

Linearly Embedding Sparse Vectors from ℓ_2 to ℓ_1 via Deterministic Dimension-Reducing Maps

Simon Foucart

Abstract This note is concerned with deterministic constructions of $m \times N$ matrices satisfying a restricted isometry property from ℓ_2 to ℓ_1 on s -sparse vectors. Similarly to the standard (ℓ_2 to ℓ_2) restricted isometry property, such constructions can be found in the regime $m \asymp s^2$, at least in theory. With effectiveness of implementation in mind, two simple constructions are presented in the less pleasing but still relevant regime $m \asymp s^4$. The first one, executing a Las Vegas strategy, is quasideterministic and applies in the real setting. The second one, exploiting Golomb rulers, is explicit and applies to the complex setting. As a stepping stone, an explicit isometric embedding from $\ell_2^n(\mathbb{C})$ to $\ell_4^{cn^2}(\mathbb{C})$ is presented. Finally, the extension of the problem from sparse vectors to low-rank matrices is raised as an open question.

1 Motivation from Sparse Vector Recovery

Almost twenty years ago [1, 2], the realization that high-dimensional but sparse vectors could be efficiently recovered from far fewer linear measurements than expected created a prolific field of research now known as compressive sensing (or compressed sensing). To be specific, vectors $x \in \mathbb{K}^N$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, are called s -sparse if

$$\|x\|_0 := |\text{supp}(x)| \leq s, \quad \text{where } \text{supp}(x) := \{j \in [1 : N] : x_j \neq 0\}.$$

Such vectors can be recovered from compressive measurements $Ax \in \mathbb{K}^m$ with m being of the order of $s \ln(N/s) \ll N$. One refers to [3] for all the nitty-gritty details. One simply mentions here that, on the one hand, the order $s \ln(N/s)$ cannot be lowered if one requires the recovery to be stable and, on the other hand, that

Simon Foucart
Texas A&M University, College Station
e-mail: foucart@tamu.edu

random measurement matrices $A \in \mathbb{K}^{m \times N}$ with $m \asymp s \ln(N/s)$ fulfill, with high probability, favorable properties that make s -sparse recovery possible. Thus, there is an abundance of matrices suitable for compressive sensing in the optimal regime $m \asymp s \ln(N/s)$, but somehow the mathematical community is unable to pinpoint a single one!

The most popular favorable property—the restricted isometry property (RIP), introduced in [4]— stipulates that the matrix $A \in \mathbb{K}^{m \times N}$ should satisfy

$$(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad \text{for all } s\text{-sparse } x \in \mathbb{K}^N. \quad (1)$$

This standard version of the RIP is ubiquitous in ensuring the success of sparse recovery via a variety of reconstruction algorithms, such as ℓ_1 -minimization (aka basis pursuit), orthogonal matching pursuit (OMP), compressive sampling matching pursuit (CoSaMP), iterative hard thresholding (IHT), hard thresholding pursuit (HTP), to name but a few. It can be interpreted as saying that the linear map $x \mapsto Ax$ provides an embedding from ℓ_2^N to ℓ_2^m with a distortion on s -sparse vectors equal to $\gamma = (1 + \delta)/(1 - \delta) \geq 1$. The latter can be made arbitrarily close to one by taking $\delta > 0$ small enough. However, the distortion need not be close to one to enable sparse recovery: a requirement $\gamma \leq \gamma^*$ for a fixed threshold $\gamma^* \geq 1$ is enough. Likewise, the embedding need not map into ℓ_2^m : any ℓ_p^m with $0 < p \leq 2$ will be convenient. This note concentrates on the case $p = 1$ —interestingly, this version appeared in [2]. Thus, instead of the standard version of the RIP, i.e., (1), one considers an ‘ ℓ_2 to ℓ_1 ’ RIP stipulating that $A \in \mathbb{K}^{m \times N}$ should satisfy

$$\alpha\|x\|_2 \leq \|Ax\|_1 \leq \beta\|x\|_2 \quad \text{for all } s\text{-sparse } x \in \mathbb{K}^N \quad (2)$$

with distortion on s -sparse vectors bounded by a fixed threshold, say $\gamma = \beta/\alpha \leq \gamma_*$. I am an advocate of this alternative version, for several reasons:

- the theory of sparse recovery can be built from (2), and not only for basis pursuit as presented in [5, Chapter 14], but also for iterative hard thresholding, see [6];
- the theory of one-bit compressive sensing can be built from (2) as well, as presented [5, Chapter 17], so long as the distortion can be made close to one, which is the case if A is a Gaussian matrix or a partial Gaussian circulant matrix [7];
- Laplace matrices (more generally subexponential random matrices) satisfy (2) in the optimal regime $m \asymp s \ln(N/s)$, see [8], while (1) would only hold in the suboptimal regime $m \asymp s \ln^2(N/s)$;
- any ensemble of random matrices $A \in \mathbb{K}^{m \times N}$ yielding (1) with exponentially small failure probability when $m \asymp s \ln(N/s)$ also yields (2) with exponentially small failure probability when $m \asymp s \ln(N/s)$, see [9] for the precise statement.

This last point suggests that the ‘ ℓ_2 to ℓ_1 ’ RIP is easier to fulfill than the standard ‘ ℓ_2 to ℓ_2 ’ RIP, so it is plausible that deterministic constructions of RIP matrices are more accessible through the ‘ ℓ_2 to ℓ_1 ’ avenue. Since deterministic constructions exist for the standard RIP with $m \asymp s^{2-\varepsilon}$ (see next section), the same is expected to hold for the ‘ ℓ_2 to ℓ_1 ’ RIP. This is indeed the case, but through a construction that may be considered inadequate, because it combines two fairly theoretical results that are, in

my view, not as explicit as hoped for. The purpose this note is to exhibit a couple of simple constructions of matrices that fulfill the ‘ ℓ_2 to ℓ_1 ’ RIP in the regime $m \asymp s^4$. This is not the desired regime, for sure, but the advantage here is the simplicity of the constructions. This simplicity is validated by the few-lines `MATLAB` implementation found in the associated reproducible file (available on the author’s webpage).

The rest of this note is organized as follows. Section 2 discusses some known facts about deterministic embeddings. Section 3 presents, in the case $\mathbb{K} = \mathbb{R}$, the first deterministic construction of an embedding of s -sparse vectors from ℓ_2^N to ℓ_1^m with $m \asymp s^4$. Section 4 uncovers, in case $\mathbb{K} = \mathbb{C}$, an explicit isometric embedding from ℓ_2^p to ℓ_4^m with $m \asymp p^2$, p being a prime number. Section 5 exploits this isometric embedding—or rather the argument leading to it—to present, in the case $\mathbb{K} = \mathbb{C}$, a second deterministic, and actually explicit, construction of an embedding of s -sparse vectors from ℓ_2^N to ℓ_1^m with $m \asymp s^4$. Section 6 briefly touches on the mostly uncharted territory of deterministic restricted isometry properties for low-rank matrices. Finally, as an aside, Section 7 recalls a connection between (almost) isometric embeddings from ℓ_2^n to ℓ_{2k}^N and (approximate) spherical designs.

