_i|\|_{L_p(\mu)} \leq K$ , there exist  $d \leq \frac{32e^2(2K)^{2p} \log n}{\varepsilon^2}$  and vectors  $y_1, \dots, y_n \in \ell_p^d$  such that

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{L_p(\mu)}^p - \varepsilon \leq \|y_i - y_j\|_{\ell_p^d}^p \leq \|x_i - x_j\|_{L_p(\mu)}^p + \varepsilon.$$

Moreover, the argument implies the existence of a greedy algorithm which outputs  $\{y_i\}_{i=1}^n$  after receiving  $\{x_i\}_{i=1}^n$  as input. The proof relies on a derandomized version of Maurey's empirical method (1981) combined with a combinatorial idea of Ball (1990) and a suitable change of measure. Motivated by the above embedding, we introduce the notion of  $\varepsilon$ -isometric dimension reduction of the unit ball  $\mathbf{B}_E$  of a normed space  $(E, \|\cdot\|_E)$  and we prove that  $\mathbf{B}_{\ell_p}$  does not admit  $\varepsilon$ -isometric dimension reduction by linear operators for any value of  $p \neq 2$ .

**Keywords** Dimension reduction ·  $\varepsilon$ -Isometric embedding · Maurey's empirical method · Change of measure

**Mathematics Subject Classification** 46B85 · 46B09 · 46E30 · 68R12

---

Editor in Charge: Kenneth Clarkson

The author was supported by a Junior Research Fellowship from Trinity College, Cambridge. A conference version of this article will be presented in SoCG 2022.

---

Alexandros Eskenazis  
alexandros.eskenazis@imj-prg.fr; ae466@cam.ac.uk

<sup>1</sup> CNRS, Institut de Mathématiques de Jussieu, Sorbonne Université, Paris, France

<sup>2</sup> Trinity College, University of Cambridge, Cambridge, UK

# 1 Introduction

## 1.1 Metric Dimension Reduction

Using standard terminology from metric embeddings (see [38]), we say that a mapping between metric spaces  $f : (M, d_M) \rightarrow (N, d_N)$  is a bi-Lipschitz embedding with distortion at most  $\alpha \in [1, \infty)$  if there exists a scaling factor  $\sigma \in (0, \infty)$  such that

$$\forall x, y \in M, \quad \sigma d_M(x, y) \leq d_N(f(x), f(y)) \leq \alpha \sigma d_M(x, y). \tag{1}$$

Throughout this paper, we shall denote by  $\ell_p^d$  the linear space  $\mathbb{R}^d$  equipped with the  $p$ -norm,

$$\forall a = (a_1, \dots, a_d) \in \mathbb{R}^d, \quad \|a\|_{\ell_p^d} = \left( \sum_{i=1}^d |a_i|^p \right)^{1/p}. \tag{2}$$

The classical Johnson–Lindenstrauss lemma [21] asserts that if  $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$  is a Hilbert space and  $x_1, \dots, x_n \in \mathcal{H}$ , then for every  $\varepsilon \in (0, 1)$  there exists  $d \leq \frac{C \log n}{\varepsilon^2}$  and  $y_1, \dots, y_n \in \ell_2^d$  such that

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{\mathcal{H}} \leq \|y_i - y_j\|_{\ell_2^d} \leq (1 + \varepsilon) \cdot \|x_i - x_j\|_{\mathcal{H}}, \tag{3}$$

where  $C \in (0, \infty)$  is a universal constant. In the above embedding terminology, the Johnson–Lindenstrauss lemma states that for every  $\varepsilon \in (0, 1)$ ,  $n \in \mathbb{N}$ , and  $d \geq \frac{C \log n}{\varepsilon^2}$ , any  $n$ -point subset of Hilbert space admits a bi-Lipschitz embedding into  $\ell_2^d$  with distortion at most  $1 + \varepsilon$ . In order to prove their result, Johnson and Lindenstrauss introduced in [21] the influential random projection method that has since had many important applications in metric geometry and theoretical computer science and kick-started the field of *metric dimension reduction* (see the recent survey [36] of Naor) which lies at the intersection of those two subjects.

Following [36], we say that an infinite dimensional Banach space  $(E, \|\cdot\|_E)$  admits bi-Lipschitz dimension reduction if there exists  $\alpha = \alpha(E) \in [1, \infty)$  such that for every  $n \in \mathbb{N}$ , there exists  $k_n = k_n(E, \alpha) \in \mathbb{N}$  satisfying

$$\lim_{n \rightarrow \infty} \frac{\log k_n}{\log n} = 0 \tag{4}$$

and such that any  $n$ -point subset  $S$  of  $E$  admits a bi-Lipschitz embedding with distortion at most  $\alpha$  in a finite-dimensional linear subspace  $F$  of  $E$  with  $\dim F \leq k_n$ . The only non-Hilbertian space that is known to admit bi-Lipschitz dimension reduction is the 2-convexification of the classical Tsirelson space, as proven by Johnson and Naor in [22]. Turning to negative results, Matoušek proved in [32] the impossibility of bi-Lipschitz dimension reduction in  $\ell_\infty$ , whereas Brinkman and Charikar [10] (see also [30] for a shorter proof) constructed an  $n$ -point subset of  $\ell_1$  which does not admit a bi-Lipschitz embedding into any  $n^{o(1)}$ -dimensional subspace of  $\ell_1$ . Their theorem

was recently refined by Naor et al. [37] who showed that the same  $n$ -point subset of  $\ell_1$  does not embed into any  $n^{o(1)}$ -dimensional subspace of the trace class  $S_1$  (see also the striking recent work [41] of Regev and Vidick, where the impossibility of polynomial almost isometric dimension reduction in  $S_1$  is established). We refer to [36, Thm. 16] for a summary of the best known bounds quantifying the aforementioned qualitative statements. Despite the lapse of almost four decades since the proof of the Johnson–Lindenstrauss lemma, the following natural question remains stubbornly open.

**Question 1.1** For which values of  $p \notin \{1, 2, \infty\}$  does  $\ell_p$  admit bi-Lipschitz dimension reduction?

## 1.2 Dimensionality and Structure

An important feature of the formalism of bi-Lipschitz dimension reduction in a Banach space  $E$  is that both the distortion  $\alpha(E)$  of the embedding and the dimension  $k_n(E, \alpha)$  of the target subspace  $F$  are independent of the given  $n$ -point subset  $\mathcal{S}$  of  $E$ . Nevertheless, there are instances in which one can construct delicate embeddings whose distortion or the dimension of their targets depends on subtle geometric parameters of  $\mathcal{S}$ . For instance, we mention an important theorem of Schechtman [42, Thm. 5] (which built on work of Klartag and Mendelson [26]) who constructed a linear embedding of an arbitrary subset  $\mathcal{S}$  of  $\ell_2$  into any Banach space  $E$  whose distortion depends only on the Gaussian width of  $\mathcal{S}$  and the  $\ell$ -norm of the identity operator  $\text{id}_E : E \rightarrow E$ . In the special case that  $E$  is a Hilbert space, a substantially richer family of such embeddings was devised in [31].

Let  $\mu$  be a probability measure. For a subset  $\mathcal{S}$  of  $L_p(\mu)$ , we shall denote

$$\mathcal{J}(\mathcal{S}) \stackrel{\text{def}}{=} \left\| \max_{x \in \mathcal{S}} |x| \right\|_{L_p(\mu)} \quad (5)$$

and we will say that  $\mathcal{S}$  is  $K$ -incompressible<sup>1</sup> if  $\mathcal{J}(\mathcal{S}) \leq K$ . The main contribution of the present paper is the following dimensionality reduction theorem for incompressible subsets of  $L_p(\mu)$  which, in contrast to all the results discussed earlier, is valid for any value of  $p \in [1, \infty)$ .

