NEAR INVARIANCE OF THE HYPERCUBE

SCOTT AARONSON AND HOI NGUYEN

ABSTRACT. We give an almost-complete description of orthogonal matrices M of order n that "rotate a non-negligible fraction of the Boolean hypercube $\mathcal{C}_n = \{-1,1\}^n$ onto itself," in the sense that

$$\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \geq n^{-C}, \text{ for some positive constant } C,$$

where $\mathbf{x}$ is sampled uniformly over $\mathcal{C}_n$. In particular, we show that such matrices M must be very close to products of permutation and reflection matrices. This result is a step toward characterizing those orthogonal and unitary matrices with large permanents, a question with applications to linear-optical quantum computing.

1. INTRODUCTION

Let $M = (m_{ij})_{1\leq i,j\leq n}$ be a square matrix of order n of real entries. Motivated by a question from linear-optical quantum computing (see [2]), the first named author and Hance [3] asked the following question.

Question 1.1. *Characterize all matrices M such that $\|M\|_2 \leq 1$ and there exists a constant $C \geq 0$ such that*

$$\mathrm{per}(M) \geq n^{-C}.$$

It is not hard to show that (see also [6]), with $\mathbf{x} = (x_1, \dots, x_n)$ where $x_1, \dots, x_n$ are iid copies of an arbitrary real random variable ξ of mean zero and variance one,

$$\mathrm{per}(M) = \mathbf{E}_{\mathbf{x}} \prod_{i=1}^{n} x_i (M\mathbf{x})_i.$$

Thus, if we choose ξ to be the Bernoulli random variable (taking values ± 1 independently with probability $1/2$), then $\mathrm{per}(M) \geq n^{-C}$ would imply that

$$\mathbf{E}_{\mathbf{x}\in\mathcal{C}_n} \prod_{i=1}^{n} |(M\mathbf{x})_i| \geq n^{-C}.$$

where $\mathcal{C}_n$ denotes the n-dimensional hypercube $\{-1,1\}^n$, and $\mathbf{x}$ is chosen uniformly in $\mathcal{C}_n$

S. Aaronson is supported by an NSF Waterman Award. H. Nguyen is supported by research grant DMS-1358648.

Furthermore, as $\prod_{i=1}^n |(M\mathbf{x})_i| \leq 1$, a simple calculation gives

$$s(M) := \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}\left(\prod_{i=1}^n |(M\mathbf{x})_i| \geq n^{-C}/2\right) \geq n^{-C}/2. \tag{1}$$

For short, the quantity $s(M)$ above is called the *score function* of M. Equation (1) motivates us to ask the following question (see also [1]).

Question 1.2. *Characterize all matrices M such that $\|M\|_2 \leq 1$ and there exists a constant C such that $s(M) \geq n^{-C}/2$.*

As the direct study of $s(M)$ seems very difficult at the moment, the goal of this paper is to focus on a simpler (but closely related) object as follows. We define the *exact score function* of M to be

$$s_0(M) := \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n).$$

Clearly, $s_0(M) \leq s(M)$. We observe that, as s_0 measures how $\mathcal{C}_n$ is preserved under M, or how far the random vector $M\mathbf{x}$ is from being a Bernoulli vector, the study of this exact score function is natural on its own.

For the sake of discussion, we will be focusing on orthogonal matrices for the rest of this section. Observe that if $M\mathbf{x} = (\varepsilon_1 x_{\pi(1)}, \dots, \varepsilon_n x_{\pi(n)})$ for any choice of signs $\varepsilon_i \in \{-1, 1\}$, and for any permutation π in S_n, then $s_0 = 1$. In other words, if M is a product of permutation and reflection matrices (or shortly permutation-reflection matrices), then $s_0(M) = 1$. We would like to study the following inverse problem.

Question 1.3. *Are permutation-reflection matrices "essentially" the only orthogonal matrices with large $s_0(M)$, say $s_0(M) \geq n^{-C}$ for some positive constant C?*

In this paper, we confirm this heuristic by showing the following.

Theorem 1.4 (Main application)**.** *Let $0 < \varepsilon < 1$ and $C > 0$ be given constants. Assume that $M \in \mathbf{O}(n)$ with $s_0(M) \geq n^{-C}$. Then all but $O(n^{1-\varepsilon})$ rows of M contain a (unique) entry of absolute value at least $1 - O(n^{-1+\varepsilon})$.*

We will see in Example 2.9 that the lower bound $1 - O(n^{-1+o(1)})$ on the large entries is tight. We believe that our approach to proving Theorem 1.4, which goes through inverse Littlewood-Offord theory initiated by Tao and Vu, will also be useful for the study of Question 1.2.

In Appendix A, we answer Question 1.1, but for *stochastic* matrices, rather than orthogonal matrices or matrices of bounded norm. In more detail, we show there that, if A is an $n \times n$ stochastic matrix with $\mathrm{Per}\,(A) \geq n^{-O(1)}$, then all but $O\,(\log n)$ of the rows of A are dominated by a single large entry (and in that sense, A is "close to a permutation matrix").

Let us mention a few interesting applications and alternative statements of Theorem 1.4.

First, Theorem 1.4 implies that there is no $n \times n$ orthogonal matrix M that maps a non-negligible fraction of *uniform superpositions* to other uniform superpositions, besides "highly degenerate" matrices (i.e., those close to permuted diagonal matrices). Here a uniform superposition is defined to be any quantum state of the form

$$|\psi\rangle = \frac{\pm |1\rangle \pm \cdots \pm |n\rangle}{\sqrt{n}}.$$

These states often arise and are of interest in quantum computing, and a-priori, one might have hoped that there would be interesting transformations that had a non-negligible probability of staying within the set of such states. We conjecture that an analogous result should hold for arbitrary *unitary* matrices, except

(1) with the condition that $M|\psi\rangle$ is a uniform superposition relaxed to the condition that $|\langle i| M |\psi\rangle| = 1/\sqrt{n}$ for all $i \in [n]$, and

(2) with the exception for matrices close to permuted diagonal matrices broadened to include matrices close to permuted *block*-diagonal matrices, with 2×2 blocks such as

$$B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

For one can check that B maps each of the four vectors $(1,1)$, $(1,-1)$, $(-1,1)$, $(-1,-1)$ to a vector both of whose entries have equal magnitude. (We thank Sumegha Garg for this observation.)

Second, it is clear that $\mathbf{x}^T M \mathbf{y} \leq n$, if M is an $n \times n$ orthogonal matrix and $\mathbf{x}$ and $\mathbf{y}$ are any vectors in $\mathcal{C}_n$. Moreover it is clear that

$$\mathbf{P}_{\mathbf{x},\mathbf{y} \in \mathcal{C}_n} \left(\mathbf{x}^T M \mathbf{y} = n \right) \leq \frac{1}{2^n},$$

since we can only have equality if $x = My$. Theorem 1.4 says that, unless M is highly degenerate (in particular, unless all but $O\left(n^{1-\varepsilon}\right)$ of M's rows are dominated by a single large entry), we have

$$\mathbf{P}_{\mathbf{x},\mathbf{y} \in \mathcal{C}_n} \left(\mathbf{x}^T M \mathbf{y} = n \right) = o\left(\frac{1}{2^n n^C} \right)$$

for all constants C.

We thank Alex Arkhipov for the third application of Theorem 1.4. Given two points $\mathbf{x}, \mathbf{y} \in \mathcal{C}_n$, let their *Hamming distance* $\Delta(\mathbf{x}, \mathbf{y})$ be the number of coordinates on which they differ. Then given a subset $S \subseteq \mathcal{C}_n$, and a function $f : S \to \mathcal{C}_n$, we call f a *bi-Lipschitz bijection of* S if

$$\Delta\left(f(\mathbf{x}), f(\mathbf{y})\right) = \Delta(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in S$. Clearly, for any S, we can produce $n! 2^n$ different bi-Lipschitz bijections of S by simply permuting and reflecting the n coordinates of the hypercube. However, one might wonder for which S's there exist bi-Lipschitz bijections that are more interesting than that. We claim that Theorem 1.4 implies that, *if* $|S| \geq 2^n / n^{O(1)}$, *then any bi-Lipschitz bijection of* S *is close (in some sense) to a permutation and reflection of the coordinates.* This is a consequence of the following proposition.