2 Known Deterministic Results

The mathematical community’s incapability to create nonrandom matrices fulfilling the standard RIP in the optimal regime is vexing. Nonrandom procedures are stuck in the quadratic regime $m \asymp s^2$ —in truth, $m \asymp s^2 \text{polylog}(N)$. There are several ways to reach this regime, mostly based on the notion of coherence. The coherence of a matrix $A \in \mathbb{K}^{m \times N}$ with unit ℓ_2 -norm columns $a_1, \dots, a_N \in \mathbb{K}^m$ is defined by

$$\mu(A) := \max_{j \neq \ell} |\langle a_j, a_\ell \rangle|.$$

Indeed, to ensure that $\delta_s(A)$, the smallest $\delta \in (0, 1)$ for which (1) holds, obeys $\delta_s(A) < \delta_*$ as soon as $m \geq Cs^2$, it is sufficient make the coherence small as per

$$\mu(A) \leq \frac{c}{\sqrt{m}}, \quad c := \sqrt{C} \delta_*,$$

by virtue of the inequality $\delta_s(A) < s\mu(A)$ (see e.g. [3, Proposition 6.2]). As examples of matrices with small coherence, let me mention

- an $m \times m^2$ matrix with columns formed by translations and modulations of the Alltop vector ([10], see also [3, Proposition 5.13]): its coherence is $\mu(A) = 1/\sqrt{m}$;
- when p is prime and $d < p$, a $p^2 \times p^{d+1}$ matrix with binary entries in $\{0, 1/\sqrt{p}\}$ ([11], see also [5, Theorem 18.5]): its coherence is $\mu(A) \leq d/p = d/\sqrt{m}$;
- when p is prime and $d < p$, a $p \times p^{d+1}$ matrix with entries

$$A_{k,f} = \frac{1}{\sqrt{p}} \exp\left(i2\pi \frac{kf(k)}{p}\right) \quad (3)$$

indexed by elements k of \mathbb{F}_p and by polynomials f over \mathbb{F}_p of degree at most d : its coherence is $\mu(A) \leq d/\sqrt{p} = d/\sqrt{m}$. The argument is simple but relies on a deep result known as Weil bound (see e.g. [12, Proposition 5.3.8]), which says that, if f is a nonconstant polynomial over \mathbb{F}_p , then

$$\left| \sum_{k \in \mathbb{F}_p} \exp\left(i2\pi \frac{f(k)}{p}\right) \right| \leq (\deg(f) - 1)\sqrt{p}.$$

There is a notable explicit construction that overcomes, albeit ever so slightly, the quadratic barrier. Indeed, the article [13] uncovered an RIP matrix with $m \asymp s^{2-\varepsilon}$ rows, where $\varepsilon > 0$ was tiny. In a practitioner's mind, this is viewed as an issue, but in fact not as the most critical one: one also has $m \geq N^{1-\varepsilon}$, making the matrix almost square and thus defeating the compressive sensing purpose of taking far fewer measurements than the high ambient dimension. A similar issue occurs in [14]: there, conditionally on a folklore conjecture in number theory, it was shown that the quadratic barrier would be overcome by the Paley matrix fulfilling the ' ℓ_2 to ℓ_2 ' RIP, but this matrix has $m = N/2$ rows.

Turning now to the deterministic ' ℓ_2 to ℓ_1 ' RIP, which has not been explored, at least to the best of my knowledge, one can follow the indirect strategy below:

- (i) consider a deterministic linear map $A' : \ell_2^N \rightarrow \ell_2^n$ with constant distortion on s -sparse vectors for $n \asymp s^\eta$;
- (ii) consider a deterministic linear map $A'' : \ell_2^n \rightarrow \ell_1^m$ with constant distortion on arbitrary vectors for $m \asymp n^\theta$.

Then, the deterministic map $A := A'' \circ A' : \ell_2^N \rightarrow \ell_1^m$ has constant distortion on s -sparse vectors for $m \asymp s^{\eta\theta}$. For (i), one can take $\eta = 2 - \varepsilon$ according to [13]. For (ii), according to [15] (or [16], which both rely on [17]), one can take $\theta = 1 + \omega$ for any $\omega > 0$. These choices yields an ' ℓ_2 to ℓ_1 ' RIP in the regime $m \asymp s^{(2-\varepsilon)(1+\omega)}$ for any $\omega > 0$, hence overcoming the quadratic barrier here, too. This is quite compelling until the hands-on stage, because the theoretical results underlying the argument are not easily implementable, despite being deterministic (which often means constructible in polynomial time).

The general strategy is still valid, though, and one can gain in explicitness by giving up on the smallest ν in $m \asymp s^\nu$. This is the spirit of Section 5, which obtains $\nu = 4$ by way of taking $\eta = 2$ in (i) and $\theta = 2$ in (ii). For the first step, one can exploit any of the small-coherence matrices listed above. For the second step, one relies on the existence of deterministic constant-distortion linear embeddings from ℓ_2^n to $\ell_1^{cn^2}$. Such an existence result is stated in [18, Lemma 3.3] by citing [19] for a construction of 4-wise independent families of vectors, which I would qualify as deterministic but not explicit. The article [16] also claims without details that such an existence result can be extracted from [20]. Section 5 will actually provide such an explicit map embedding ℓ_2^n into $\ell_1^{cn^2}$. I suspect this construction to be very close to what should have been extracted from [20]. It will yield an explicit embedding on sparse vectors taking place in the complex setting. In the real setting, the construction presented next in Section 3 is more direct, as it bypasses (i)-(ii), but I would qualify

it as quasideterministic rather than explicit. Both constructions pertain to the regime $m \asymp s^4$. This should definitely be improved to $m \asymp s^\nu$ with $\nu \leq 2$, but for the moment there is comfort in the simplicity of the constructions.

3 Simple embedding of s -sparse vectors from $\ell_2^N(\mathbb{R})$ to $\ell_1^{cs^4}(\mathbb{R})$

The construction proposed in this section is not explicit, but rather (quasi)deterministic, in the sense that it is outputted by the following Las Vegas algorithm:

for $t = 1, 2, \dots$

- draw $A \in \mathbb{R}^{m \times N}$ populated with independent Rademacher random variables,
- check if conditions (a)-(b) of Theorem 1 below are satisfied for $\kappa = \sqrt{8 \ln(N)}$,

and if they are, then return A .

The success of this procedure is guaranteed by the fact that conditions (a)-(b) certify that A yields an embedding on s -sparse vectors from ℓ_2^N to ℓ_1^m (by Theorem 1), together with the fact that each draw of A satisfies (a)-(b) for $\kappa = \sqrt{8 \ln(N)}$ with failure probability at most $1/3$, so the failure probability after T independent rounds is at most $(1/3)^T$. This vanishing quantity is never exactly zero, hence the procedure cannot technically be qualified as deterministic, but quasideterministic sounds like a suitable designation. Note that the verification of (a)-(b) can be performed in polynomial time, precisely in $O(N^4 m)$ multiplications.

Theorem 1 *Let $A \in \mathbb{R}^{m \times N}$ be populated with entries $A_{j,k} = \pm 1$. Assume that*

- (a) $\left| \sum_{j=1}^m A_{j,k} A_{j,k'} \right| \leq \kappa \sqrt{m}$ for all distinct $k, k' \in [1 : N]$,
- (b) $\left| \sum_{j=1}^m A_{j,k} A_{j,k'} A_{j,\ell} A_{j,\ell'} \right| \leq \kappa \sqrt{m}$ for all distinct $k, k', \ell, \ell' \in [1 : N]$.

Then, for any $\delta \in (0, 1)$, the linear map $A : \ell_2^N \rightarrow \ell_1^m$ has the property that

$$\alpha m \|x\|_2 \leq \|Ax\|_1 \leq \beta m \|x\|_2 \quad \text{for all } s\text{-sparse } x \in \mathbb{R}^N,$$

with distortion $\gamma = \beta/\alpha \leq \sqrt{3}((1+\delta)/(1-\delta))^{3/2}$ as soon as $m \geq \kappa^2 \delta^{-2} s^4$.