**Theorem 1.2** ( $\varepsilon$ -isometric dimension reduction for incompressible subsets of  $L_p(\mu)$ )  
Fix parameters  $p \in [1, \infty)$ ,  $n \in \mathbb{N}$ ,  $K \in (0, \infty)$  and let  $\{x_i\}_{i=1}^n$  be a  $K$ -incompressible family of vectors in  $L_p(\mu)$  for some probability measure  $\mu$ . Then for every  $\varepsilon \in (0, 1)$ , there exist  $d \in \mathbb{N}$  with  $d \leq \frac{32e^2(2K)^{2p} \log n}{\varepsilon^2}$  and points  $y_1, \dots, y_n \in \ell_p^d$  such that

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{L_p(\mu)}^p - \varepsilon \leq \|y_i - y_j\|_{\ell_p^d}^p \leq \|x_i - x_j\|_{L_p(\mu)}^p + \varepsilon. \quad (6)$$

<sup>1</sup> The terminology is borrowed by the standard use of the term “incompressible vector” from random matrix theory, which refers to points on the unit sphere of  $\mathbb{R}^n$  which are far from the coordinate vectors  $e_1, \dots, e_n$ .

Besides the appearance of the incompressibility parameter  $K$  in the bound for the dimension  $d$  of the target space, Theorem 1.2 differs from the Johnson–Lindenstrauss lemma in that the error in (6) is *additive* rather than *multiplicative*. Recall that a map between metric spaces  $f : (M, d_M) \rightarrow (N, d_N)$  is called an  $\varepsilon$ -isometric embedding if

$$\forall x, y \in M, \quad |d_N(f(x), f(y)) - d_M(x, y)| \leq \varepsilon. \quad (7)$$

Embeddings with additive errors occur naturally in metric geometry and, more specifically, in metric dimension reduction (see e.g. [44, Sect. 9.3]). We mention for instance a result [40, Thm. 1.5] of Plan and Vershynin who showed that any subset  $\mathcal{S}$  of the unit sphere in  $\ell_2^n$  admits a  $\delta$ -isometric embedding into the  $d$ -dimensional Hamming cube  $(\{-1, 1\}^d, \|\cdot\|_1)$ , where  $d$  depends polynomially on  $\delta^{-1}$  and the Gaussian width of  $\mathcal{S}$ . In the above embedding terminology and in view of the elementary inequality  $|\alpha - \beta| \leq |\alpha^p - \beta^p|^{1/p}$  which holds for every  $\alpha, \beta > 0$ , Theorem 1.2 asserts that any  $n$ -point  $K$ -incompressible subset of  $L_p(\mu)$  admits an  $\varepsilon^{1/p}$ -isometric embedding into  $\ell_p^d$  for the above choice of dimension  $d$ . For further occurrences of  $\varepsilon$ -isometric embeddings in the dimensionality reduction and compressed sensing literatures, we refer to [8, 19, 20, 31, 40, 44] and the references therein.

### 1.3 Method of Proof

A large part of the (vast) literature on metric dimension reduction focuses on showing that a typical low-rank linear operator chosen randomly from a specific ensemble acts as an approximate isometry on a given set  $\mathcal{S}$  with high probability. For subsets  $\mathcal{S}$  of Euclidean space, this principle has been confirmed for random projections [12, 14, 21, 36], matrices with Gaussian [15, 16, 42], Rademacher [1, 5], and subgaussian [13, 17, 26, 31] entries, randomizations of matrices with the RIP [27] as well as more computationally efficient models [2, 3, 9, 24, 33] which are based on sparse matrices. Beyond its inherent interest as an  $\ell_p$ -dimension reduction theorem (albeit, for specific configurations of points), Theorem 1.2 also differs from the aforementioned works in its method of proof. The core of the argument, rather than sampling from a random matrix ensemble, relies on Maurey's empirical method [39] (see Sect. 2.1) which is a dimension-free way to approximate points in bounded convex subsets of Banach spaces by convex combinations of extreme points with prescribed length. An application of the method to the positive cone of  $L_p$ -distance matrices (the use of which in this context is inspired by classical work of Ball [6]) equipped with the supremum norm allows us to deduce (see Proposition 2.1) the conclusion of Theorem 1.2 under the stronger assumption that

$$K \geq \max_{i \in \{1, \dots, n\}} \|x_i\|_{L_\infty(\mu)}. \quad (8)$$

While Maurey's empirical method is an a priori existential statement that is proven via the probabilistic method, recent works (see [7, 18]) have focused on derandomizing its proof for specific Banach spaces. In the setting of Theorem 1.2, we can use these

tools to show (see Corollary 2.7) that there exists a greedy algorithm which receives as input the high-dimensional data  $\{x_i\}_{i=1}^n$  and produces as output the low-dimensional points  $\{y_i\}_{i=1}^n$ . Finally, using a suitable change of measure [34] (see Sect. 2.3) we are able to relax the stronger assumption (8) to that of  $K$ -incompressibility and derive the conclusion of Theorem 1.2. Finally, we emphasize that, in contrast to most of the dimension reduction algorithms (randomized or not) discussed earlier, the one which gives Theorem 1.2 is not *oblivious* but is rather tailored to the specific configuration of points  $\{x_i\}_{i=1}^n$  as it relies on the use of Maurey's empirical method.

#### 1.4 $\varepsilon$ -Isometric Dimension Reduction

Given two moduli  $\omega, \Omega : [0, \infty) \rightarrow [0, \infty)$ , we say (following [36]) that a Banach space  $(E, \|\cdot\|_E)$  admits metric dimension reduction with moduli  $(\omega, \Omega)$  if for any  $n \in \mathbb{N}$  there exists  $k_n = k_n(E) \in \mathbb{N}$  with  $k_n = n^{o(1)}$  as  $n \rightarrow \infty$  such that for any  $x_1, \dots, x_n \in E$ , there exist a subspace  $F$  of  $E$  with  $\dim F \leq k_n$  and  $y_1, \dots, y_n \in F$  satisfying

$$\forall i, j \in \{1, \dots, n\}, \quad \omega(\|x_i - x_j\|_E) \leq \|y_i - y_j\|_E \leq \Omega(\|x_i - x_j\|_E). \quad (9)$$

In view of Theorem 1.2, we would be interested in formulating a suitable notion of dimension reduction via  $\varepsilon$ -isometric embeddings which would be fitting to the moduli appearing in (6).

**Remark 1.3** Let  $a, b \in (0, \infty)$ , suppose that  $\omega, \Omega : [0, \infty) \rightarrow [0, \infty)$  are two moduli satisfying

$$\lim_{t \rightarrow \infty} \frac{\omega(t)}{t} = a \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{\Omega(t)}{t} = b \quad (10)$$

and that the Banach space  $(E, \|\cdot\|_E)$  admits metric dimension reduction with moduli  $(\omega, \Omega)$ . Fix  $n \in \mathbb{N}$  and  $x_1, \dots, x_n \in E$ . Applying the assumption (9) to the points  $sx_1, \dots, sx_n$  where  $s \gg 1$ , we deduce that there exist points  $y_1(s), \dots, y_n(s)$  in a  $k_n$ -dimensional subspace  $F(s)$  of  $E$  such that

$$\forall i, j \in \{1, \dots, n\}, \quad \omega(s\|x_i - x_j\|_E) \leq \|y_i(s) - y_j(s)\|_E \leq \Omega(s\|x_i - x_j\|_E). \quad (11)$$

For any  $\eta \in (0, 1)$ , we can then choose  $s$  large enough (as a function of  $\eta$  and the  $x_i$ ) such that

$$\forall i, j \in \{1, \dots, n\}, \quad (1 - \eta)a\|x_i - x_j\|_E \leq \frac{\|y_i(s) - y_j(s)\|_E}{s} \leq (1 + \eta)b\|x_i - x_j\|_E. \quad (12)$$

Therefore, we conclude that  $E$  also admits bi-Lipschitz dimension reduction (with distortion  $b/a$ ).

This simple scaling argument suggests that any reasonable notion of  $\varepsilon$ -isometric dimension reduction can differ from the corresponding bi-Lipschitz theory only in small scales, thus motivating the following definition. We denote by  $\mathbf{B}_E$  the unit ball of a normed space  $(E, \|\cdot\|_E)$ .