Proposition 1.5. *Given any bi-Lipschitz bijection* $f : S \to \mathcal{C}_n$, *there exists an orthogonal matrix* M *such that* $M\mathbf{x} = f(\mathbf{x})$ *for all* $\mathbf{x} \in S$.

Proof. (of Proposition 1.5) Note that, if we interpret $\mathbf{x}, \mathbf{y} \in \mathcal{C}_n$ as points in $\mathbf{R}^n$, then

$$\Delta(\mathbf{x}, \mathbf{y}) = \frac{\|\mathbf{x} - \mathbf{y}\|_2^2}{4}.$$

Thus,

$$\|f(\mathbf{x}) - f(\mathbf{y})\|_2 = \|\mathbf{x} - \mathbf{y}\|_2$$

for all $\mathbf{x}, \mathbf{y} \in S$. Clearly we also have $\|f(\mathbf{x})\|_2 = \|\mathbf{x}\|_2 (= \sqrt{n})$ for all $\mathbf{x} \in S$. This means that the set $f(S)$ is a rigid rotation and/or reflection of the set S, so it must be possible to get from one to the other by applying an orthogonal matrix. □

2. Characterization of matrices with large score function

In line with Question 1.1 and 1.2, it is natural to study the exact score function for more general matrices.

Question 2.1. *Assume that $M \in \mathbf{M}(n)$ with $s_0(M) \geq n^{-C}$ for some constant $C > 0$. What can we say about the structure of M?*

Let us consider some natural candidates for M.

Example 2.2 (special matrices of $\{-1, 0, 1\}$ entries)**.** *It is clear that if M is a $\{-1, 0, 1\}$ matrix where each row contains exactly one non-zero entry, then $s_0(M) = 1$. More generally, one can construct M satisfying $s_0(M) \geq n^{-C}$ from such matrices of size close to n together with other block matrices of extremely small size.*

For further examples, we introduce the notion of generalized arithmetic progression (GAP).

Definition 2.3. A set $Q \subset Z$, where Z is an abelian torsion-free group, is a *GAP of rank r* if it can be expressed in the form

$$Q = \{\mathbf{g}_0 + k_1\mathbf{g}_1 + \cdots + k_r\mathbf{g}_r : K_i \leq k_i \leq K_i', k_i \in \mathbf{Z} \text{ for all } 1 \leq i \leq r\}$$

for some $\mathbf{g}_0, \ldots, \mathbf{g}_r \in Z$, and some integers $K_1, \ldots, K_r, K_1', \ldots, K_r'$.

The elements $\mathbf{g}_i \in Z$ are the *generators* of Q, the numbers K_i' and K_i are the *dimensions* of Q. We say that Q is *proper* if $|Q| = \prod(K_i' - K_i + 1)$. If $g_0 = 0$ and $-K_i = K_i'$ for all $i \geq 1$, we say that Q is *symmetric*.

Example 2.4 (Additively perturbed matrices)**.** *One can perturb a matrix from Example 2.2 by elements from a GAP to obtain a matrix $M \in \mathbf{M}(n)$ with large score function. Indeed, let F_0 be any matrix of size n from Example 2.2, and let $\mathbf{g}_1, \ldots, \mathbf{g}_r \in \mathbf{R}^n$ be r vectors in $\mathbf{R}^n$, where $r = O(1)$. Consider a GAP*

$$Q = \{k_1\mathbf{g}_1 + \cdots + k_r\mathbf{g}_r : |k_i| \leq K_i\},$$

where $\prod_{i=1}^r (2K_i + 1) = n^{O(1)}$.

Choose any $n-1$ elements $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ from Q. By a standard deviation principle and by the pigeonhole principle, there exists $\mathbf{u}_n \in (10\sqrt{n})Q = \{k_1\mathbf{g}_1 + \dots + k_r\mathbf{g}_r : |k_i| \le 10\sqrt{n}K_i\}$ such that

$$\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}\left(\sum_{i=1}^{n-1} x_i\mathbf{u}_i + \mathbf{u}_n = 0\right) \ge n^{-O(1)}.$$

Let U be the matrix of the column vectors $\mathbf{u}_i$ and set $M := F_0 + U$. By definition,

$$s_0(M) = \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \ge \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(U\mathbf{x} = 0) \ge \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}\left(\sum_{i=1}^{n-1} x_i\mathbf{u}_i + \mathbf{u}_n = 0\right) \ge n^{-O(1)}.$$

It is thus natural to conjecture that the matrices from Example 2.2 and Example 2.4 are essentially the only ones that have large score function. We support this conjecture by showing the following.

Definition 2.5. For integers a, b, let $\mathcal{F}_{ab}$ denote the collection of all $\{-1, 0, 1\}$-matrices of size $a \times b$ where each row contains *at most* one non-zero entry.

Theorem 2.6 (Characterization of general matrices, main result)**.** *Let $0 < \varepsilon < 1$ and C be positive constants. Suppose that $M = (m_{ij})_{1\le i,j\le n} \in \mathbf{M}(n)$ satisfies $s_0(M) \ge n^{-C}$ for some positive constant $C > 0$. Then there exists a submatrix M' of M of size $n_1 \times n_2$, with $n_1, n_2 = n - O_{C,\varepsilon}(n^{1-\varepsilon})$, and a set of $r = O_{C,\varepsilon}(1)$ vectors $\mathbf{g}_1, \dots, \mathbf{g}_r \in \mathbf{R}^{n_1}$ such that M' can be written as*

$$M' = M'' + F,$$

where $F \in \mathcal{F}_{n_1 n_2}$, and the columns of M'' belong to a GAP of size $n^{O_{C,\varepsilon}(1)}$ generated by $\mathbf{g}_1, \dots, \mathbf{g}_r$.

When M has nearly full rank, one can easily deduce the following consequence with much more information on F.

Corollary 2.7. *Suppose that M satisfies the condition of Theorem 2.6 as well as $\operatorname{rank}(M) = n - o(n)$. Then one can also assume that each row of F contains exactly one non-zero entry which is either 1 or -1. In other words, F is nearly a permutation-reflection matrix modulo a low-rank perturbation.*

We next focus on orthogonal matrices. Similarly to Question 2.1, one would like to characterize orthogonal matrices with large score function.

Question 2.8. *Assume that M is an orthogonal matrix satisfying $s_0(M) \ge n^{-C}$ for some constant $C > 0$. What can we say about the structure of M?*

As suggested by Theorem 2.6 and Corollary 2.7, it is natural to search for M from the matrices of Example 2.4: thus are there non-trivial low rank perturbations of an orthogonal matrix which are again orthogonal and $s_0(M) \geq n^{-C}$? The answer is positive, even for rank-one perturbation.

Example 2.9. *Let $\mathbf{u}_1 = (x, t_2 x, \dots, t_n x) \in \mathbf{R}^n$, where $x \neq 0$ and $t_2, \dots, t_n \in \mathbf{R}$ are to be chosen. We will select $\mathbf{u}_1$ so that $\mathbf{c}_1 = \mathbf{u}_1 + \mathbf{e}_1$ has norm 1, which hence requires*

$$(x+1)^2 + x^2(t_2^2 + \dots + t_n^2) = 1, \text{ or equivalently, } \mathbf{u}_1 \cdot \mathbf{u}_1 + 2x = 0. \tag{2}$$

For $2 \leq i \leq n$, define $\mathbf{u}_i := t_i \mathbf{u}_1$. Thus if $t_i = k_i/N$ with $k_i \in \mathbf{Z}, |k_i| \leq n^{O(1)}$ then $\mathbf{u}_i$ belongs to the rank-one arithmetic progression $\{k\mathbf{u}_1/N, |k| \leq n^{O(1)}\}$.