Before justifying this theorem, it is worth observing that its assumptions (a)-(b) are indeed fulfilled when $A \in \mathbb{R}^{m \times N}$ is a Rademacher random matrix and $\kappa = \sqrt{8 \ln(N)}$. Fixing distinct k, k' in $[1 : N]$, the $A_{j,k} A_{j,k'}$, $j \in [1 : m]$, are independent Rademacher variables, so by Hoeffding inequality (see e.g. [3, Corollary 8.8])

$$\mathbb{P} \left[\left| \sum_{i=1}^m A_{i,j_1} A_{i,j_2} \right| \geq \kappa \sqrt{m} \right] \leq 2 \exp \left(-\frac{\kappa^2}{2} \right).$$

Hence, by a union bound, one derives

$$\mathbb{P}[(\text{a}) \text{ fails}] \leq \binom{N}{2} 2 \exp\left(-\frac{\kappa^2}{2}\right) \leq N^2 \exp\left(-\frac{\kappa^2}{2}\right) = \frac{1}{N^2}.$$

Likewise, fixing distinct $k, k', \ell, \ell' \in [1 : N]$, the $A_{j,k} A_{j,k'} A_{j,\ell} A_{j,\ell'}$, $j \in [1 : m]$, are independent Rademacher variables, so Hoeffding inequality followed by a union bound once again yields

$$\mathbb{P}[(\text{b}) \text{ fails}] \leq \binom{N}{4} 2 \exp\left(-\frac{\kappa^2}{2}\right) \leq \frac{N^4}{12} \exp\left(-\frac{\kappa^2}{2}\right) = \frac{1}{12}.$$

Consequently, (a) and (b) are indeed both fulfilled with failure probability bounded, for $N \geq 2$, as

$$\mathbb{P}[(\text{a}) \text{ or } (\text{b}) \text{ fail}] \leq \frac{1}{N^2} + \frac{1}{12} \leq \frac{1}{3}.$$

Before turning the attention to the proof of Theorem 1, it is also worth isolating two key ingredients as separate lemmas, one being a technical calculation to be reused later and the other one being a way to estimate the ℓ_1 -norm from below via the ℓ_2 - and ℓ_4 -norms.

Lemma 1 *Let $B \in \mathbb{C}^{q \times r}$ with $|B_{j,k}| = 1$ for all $j \in [1 : q]$ and $k \in [1 : r]$. Then, for any $x \in \mathbb{C}^r$, one has*

$$\|Bx\|_2^2 = q\|x\|_2^2 + \sum_{1 \leq k \neq k' \leq r} \left(\sum_{j=1}^q \overline{B_{j,k}} B_{j,k'} \right) \overline{x_k} x_{k'}, \quad (4)$$

$$\|Bx\|_4^4 = 2\|x\|_2^2 \|Bx\|_2^2 - q\|x\|_4^4 + \Sigma_1 \quad (5)$$

$$= 2\|x\|_2^2 \|Bx\|_2^2 - q\|x\|_4^4 + \sum_{k \neq k'} \left(\sum_{j=1}^q \overline{B_{j,k}}^2 B_{j,k'}^2 \right) \overline{x_k}^2 x_{k'}^2 + \Sigma_2, \quad (6)$$

where the quantities Σ_1 and Σ_2 are given by

$$\Sigma_1 := \sum_{\substack{(k \neq k') \neq (\ell \neq \ell')}} \left(\sum_{j=1}^q \overline{B_{j,k}} B_{j,k'} B_{j,\ell} \overline{B_{j,\ell'}} \right) \overline{x_k} x_{k'} x_{\ell} \overline{x_{\ell'}},$$

$$\Sigma_2 := \sum_{\substack{(k \neq k') \neq (\ell \neq \ell') \\ (k \neq k') \neq (\ell' \neq \ell)}} \left(\sum_{j=1}^q \overline{B_{j,k}} B_{j,k'} B_{j,\ell} \overline{B_{j,\ell'}} \right) \overline{x_k} x_{k'} x_{\ell} \overline{x_{\ell'}}.$$

Proof For the first identity, one writes

$$\begin{aligned}
\|Bx\|_2^2 &= \sum_{j=1}^q \overline{(Bx)_j} (Bx)_j = \sum_{j=1}^q \left(\sum_{k=1}^r \overline{B_{j,k} \bar{x}_k} \right) \left(\sum_{k'=1}^r B_{j,k'} x_{k'} \right) \\
&= \sum_{j=1}^q \left(\sum_{k=1}^r \overline{B_{j,k} \bar{x}_k} B_{j,k} x_k + \sum_{k \neq k'} \overline{B_{j,k} \bar{x}_k} B_{j,k'} x_{k'} \right) \\
&= \sum_{j=1}^q \sum_{k=1}^r |B_{j,k}|^2 |x_k|^2 + \sum_{j=1}^q \sum_{k \neq k'} \overline{B_{j,k} \bar{x}_k} B_{j,k'} x_{k'},
\end{aligned}$$

which, by virtue of $|B_{j,k}| = 1$, reduces to (4) after changing the order of summation.

For the second identity, one writes

$$\begin{aligned}
\|Bx\|_4^4 &= \sum_{j=1}^q \left[\overline{(Bx)_j} (Bx)_j \right]^2 = \sum_{j=1}^q \left[\left(\sum_{k=1}^r \overline{B_{j,k} \bar{x}_k} \right) \left(\sum_{k'=1}^r B_{j,k'} x_{k'} \right) \right]^2 \\
&= \sum_{j=1}^q \left[\sum_{k=1}^r |B_{j,k}|^2 |x_k|^2 + S_j \right]^2 \quad \text{where } S_j := \sum_{k \neq k'} \overline{B_{j,k} \bar{x}_k} B_{j,k'} x_{k'} \\
&= \sum_{j=1}^q \left[\|x\|_2^2 + S_j \right]^2 = q \|x\|_2^4 + 2 \|x\|_2^2 \sum_{j=1}^q S_j + \sum_{j=1}^q S_j^2. \tag{7}
\end{aligned}$$

Taking (4) into account, one notices that

$$\sum_{j=1}^q S_j = \|Bx\|_2^2 - q \|x\|_2^2. \tag{8}$$

Next, exploiting the fact that $S_j \in \mathbb{R}$ and separating the cases $(k \neq k') = (\ell \neq \ell')$ and $(k \neq k') \neq (\ell \neq \ell')$, one obtains

$$\begin{aligned}
S_j^2 &= S_j \overline{S_j} \\
&= \sum_{k \neq k'} |B_{j,k}|^2 |B_{j,k'}|^2 |x_k|^2 |x_{k'}|^2 + \sum_{(k \neq k') \neq (\ell \neq \ell')} \overline{B_{j,k} \bar{x}_k} B_{j,k'} x_{k'} \overline{B_{j,\ell} \bar{x}_\ell} B_{j,\ell} x_\ell \overline{x_{\ell'}},
\end{aligned}$$

from where it follows that

$$\sum_{j=1}^q S_j^2 = \sum_{j=1}^q \left(\sum_{k,k'} |x_k|^2 |x_{k'}|^2 - \sum_k |x_k|^4 \right) + \Sigma_1 = q \left(\|x\|_2^4 - \|x\|_4^4 \right) + \Sigma_1. \tag{9}$$

It remains to substitute (8) and (9) into (7) to arrive at (5).