**Definition 1.4** ( *$\varepsilon$ -isometric dimension reduction*) Fix  $\varepsilon \in (0, 1)$ ,  $r \in (0, \infty)$  and let  $(E, \|\cdot\|_E)$  be an infinite-dimensional Banach space. We say that  $\mathbf{B}_E$  admits  $\varepsilon$ -isometric dimension reduction with power  $r$  if for every  $n \in \mathbb{N}$  there exists  $k_n = k_n^r(E, \varepsilon) \in \mathbb{N}$  with  $k_n = n^{o(1)}$  as  $n \rightarrow \infty$  for which the following condition holds. For every  $n$  points  $x_1, \dots, x_n \in \mathbf{B}_E$  there exist a linear subspace  $F$  of  $E$  with  $\dim F \leq k_n$  and points  $y_1, \dots, y_n \in F$  satisfying

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_E^r - \varepsilon \leq \|y_i - y_j\|_E^r \leq \|x_i - x_j\|_E^r + \varepsilon. \quad (13)$$

The fact that even high-dimensional infinite subsets of Euclidean space  $\ell_2$  may admit  $\varepsilon$ -isometric embeddings into low-dimensional subspaces follows from the additive version of the Johnson–Lindenstrauss lemma, first proven by Liaw, Mehrabian, Plan, and Vershynin [31] (see also [44, Prop. 9.3.2]). In contrast to that, combining the scaling argument of Remark 1.3 with the fact that any  $d$ -dimensional subspace of  $\ell_2$  is isometric to  $\ell_2^d$ , we deduce that if  $k_n(\varepsilon)$  is the least dimension such that any  $n$  points in  $\ell_2$  embed  $\varepsilon$ -isometrically in  $\ell_2^{k_n(\varepsilon)}$ , then  $k_n(\varepsilon) = n - 1$ . This justifies the restriction of Definition 1.4 to the unit ball  $\mathbf{B}_E$  of  $E$ .

It is clear from the definitions that if a Banach space  $E$  admits bi-Lipschitz dimension reduction with distortion  $\frac{1+\varepsilon}{1-\varepsilon}$ , where  $\varepsilon \in (0, 1)$ , then  $\mathbf{B}_E$  admits  $2\varepsilon$ -isometric dimension reduction with power  $r = 1$ . The  $\varepsilon$ -isometric analogue of Question 1.1 deserves further investigation.

**Question 1.5** For which values of  $p \neq 2$  does  $\mathbf{B}_{\ell_p}$  admit  $\varepsilon$ -isometric dimension reduction?

Even though the  $K$ -incompressibility assumption of Theorem 1.2 may a priori seem restrictive, it is satisfied for *most* configurations of points in  $\mathbf{B}_{\ell_p}$ . Suppose that  $n, N \in \mathbb{N}$  such that  $N$  is polynomial<sup>2</sup> in  $n$ . Then, standard considerations show that with high probability, a uniformly chosen  $n$ -point subset  $\mathcal{S}$  of  $N^{1/p} \mathbf{B}_{\ell_p^N}$  is  $O(\log n)^{1/p}$ -incompressible. We refer to Remark 2.4 for more information on this and related generic properties of finite subsets of rescaled  $p$ -balls.

### 1.5 $\varepsilon$ -Isometric Dimension Reduction by Linear Maps

A close inspection of the proof of Theorem 1.2 (see Remark 2.6) reveals that in fact the low-dimensional points  $\{y_i\}_{i=1}^n$  can be realized as images of the initial data  $\{x_i\}_{i=1}^n$  under a carefully chosen linear operator. Nevertheless, we will show that for any  $p \neq 2$  and  $n$  large enough, there exists an  $n$ -point subset of  $\mathbf{B}_{\ell_p}$  whose image under any fixed

<sup>2</sup> This relation between the parameters  $n, N$  is natural as any  $n$ -point subset of  $\ell_p$  embeds isometrically in  $\ell_p^N$  with  $N = \binom{n}{2} + 1$  by Ball’s isometric embedding theorem [6].

linear  $\varepsilon$ -isometric embedding has rank which is linear in  $n$ . In fact, we shall prove the following more general statement which refines a theorem that Lee, Mendel and Naor proved in [29] for bi-Lipschitz embeddings.

**Theorem 1.6** (Impossibility of linear dimension reduction in  $\mathbf{B}_{\ell_p}$ ) *Fix  $p \neq 2$  and two moduli  $\omega, \Omega : [0, \infty) \rightarrow [0, \infty)$  with  $\omega(1) > 0$ . For arbitrarily large  $n \in \mathbb{N}$ , there exists an  $n$ -point subset  $\mathcal{S}_{n,p}$  of  $\mathbf{B}_{\ell_p}$  such that the following holds. If  $T : \text{span}(\mathcal{S}_{n,p}) \rightarrow \ell_p^d$  is a linear operator satisfying*

$$\forall x, y \in \mathcal{S}_{n,p}, \quad \omega(\|x - y\|_{\ell_p}) \leq \|Tx - Ty\|_{\ell_p^d} \leq \Omega(\|x - y\|_{\ell_p}), \tag{14}$$

then  $d \geq \left(\frac{\omega(1)}{\Omega(1)}\right)^{\frac{2p}{|p-2|}} \cdot \frac{n-1}{2}$ .

## 2 Proof of Theorem 1.2

We say that a normed space  $(E, \|\cdot\|_E)$  has Rademacher type  $p$  if there exists a universal constant  $T \in (0, \infty)$  such that for every  $n \in \mathbb{N}$  and every  $x_1, \dots, x_n \in E$ ,

$$\frac{1}{2^n} \sum_{\varepsilon \in \{-1,1\}^n} \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|_E^p \leq T^p \sum_{i=1}^n \|x_i\|_E^p. \tag{15}$$

The least constant  $T$  such that (15) is satisfied is denoted by  $T_p(E)$ . A standard symmetrization argument (see [28, Prop. 9.11]) shows that if  $X_1, \dots, X_n$  are independent  $E$ -valued random variables with  $\mathbb{E}[X_i] = 0$  for every  $i \in \{1, \dots, n\}$ , then

$$\mathbb{E} \left\| \sum_{i=1}^n X_i \right\|_E^p \leq (2T_p(E))^p \sum_{i=1}^n \mathbb{E} \|X_i\|_E^p. \tag{16}$$

### 2.1 Maurey’s Empirical Method and Its Algorithmic Counterparts

A classical theorem of Carathéodory asserts that if  $\mathcal{T}$  is a subset of  $\mathbb{R}^m$ , then any point  $z$  in the convex hull  $\text{conv}(\mathcal{T})$  (that is, a convex combination of finitely many elements of  $\mathcal{T}$ ) can be expressed as a convex combination of at most  $m + 1$  points of  $\mathcal{T}$ . Maurey’s empirical method is a powerful dimension-free approximate version of Carathéodory’s theorem, first popularized in [39], that has numerous applications in geometry and theoretical computer science. Let  $(E, \|\cdot\|_E)$  be a Banach space, consider a bounded subset  $\mathcal{T}$  of  $E$  and fix  $z \in \text{conv}(\mathcal{T})$ . Since  $z$  is a convex combination of elements of  $\mathcal{T}$ , there exist  $m \in \mathbb{N}, \lambda_1, \dots, \lambda_m \in (0, \infty)$ , and  $t_1, \dots, t_m \in \mathcal{T}$  such that

$$\sum_{k=1}^m \lambda_k = 1 \quad \text{and} \quad z = \sum_{k=1}^m \lambda_k t_k. \tag{17}$$