Set $\mathbf{c}_i := \mathbf{u}_i + \mathbf{e}_i$, and let M be the matrix of the column vectors $\mathbf{c}_i$. In other words,

$$M = (\delta_{ij} + x t_i t_j)_{1 \leq i,j \leq n}. \tag{3}$$

By definition, it can be verified that

$$\|\mathbf{c}_i\|_2 = 1, \text{ and } \mathbf{c}_i \cdot \mathbf{c}_j = 0, \forall i \neq j.$$

Indeed, for the first part, using (2)

$$\|\mathbf{c}_i\|_2^2 = \mathbf{u}_i \cdot \mathbf{u}_i + 2\mathbf{u}_i \cdot \mathbf{e}_i + 1 = t_i^2 \mathbf{u}_1 \cdot \mathbf{u}_1 + 2t_i^2 x + 1 = 1.$$

For the second part, similarly,

$$\mathbf{c}_i \cdot \mathbf{c}_j = \mathbf{u}_i \cdot \mathbf{u}_j + \mathbf{u}_i \cdot \mathbf{e}_j + \mathbf{u}_j \cdot \mathbf{e}_i = t_i t_j \mathbf{u}_1 \cdot \mathbf{u}_1 + 2x t_i t_j = 0.$$

Notice that the specific choice of $x = -2/n$ and $t_i = 1$ would fulfill (2)*, and hence confirms the asymptotic sharpness of Theorem 1.4.*

In what follows we describe another natural way to sample the t_i.

- *First, select integers $k_2, \dots, k_n \in \mathbf{Z}$ within the range $|k_i| \leq n^{O(1)}$ so that $\mathbf{P}_{\mathbf{x} \in \mathcal{C}_n}(x_1 + \sum_{i=2}^n x_i k_i = 0) \geq n^{-O(1)}$. (Although this is not true for all choices of k_i, it holds for many natural families.)*
- *Next, choose $x = -1$ and set $t_i := k_i/N$, where $N^2 := \sum_{i=2}^n k_i^2$.*

It is clear that (2) *is fulfilled, and hence by definition,*

$$s_0(M) = \mathbf{P}_{\mathbf{x} \in \mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \geq \mathbf{P}_{\mathbf{x} \in \mathcal{C}_n}(U\mathbf{x} = \mathbf{0}) \geq n^{-O(1)}.$$

As an extension of Theorem 2.6 and Corollary 2.7, we show that if M is as in Question 2.8, then it essentially resembles the matrices of Example 2.9.

Theorem 2.10 (Characterization of orthogonal matrices, main result)**.** *Let $0 < \varepsilon < 1$ and C be positive constants. Suppose that $M = (m_{ij})_{1 \le i,j \le n} \in \mathbf{O}(n)$ satisfies $s_0(M) \ge n^{-C}$ for some positive constant $C > 0$. Then there exists a submatrix M' of M of size $n \times n_2$, with $n_2 = n - O_{C,\varepsilon}(n^{1-\varepsilon})$, and a set of $r = O_{C,\varepsilon}(1)$ vectors $\mathbf{g}_1, \dots, \mathbf{g}_r \in \mathbf{R}^n$ such that M' can be written as*

$$M' = M'' + F,$$

where $F \in \mathcal{F}_{nn_2}$ and the columns of M'' belong to a GAP of small size generated by $\mathbf{g}_1, \dots, \mathbf{g}_r$.

Furthermore, the matrix $M'' + F$ (modulo appropriate row permutations) when restricting to the first n_2 rows can be written as

$$\begin{pmatrix} U & UD^T \\ DU & DUD^T \end{pmatrix} + I \tag{4}$$

where U is a square matrix of size r and D is an $(n_2 - r) \times r$ matrix, and I is a diagonal square matrix of size n_2 of entries from $\{-1, 1\}$.

We remark that (4) is a generalization of (3), and it is not hard to construct small rank perturbations of type (4) of permutation-reflection matrices which are again orthogonal and have large score function.

One of the main tools of our treatment is an inverse-type Erdős-Littlewood-Offord result. The classical result of Erdős states as follows.

Theorem 2.11. [5] *Let $\mathbf{a}_1, \dots, \mathbf{a}_n$ be elements of an abelian torsion-free group Z such that at least n' of them are non-zero, then*

$$\sup_{\mathbf{a} \in Z} \mathbf{P}_{\mathbf{x} \in \mathcal{C}_n} \left(\sum_{i=1}^n x_i \mathbf{a}_i = \mathbf{a} \right) \le \binom{n'}{n'/2} / 2^{n'} = O(1/\sqrt{n'}).$$

This result would be sufficient for characterizing M with $s_0(M) \ge \Omega(1)$. However, in the polynomial regime $s_0(M) \ge n^{-C}$, one will need stronger inverse results. We will invoke the following from [7, Theorem 2.5], which was first proved in [9, 10] by Tao and Vu.

Theorem 2.12. *Let $0 < \varepsilon < 1$ and C be positive constants. Assume that $\mathbf{a}_1, \dots, \mathbf{a}_n$ are elements of an abelian torsion-free group Z such that*

$$\rho(\mathbf{a}_1, \dots, \mathbf{a}_n) := \sup_{\mathbf{a} \in Z} \mathbf{P}_{\mathbf{x} \in \mathcal{C}_n} \left(\sum_{i=1}^n x_i \mathbf{a}_i = \mathbf{a} \right) \ge n^{-C}.$$

Then there exists a proper symmetric GAP $Q \subset Z$ which contains all but at most n^ε elements of $\mathbf{a}_1, \dots, \mathbf{a}_n$. Furthermore,

- *the rank r of Q is bounded by a constant, $r = O_{C,\varepsilon}(1)$,*
- *the size of Q is small, $|Q| = O_{C,\varepsilon}(\rho^{-1} n^{-\varepsilon r/2})$.*

It is crucial to remark that the implied constants are independent of the group Z. Furthermore, by following the proof of [7, Theorem 2.5], these constants are $O\left(2^{2^{2^{\alpha(C/\varepsilon)}}}\right)$ at most, where α is an absolute constant. Consequently, the dependent constants under $O_{C,\varepsilon}(.)$ in Theorem 2.6 and Theorem 2.10 can be taken to be $O\left(2^{2^{2^{\alpha(C/\varepsilon)}}}\right)$. However, as these dependencies are not our current focus, we will skip the details to ease the presentation.

Notation.

All of the implied constants in our $O(.), o(.)$ notations, if not specified, will depend on the common parameters C and ε. As such, allow us to skip these subscripts for brevity.

For a matrix M, $\mathbf{r}_i(M), \mathbf{c}_j(M)$ denote the i-th row and j-th column respectively. For a vector $\mathbf{v} \in \mathbf{R}^n$, $(\mathbf{v})_i$ denotes its i-th component. For an index set $T \subset [n]$, $\mathbf{v}^{[T]}$ represents the subvector of $\mathbf{v}$ of components indexing by T.

For short, we say that a symmetric proper GAP $P = \{\sum_{i=1}^r k_i \mathbf{g}_i\}$ has small size and bounded rank if $|P| = n^{O(1)}$ and $\mathrm{rank}(P) = O(1)$.

We say that r elements $\mathbf{x}_1 = k_{11}\mathbf{g}_1 + \dots + k_{1r}\mathbf{g}_r, \dots, \mathbf{x}_r = k_{r1}\mathbf{g}_1 + \dots, k_{rr}\mathbf{g}_r$ span P if the corresponding vectors

$$(k_{11}, \dots, k_{1r}), \dots, (k_{r1}, \dots, k_{rr})$$

have full rank in $\mathbf{R}^r$.