For the last identity, one simply writes

$$\begin{aligned}\Sigma_1 &= \left\{ \sum_{\substack{(k \neq k') \neq (\ell \neq \ell') \\ (k \neq k') = (\ell' \neq \ell)}} + \sum_{\substack{(k \neq k') \neq (\ell \neq \ell') \\ (k \neq k') \neq (\ell' \neq \ell)}} \right\} \left(\sum_{j=1}^q \overline{B_{j,k} B_{j,k'} B_{j,\ell} B_{j,\ell'}} \right) \overline{x_k x_{k'} x_\ell x_{\ell'}} \\ &= \sum_{k \neq k'} \left(\sum_{j=1}^q \overline{B_{j,k}^2 B_{j,k'}^2} \right) \overline{x_k^2 x_{k'}^2} + \Sigma_2,\end{aligned}$$

and substituting the latter into (5) directly leads to (6). \square

As for the lower estimate of the ℓ_1 -norm, this is achieved via an upper estimate of the ℓ_4 -norm, as stated below.

Lemma 2 For any $y \in \mathbb{C}^m$,

$$\|y\|_1 \geq \frac{\|y\|_2^3}{\|y\|_4^2}.$$

Proof This is Hölder inequality in disguise. Namely, one can easily rearrange

$$\|y\|_2^2 = \sum_{i=1}^m |y_i|^{1 \times \frac{2}{3} + 4 \times \frac{1}{3}} \leq \left(\sum_{i=1}^m |y_i|^1 \right)^{\frac{2}{3}} \left(\sum_{i=1}^m |y_i|^4 \right)^{\frac{1}{3}} = \|y\|_1^{\frac{2}{3}} \|y\|_4^{\frac{4}{3}}$$

into the announced inequality. \square

One is now ready to justify the main result of this section.

Proof (of Theorem 1) Let an s -sparse vector $x \in \mathbb{R}^N$ be fixed throughout the proof. Applying Lemma 1 with the real matrix $A \in \mathbb{R}^{m \times N}$ taking the role of the complex matrix $B \in \mathbb{C}^{q \times r}$, one first observes, from (4), that

$$\|Ax\|_2^2 - m\|x\|_2^2 = \sum_{k \neq k'} \left(\sum_{j=1}^m A_{j,k} A_{j,k'} \right) x_k x_{k'}.$$

Taking the bound (a) into consideration, as well as the sparsity of x , yields

$$\left| \|Ax\|_2^2 - m\|x\|_2^2 \right| \leq \sum_{k \neq k'} \kappa \sqrt{m} |x_k| |x_{k'}| \leq \kappa \sqrt{m} \|x\|_1^2 \leq \frac{\kappa s}{\sqrt{m}} m \|x\|_2^2.$$

For $m \geq \kappa^2 \delta^{-2} s^2$ (which is the case since $m \geq \kappa^2 \delta^{-2} s^4$), this implies

$$(1 - \delta) m \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta) m \|x\|_2^2. \quad (10)$$

As a side note, this is the standard RIP for the renormalized matrix A/\sqrt{m} . It was derived solely from (a), which is nothing but a coherence assumption for this matrix. Now, using the standard comparison of the ℓ_1 -norm and ℓ_2 -norm, one arrives at

$$\|Ax\|_1 \leq \sqrt{m} \|Ax\|_2 \leq (1 + \delta)^{1/2} m \|x\|_2. \quad (11)$$

Next, using (6) in Lemma 1 while taking into account that $(k \neq k') \neq (\ell \neq \ell')$ and $(k \neq k') \neq (\ell' \neq \ell)$ means that k, k', ℓ, ℓ' are all distinct, and using assumption (b) as well, one can write

$$\begin{aligned} \|Ax\|_4^4 &\leq 2\|x\|_2^2\|Ax\|_2^2 + m\|x\|_2^4 + \sum_{\substack{k,k',\ell,\ell' \\ \text{all distinct}}} \left| \sum_{j=1}^m A_{j,k}A_{j,k'}A_{j,\ell}A_{j,\ell'} \right| |x_k||x_{k'}||x_\ell||x_{\ell'}| \\ &\leq (2(1+\delta) + 1)m\|x\|_2^4 + \kappa\sqrt{m}\|x\|_1^4 \leq (3+2\delta)m\|x\|_2^4 + \kappa\sqrt{m}s^2\|x\|_2^4 \\ &= \left(3+2\delta + \frac{\kappa s^2}{\sqrt{m}}\right)m\|x\|_2^4 \leq (3+3\delta)m\|x\|_2^4, \end{aligned}$$

where the last step exploited $m \geq \kappa^2\delta^{-2}s^4$. This upper bound on $\|Ax\|_4$, combined with the lower bound on $\|Ax\|_2$ from (10), ensures, according to Lemma 2, that

$$\|Ax\|_1 \geq \frac{(\|Ax\|_2^2)^{3/2}}{(\|Ax\|_4^4)^{1/2}} \geq \frac{((1-\delta)m\|x\|_2^2)^{3/2}}{(3(1+\delta)m\|x\|_2^4)^{1/2}} = \left(\frac{(1-\delta)^3}{3(1+\delta)}\right)^{1/2} m\|x\|_2. \quad (12)$$

The estimates (11) and (12) together establish the result, noting in particular the expression of the distortion. \square

Remark 1 It is unclear if the above argument can be refined to improve the exponent $\nu = 4$ for the regime $m \asymp s^\nu$ where the ' ℓ_2 to ℓ_1 ' RIP provably holds. One point is certain, however: for a matrix $A \in \mathbb{K}^{m \times N}$ whose entries all have modulus/absolute value equal to one, one cannot beat $\nu = 2$ if one estimates the ℓ_1 -norm from below via an upper bound on the ℓ_4 -norm of the form $\|Ax\|_4^4 \leq Cm\|x\|_2^4$ for all s -sparse vectors $x \in \mathbb{K}^N$. Indeed, for a fixed $i \in [1 : m]$, if $x = A_{S,i}$ represents the i th rows of A restricted to a set $S \subseteq [1 : N]$ of size s , then

$$Cms^2 = Cm\|x\|_2^4 \geq \|Ax\|_4^4 = \sum_{j=1}^m |\langle A_{\cdot,j}, A_{S,i} \rangle|^4 \geq |\langle A_{\cdot,i}, A_{S,i} \rangle|^4 = s^4,$$

which forces $m \geq C^{-1}s^2$.

4 Explicit isometric embedding from $\ell_2^p(\mathbb{C})$ into $\ell_4^{cp^2}(\mathbb{C})$

As mentioned earlier, deterministic linear embeddings from ℓ_2^n to $\ell_1^{cn^2}$ with constant distortion have been claimed to exist in different places without full details, e.g. stating that a construction can be extracted from [20]. In the latter, a central role was played by Sidon sets, which are the same as Golomb rulers. In what follows, I propose a Golomb-ruler argument which likely coincides with what should have been extracted from [20]. Taking a detour via the ℓ_4 -norm, the argument actually uncovers a deterministic embedding from ℓ_2^p (p being prime) to $\ell_4^{cp^2}$ whose distortion

is exactly equal to one. This isometric embedding is singled out in this section. The main ingredient, due to [21], is reproduced here for completeness.