Let  $X$  be an  $E$ -valued discrete random variable with  $\mathbb{P}\{X = t_k\} = \lambda_k$  for all  $k \in \{1, \dots, m\}$  and consider  $X_1, \dots, X_d$  i.i.d. copies of  $X$ . Then, conditions (17) ensure that  $X$  is well defined and  $\mathbb{E}[X] = z$ . Therefore, applying the Rademacher type condition (16) to the centered random variables  $\{X_s - z\}_{s=1}^d$  and normalizing, we get

$$\mathbb{E} \left\| \frac{1}{d} \sum_{s=1}^d X_s - z \right\|_E^p \leq \frac{(2T_p(E))^p}{d^{p-1}} \mathbb{E} \|X - z\|_E^p. \tag{18}$$

Since  $X$  takes values in  $\mathcal{T}$ , if  $\mathcal{T} \subseteq \mathbf{RB}_E$ , we then deduce that there exist  $x_1, \dots, x_d \in \mathcal{T}$  such that

$$\left\| \frac{1}{d} \sum_{s=1}^d x_s - z \right\|_E \leq \frac{4RT_p(E)}{d^{1-1/p}}. \tag{19}$$

While the above argument is probabilistic, recent works have focused on derandomizing Maurey’s sampling lemma for smaller classes of Banach spaces, thus constructing deterministic algorithms which output the empirical approximation  $\frac{x_1 + \dots + x_d}{d}$  of  $z$ . The first result in this direction is due to Barman [7] who treated the case that  $E$  is an  $L_r(\mu)$ -space,  $r \in (1, \infty)$ . This assumption was recently generalized by Ivanov in [18] who built a greedy algorithm which constructs the desired empirical mean in an arbitrary  $p$ -uniformly smooth space.

### 2.2 Dimension Reduction in $L_p(\mu)$ for Uniformly Bounded Vectors

With Maurey’s empirical method at hand, we are ready to proceed to the first part of the proof of Theorem 1.2, namely the  $\varepsilon$ -isometric dimension reduction property of  $L_p(\mu)$  under the strong assumption that the given point set consists of functions which are bounded in  $L_\infty(\mu)$ .

**Proposition 2.1** *Fix  $p \in [1, \infty)$ ,  $n \in \mathbb{N}$  and let  $\{x_i\}_{i=1}^n$  be a family of vectors in  $L_p(\mu)$  for some probability measure  $\mu$ . Denote by  $L = \max_{i \in \{1, \dots, n\}} \|x_i\|_{L_\infty(\mu)} \in [0, \infty]$ . Then for every  $\varepsilon \in (0, 1)$ , there exist  $d \in \mathbb{N}$  with  $d \leq \frac{32e^2(2L)^{2p} \log n}{\varepsilon^2}$  and  $y_1, \dots, y_n \in \ell_p^d$  such that*

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{L_p(\mu)}^p - \varepsilon \leq \|y_i - y_j\|_{\ell_p^d}^p \leq \|x_i - x_j\|_{L_p(\mu)}^p + \varepsilon. \tag{20}$$

**Proof** We shall identify  $\ell_\infty^{\binom{n}{2}}$  with the vector space of all symmetric  $n \times n$  real matrices with 0 on the diagonal equipped with the supremum norm. Consider the set

$$\mathcal{C}_p = \left\{ (\|z_i - z_j\|_{L_p(\rho)}^p)_{i,j=1,\dots,n} : \rho \text{ is a probability measure and } z_1, \dots, z_n \in L_p(\rho) \right\} \subseteq \ell_\infty^{\binom{n}{2}}. \tag{21}$$

It is obvious that  $\mathcal{C}_p$  is a cone in the sense that  $\mathcal{C}_p = \lambda \mathcal{C}_p$  for every  $\lambda > 0$  but moreover  $\mathcal{C}_p$  is convex. To see this, consider  $A, B \in \mathcal{C}_p$ , probability spaces  $(\Omega_1, \rho_1), (\Omega_2, \rho_2)$ , and vectors  $\{z_i\}_{i=1}^n, \{w_i\}_{i=1}^n$  in  $L_p(\rho_1)$  and  $L_p(\rho_2)$  respectively such that

$$\forall i, j \in \{1, \dots, n\}, \quad A_{ij} = \|z_i - z_j\|_{L_p(\rho_1)}^p \quad \text{and} \quad B_{ij} = \|w_i - w_j\|_{L_p(\rho_2)}^p. \quad (22)$$

Fix  $\lambda \in (0, 1)$  and consider the disjoint union  $\Omega_1 \sqcup \Omega_2$  of  $\Omega_1$  and  $\Omega_2$  equipped with the probability measure  $\rho(\lambda) = \lambda\rho_1 + (1 - \lambda)\rho_2$ . Then, by (22) the functions  $\zeta_i : \Omega_1 \sqcup \Omega_2 \rightarrow \mathbb{R}$  given by  $\zeta_i|_{\Omega_1} = z_i$  and  $\zeta_i|_{\Omega_2} = w_i$ , where  $i \in \{1, \dots, n\}$ , belong to  $L_p(\rho(\lambda))$  and satisfy the conditions

$$\begin{aligned} \forall i, j \in \{1, \dots, n\}, \quad \|\zeta_i - \zeta_j\|_{L_p(\rho(\lambda))}^p &= \lambda \|z_i - z_j\|_{L_p(\rho_1)}^p + (1 - \lambda) \|w_i - w_j\|_{L_p(\rho_2)}^p \\ &= \lambda A_{ij} + (1 - \lambda) B_{ij}, \end{aligned} \quad (23)$$

which ensure that  $\lambda A + (1 - \lambda)B \in \mathcal{C}_p$ , making  $\mathcal{C}_p$  a convex cone. Consider the embedding  $\mathcal{M} : L_p(\mu)^n \rightarrow \mathcal{C}_p$  mapping a vector  $z = (z_1, \dots, z_n)$  to the corresponding distance matrix, i.e.

$$\forall i, j \in \{1, \dots, n\}, \quad \mathcal{M}(z)_{ij} = \|z_i - z_j\|_{L_p(\mu)}^p. \quad (24)$$

By Ball’s isometric embedding theorem [6],  $x_1, \dots, x_n$  have isometric images in  $\ell_p^N$  with  $N = \binom{n}{2} + 1$ . Without loss of generality we will thus assume that the given points  $x_1, \dots, x_n \in L_p(\mu)$  are simple functions (that is, each of them takes only finitely many values) with  $\|x_i\|_{L_\infty(\mu)} \leq L$ . Let  $\{S_1, \dots, S_m\}$  be a partition of the underlying measure space such that each function  $x_i$  is constant on each  $S_k$  and suppose that  $x_i|_{S_k} = a(i, k) \in [-L, L]$  for  $i \in \{1, \dots, n\}$  and  $k \in \{1, \dots, m\}$ . Then, for every  $i, j \in \{1, \dots, n\}$ , we have

$$\begin{aligned} \mathcal{M}(x)_{ij} &= \sum_{k=1}^m \int_{S_k} |x_i - x_j|^p \, d\mu = \sum_{k=1}^m \mu(S_k) \cdot |a(i, k) - a(j, k)|^p \\ &= \sum_{k=1}^m \mu(S_k) \mathcal{M}(y(k))_{ij}, \end{aligned} \quad (25)$$

where  $y(k) \stackrel{\text{def}}{=} (a(1, k), \dots, a(n, k)) \in L_p(\mu)^n$  is a vector whose components are constant functions. As  $\mu$  is a probability measure and  $\{S_1, \dots, S_m\}$  is a partition, identity (25) implies that

$$\mathcal{M}(x) \in \text{conv}\{\mathcal{M}(y(k)) : k \in \{1, \dots, m\}\} \subseteq \ell_\infty^{\binom{n}{2}}. \quad (26)$$

Observe that since  $a(i, k) \in [-L, L]$  for every  $i \in \{1, \dots, n\}$  and  $k \in \{1, \dots, m\}$ , we have

$$\forall k \in \{1, \dots, m\}, \quad \|\mathcal{M}(y(k))\|_{\ell_\infty^{\binom{n}{2}}} = \max_{i, j \in \{1, \dots, n\}} |a(i, k) - a(j, k)|^p \leq (2L)^p. \quad (27)$$