The rest of the paper is organized as follows. We will introduce some key lemmas in Section 3, and prove Theorem 2.6 and Theorem 2.10 in Section 4 and Section 5 respectively. Before concluding the paper with a problem section (Section 8) and a remark on stochastic matrices (Appendix A), we give two applications: one shows that general matrices with sufficiently small entries are not near invariant with respect to the hypercube (Section 6), and one deduces Theorem 1.4 (Section 7).

3. Structural relation between rows and columns

Assume that we were in the strong case that $s_0(M) = \Omega(1)$. Then by Theorem 2.11, for any row $\mathbf{r}_i = (m_{i1}, \dots, m_{in})$ of M, all but $O(1)$ of the components m_{ij} would be zero. It turns out that, if $s_0(M) \geq n^{-C}$ as in Theorem 2.6, we can conclude similarly. Namely, by Theorem 2.12, it is not hard to show that for any row $\mathbf{r}_i = (m_{i1}, \dots, m_{in})$ of M, all but n^ε of the components m_{ij} belong to a GAP of bounded rank and small size. In the lemma below we slightly improve this result for collections of several rows.

For any d indices $1 \le i_1 < \cdots < i_d \le n$, let $\mathbf{c}_1^{[i_1,\ldots,i_d]}, \ldots, \mathbf{c}_n^{[i_1,\ldots,i_d]}$ be the column vectors of the $d \times n$ submatrix spanned by $\mathbf{r}_{i_1}(M), \ldots, \mathbf{r}_{i_d}(M)$ of M. We will prove the following.

Lemma 3.1 (Structure for row and column vectors I.)**.** *Let $0 < \varepsilon < 1$ and C be positive constants. Let M be a matrix with $s_0(M) \ge n^{-C}$. Assume that $1 \le d \le c \log n$, where $c > 0$ is a constant, and let $1 \le i_1 < \cdots < i_d \le n$ be any d indices. Then there exist an exceptional index set $I_{i_1,\ldots,i_d} \subset [n]$ of size at most n^ε, and a GAP $Q_{i_1,\ldots i_d} \subset \mathbf{R}^d$ which contains all $\mathbf{c}_i^{[i_1,\ldots,i_d]}, i \in \bar{I}_{i_1,\ldots,i_d}$, such that*

$$|Q_{i_1,\ldots,i_d}| = O(2^d n^C / n^{r\varepsilon/2}).$$

As a consequence,

(i) the rank r of $Q_{i_1,\ldots,i_d}$ is bounded, $\operatorname{rank}(Q_{i_1,\ldots,i_d}) = O(1)$*;*

(ii) the rows of the matrix generated by $\mathbf{c}_i^{[i_1,\ldots,i_d]}, i \in \bar{I}_{i_1,\ldots,i_d}$, span a subspace of dimension at most $\operatorname{rank}(Q_{i_1,\ldots,i_d})$*.*

Proof. (of Lemma 3.1) As $\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \ge n^{-C}$, by applying the projection $\pi_{i_1,\ldots,i_d}$ (mapping $\mathbf{R}^n$ onto the components of indices $i_1, \ldots, i_d$), the column vectors $\mathbf{c}_1^{[i_1,\ldots,i_d]}, \ldots, \mathbf{c}_n^{[i_1,\ldots,i_d]}$ have large concentration probability ρ,

$$\rho(\mathbf{c}_1^{[i_1,\ldots,i_d]}, \ldots, \mathbf{c}_n^{[i_1,\ldots,i_d]}) \ge n^{-C} 2^{-d}.$$

The conclusion of Lemma 3.1 then follows by applying Theorem 2.12 to these column vectors. □

We next show that the result of Lemma 3.1 can be extended to collections of as many as $n^{1-\varepsilon}$ rows. Let $r_0 = O_{C,\varepsilon}(1)$ be an upper bound for all $\operatorname{rank}(Q)$ (applied to all d indices $i_1, \ldots, i_d$) from Lemma 3.1.

Lemma 3.2 (Structure for row and column vectors II)**.** *Let $1 \le k \le n^{1-\varepsilon}$ and $r_0 \le d \le c \log n$. Consider any m indices $1 \le i_1 < \cdots < i_m \le n$, where $m = k(d - r_0) + d$. Then for the rows $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_m}$ of M, there exist an exceptional index set $I_{i_1\ldots i_m} \subset [n]$ of size at most $(k+1)n^\varepsilon$ and a GAP $Q_{i_1,\ldots,i_m} \subset \mathbf{R}^m$ of rank $r \le r_0$ such that the following holds.*

(i) $Q_{i_1,\ldots,i_m} \subset \mathbf{R}^m$ contains all $\mathbf{c}_i^{[i_1,\ldots,i_m]}, i \in \bar{I}_{i_1,\ldots,i_m}$, and

$$|Q_{i_1,\ldots,i_m}| = O(2^d n^C / n^{r\varepsilon/2}).$$

(ii) The row vectors of the submatrix spanned by $\mathbf{c}_i^{[i_1,\ldots,i_m]}, i \in \bar{I}_{i_1,\ldots,i_m}$, span a subspace of dimension at most r_0.

Proof. (of Lemma 3.2) For (ii), we first apply Lemma 3.1 to the first d rows $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_d}$ to obtain an exceptional set $I_{i_1,\ldots,i_d}$ and a collection of r_0 rows which span the subspace of the rows of the matrix generated by $\mathbf{c}_i^{[i_1,\ldots,i_d]}, i \in \bar{I}_{i_1,\ldots,i_d}$. We next add to these r_0 rows another

$d - r_0$ ones among the remaining $\mathbf{r}_{i_{d+1}}, \dots, \mathbf{r}_{i_m}$ and apply Lemma 3.1 again. Iterate this process $k+1$ times and let $I = I_{i_1,\dots,i_m}$ be the union of the exceptional sets from each step. Hence $|I| \leq (k+1)n^{\varepsilon}$. One can check that by definition the row vectors of the matrix spanned by $\mathbf{c}_i^{[i_1,\dots,i_m]}, i \in \bar{I}_{i_1,\dots,i_m}$, span a subspace of dimension $r \leq r_0$.

For (i), we will show that the column vectors $\mathbf{c}_i^{[i_1,\dots,i_m]}$ belong to a GAP by restricting to the projection of $\mathbf{R}^n$ onto the components of indices from $\bar{I}_{i_1,\dots,i_m}$. At the last step of the above process, choose $r \leq r_0$ rows $\mathbf{r}_{j_1}, \dots, \mathbf{r}_{j_r}$ that span the subspace generated by the lastly considered d rows, and hence by definition of the process, they also span the subspace generated by all m rows (restricting to the components of indices from $\bar{I}_{i_1,\dots,i_m}$).