Lemma 3 *For a prime number $p \geq 3$, the integers*

$$g(k) = 2pk + (k^2)_p, \quad k \in [0 : p - 1],$$

where $(k^2)_p$ denotes the integer $t \in [0 : p - 1]$ such that $k^2 \equiv t \pmod{p}$, form a Golomb ruler, in the sense that

$$g(k) - g(k') \neq g(\ell) - g(\ell') \quad \text{whenever } (k \neq k') \neq (\ell \neq \ell').$$

Proof Let $(k \neq k') \neq (\ell \neq \ell')$ and suppose $g(k) - g(k') = g(\ell) - g(\ell')$. Writing $k' = k + \sigma$ and $\ell' = \ell + \tau$ for some $\sigma, \tau \in [-p + 1 : p - 1] \setminus \{0\}$, this reads $2pk + (k^2)_p - 2p(k + \sigma) - ((k + \sigma)^2)_p = 2p\ell + (\ell^2)_p - 2p(\ell + \tau) - ((\ell + \tau)^2)_p$, i.e.,

$$(k^2)_p - ((k + \sigma)^2)_p - (\ell^2)_p + ((\ell + \tau)^2)_p = 2p(\sigma - \tau). \quad (13)$$

Since the left-hand side, in absolute value, is at most $2(p - 1) < p$, the right-hand side must be zero, so that $\sigma = \tau$. Taking $((k + \sigma)^2)_p \equiv (k^2)_p + (2\sigma k)_p + (\sigma^2)_p \pmod{p}$ into account, as well as $((\ell + \tau)^2)_p \equiv ((\ell + \sigma)^2)_p \equiv (\ell^2)_p + (2\sigma\ell)_p + (\sigma^2)_p \pmod{p}$, looking at (13) modulo p yields

$$(2\sigma\ell)_p - (2\sigma k)_p \equiv 0 \pmod{p}, \quad \text{i.e.,} \quad 2\sigma(\ell - k) \equiv 0 \pmod{p}.$$

But since $2 \not\equiv 0$ and $\sigma \not\equiv 0 \pmod{p}$, one deduces that $\ell - k \equiv 0 \pmod{p}$ and hence that $k = \ell$. In view of $k' = k + \sigma$ and $\ell' = \ell + \tau$ with $\sigma = \tau$, it follows that $(k \neq k') = (\ell \neq \ell')$. This leads to a contradiction, showing that g indeed generates a Golomb ruler. \square

The coveted isometric embedding from ℓ_2^p to $\ell_4^{c p^2}$ will be obtained by combining Lemma 1 with Lemma 3. Before that, it is worth pausing to remark that g maps $[0 : p - 1]$ into $[0 : q - 1]$, where $q = 3p(p - 1) + 1$ is quadratic in p , and that such a quadratic order is optimal. Indeed, for any Golomb ruler g from $[0 : p - 1]$ into $[0 : q - 1]$, all the distinct $g(k) - g(k')$ indexed by ordered pairs $(k \neq k')$ are contained in $[-q + 1 : q - 1]$, and as such $2q - 1 \geq p(p - 1)$.

Theorem 2 *For a prime number $p \geq 3$, let $m = 6p^2 - 6p + 1$. Consider the matrix $A'' \in \mathbb{C}^{m \times p}$ with entries*

$$A''_{j,k} = \exp\left(i2\pi \frac{jg(k)}{m}\right), \quad j \in [0 : m - 1], \quad k \in [0 : p - 1]. \quad (14)$$

Then the matrix $M \in \mathbb{C}^{(m+p) \times p}$ defined as

$$M := \begin{bmatrix} \frac{1}{(2m)^{1/4}} A'' \\ \frac{1}{2^{1/4}} I_p \end{bmatrix} \quad (15)$$

provides an isometric embedding from $\ell_2^P(\mathbb{C})$ into $\ell_2^{m+P}(\mathbb{C})$, i.e.,

$$\|Mx\|_4 = \|x\|_2 \quad \text{for all } x \in \mathbb{C}^P.$$

Proof Let a vector $x \in \mathbb{C}^P$ be fixed throughout the proof. For any $k \neq k'$, one observes that

$$\sum_{j=0}^{m-1} \overline{A''_{j,k}} A''_{j,k'} = \sum_{j=0}^{m-1} \exp\left(-i2\pi \frac{j(g(k) - g(k'))}{m}\right) = 0,$$

owing to $g(k) - g(k') \in [-q+1 : q-1] \subseteq [-m+1 : m-1]$ being nonzero (otherwise, if $g(k) = g(k')$, choosing $k'' \notin \{k, k'\}$ would yield $g(k) - g(k'') = g(k') - g(k'')$). According to Lemma 1, and specifically to (4), one therefore has

$$\|A''x\|_2^2 = m\|x\|_2^2.$$

Next, for any $(k \neq k') \neq (\ell \neq \ell')$, one observes that

$$\sum_{j=0}^{m-1} \overline{A''_{j,k}} A''_{j,k'} \overline{A''_{j,\ell}} A''_{j,\ell'} = \exp\left(-i2\pi \frac{j((g(k) - g(k')) - (g(\ell) - g(\ell')))}{m}\right) = 0,$$

owing to $(g(k) - g(k')) - (g(\ell) - g(\ell')) \in [-2q+2 : 2q-2] = [-m+1 : m-1]$ being nonzero. According to Lemma 1 again, and specifically to (5), one obtains

$$\|A''x\|_4^4 = 2\|x\|_2^2 \|A''x\|_2^2 - m\|x\|_4^4 = 2m\|x\|_2^4 - m\|x\|_4^4.$$

From here, it easily follows that

$$\|Mx\|_4^4 = \frac{1}{2m} \|A''x\|_4^4 + \frac{1}{2} \|x\|_4^4 = \|x\|_2^4,$$

which is the desired result. \square

Remark 2 Isometric embeddings from ℓ_2 to ℓ_4 can alternatively be viewed through the lenses of spherical designs and of tensors. In order not to be diverted from the main goal, this connection will be brought forward much later, in Section 7.

5 Explicit embedding of s -sparse vectors from $\ell_2^N(\mathbb{C})$ to $\ell_1^{cs^4}(\mathbb{C})$

Based on Theorem 2 in the previous section, a linear embedding from $\ell_2^P(\mathbb{C})$ to $\ell_1^{cP^2}(\mathbb{C})$ with constant distortion can easily be generated. However, instead of using the matrix M from (15), which provided an isometric embedding from $\ell_2^P(\mathbb{C})$ to $\ell_4^{cP^2}(\mathbb{C})$, the matrix A'' from (14) is preferred, as it leads to a nicer expression for the distortion. Recall that the proof of Theorem 2 revealed that, for any $x \in \mathbb{C}^P$,

$$\|A''x\|_2 = \sqrt{m}\|x\|_2 \quad \text{and} \quad \|A''x\|_4 \leq (2m)^{1/4}\|x\|_2.$$

Theorem 3 For a prime number $p \geq 3$, let $m = 6p^2 - 6p + 1$. The matrix $A'' \in \mathbb{C}^{m \times p}$ defined in (14) provides an embedding from $\ell_2^p(\mathbb{C})$ into $\ell_1^m(\mathbb{C})$ with distortion at most $\sqrt{2}$. Precisely, one has

$$\frac{m}{\sqrt{2}}\|x\|_2 \leq \|A''x\|_1 \leq m\|x\|_2 \quad \text{for all } x \in \mathbb{C}^p.$$

Proof Let a vector $x \in \mathbb{C}^p$ be fixed throughout this short proof. On the one hand, comparing ℓ_1 - and ℓ_2 -norms yields

$$\|A''x\|_1 \leq \sqrt{m}\|A''x\|_2 = m\|x\|_2.$$

On the other hand, according to Lemma 2, one has

$$\|A''x\|_1 \geq \frac{\|A''x\|_2^3}{\|A''x\|_4^2} \geq \frac{\sqrt{m}^3\|x\|_2^3}{\sqrt{2m}\|x\|_2^2} = \frac{m}{\sqrt{2}}\|x\|_2.$$

These two inequalities together justify the announced embedding. \square

It is now time for the main result of this section, namely the awaited explicit linear embedding of s -sparse vectors from $\ell_2^N(\mathbb{C})$ to $\ell_1^{cs^4}(\mathbb{C})$. It repeats the general strategy (i)-(ii) outlined in Section 2.