Moreover,  $\ell_\infty^{\binom{n}{2}}$  is  $e$ -isomorphic to  $\ell_{p_n}^{\binom{n}{2}}$  where  $p_n = \log \binom{n}{2}$ . It is well-known (see [28, Chap. 9]) that  $T_2(\ell_p) \leq \sqrt{p-1}$  for every  $p \geq 2$  and thus

$$T_2(\ell_\infty^{\binom{n}{2}}) \leq e\sqrt{p_n - 1} < \sqrt{2e^2 \log n}. \tag{28}$$

Applying Maurey’s sampling lemma (Sect. 2.1) while taking into account (27) and (28), we deduce that for every  $d \geq 1$  there exist  $k_1, \dots, k_d \in \{1, \dots, m\}$  such that

$$\left\| \frac{1}{d} \sum_{s=1}^d \mathcal{M}(y(k_s)) - \mathcal{M}(x) \right\|_{\ell_\infty^{\binom{n}{2}}} \leq \frac{2^{p+\frac{5}{2}} e L^p \sqrt{\log n}}{\sqrt{d}}. \tag{29}$$

Therefore, if  $\varepsilon \in (0, 1)$  is such that  $d \geq \frac{32e^2(2L)^{2p} \log n}{\varepsilon^2}$  we then have

$$\forall i, j \in \{1, \dots, n\}, \quad \left| \frac{1}{d} \sum_{s=1}^d |a(i, k_s) - a(j, k_s)|^p - \|x_i - x_j\|_{L_p(\mu)}^p \right| \leq \varepsilon. \tag{30}$$

Finally, consider for each  $i \in \{1, \dots, n\}$  a vector  $y_i = (y_i(1), \dots, y_i(d)) \in \ell_p^d$  given by

$$\forall s \in \{1, \dots, d\}, y_i(s) = \frac{a(i, k_s)}{d^{1/p}} \tag{31}$$

and notice that (30) can be equivalently rewritten as

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{L_p(\mu)}^p - \varepsilon \leq \|y_i - y_j\|_{\ell_p^d}^p \leq \|x_i - x_j\|_{L_p(\mu)}^p + \varepsilon, \tag{32}$$

concluding the proof of the proposition. □

**Remark 2.2** It is worth emphasizing that the coordinates of the vectors  $y_1, \dots, y_n$  produced in Proposition 2.1 consist (up to rescaling) of values of the functions  $x_1, \dots, x_n$ . Such low-dimensional embeddings via *sampling* are a central object of study in approximation theory, see e.g. the recent survey [25] and the references therein.

The additive version of the Johnson–Lindenstrauss lemma, first observed in [31] as a consequence of a deep matrix deviation inequality (see also [44, Chap. 9]), asserts that for every  $n$ -point subset  $\mathcal{X} = \{x_1, \dots, x_n\}$  of a Hilbert space  $\mathcal{H}$  and every  $\varepsilon \in (0, 1)$ , there exist  $d \leq \frac{Cw(\mathcal{X})^2}{\varepsilon^2}$  and points  $y_1, \dots, y_n \in \ell_2^d$  such that

$$\forall i, j \in \{1, \dots, n\}, \quad \|x_i - x_j\|_{\mathcal{H}} - \varepsilon \leq \|y_i - y_j\|_{\ell_2^d} \leq \|x_i - x_j\|_{\mathcal{H}} + \varepsilon, \tag{33}$$

where  $w(\mathcal{X})$  is the mean width of  $\mathcal{X}$ . We will now observe that the spherical symmetry of  $\mathbf{B}_{\ell_2}$  allows us to deduce a similar conclusion for points in  $\mathbf{B}_{\mathcal{H}}$  by removing the

incompressibility assumption from Proposition 2.1 when  $p = 2$ . We shall use the standard notation  $L_p^N$  for the space  $L_p(\mu_N)$  where  $\mu_N$  is the normalized counting measure on the finite set  $\{1, \dots, N\}$ , that is

$$\forall a = (a_1, \dots, a_N) \in \mathbb{R}^N, \|a\|_{L_p^N} \stackrel{\text{def}}{=} \left( \frac{1}{N} \sum_{i=1}^N |a_i|^p \right)^{1/p}. \tag{34}$$

Observe that for  $0 < p < q \leq \infty$ , we have  $\mathbf{B}_{L_q^N} \subseteq \mathbf{B}_{L_p^N}$ .

**Corollary 2.3** *There exists a universal constant  $C \in (0, \infty)$  such that the following statement holds. Fix  $n \in \mathbb{N}$  and let  $\{x_i\}_{i=1}^n$  be a family of vectors in  $\mathbf{B}_{\mathcal{H}}$  for some Hilbert space  $\mathcal{H}$ . Then for every  $\varepsilon \in (0, 1)$ , there exist  $d \in \mathbb{N}$  with  $d \leq \frac{C(\log n)^3}{\varepsilon^4}$  and points  $y_1, \dots, y_n \in \ell_2^d$  such that*

$$\forall i, j \in \{1, \dots, n\}, \|x_i - x_j\|_{\mathcal{H}} - \varepsilon \leq \|y_i - y_j\|_{\ell_2^d} \leq \|x_i - x_j\|_{\mathcal{H}} + \varepsilon. \tag{35}$$

Before proceeding to the derivation of (35) we emphasize that since the given points  $\{x_i\}_{i=1}^n$  belong to  $\mathbf{B}_{\mathcal{H}}$ , Corollary 2.3 is formally weaker than the Johnson–Lindenstrauss lemma. However we include it here since it differs from [21] in that the low-dimensional point set  $\{y_i\}_{i=1}^n$  is not obtained as an image of  $\{x_i\}_{i=1}^n$  under a typical low-rank matrix from a specific ensemble.

**Proof of Corollary 2.3** Since any  $n$ -point subset  $\{x_1, \dots, x_n\}$  of  $\mathcal{H}$  embeds linearly and isometrically in  $L_2^n$ , we assume that  $x_1, \dots, x_n \in \mathbf{B}_{L_2^n}$ . We will need the following claim.

*Claim.* Suppose that  $X_1, \dots, X_n$  are (not necessarily independent) random vectors, each uniformly distributed on the unit sphere  $\mathbb{S}^{n-1}$  of  $L_2^n$ . Then, for some universal constant  $S \in (0, \infty)$ ,

$$\mathbb{E} \left[ \max_{i \in \{1, \dots, n\}} \|X_i\|_{L_\infty^n} \right] \leq S\sqrt{\log n}, \tag{36}$$

**Proof of the Claim** By a standard estimate of Schechtman and Zinn [43, Thm. 3], for a uniformly distributed random vector  $X$  on the unit sphere  $\mathbb{S}^{n-1}$  of  $L_2^n$ , we have

$$\forall t \geq \gamma_1\sqrt{\log n}, \quad \mathbb{P}\{\|X\|_{L_\infty^n} > t\} \leq e^{-\gamma_2 t^2} \tag{37}$$

for some absolute constants  $\gamma_1, \gamma_2 \in (0, \infty)$ . Let  $W \stackrel{\text{def}}{=} \max_{i \in \{1, \dots, n\}} \|X_i\|_{L_\infty^n}$  and notice that

$$\forall K \in (\gamma_1, \infty), \quad \mathbb{E}[W] = \int_0^\infty \mathbb{P}\{W > t\} dt \leq K\sqrt{\log n} + \int_{K\sqrt{\log n}}^\infty \mathbb{P}\{W > t\} dt. \tag{38}$$

By the union bound, we have

$$\forall t > 0, \quad \mathbb{P}\{W > t\} \leq \sum_{i=1}^n \mathbb{P}\{X_i > t\} = n\mathbb{P}\{X_1 > t\}. \tag{39}$$

Combining (38) and (39), we therefore get

$$\begin{aligned} \mathbb{E}[W] &\leq K\sqrt{\log n} + n \int_{K\sqrt{\log n}}^{\infty} \mathbb{P}\{X_1 > t\} dt \stackrel{(37)}{\leq} K\sqrt{\log n} + n \int_{K\sqrt{\log n}}^{\infty} e^{-\gamma_2 t^2} dt \\ &= K\sqrt{\log n} + n\sqrt{\log n} \int_K^{\infty} n^{-\gamma_2 u^2} du \\ &= K\sqrt{\log n} + \sqrt{\log n} \int_K^{\infty} n^{1-\gamma_2 u^2} du. \end{aligned} \tag{40}$$

Choosing  $K > \gamma_1$  such that  $K^2\gamma_2 > 1$ , the exponent in the last integrand becomes negative, thus

$$\mathbb{E}[W] \leq K\sqrt{\log n} + 2\sqrt{\log n} \int_K^{\infty} 2^{-\gamma_2 u^2} du \leq S\sqrt{\log n} \tag{41}$$

for a large enough constant  $S \in (0, \infty)$  and the claim follows.