After the application of Lemma 3.1 at this last step, the columns $\mathbf{c}_i^{[j_1,\dots,j_r]}, i \in \bar{I}_{i_1,\dots,i_m}$, of the matrix corresponding to $\mathbf{r}_{j_1}, \dots, \mathbf{r}_{j_r}$ belong to a GAP $Q_r \subset \mathbf{R}^r$ of rank $r \leq r_0$ and of small size. Let $\mathbf{g}_1^{[j_1,\dots,j_r]}, \dots, \mathbf{g}_r^{[j_1,\dots,j_r]} \in \mathbf{R}^r$ be the generators of Q_r. We claim that this structure can be extended to $\mathbf{R}^m$ to contain all the extension vectors $\mathbf{c}_i^{[i_1,\dots,i_m]}$ of $\mathbf{c}_i^{[j_1,\dots,j_r]}$. Recall that

(1) As by (ii), for any $i \in \{i_1, \dots, i_m\}$, there exist real coefficients $\alpha_{i1}, \dots, \alpha_{ir} \in \mathbf{R}$ such that the restricted row vector $\mathbf{r}_i^{[\bar{I}_{i_1,\dots,i_m}]} \in \mathbf{R}^{n-|I_{i_1,\dots,i_m}|}$ satisfies

$$\mathbf{r}_i^{[\bar{I}_{i_1,\dots,i_m}]} = \alpha_{i1}\mathbf{r}_{j_1}^{[\bar{I}_{i_1,\dots,i_m}]} + \dots + \alpha_{ir}\mathbf{r}_{j_r}^{[\bar{I}_{i_1,\dots,i_m}]}.$$

(2) For any $i \in \bar{I}_{i_1,\dots,i_m}$, there exist integral coefficients $k_{i1}, \dots, k_{ir} \in \mathbf{R}$ with $|k_{i1} \dots k_{ir}| \leq |Q_r|$ such that the column vector $\mathbf{c}_i^{[j_1,\dots,j_r]}$ satisfies

$$\mathbf{c}_i^{[j_1,\dots,j_r]} = k_{i1}\mathbf{g}_1^{[j_1,\dots,j_r]} + \dots + k_{ir}\mathbf{g}_r^{[j_1,\dots,j_r]}.$$

Claim 3.3 (Structure extension)**.** *There exists a GAP Q_m in $\mathbf{R}^m$ with the following properties*

- *Q_m has the same rank and size with Q_r;*
- *Q_m contains the extension $\mathbf{c}_i^{[i_1,\dots,i_m]}$ of $\mathbf{c}_i^{[j_1,\dots,j_r]}$, where $i \in \bar{I}_{i_1,\dots,i_m}$.*

Proof. (of Claim 3.3) First, by using the coefficients α_{ij} from (1), we can extend the vectors $\mathbf{g}_i^{[j_1,\dots,j_r]}$ to the corresponding vector $\mathbf{g}_i^{[i_1,\dots,i_m]}$ in $\mathbf{R}^m$; these new vectors in $\mathbf{R}^m$ will serve as the generators of Q_m.

Let $\mathbf{c}_i^{[i_1,\dots,i_m]}$ be any column vector, where $i \in \bar{I}_{i_1,\dots,i_m}$. By (2),

$$\mathbf{c}_i^{[j_1,\dots,j_r]} = k_{i1}\mathbf{g}_1^{[j_1,\dots,j_r]} + \dots + k_{ir}\mathbf{g}_r^{[j_1,\dots,j_r]}.$$

By the definition of extension, we also have

$$\mathbf{c}_i^{[i_1,\dots,i_m]} = k_{i1}\mathbf{g}_1^{[i_1,\dots,i_m]} + \dots + k_{ir}\mathbf{g}_r^{[i_1,\dots,i_m]}.$$

□

It is clear that Q_m has the same rank and size as Q_r, and thus $r = \text{rank}(Q_m) \leq r_0$ and

$$|Q_m| = O(2^d n^C / n^{r\varepsilon/2}).$$

□

We now deduce a useful corollary of Lemma 3.2. Let H be the subspace obtained from (ii) of Lemma 3.2, and let $\mathbf{r}_{l_1}^{[\bar{I}_{i_1,\ldots,i_m}]}, \ldots, \mathbf{r}_{l_r}^{[\bar{I}_{i_1,\ldots,i_m}]}, l_1, \ldots, l_r \in \{i_1, \ldots, i_m\}, r = \dim(H)$ be any vectors that span H; we will refer to them as *base vectors*. By definition, we have the following.

Claim 3.4. *For all $\mathbf{r}_i, i \in \{i_1, \ldots, i_m\}$, the following holds: there exist real numbers $t_{i1}, \ldots, t_{ir}$ such that*

$$\mathbf{r}_i^{[\bar{I}_{i_1,\ldots,i_m}]} = \sum_{j=1}^{r} t_{ij} \mathbf{r}_{l_j}^{[\bar{I}_{i_1,\ldots,i_m}]}. \tag{5}$$

One can also rewrite (5) in a simple matrix form. Let $M_{i_1,\ldots,i_m}$ be the invertible matrix of order n obtained from I_n by replacing its i-th rows, with $i \in \{i_1, \ldots, i_m\} \backslash \{l_1, \ldots, l_r\}$, by the vector

$$(0, \ldots, 0, -t_{i1}, 0, \ldots, 0, \ldots, -t_{ir}, 0, \ldots, 0, 1, 0, \ldots, 0).$$

In other words, the matrix $M_{i_1,\ldots,i_m}$ acts on M by fixing every row except those $\mathbf{r}_i$ with $i \in \{i_1, \ldots, i_m\} \backslash \{l_1, \ldots, l_r\}$ in which case

$$M_{i_1,\ldots,i_m} : \mathbf{r}_i(M) \to \mathbf{r}_i(M) - \sum_{j=1}^{r} t_{ij} \mathbf{r}_{l_j}(M).$$

Corollary 3.5. *With the definition of $M_{i_1,\ldots,i_m}$ as above, for any $i \in \{i_1, \ldots, i_m\}$, the projections of the row vectors of the product matrix $M_{i_1,\ldots,i_m} M$ onto their components of indices in $\bar{I}_{i_1,\ldots,i_n}$ vanish:*

$$\mathbf{r}_i^{[\bar{I}_{i_1,\ldots,i_m}]}(M_{i_1,\ldots,i_m} M) = \mathbf{0}, \forall i \in \{i_1, \ldots, i_m\}.$$

4. Treatment for general matrices: proof of Theorem 2.6

4.1. **A proof without additive structure.** We first show an easier variant of Theorem 2.6 where the additive structure is omitted.

Theorem 4.2. *Let $0 < \varepsilon < 1$ and C be positive constants. Let $M = (m_{ij})_{1 \le i,j \le n} \in \mathbf{M}(n)$ be a matrix with $s_0(M) \ge n^{-C}$. Then there exists a submatrix M' of M of size $n_1 \times n_2$, with $n_1, n_2 = n - O(n^{1-\varepsilon})$, and a set of $r = O(1)$ vectors $\mathbf{g}_1, \dots, \mathbf{g}_r \in \mathbf{R}^{n_2}$ such that M' can be written as $M'' + F$, where $F \in \mathcal{F}_{n_1 n_2}$ and the rows of M'' are generated by $\mathbf{g}_1, \dots, \mathbf{g}_r$.*

As the subspace generated by the columns of M'' also has dimension at most r, we can restate Theorem 4.2 as follows.

Theorem 4.3. *There exist a submatrix M' of M of size $n_1 \times n_2$, with $n_1, n_2 = n - O(n^{1-\varepsilon})$, and a set of $r = O(1)$ vectors $\mathbf{g}_1, \dots, \mathbf{g}_r \in \mathbf{R}^{n_1}$ such that such that M' can be written as $M'' + F$, where $F \in \mathcal{F}_{n_1 n_2}$ and the columns of M'' are generated by $\mathbf{g}_1, \dots, \mathbf{g}_r$.*

We now prove Theorem 4.2. As the property $\mathbf{P}_{\mathbf{x} \in \mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \ge n^{-C}$ does not change by swapping the rows and columns of M, we will apply these swaps whenever necessary to simplify the presentation.

We will apply Lemma 3.2 and Corollary 3.5 to blocks of $m = r_0 + n^{2\varepsilon}$ consecutive rows of M. In each block B_i, we will keep track of the base vectors (by adding at most $r_0 - 1$ vectors if needed, we always assume that there are exactly r_0 base vectors) and $n^{2\varepsilon}$ others that belong to the subspace generated by these vectors. By swapping the rows if necessary, we obtain the following.