Theorem 4 Given integers $N, s \geq 1$ with $N \gg s^4$, let $p \geq 3$ be a prime number between $9s^2 \lceil \ln^2(N) \rceil$ and $18s^2 \lceil \ln^2(N) \rceil$ and let $m = 6p^2 - 6p + 1 \asymp s^4 \ln^2(N)$. Then the matrix $A \in \mathbb{C}^{m \times N}$ indexed by $[0 : m - 1]$ and by an arbitrary N -set \mathcal{F} of polynomials over \mathbb{F}_p of degree at most $d = \lceil \ln(N/p) / \ln(p) \rceil$ and with entries

$$A_{k,f} = \frac{1}{\sqrt{p}} \sum_{k=0}^{p-1} \exp \left(i2\pi \left(j \frac{2pk + (k^2)_p}{m} + \frac{kf(k)}{p} \right) \right), \quad k \in [0 : m - 1], f \in \mathcal{F}, \quad (16)$$

provides an explicit embedding from $\ell_2^N(\mathbb{C})$ into $\ell_1^m(\mathbb{C})$ with distortion on s -sparse vectors at most 2, namely

$$\frac{m}{\sqrt{3}}\|x\|_2 \leq \|Ax\|_1 \leq \frac{2m}{\sqrt{3}}\|x\|_2 \quad \text{for all } s\text{-sparse } x \in \mathbb{C}^N. \quad (17)$$

Proof Picking a prime number p between $9s^2 \lceil \ln^2(N) \rceil$ and $18s^2 \lceil \ln^2(N) \rceil$ is possible by Bertrand postulate. One considers the smallest integer d such that $p^{d+1} \geq N$, i.e., $d = \lceil \ln(N/p) / \ln(p) \rceil$. Note that $1 < p < d$ since $p^2 \lesssim s^4 \text{polylog}(N) < N$ and $p^p \geq e^p \geq e^{\ln(N)} = N$. From $p^d < N$, one also deduces that $d \leq d \ln(p) < \ln(N)$. Let then $A' \in \mathbb{C}^{p \times p^{d+1}}$ be the ‘Weil’ matrix with entries defined in (3), recalling that $\delta_s(A') < s\mu(A) < sd/\sqrt{p} < s \ln(N) / (3s \ln(N)) = 1/3$, so that

$$\|A'x\|_2^2 \begin{cases} \leq (1 + \delta_s(A'))\|x\|_2^2 \leq \frac{4}{3}\|x\|_2^2 \\ \geq (1 - \delta_s(A'))\|x\|_2^2 \geq \frac{2}{3}\|x\|_2^2 \end{cases} \quad \text{for all } s\text{-sparse } x \in \mathbb{C}^{p^{d+1}}.$$

Let also $A'' \in \mathbb{C}^{m \times p}$ be the ‘Golomb’ matrix with entries defined in (14), recalling that

$$\|A''x\|_1 \begin{cases} \leq m\|x\|_2 \\ \geq \frac{m}{\sqrt{2}}\|x\|_2 \end{cases} \quad \text{for all } x \in \mathbb{C}^p.$$

As a result, for the matrix $\tilde{A} := A'' \times A' \in \mathbb{C}^{m \times p^{d+1}}$ and for any s -sparse $x \in \mathbb{C}^{p^{d+1}}$,

$$\|\tilde{A}x\|_1 = \|A''(A'x)\|_1 \begin{cases} \leq m\|A'x\|_2 \leq m\sqrt{\frac{4}{3}}\|x\|_2 = \frac{2m}{\sqrt{3}}\|x\|_2, \\ \geq \frac{m}{\sqrt{2}}\|A'x\|_2 \geq \frac{m}{\sqrt{2}}\sqrt{\frac{2}{3}}\|x\|_2 = \frac{m}{\sqrt{3}}\|x\|_2. \end{cases}$$

This strongly resembles the desired inequalities (17), except that the matrix \tilde{A} has p^{d+1} columns, while the matrix A should have N columns. But one can simply remove $p^{d+1} - N$ arbitrary columns from $\tilde{A} \in \mathbb{C}^{m \times p^{d+1}}$ to create a matrix $A \in \mathbb{C}^{m \times N}$ satisfying (17). The entries of this matrix, indexed by j in $[0 : m - 1]$ and by f in a set \mathcal{F} of polynomials over \mathbb{F}_p of degree at most d with size N , are given by

$$A_{j,f} = \sum_{k=0}^{p-1} A''_{j,k} A'_{k,f} = \sum_{k=0}^{p-1} \exp\left(i2\pi \frac{jg(k)}{m}\right) \frac{1}{\sqrt{p}} \exp\left(i2\pi \frac{kf(k)}{p}\right),$$

which reduces to the expression announced in (16). \square

6 Outlook into low-rank recovery

The theory of compressive sensing also deals with objects of nominally high but intrinsically low dimension beyond s -sparse vectors $x \in \mathbb{K}^N$, prototypically with matrices $X \in \mathbb{K}^{n \times n}$ of rank at most r . In this scenario, too, one can recover such objects from their compressive measurements $\mathcal{A}(X) \in \mathbb{K}^m$ with m being of the order of $nr \ll n^2$. This coup can be achieved (see [22]) by nuclear norm minimization when the linear map $\mathcal{A} : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^m$ satisfies an RIP of the form

$$(1 - \delta)\|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1 + \delta)\|X\|_F^2 \quad \text{for all rank-}r \text{ } X \in \mathbb{K}^{n \times n}. \quad (18)$$

For $m \asymp nr$, this RIP is fulfilled when $A_1, \dots, A_m \in \mathbb{K}^{n \times n}$ in $\mathcal{A}(Z)_i = \langle A_i, X \rangle_F$ are independent random matrices populated with independent properly normalized

gaussian entries, see [23]. But, as in the vector case, no deterministic linear map $\mathcal{A} : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^m$ is known in this optimal regime $m \asymp nr$. Worst, I am not aware of simple deterministic constructions overcoming the trivial regime $m \asymp n^2$ and I have not seen an effective analog of the notion of coherence.

Interestingly, the theory of low-rank recovery can also be built from a modification of the RIP featuring the ℓ_1 -norm as the inner norm, see [5, Chapter 16], namely from

$$(1 - \delta)\|X\|_F \leq \|\mathcal{A}(X)\|_1 \leq (1 + \delta)\|X\|_F \quad \text{for all rank-}r \text{ } X \in \mathbb{K}^{n \times n}. \quad (19)$$

This alternative version seems even more relevant in the present scenario. Indeed, when $m \asymp mr$, the rank-one measurements

$$\mathcal{A}(Z)_i = \langle b_i, Za_i \rangle = \langle A_i, Z \rangle_F, \quad A_i := b_i a_i^*,$$

with independent properly normalized gaussian vectors $a_1, \dots, a_m, b_1, \dots, b_m \in \mathbb{K}^n$ do not lead to the RIP (18) but to the RIP (19), while the latter still enables low-rank recovery via nuclear norm minimization, see [24]. It also enables low-rank recovery via iterative-thresholding-type algorithms, see [25]. But the possibility of fulfilling this modified RIP in a regime $m \asymp n^\lambda r^\mu \ll n^2$ with deterministic matrices $A_1, \dots, A_m \in \mathbb{K}^{n \times n}$ —of rank one or even unrestricted—is a wide open question. This question is hereby set as a challenge to the readers.

7 Addendum: isometric embeddings, spherical designs, tensors

As mentioned in Remark 2, isometric embedding from $\ell_2^N(\mathbb{K})$ to $\ell_{2k}(\mathbb{K})$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, have connections with spherical designs and tensors. This is made precise by the following result, found in [26] for the case $\mathbb{K} = \mathbb{R}$.