Now let  $U \in \mathcal{O}(n)$  be a uniformly chosen random rotation on  $\mathbb{R}^n$ . The aforementioned claim shows that since  $\|x_i\|_{L_2^n} \leq 1$  for every  $i \in \{1, \dots, n\}$ , writing  $\hat{x}_i = \frac{x_i}{\|x_i\|_{L_2^n}}$ , we have the estimate

$$\mathbb{E}\left[\max_{i \in \{1, \dots, n\}} \|Ux_i\|_{L_\infty^n}\right] \leq \mathbb{E}\left[\max_{i \in \{1, \dots, n\}} \|U\hat{x}_i\|_{L_\infty^n}\right] \leq S\sqrt{\log n}. \tag{42}$$

Therefore, by (42) and Proposition 2.1 there exist a constant  $C \in (0, \infty)$  and a rotation  $U \in \mathcal{O}(n)$  such that for every  $\varepsilon \in (0, 1)$  there exist  $d \leq \frac{C(\log n)^3}{\varepsilon^4}$  and points  $y_1, \dots, y_n \in \ell_2^d$  for which

$$\forall i, j \in \{1, \dots, n\}, \quad \|Ux_i - Ux_j\|_{L_2^n}^2 - \varepsilon^2 \leq \|y_i - y_j\|_{\ell_2^d}^2 \leq \|Ux_i - Ux_j\|_{L_2^n}^2 + \varepsilon^2. \tag{43}$$

Since  $\|Ua - Ub\|_{L_2^n} = \|a - b\|_{L_2^n}$  for every  $a, b \in L_2^n$ , the conclusion follows by the elementary inequality  $|\alpha - \beta| \leq \sqrt{|\alpha^2 - \beta^2|}$  which holds for every positive numbers  $\alpha, \beta \in (0, \infty)$ . □

**Remark 2.4** Fix  $p \in [1, \infty)$ . The isometric embedding theorem of Ball [6] asserts that any  $n$ -point subset of  $\ell_p$  admits an isometric embedding into  $\ell_p^N$  where  $N = \binom{n}{2} + 1$ . Suppose, more generally, that  $n, N \in \mathbb{N}$  are such that  $N$  is polynomial in  $n$ . Considerations in the spirit of the proof of Corollary 2.3 (e.g. relying on [43]) then show

that if  $x_1, \dots, x_n$  are independent uniformly random points in  $\mathbf{B}_{L_p^N}$ , then the random set  $\{x_1, \dots, x_n\}$  is  $O(\log n)^{1/p}$ -incompressible. In other words, incompressibility is a *generic* property of random  $n$ -point subsets of  $\mathbf{B}_{L_p^N}$ . On the other hand, a typical  $n$ -point subset of  $\mathbf{B}_{L_p^N}$  is known to be approximately a simplex due to work of Arias-de-Reyna, Ball, and Villa [4] and so, in particular, it can be bi-Lipschitzly embedded in  $O(\log n)$  dimensions.

### 2.3 Factorization and Proof of Theorem 1.2

Observe that Proposition 2.1 is rather non-canonical as the conclusion depends on the pairwise distances between the points  $\{x_i\}_{i=1}^n$  in  $L_p(\mu)$  whereas the bound on the dimension depends on  $L = \max_i \|x_i\|_{L_\infty(\mu)}$ . In order to deduce Theorem 1.2 from this (a priori weaker) statement we shall leverage the fact that Proposition 2.1 holds for *any* probability measure  $\mu$  by optimizing this parameter  $L$  over all lattice-isomorphic images of  $\{x_i\}_{i=1}^n$ . The optimal such *change of measure* which allows us to replace  $L$  by  $\|\max_i |x_i|\|_{L_p(\mu)}$  is a special case of a classical factorization theorem of Maurey (see [34] or [23, Thm. 5] for the general statement), whose short proof we include for completeness.

**Proposition 2.5** *Fix  $n \in \mathbb{N}$ ,  $p \in (0, \infty)$ , and a probability space  $(\Omega, \mu)$ . For every points  $x_1, \dots, x_n \in L_p(\mu)$ , there exists a nonnegative density function  $f : \Omega \rightarrow \mathbb{R}_+$  supported on the support of  $\max_i |x_i|$  such that if  $\nu$  is the probability measure on  $\Omega$  given by  $\frac{d\nu}{d\mu} = f$ , then*

$$\max_{i \in \{1, \dots, n\}} \|x_i f^{-1/p}\|_{L_\infty(\nu)} \leq \left\| \max_{i \in \{1, \dots, n\}} |x_i| \right\|_{L_p(\mu)}. \tag{44}$$

**Proof** Let  $V = \text{supp}(\max_i |x_i|) \subseteq \Omega$  and define the change of measure  $f$  as

$$\forall \omega \in \Omega, \quad f(\omega) \stackrel{\text{def}}{=} \frac{\max_{i \in \{1, \dots, n\}} |x_i(\omega)|^p}{\int_\Omega \max_{i \in \{1, \dots, n\}} |x_i(\theta)|^p d\theta}. \tag{45}$$

Then, (44) is elementary to check. □

We are now ready to complete the proof of Theorem 1.2.

**Proof of Theorem 1.2** Fix a  $K$ -incompressible family of vectors  $x_1, \dots, x_n \in L_p(\Omega, \mu)$  and let  $V = \text{supp}(\max_i |x_i|) \subseteq \Omega$ . Denote by  $f : \Omega \rightarrow \mathbb{R}_+$  the change of density from Proposition 2.5. If  $\frac{d\nu}{d\mu} = f$ , then the linear operator  $T : L_p(V, \mu) \rightarrow L_p(\Omega, \nu)$  given by  $Tg = f^{-1/p}g$  is (trivially) a linear isometry. Therefore, Proposition 2.1 and (44) show that there exist  $d \in \mathbb{N}$  with  $d \leq \frac{32e^2(2K)^{2p} \log n}{\varepsilon^2}$  and points  $y_1, \dots, y_n \in \ell_d^p$  such that the condition

$$\begin{aligned} \|x_i - x_j\|_{L_p(\mu)}^p - \varepsilon &= \|Tx_i - Tx_j\|_{L_p(\nu)}^p - \varepsilon \leq \|y_i - y_j\|_{\ell_d^p}^p \\ &\leq \|Tx_i - Tx_j\|_{L_p(\nu)}^p + \varepsilon = \|x_i - x_j\|_{L_p(\mu)}^p + \varepsilon, \end{aligned} \tag{46}$$

is satisfied for every  $i, j \in \{1, \dots, n\}$ . This concludes the proof of Theorem 1.2.