Claim 4.4. *The row set of M (with an exception of at most $r_0 + n^{2\varepsilon} - 1$ last rows) can be decomposed into n_0 consecutive blocks of size $r_0 + n^{2\varepsilon}$ each, here $n_0 \ge n/(r_0 + n^{2\varepsilon}) - 1$, such that in each block, the first r_0 rows serve as the base vectors and the next $n^{2\varepsilon}$ rows belong to the subspace generated by these r_0 base vectors.*

As such, for each block matrix B_i (generated by the rows in the i-th block), the matrix M_i obtained in Corollary 3.5 is a lower triangular matrix with 1's on its main diagonal of the form

$$M_i = \begin{pmatrix} I_{(i-1)(r_0+n^{2\varepsilon})} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & T_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_{n-i(r_0+n^{2\varepsilon})} \end{pmatrix},$$

where T_i is the corresponding lower triangular square matrix of order $r_0 + n^{2\varepsilon}$.

Observe also that for any $\mathbf{y} = (y_1, \dots, y_n)$, $M_i\mathbf{y}$ just changes the components y_{i_0} with $i_0 \in \{(i-1)(r_0 + n^{2\varepsilon}) + r_0 + 1, \dots, i(r_0 + n^{2\varepsilon})\}$,

$$(M_i\mathbf{y})_{i_0} = y_{i_0} - \sum_{j=1}^{r_0} t_{i_0 j} y_{(i-1)n^{2\varepsilon}+j}. \tag{6}$$

Set

$$M_0 := M_{n_0} \cdots M_1 \text{ and } M_0' := M_0 M. \tag{7}$$

Let $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{y} = M\mathbf{x}$. By definition, the components of $\mathbf{y} \in \mathbf{R}^n$ can be decomposed into blocks of $r_0+n^{2\varepsilon}$ consecutive components $\mathbf{y}_1 = (y_1, \dots, y_{n^{2\varepsilon}+r_0}), \mathbf{y}_2 = (y_{n^{2\varepsilon}+r_0+1}, \dots, y_{2n^{2\varepsilon}+2r_0}), \dots$ (except at most $n^{2\varepsilon} + r_0 - 1$ last components) such that

$$M_0\mathbf{y} = (T_1\mathbf{y}_1, T_2\mathbf{y}_2, \dots).$$

Also, by definition,

$$M_0' = \begin{pmatrix} T_1B_1 \\ T_2B_2 \\ \cdots \end{pmatrix},$$

where we recall that B_i is the $(r_0+n^{2\varepsilon})\times n$ matrix generated by $\mathbf{r}_j(M), (i-1)(r_0+n^{2\varepsilon})+1 \leq j \leq i(r_0+n^{2\varepsilon})$. Furthermore, by Corollary 3.5

$$T_iB_i = \begin{pmatrix} X_i \\ Y_i \end{pmatrix},$$

where X_i is an $r_0 \times n$ matrix and Y_i is an $n^{2\varepsilon} \times n$ matrix satisfying the following property: there exists an exceptional index set $I_i \subset [n]$ of size at most $(k+1)n^{\varepsilon} < n^{3\varepsilon}$ such that for any $j \in \bar{I}_i$,

$$\mathbf{c}_j(Y_i) = \mathbf{0}. \tag{8}$$

Now we analyze the event $\mathbf{y} = M\mathbf{x} \in \mathcal{C}_n$. Rewrite as

$$M_0'\mathbf{x} = M_0\mathbf{y}. \tag{9}$$

Projecting (9) onto the components of indices from $\{(i-1)(r_0+n^{2\varepsilon})+r_0+1, \dots, i(r_0+n^{2\varepsilon})\}$, we obtain via (6)

$$Y_i\mathbf{x} = \Big(y_{i_0} - \sum_{j=1}^{r_0} t_{i_0 j} y_{(i-1)(r_0+n^{2\varepsilon})+j}\Big)_{i_0 \in \{(i-1)(r_0+n^{2\varepsilon})+r_0+1,\dots,i(r_0+n^{2\varepsilon})\}}. \tag{10}$$

Furthermore, when $\mathbf{y} \in \mathcal{C}_n$,

$$(y_{i_0})_{i_0 \in \{(i-1)(r_0+n^{2\varepsilon})+r_0+1,\dots,i(r_0+n^{2\varepsilon})\}} \in \mathcal{C}_{n^{2\varepsilon}} \text{ and } (y_{(i-1)(r_0+n^{2\varepsilon})+1}, \dots, y_{(i-1)(r_0+n^{2\varepsilon})+r_0}) \in \mathcal{C}_{r_0}.$$

Define L_i to be the $n^{2\varepsilon} \times r_0$ matrix

$$L_i \mathbf{z} := (-\sum_{j=1}^{r_0} t_{i_0 j} z_j)_{i_0 \in \{(i-1)(r_0+n^{2\varepsilon})+r_0+1,\dots,i(r_0+n^{2\varepsilon})\}}.$$

We obtain from (10) the following useful bound.

Lemma 4.5 (Block structure I)**.** *Assume that* $\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(M\mathbf{x} \in \mathcal{C}_n) \geq n^{-C}$*. Then,*

$$\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(\wedge_{1\leq i\leq n_0} Y_i\mathbf{x} \in L_i\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \geq n^{-C}.$$

Independence decomposition. As a consequence of Lemma 4.5, for any $j_0 \leq n_0$ and any j_0 indices $i_1 < \cdots < i_{j_0}$, one also has

$$\mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(\wedge_{1\leq j\leq j_0} Y_{i_j}\mathbf{x} \in L_{i_j}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \geq n^{-C}. \tag{11}$$

Recall from (8) that the number of non-zero columns of each Y_{i_j} is $|I_{i_j}| \leq n^{3\varepsilon}$, and thus they are extremely sparse. Furthermore, if the index sets $I_{i_1}, \dots I_{i_{j_0}}$ were disjoint for some $j_0 \gg \log n$, then the events $Y_{i_j}\mathbf{x} \in L_{i_j}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}$ would be independent, and so (11) would imply that most of the events $Y_{i_j}\mathbf{x} \in L_{i_j}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}$ hold with probability very close to one. We will make this observation rigorous in the next step.

Let $Y_{i_j|i_1,\dots,i_{j-1}}$ be the (possibly empty) submatrix of Y_{i_j} of columns indexing from $I_{i_j|i_1,\dots i_{j-1}} := I_{i_j}\backslash(I_{i_1} \cup \dots I_{i_{j-1}})$, we can rewrite (11) in terms of conditional probability

$$\begin{aligned}
n^{-C} &\leq \mathbf{P}_{\mathbf{x}\in\mathcal{C}_n}(\wedge_{1\leq j\leq j_0} Y_{i_j}\mathbf{x} \in L_{i_j}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \\
&= \mathbf{P}_{\mathbf{x}_{I_{i_1}}}(Y_{i_1}\mathbf{x}_{I_{i_1}} \in L_{i_1}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \times \mathbf{P}_{\mathbf{x}_{I_{i_2}}}(Y_{i_2}\mathbf{x}_{I_{i_2}} \in L_{i_2}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}} | A(\mathbf{x}_{I_{i_1}})) \times \dots \\
&\times \mathbf{P}_{\mathbf{x}_{I_{i_{j_0}}}}(Y_{i_{j_0}}\mathbf{x}_{I_{i_{j_0}}} \in L_{i_{j_0}}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}} | A(\mathbf{x}_{I_{i_1}}) \wedge \dots \wedge A(\mathbf{x}_{I_{i_{j_0-1}}})) \\
&\leq \mathbf{P}_{\mathbf{x}_{I_{i_1}}\in\mathcal{C}_{|I_{i_1}|}}(Y_{i_1}\mathbf{x}_{I_{i_1}} \in L_{i_1}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \times \sup_{\mathbf{a}_2} \mathbf{P}_{\mathbf{x}_{I_{i_2|i_1}} \in \mathcal{C}_{|I_{i_2|i_1}|}}(Y_{i_2|i_1}\mathbf{x}_{I_{i_2|i_1}} \in \mathbf{a}_2 + L_{i_2}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \times \dots \\
&\times \sup_{\mathbf{a}_{j_0}} \mathbf{P}_{\mathbf{x}_{I_{j_0|i_1\dots i_{j_0-1}}} \in \mathcal{C}_{|I_{i_{j_0}|i_1\dots i_{j_0-1}}|}}(Y_{i_{j_0}|i_1,\dots,i_{j_0-1}}\mathbf{x}_{I_{j_0|i_1\dots i_{j_0-1}}} \in \mathbf{a}_{j_0} + L_{i_{j_0}}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}),
\end{aligned} \tag{12}$$

where $A(\mathbf{x}_{I_{i_j}})$ are the events $Y_{i_j}\mathbf{x}_{I_{i_j}} \in L_{i_j}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}$.