Theorem 5 *Let $2k \geq 2$ be an even integer. The following properties are equivalent:*

1) *There exists $A \in \mathbb{K}^{N \times n}$ providing an isometric embedding from ℓ_2^N into ℓ_{2k}^n , i.e.,*

$$\|Ax\|_{2k} = \|x\|_2 \quad \text{for all } x \in \mathbb{K}^n;$$

2) *There exist $x_1, \dots, x_N \in S_2^n$ and $\tau_1, \dots, \tau_N \geq 0$ with $\tau_1 + \dots + \tau_N = 1$ such that*

$$\sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} = \int_{S_2^n \times S_2^n} |\langle x, y \rangle|^{2k} d\sigma(x) d\sigma(y),$$

where σ is the normalized standard measure on the unit sphere S_2^n of ℓ_2^n ;

3) *There exist $x_1, \dots, x_N \in S_2^n$ and $\tau_1, \dots, \tau_N \geq 0$ with $\tau_1 + \dots + \tau_N = 1$ such that*

$$\sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) = \int_{S_2^n} \otimes^k (x \otimes \bar{x}) d\sigma(x).$$

Since RIPs are not genuine isometric embeddings, but almost isometric ones, this result will be established in a slightly stronger form for the sake of completeness. Towards this end, some pieces of notations and some identities are brought forth as a preamble. First, one considers the quantity δ (independent of $x \in S_2^n$) defined by

$$\delta = \delta_{n,2k} := \int_{S_2^n} |\langle x, y \rangle|^{2k} d\sigma(y) = \begin{cases} \frac{(2k-1) \cdots 3 \cdot 1}{(n+2k-2) \cdots (n+2) \cdot n} & \text{for } \mathbb{K} = \mathbb{R}, \\ \frac{k!}{(n+k-1) \cdots (n+1) \cdot n} & \text{for } \mathbb{K} = \mathbb{C}, \end{cases}$$

whose numerical value is given in [26] for $\mathbb{K} = \mathbb{R}$ and in [27] for $\mathbb{K} = \mathbb{C}$. Note that δ coincides with the double integral appearing in 2). As for the integral appearing in 3), called distribution $2k$ -tensor (when $\mathbb{K} = \mathbb{R}$), it shall be denoted by D . Thus,

$$D = D_{n,2k} := \int_{S_2^n} \otimes^k (y \otimes \bar{y}) d\sigma(y).$$

Using the standard notion of inner product on tensor spaces, one verifies below that, for any $x \in S_2^n$,

$$\langle \otimes^k (x \otimes \bar{x}), D \rangle = \delta \quad \text{and} \quad \langle D, D \rangle = \delta. \quad (20)$$

To justify the leftmost identity of (20), it suffices to write

$$\begin{aligned} \langle \otimes^k (x \otimes \bar{x}), D \rangle &= \left\langle \otimes^k (x \otimes \bar{x}), \int_{S_2^n} \otimes^k (y \otimes \bar{y}) d\sigma(y) \right\rangle \\ &= \int_{S_2^n} \langle \otimes^k (x \otimes \bar{x}), \otimes^k (y \otimes \bar{y}) \rangle d\sigma(y) \\ &= \int_{S_2^n} \langle x \otimes \bar{x}, y \otimes \bar{y} \rangle^k d\sigma(y) = \int_{S_2^n} (|\langle x, y \rangle|^2)^k d\sigma(y) \\ &= \delta. \end{aligned}$$

To justify the rightmost identity of (20), it suffices to write

$$\begin{aligned} \langle D, D \rangle &= \left\langle \int_{S_2^n} \otimes^k (x \otimes \bar{x}) d\sigma(x), \int_{S_2^n} \otimes^k (y \otimes \bar{y}) d\sigma(y) \right\rangle \\ &= \int_{S_2^n} \langle \otimes^k (x \otimes \bar{x}), D \rangle d\sigma(x) = \int_{S_2^n} \delta d\sigma(x) \\ &= \delta. \end{aligned}$$

From here, an important identity follows easily, namely: for any $x_1, \dots, x_N \in S_2^n$ and any $\tau_1, \dots, \tau_N \geq 0$ with $\tau_1 + \dots + \tau_N = 1$,

$$\left\| \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) - D \right\|_2^2 = \sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} - \delta. \quad (21)$$

Indeed, the justification of (21) simply reads

$$\begin{aligned}
& \left\| \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) - D \right\|_2^2 \\
&= \sum_{i,j=1}^N \tau_i \tau_j \langle \otimes^k (x_i \otimes \bar{x}_i), \otimes^k (x_j \otimes \bar{x}_j) \rangle - 2 \operatorname{Re} \left\langle \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i), D \right\rangle + \langle D, D \rangle \\
&= \sum_{i,j=1}^N \tau_i \tau_j \langle x_i \otimes \bar{x}_i, x_j \otimes \bar{x}_j \rangle^k - 2 \sum_{i=1}^N \tau_i \operatorname{Re} \langle \otimes^k (x_i \otimes \bar{x}_i), D \rangle + \langle D, D \rangle \\
&= \sum_{i,j=1}^N \tau_i \tau_j (|\langle x_i, x_j \rangle|^2)^k - 2 \sum_{i=1}^N \tau_i \operatorname{Re}(\delta) + \delta,
\end{aligned}$$

where both identities of (20) were used in the last step. Taking $\sum_{i=1}^N \tau_i = 1$ into account then leads to the desired identity (21). Note that it implies the so-called Sidelnikov inequality, namely

$$\sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} \geq \delta \quad \text{whenever } x_1, \dots, x_N \in S_2^n, \tau_1, \dots, \tau_N \geq 0, \sum_{i=1}^N \tau_i = 1.$$

After this preparatory work, one can now state and prove the slight generalization of Theorem 5, which is retrieved as the special case $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 0$.

Theorem 6 *Let $2k \geq 2$ be an even integer. For $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$ that depend on each other, the following properties are equivalent:*

1') *There exists a matrix $A \in \mathbb{K}^{N \times n}$ for which*

$$(1 - \varepsilon_1) \|x\|_2^{2k} \leq \|Ax\|_{2k}^{2k} \leq (1 + \varepsilon_1) \|x\|_2^{2k} \quad \text{for all } x \in \mathbb{K}^n;$$

2') *There exist $x_1, \dots, x_N \in S_2^n$ and $\tau_1, \dots, \tau_N \geq 0$ with $\tau_1 + \dots + \tau_N = 1$ such that*

$$\sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} \leq \delta + \varepsilon_2;$$

3') *There exist $x_1, \dots, x_N \in S_2^n$ and $\tau_1, \dots, \tau_N \geq 0$ with $\tau_1 + \dots + \tau_N = 1$ such that*

$$\left\| \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) - D \right\|_2 \leq \varepsilon_3.$$

Proof 2') \Leftrightarrow 3') with $\varepsilon_3 = \sqrt{\varepsilon_2}$. This equivalence results from identity (21).