**Remark 2.6** A careful inspection of the proof of Theorem 1.2 reveals that the low-dimensional points  $\{y_i\}_{i=1}^n$  can be obtained as images of the given points  $\{x_i\}_{i=1}^n$  under a linear transformation. Indeed, starting from a  $K$ -incompressible family of points  $\{x_i\}_{i=1}^n$  in  $L_p(\Omega, \mu)$ , we use Proposition 2.5 to find a change of measure  $T : L_p(V, \mu) \rightarrow L_p(\Omega, \nu)$  such that  $\{Tx_i\}_{i=1}^n$  satisfy the stronger assumption of Proposition 2.1. Then, for some  $d \in \mathbb{N}$  with  $d \leq \frac{32e^2(2K)^{2p} \log n}{\varepsilon^2}$  we find pairwise disjoint measurable subsets  $S_1, \dots, S_d$  of  $\Omega$ , each with positive measure, such that if  $S : L_p(\Omega, \nu) \rightarrow \ell_p^d$  is the linear map

$$\forall z \in L_p(\Omega, \nu), \quad Sz \stackrel{\text{def}}{=} \left( \frac{1}{d^{1/p}} \left( \frac{1}{\mu(S_1)} \int_{S_1} z \, d\nu, \dots, \frac{1}{\mu(S_d)} \int_{S_d} z \, d\nu \right) \right) \in \ell_p^d, \quad (47)$$

then the points  $\{y_i\}_{i=1}^n = \{(S \circ T)x_i\}_{i=1}^n \subseteq \ell_p^d$  satisfy the desired conclusion (6).

We conclude this section by observing that the argument leading to Theorem 1.2 is constructive.

**Corollary 2.7** *In the setting of Theorem 1.2, there exists a greedy algorithm which receives as input the high-dimensional points  $\{x_i\}_{i=1}^n$  and produces as output the low-dimensional points  $\{y_i\}_{i=1}^n$ .*

**Proof** As the density (45) is explicitly defined, the linear operator  $T : L_p(V, \mu) \rightarrow L_p(\Omega, \nu)$  can also be efficiently constructed. On the other hand, in order to construct the operator  $S$  defined by (47) one needs to find the corresponding partition  $\{S_1, \dots, S_d\}$  and this was achieved in Proposition 2.1 via an application of Maurey’s sampling lemma to the cone  $\mathcal{C}_p \subseteq \ell_\infty^N$  where  $N = \binom{n}{2}$ . As  $\ell_\infty^N$  is  $e$ -isomorphic to the 2-uniformly smooth space  $\ell_{\log N}^N$ , Ivanov’s result from [18] implies that the construction can be implemented by a greedy algorithm.  $\square$

**Analysis of the algorithm.** The only nontrivial algorithmic task in our dimensionality reduction result is the implementation of Maurey’s approximate Carathéodory theorem. In the special case of  $\ell_p$  spaces, various constructive proofs of Maurey’s lemma are known [7, 11, 35], each of which allows for an analysis of the algorithm’s running time. Assume that the initial points  $x_1, \dots, x_n \in \mathbf{B}_{\ell_p}^m$  for some finite  $m$ . Implementing, for instance, the mirror descent algorithm of [35, Thm. 3.5] on the convex hull of  $\mathcal{M}(y(1)), \dots, \mathcal{M}(y(m))$  appearing in the proof of Theorem 1.2, the corresponding indices  $k_1, \dots, k_d$  can be produced in time  $O(mn^2 \log n/\varepsilon^2)$ . Therefore, assuming that the points  $x_1, \dots, x_n$  a priori lie in a  $\text{poly}(n)$ -dimensional space (as is reasonable by Ball’s embedding theorem), the output points  $y_1, \dots, y_n$  can be constructed in time  $\text{poly}(n, 1/\varepsilon)$ .

### 3 Proof of Theorem 1.6

In this section we prove Theorem 1.6. The constructed subset of  $\mathbf{B}_{\ell_p}$  which does not embed linearly into  $\ell_p^d$  for small  $d$  is a slight modification of the one considered in [29].

**Proof of Theorem 1.6** Fix  $m \in \mathbb{N}$  and denote by  $\{w_i\}_{i=1}^{2^m}$  the rows of the  $2^m \times 2^m$  Walsh matrix and by  $\{e_i\}_{i=1}^{2^m}$  the coordinate basis vectors of  $\mathbb{R}^{2^m}$ . Consider the  $n$ -point set

$$\mathcal{S}_{n,p} = \{0\} \cup \{e_1, \dots, e_{2^m}\} \cup \left\{ \frac{w_1}{2^{m/p}}, \dots, \frac{w_{2^m}}{2^{m/p}} \right\} \subseteq \mathbf{B}_{\ell_p^{2^m}} \tag{48}$$

where  $n = 2^{m+1} + 1$  and suppose that  $T : \ell_p^{2^m} \rightarrow \ell_p^d$  is a linear operator such that

$$\forall x, y \in \mathcal{S}_{n,p}, \omega(\|x - y\|_{\ell_p^{2^m}}) \leq \|Tx - Ty\|_{\ell_p^d} \leq \Omega(\|x - y\|_{\ell_p^{2^m}}). \tag{49}$$

Assume first that  $1 \leq p < 2$ . If we write  $w_i = \sum_{j=1}^{2^m} w_i(j)e_j$  then by orthogonality of  $\{w_i\}_{i=1}^{2^m}$ ,

$$\begin{aligned} \sum_{i=1}^{2^m} \|Tw_i\|_{\ell_p^d}^2 &= \sum_{i=1}^{2^m} \left\| \sum_{j=1}^{2^m} w_i(j)Te_j \right\|_{\ell_p^d}^2 = \sum_{j,k=1}^{2^m} \langle w_j, w_k \rangle \langle Te_j, Te_k \rangle \\ &= 2^m \sum_{j=1}^{2^m} \|Te_j\|_{\ell_p^d}^2. \end{aligned} \tag{50}$$

By assumption (49) on  $T$ , we have

$$\forall j \in \{1, \dots, 2^m\}, \quad \|Te_j\|_{\ell_p^d}^2 \leq \|Te_j\|_{\ell_p^2}^2 \leq \Omega(1)^2 \tag{51}$$

and

$$\forall j \in \{1, \dots, 2^m\}, \quad \|Tw_j\|_{\ell_p^d}^2 \geq 2^{\frac{2m}{p}} d^{-\frac{2-p}{p}} \left\| T\left(\frac{w_j}{2^{m/p}}\right) \right\|_{\ell_p^d}^2 \geq 2^{\frac{2m}{p}} d^{-\frac{2-p}{p}} \omega(1)^2. \tag{52}$$

Combining (50), (51), and (52) we deduce that

$$2^{m(1+\frac{2}{p})} d^{-\frac{2-p}{p}} \omega(1)^2 \leq 4^m \Omega(1)^2, \tag{53}$$

which is equivalent to  $d \geq \left(\frac{\omega(1)}{\Omega(1)}\right)^{\frac{2p}{2-p}} 2^m = \left(\frac{\omega(1)}{\Omega(1)}\right)^{\frac{2p}{|p-2|}} \cdot \frac{n-1}{2}$ . The case  $p > 2$  is treated similarly.

**Remark 3.1** The point set  $\mathcal{S}_{n,p}$  considered in the proof of Theorem 1.6 for  $p \neq 2$  is  $O(n^{1/p})$  incompressible and does not admit a linear  $\frac{1}{2}$ -isometric embedding in fewer than  $\Omega(n)$  dimensions. This shows that the dimension of the linear embedding exhibited in Theorem 1.2 has to be of order at least  $\Omega(K^p)$  up to lower order terms. This should be compared with the  $O(K^{2p} \log n)$  upper bound of Theorem 1.2.