Definition 4.6. Motivated by (12), we say that a subsequence $\{Y_{i_1}, \dots, Y_{i_l}\}$ is *bad* if for all $1 \leq k \leq l$,

$$\sup_{\mathbf{a}} \mathbf{P}_{\mathbf{x}_{I_{i_k|i_1,\dots,i_{k-1}}} \in \mathcal{C}_{I_{i_k|i_1,\dots,i_{k-1}}}}(Y_{i_k|i_1,\dots,i_{k-1}}\mathbf{x}_{I_{i_k|i_1,\dots,i_{k-1}}} \in \mathbf{a} + L_{i_k}\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \leq 1 - \varepsilon.$$

Note that this definition trivially implies

$$|I_{i_k|i_1,\dots,i_{k-1}}| = |I_{i_k}\backslash(I_{i_1} \cup I_{i_2} \cup \dots \cup I_{i_{k-1}})| \geq 1.$$

Claim 4.7. *If the subsequence* $\{Y_{i_1}, \dots, Y_{i_l}\}$ *is bad, then*

$$l \leq (2C/\varepsilon) \log n.$$

Proof. This follows directly from (12),

$$n^{-C} \leq (1-\varepsilon)^l.$$

□

In our next step, choose a longest possible bad sequence. Without loss of generality, we assume that this consists of $Y_1, \dots, Y_l$. Next, for $l+1 \leq i \leq n_0$, call i *suitable* if

$$|J_i| := |I_i \backslash I_1 \cup \dots \cup I_l| \geq 1.$$

Let $\mathcal{I}_g$ be the collection of suitable indices. Choose any i from $\mathcal{I}_g$, we will be focusing on the structure of the matrix $Y_{i|i_1,\dots,i_l}$.

Lemma 4.8 (Block structure II)**.** *Let A be a matrix of size $n^{2\varepsilon} \times k$, where $k \geq 1$ and*

$$\sup_{\mathbf{a}} \mathbf{P}_{\mathbf{x} \in \mathcal{C}_k}(A\mathbf{x} \in \mathbf{a} + L\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \geq 1/2,$$

for some deterministic $n^{2\varepsilon} \times r_0$ matrix L. Then there exists an index set $I_A \subset [k]$ of size $O_{r_0}(1)$ such that the submatrix of A generated by columns indexing from $\bar{I}(A)$ has the following property: every row vector is either zero or contains exactly one ± 1 entry.

In our analysis A will play the role of the matrices $Y_{i|i_1,\dots,i_l}$.

Proof. (of Lemma 4.8) The assumption $\sup_{\mathbf{a}} \mathbf{P}_{\mathbf{x} \in \mathcal{C}_k}(A\mathbf{x} \in \mathbf{a} + L\mathcal{C}_{r_0} + \mathcal{C}_{n^{2\varepsilon}}) \geq 1/2$ implies

$$\sup_{\mathbf{a}} \mathbf{P}_{\mathbf{x} \in \mathcal{C}_k}(A\mathbf{x} \in \mathbf{a} + \mathcal{C}_{n^{2\varepsilon}}) \geq 1/2^{r_0+1}. \tag{13}$$

We apply Theorem 2.11 for any row $\mathbf{r}_i = (a_{i1}, \dots, a_{ik})$ of A. It is implied that all but $O_{r_0}(1)$ components a_{ij} are zero. For each $1 \leq i \leq n^{2\varepsilon}$, let $E_i \subset [n]$ denote the index set of the non-zero elements $a_{ij}, 1 \leq j \leq k$. Similarly to what we have done so far, call a sequence $E_{i_1}, \dots, E_{i_s}$ *ill* if one of the following holds:

- either $|E_{i_j} \backslash E_{i_1} \cup \dots \cup E_{i_{j-1}}| \geq 2$,
- or $|E_{i_j} \backslash E_{i_1} \cup \dots \cup E_{i_{j-1}}| = 1$, and the corresponding unique non-zero element $a_{i_j *}$ (which belongs to E_{i_j} but not to $E_{i_1} \cup \dots \cup E_{i_{j-1}}$) is different from ± 1.

If all of the E_i have size at most 1 and the corresponding non-zero elements (if any) are either 1 or -1, then we are done. Otherwise, choose a longest possible ill subsequence, and without loss of generality, assume that this sequence consists of $E_1, \dots, E_s$.

Claim 4.9. *We have*

$$s \leq r_0 + 1.$$

Proof. (of Claim 4.9) Observe that if the sequence $a_1, \dots, a_k \in \mathbf{R}$ has at least two non-zero entries or one non-zero entry different from ± 1, then $\mathbf{P}(\sum_i \varepsilon_i a_i \in a + \{-1, 1\}) \leq 1/2$. To complete the proof one just needs to rewrite (13) as product of conditional probabilities as in (12),

$$(1/2)/2^{r_0} \leq (1/2)^s.$$

□

We now use Claim 4.9 to complete the proof of Lemma 4.8. Let $s+1 \leq i \leq n^{2\varepsilon}$ be an arbitrary index. Then, as $E_1, \dots, E_s$ is longest possible, $|E_i \backslash E_1 \cup \dots \cup E_s| \leq 1$, and if equality holds then the corresponding non-zero element a_{i*} must be either 1 or -1. Set

$$I(A) := \cup_{1 \leq i \leq s} E_i.$$

Then $|I(A)| \leq s \times O_{r_0}(1) = O_{r_0}(1)$, completing the proof. □

To proceed further, we apply Lemma 4.8 to each $A = Y_{i|i_1,\dots,i_l}, i \in \mathcal{I}_g$, and set

$$I(M_0') := I_1 \cup \dots \cup I_l \cup_{i \in \mathcal{I}_g} I(A).$$

It is clear that this index set has size at most

$$|I(M_0')| \leq O(n^{3\varepsilon} \log n) + n_0 O_{r_0}(1) = O(n^{1-2\varepsilon}).$$

Putting together, we have obtained the following: for every vector $\mathbf{r}_{i_0}(M_0')$ with $i_0 \in \{(i-1)(r_0+n^{2\varepsilon})+r+1, \dots, i(r_0+n^{2\varepsilon}), 1 \leq i \leq n_0\}$, its restriction over the components indexing from $\bar{I}(M_0')$, $\mathbf{r}_{i_0}^{[\bar{I}(M_0')]}$, is either zero or contains exactly one element from ± 1. Recall the definition of M_0' from (7), we can restate the result in terms of M as follows.