3') \Rightarrow 1') with $\varepsilon_1 = \varepsilon_3/\delta$. For any $x \in \mathbb{K}^n$, one observes that the leftmost identity of (20) yields $\langle D, \otimes^k (x \otimes \bar{x}) \rangle = \delta \|x\|_2^{2k}$ by homogeneity. One also observes that

$$\begin{aligned} \left\langle \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i), \otimes^k (x \otimes \bar{x}) \right\rangle &= \sum_{i=1}^N \tau_i \langle \otimes^k (x_i \otimes \bar{x}_i), \otimes^k (x \otimes \bar{x}) \rangle \\ &= \sum_{i=1}^N \tau_i (|\langle x_i, x \rangle|^2)^k. \end{aligned}$$

From both these observations, it follows that

$$\begin{aligned} \left| \sum_{i=1}^N \tau_i |\langle x_i, x \rangle|^{2k} - \delta \|x\|_2^{2k} \right| &= \left| \left\langle \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) - D, \otimes^k (x \otimes \bar{x}) \right\rangle \right| \\ &\leq \left\| \sum_{i=1}^N \tau_i \otimes^k (x_i \otimes \bar{x}_i) - D \right\|_2 \left\| \otimes^k (x \otimes \bar{x}) \right\|_2 \\ &\leq \varepsilon_3 \|x\|_2^{2k}. \end{aligned}$$

Therefore, setting $a_i := (\tau_i/\delta)^{\frac{1}{2k}} x_i$ for each $i \in [1 : N]$, the above becomes

$$\left| \sum_{i=1}^N |\langle a_i, x \rangle|^{2k} - \|x\|_2^{2k} \right| \leq \varepsilon_1 \|x\|_2^{2k},$$

which reduces to 1') for the matrix $A \in \mathbb{K}^{N \times n}$ with rows a_1^*, \dots, a_N^* .

1') \Rightarrow (2') with $\varepsilon_2 = 4\varepsilon_1\delta$ when $\varepsilon_1 \leq 1/2$. With a_1^*, \dots, a_N^* denoting the rows of the matrix $A \in \mathbb{K}^{N \times n}$, the almost isometric embedding takes the form

$$(1 - \varepsilon_1) \|x\|_2^{2k} \leq \sum_{i=1}^N |\langle a_i, x \rangle|^{2k} \leq (1 + \varepsilon_1) \|x\|_2^{2k} \quad \text{for all } x \in \mathbb{K}^n. \quad (22)$$

First, integrating (22) over $x \in S_2^n$ implies that $1 - \varepsilon_1 \leq \sum_{i=1}^N \|a_i\|_2^{2k} \delta \leq 1 + \varepsilon_1$. Setting $S := \sum_{i=1}^N \|a_i\|_2^{2k}$, this means that

$$\frac{\delta}{1 + \varepsilon_1} \leq \frac{1}{S} \leq \frac{\delta}{1 - \varepsilon_1}. \quad (23)$$

Second, setting $x_i := a_i / \|a_i\|_2 \in S_2^n$ for each $i \in [1 : N]$ and selecting $x = x_j$ in (22) yields

$$1 - \varepsilon_1 \leq \sum_{i=1}^N \|a_i\|_2^{2k} |\langle x_i, x_j \rangle|^{2k} \leq 1 + \varepsilon_1.$$

Defining $\tau_i := \|a_i\|_2^{2k} / S$, multiplying the latter by τ_j and summing over $j \in [1 : N]$ leads to

$$\frac{1 - \varepsilon_1}{S} \leq \sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} \leq \frac{1 + \varepsilon_1}{S}. \quad (24)$$

From the estimates (23) and (24), one finally concludes that

$$\sum_{i,j=1}^N \tau_i \tau_j |\langle x_i, x_j \rangle|^{2k} \leq \frac{1 + \varepsilon_1}{1 - \varepsilon_1} \delta \leq (1 + 4\varepsilon_1) \delta,$$

where the last step used $\varepsilon_1 \leq 1/2$. This is the desired inequality with $\varepsilon_2 = 4\varepsilon_1 \delta$. \square

Acknowledgements S. F. is partially supported by grants from the NSF (DMS-2053172) and from the ONR (N00014-20-1-2787).

References

1. Candès, E., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* **52**, 489–509 (2006).
2. Donoho, D.: For most large underdetermined systems of linear equations the minimal ℓ_1 solution is also the sparsest solution. *Communications on Pure and Applied Mathematics* **59**, 797–829 (2006).
3. Foucart, S., Rauhut, H.: *A Mathematical Introduction to Compressive Sensing*. Birkhäuser (2013).
4. Candès, E., Tao, T.: Decoding by linear programming. *IEEE Transactions on Information Theory* **51**, 4203–4215 (2005).
5. Foucart, S.: *Mathematical Pictures at a Data Science Exhibition*. Cambridge University Press (2022).
6. Foucart, S., Lecué, G.: An IHT algorithm for sparse recovery from subexponential measurements. *IEEE Signal Processing Letters* **24**, 1280–1283 (2017).
7. Dirksen, S., Jung, H. C., Rauhut, H.: One-bit compressed sensing with partial Gaussian circulant matrices. *Information and Inference* **9**, 601–626 (2020).
8. Foucart, S., Lai, M.-J.: Sparse recovery with pre-Gaussian random matrices. *Studia Mathematica* **200**, 91–102 (2010).
9. Foucart, S.: The sparsity of LASSO-type minimizers. *Applied and Computational Harmonic Analysis* **62**, 441–452 (2023).
10. Strohmer, T., Heath Jr., R. W.: Grassmannian frames with applications to coding and communication. *Applied and Computational Harmonic Analysis* **14**, 257–275 (2003).
11. DeVore, R. A.: Deterministic constructions of compressed sensing matrices. *Journal of Complexity* **23**, 918–925 (2007).
12. Niederreiter, H., Winterhof, A.: *Applied Number Theory*. Springer (2015).
13. Bourgain, J., Dilworth, S., Ford, K., Konyagin, S., Kutzarova, D.: Explicit constructions of RIP matrices and related problems. *Duke Mathematical Journal* **159**, 145–185 (2011).
14. Bandeira, A. S., Mixon, D. G., Moreira, J.: A conditional construction of restricted isometries. *International Mathematics Research Notices* **2017**, 372–381 (2017).
15. Indyk, P.: Uncertainty principles, extractors, and explicit embeddings of ℓ_2 into ℓ_1 . *Proceedings of the thirty-ninth annual ACM symposium on Theory of Computing* (2007).
16. Guruswami, V., Lee, J. R., Razborov, A.: Almost Euclidean subspaces of ℓ_1^N via expander codes. *Combinatorica* **30**, 47–68 (2010).
17. Guruswami, V., Umans, C., Vadhan, S.: Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Journal of the ACM* **56**, 20 (2009).
18. Linial, N., London, E., and Rabinovich, Y.: The geometry of graphs and some of its algorithmic applications. *Combinatorica* **15**, 215–245 (1995).
19. Berger, B.: The fourth moment method. *SIAM Journal on Computing* **26**, 1188–1207 (1997).
20. Rudin, W.: Trigonometric series with gaps. *Journal of Mathematics and Mechanics* **9**, 203–227 (1960).

21. Erdős, P., Turán, P.: On a problem of Sidon in additive number theory, and on some related problems. *Journal of the London Mathematical Society* **16**, 212–215 (1941).
22. Recht, B., Fazel, M., Parrilo, P. A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review* **52**, 471–501 (2010).
23. Candès, E., Plan, Y.: Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Transactions on Information Theory* **57**, 2342–2359 (2011).
24. Cai, T., Zhang, A.: ROP: Matrix recovery via rank-one projections. *The Annals of Statistics* **43**, 102–138 (2015).
25. Foucart, S., Subramanian, S.: Iterative hard thresholding for low-rank recovery from rank-one projections. *Linear Algebra and its Applications* **572**, 117–134 (2019).
26. Seidel, J. J.: Spherical designs and tensors. In: *Progress in Algebraic Combinatorics* **24**, 309–322. Mathematical Society of Japan, 1996.
27. Kotelina, N. O., Pevnyi, A. B.: Complex spherical semi-designs. *Russian Mathematics* **61**, 46–51 (2017).