**Acknowledgements** I am grateful to Keith Ball, Assaf Naor, and Pierre Youssef for insightful discussions and useful feedback.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Achlioptas, D.: Database-friendly random projections: Johnson–Lindenstrauss with binary coins. *J. Comput. Syst. Sci.* **66**, 671–687 (2003)
2. Ailon, N., Chazelle, B.: The fast Johnson–Lindenstrauss transform and approximate nearest neighbors. *SIAM J. Comput.* **39**(1), 302–322 (2009)
3. Ailon, N., Liberty, E.: An almost optimal unrestricted fast Johnson–Lindenstrauss transform. *ACM Trans. Algorithms* **9**(3), # 21 (2013)
4. Arias-de Reyna, J., Ball, K., Villa, R.: Concentration of the distance in finite-dimensional normed spaces. *Mathematika* **45**(2), 245–252 (1998)
5. Arriaga, R.I., Vempala, S.: An algorithmic theory of learning: robust concepts and random projection. In: 40th Annual Symposium on Foundations of Computer Science (New York 1999), pp. 616–623. IEEE Computer Society, Los Alamitos (1999)
6. Ball, K.: Isometric embedding in  $l_p$ -spaces. *Eur. J. Combin.* **11**(4), 305–311 (1990)
7. Barman, S.: Approximating Nash equilibria and dense subgraphs via an approximate version of Carathéodory's theorem. *SIAM J. Comput.* **47**(3), 960–981 (2018)
8. Bartal, Y., Gottlieb, L.-A.: Approximate nearest neighbor search for  $\ell_p$ -spaces ( $2 < p < \infty$ ) via embeddings. *Theor. Comput. Sci.* **757**, 27–35 (2019)
9. Bourgain, J., Dirksen, S., Nelson, J.: Toward a unified theory of sparse dimensionality reduction in Euclidean space. *Geom. Funct. Anal.* **25**(4), 1009–1088 (2015)
10. Brinkman, B., Charikar, M.: On the impossibility of dimension reduction in  $l_1$ . *J. ACM* **52**(5), 766–788 (2005)
11. Combettes, C.W., Pokutta, S.: Revisiting the approximate Carathéodory problem via the Frank–Wolfe algorithm. *Math. Program.* **197**(1), 191–214 (2023)
12. Dasgupta, S., Gupta, A.: An elementary proof of a theorem of Johnson and Lindenstrauss. *Random Struct. Algorithms* **22**(1), 60–65 (2003)
13. Dirksen, S.: Dimensionality reduction with subgaussian matrices: a unified theory. *Found. Comput. Math.* **16**(5), 1367–1396 (2016)
14. Frankl, P., Maehara, H.: The Johnson–Lindenstrauss lemma and the sphericity of some graphs. *J. Combin. Theory Ser. B* **44**(3), 355–362 (1988)
15. Gordon, Y.: On Milman's inequality and random subspaces which escape through a mesh in  $R^n$ . In: *Geometric Aspects of Functional Analysis (1986/1987)*. Lecture Notes in Mathematics, vol. 1317, pp. 84–106. Springer, Berlin (1988)
16. Indyk, P., Motwani, R.: Approximate nearest neighbors: towards removing the curse of dimensionality. In: 30th Annual ACM Symposium on Theory of Computing (Dallas 1998), pp. 604–613. ACM, New York (1999)
17. Indyk, P., Naor, A.: Nearest-neighbor-preserving embeddings. *ACM Trans. Algorithms* **3**(3), # 31 (2007)
18. Ivanov, G.: Approximate Carathéodory's theorem in uniformly smooth Banach spaces. *Discrete Comput. Geom.* **66**(1), 273–280 (2021)
19. Jacques, L.: A quantized Johnson–Lindenstrauss lemma: the finding of Buffon's needle. *IEEE Trans. Inf. Theory* **61**(9), 5012–5027 (2015)

20. Jacques, L.: Small width, low distortions: quantized random embeddings of low-complexity sets. *IEEE Trans. Inf. Theory* **63**(9), 5477–5495 (2017)
21. Johnson, W.B., Lindenstrauss, J.: Extensions of Lipschitz mappings into a Hilbert space. In: *Conference in Modern Analysis and Probability (New Haven 1982)*. Contemporary Mathematics, vol. 26, pp. 189–206. American Mathematical Society, Providence (1984)
22. Johnson, W.B., Naor, A.: The Johnson–Lindenstrauss lemma almost characterizes Hilbert space, but not quite. *Discrete Comput. Geom.* **43**(3), 542–553 (2010)
23. Johnson, W.B., Schechtman, G.: Finite dimensional subspaces of  $L_p$ . In: *Handbook of the Geometry of Banach Spaces*, vol. I, pp. 837–870. North-Holland, Amsterdam (2001)
24. Kane, D.M., Nelson, J.: Sparser Johnson–Lindenstrauss transforms. *J. ACM* **61**(1), # 4 (2014)
25. Kashin, B., Kosov, E., Limonova, I., Temlyakov, V.: Sampling discretization and related problems. *J. Complex.* **71**, # 101653 (2022)
26. Klartag, B., Mendelson, S.: Empirical processes and random projections. *J. Funct. Anal.* **225**(1), 229–245 (2005)
27. Krahmer, F., Ward, R.: New and improved Johnson–Lindenstrauss embeddings via the restricted isometry property. *SIAM J. Math. Anal.* **43**(3), 1269–1281 (2011)
28. Ledoux, M., Talagrand, M.: *Probability in Banach Spaces. Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 23. Springer, Berlin (1991)
29. Lee, J.R., Mendel, M., Naor, A.: Metric structures in  $L_1$ : dimension, snowflakes, and average distortion. *Eur. J. Combin.* **26**(8), 1180–1190 (2005)
30. Lee, J.R., Naor, A.: Embedding the diamond graph in  $L_p$  and dimension reduction in  $L_1$ . *Geom. Funct. Anal.* **14**(4), 745–747 (2004)
31. Liaw, C., Mehrabian, A., Plan, Y., Vershynin, R.: A simple tool for bounding the deviation of random matrices on geometric sets. In: *Geometric Aspects of Functional Analysis. Lecture Notes in Mathematics*, vol. 2169, pp. 277–299. Springer, Cham (2017)
32. Matoušek, J.: On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.* **93**, 333–344 (1996)
33. Matoušek, J.: On variants of the Johnson–Lindenstrauss lemma. *Random Struct. Algorithms* **33**(2), 142–156 (2008)
34. Maurey, B.: *Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces  $L^p$ . Astérisque*, vol. 11. Société Mathématique de France, Paris (1974)
35. Mirrokni, V., Leme, R.P., Vladu, A., Wong, S.-C.: Tight bounds for approximate Carathéodory and beyond. In: *34th International Conference on Machine Learning (Sydney 2017)*. Proceedings of Machine Learning Research, vol. 70, pp. 2440–2448 (2017). <http://JMLR.org>
36. Naor, A.: Metric dimension reduction: a snapshot of the Ribe program. In: *International Congress of Mathematicians (Rio de Janeiro 2018)*, vol. I. Plenary Lectures, pp. 759–837. World Scientific, Hackensack (2018)
37. Naor, A., Pisier, G., Schechtman, G.: Impossibility of dimension reduction in the nuclear norm. *Discrete Comput. Geom.* **63**(2), 319–345 (2020)
38. Ostrovskii, M.I.: *Metric Embeddings: Bilipschitz and Coarse Embeddings into Banach Spaces. De Gruyter Studies in Mathematics*, vol. 49. De Gruyter, Berlin (2013)
39. Pisier, G.: Remarques sur un résultat non publié de B. Maurey. In: *Seminar on Functional Analysis, 1980–1981*, # 5. École Polytech., Palaiseau (1981)
40. Plan, Y., Vershynin, R.: Dimension reduction by random hyperplane tessellations. *Discrete Comput. Geom.* **51**(2), 438–461 (2014)
41. Regev, O., Vidick, T.: Bounds on dimension reduction in the nuclear norm. In: *Geometric Aspects of Functional Analysis*, vol. II. Lecture Notes in Mathematics, vol. 2266, pp. 279–299. Springer, Cham (2020)
42. Schechtman, G.: Two observations regarding embedding subsets of Euclidean spaces in normed spaces. *Adv. Math.* **200**(1), 125–135 (2006)
43. Schechtman, G., Zinn, J.: On the volume of the intersection of two  $L_p^n$  balls. *Proc. Am. Math. Soc.* **110**(1), 217–224 (1990)
44. Vershynin, R.: *High-Dimensional Probability. Cambridge Series in Statistical and Probabilistic Mathematics*, vol. 47. Cambridge University Press, Cambridge (2018)