Lemma 4.10. *There exists a set $I(M)(= I(M_0'))$ of exceptional indices with $|I(M)| = O(n^{1-2\varepsilon})$ such that the following holds for each $\mathbf{r}_{i_0}(M)$ with $i_0 \in \{(i-1)(r_0+n^{2\varepsilon})+r_0+1, \dots, i(r_0+n^{2\varepsilon}), 1 \leq i \leq n_0\}$: the row vectors $\mathbf{r}_{i_0}^{[\bar{I}(M)]} - \sum_{k=1}^{r_0} t_{i_0 k} \mathbf{r}_{(i-1)(r_0+n^{2\varepsilon})+k}^{[\bar{I}(M)]}$ are either zero or contain exactly one element from ± 1.*

To complete the proof of Theorem 4.2, we show that there is a projection on to $n-O(n^{1-\varepsilon})$ components for which the base vectors $\mathbf{r}_{(i-1)(r_0+n^{2\varepsilon})+k}, 1 \leq i \leq n_0, 1 \leq k \leq r_0$, generate a subspace of dimension at most r_0. For convenience, write $\mathbf{r}_{ik} := \mathbf{r}_{(i-1)(r_0+n^{2\varepsilon})+k}, 1 \leq i \leq n_0, 1 \leq k \leq r_0$. Lemma 3.2 (ii) applied to these m vectors, with $m = n_0 r_0$, implies that there exist s vectors $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_s}$ among the base vectors $\mathbf{r}_{ik}$, where $s \leq r_0$, and an index set $J = J(M)$ of size $|J| = O(mn^{\varepsilon}) = O(n^{1-\varepsilon})$ such that the following holds for any $\mathbf{r}_{ik}$:

$$\mathbf{r}_{ik}^{[\bar{J}]} \in span(\mathbf{r}_{i_1}^{[\bar{J}]}, \ldots, \mathbf{r}_{i_s}^{[\bar{J}]}).$$

Let N_1 be the set of row indices from $\{(i-1)(r_0+n^{2\varepsilon})+r_0+1, \ldots, i(r_0+n^{2\varepsilon}), 1 \leq i \leq n_0\}$, and N_2 be the set of column indices from $\bar{I}(M) \cap \bar{J}(M)$. The proof of Theorem 4.2 is completed by setting M' to be the submatrix of M generated by the rows indexing in N_1 and by the columns indexing in N_2.

4.11. **Proof of Theorem 2.6.** We will mainly be focusing on the matrix M' obtained in Theorem 4.2. Assume that $r = \mathrm{rank}(M'')$, where $r \leq r_0$. Assume also that $\mathbf{r}_{i_1}(M''), \ldots, \mathbf{r}_{i_r}(M'')$, $i_1, \ldots, i_r \in N_1$, span the whole row space of M''. Consider the corresponding row vectors $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_r}$ of M. Note that, by the definition of $\mathcal{F}_{n_1 n_2}$, when restricting $\mathbf{r}_{i_j}(M)$ to M', the vectors $\mathbf{r}_{i_j}(M')$ are different from $\mathbf{r}_{i_j}(M'')$ in at most one component. Let $I(N_2) \subset N_2$ be the set of column indices of the components where $\mathbf{r}_{i_j}(M')$ differ from $\mathbf{r}_{i_j}(M''), 1 \leq j \leq r$. Then $|I(N_2)| \leq r$.

Lemma 3.2 (i) applied to the vectors $\mathbf{r}_{i_1}(M), \ldots, \mathbf{r}_{i_r}(M)$ implies a GAP $Q_r \subset \mathbf{R}^r$ of small size and bounded rank and a small set $I_{i_1,\ldots,i_r}$ of exceptional indices such that that Q_r contains all restricted columns $\mathbf{c}_i^{[i_1,\ldots,i_r]}(M), i \in \bar{I}_{i_1,\ldots,i_r}$. In particular, Q_r contains all columns $\mathbf{c}_i^{[i_1,\ldots,i_r]}(M'')$, where $i \in N_2' := \bar{I}_{i_1,\ldots,i_r} \cap N_2 \cap \bar{I}(N_2)$. By passing to a GAP of smaller rank (and still of small size) if needed, one can assume that these restricted column vectors indeed span Q_r (see for instance [8, Section 8]), where we recall the notion of spanning from Section 3.

To this end, because $\mathbf{r}_{i_1}(M''), \ldots, \mathbf{r}_{i_r}(M'')$ span the whole row space of M'', we just follow the proof of Lemma 3.2 (i) identically to show that $Q_r \subset \mathbf{R}^r$ can be extended to another GAP $Q \subset \mathbf{R}^{n_1}$ of the same rank and size which contains all of the columns of M'' of indexing in N_2'. Finally, one deletes from M' the columns indexing in $\bar{N}_2'$ to obtain the new M', which clearly satisfies all of the desired properties (noting that columns deletion does not affect the property of $\mathcal{F}$.)

5. Treatment for orthogonal matrices: proof of Theorem 2.10

We first show that, for orthogonal matrices, the structures in Theorem 2.6 can be extended to the whole columns without significant increase of the sets of exceptional indices.

Theorem 5.1. *Let $0 < \varepsilon < 1$ and C be positive constants. Let $M = (m_{ij})_{1 \leq i,j \leq n} \in \mathbf{O}(n)$ be an orthogonal matrix with $s_0(M) \geq n^{-C}$ for some positive constant $C > 0$. Then there exist a submatrix M' of M of size $n \times n_2$, with $n_2 = n - O(n^{1-\varepsilon})$, and a set of $r = O(1)$*

vectors $\mathbf{g}_1, \dots, \mathbf{g}_r \in \mathbf{R}^n$ *such that* M' *can be written as* $M'' + F$*, where* $F \in \mathcal{F}_{nn_2}$ *and the columns of* M'' *belong to a GAP of small size generated by* $\mathbf{g}_1, \dots, \mathbf{g}_r$*.*

Proof. (of Theorem 5.1) We first apply Theorem 2.6. Without loss of generality, assume that M' consists of the first n_1 rows and n_2 columns of M, and $n_2 \le n_1$ (otherwise we just need to throw away a few columns so that $n_2 = n_1$). For $1 \le i \le n_2$, consider the $\mathbf{R}^{n_1}$ column vectors of M', $\mathbf{c}_i(M') = (m_{1i}, \dots, m_{n_1 i})$ and the $\mathbf{R}^{n-n_1}$ vectors $\mathbf{u}_i = (m_{(n_1+1)i}, \dots, m_{ni})$. Using the definition from the proof of Theorem 2.6, one can write

$$\mathbf{c}_i(M') = \mathbf{c}_i(M'') + \mathbf{c}_i(F),$$

where $r = \mathrm{rank}(M'') = O(1)$ and $F \in \mathcal{F}_{n_1 n_2}$.

Note that, as $\mathrm{rank}(F) \ge \mathrm{rank}(M') - \mathrm{rank}(M'') \ge n - O(n^{1-\varepsilon}) - r$, and as F contain at most $n - O(n^{1-\varepsilon})$ non-zero entries, by deleting at most $O(n^{1-\varepsilon})$ columns of M' if needed, one can assume that each column of F contains at most one non-zero entry (or exactly one, but we don't need this fact).

Similarly, as $\dim(span(\mathbf{c}_{i_1}(M), \dots, \mathbf{c}_{i_k}(M))) = k$, we have

$$\dim(span(\mathbf{c}_{i_1}(M'), \dots, \mathbf{c}_{i_k}(M'))) \ge k - r - (n - n_1) = k - r - O(n^{1-\varepsilon}). \tag{14}$$

Roughly speaking, our next move consists of two main steps.

(1) Starting from M', we showed that there exists an index set T of size $n - O(n^{1-\varepsilon})$ such that when restricting the low rank part M'' of M' onto T,
$$\mathbf{r}_i^{[T]}(M) \in span\big(\mathbf{r}_j^{[T]}(M''), 1 \le j \le n_2\big), \forall n_1 + 1 \le i \le n.$$

(2) Hence, by the argument from the proof of (i) of Lemma 3.2, the GAP containing the columns $\mathbf{c}_i(M''), i \in T$ can be extended to a GAP that contains the column vectors $\mathbf{c}_i(M), i \in T$.

We next explain these ideas in more detail.

Step 1. Consider the following process. At step 0, the row index set $S \subset [n_1]$ is set to be empty. At step $1 \le i$, if possible, we choose s column indices $i_1, \dots, i_s \in [n_2]$ that have never been used before, for some $s \le r + 1$, so that the following holds.

(i) The row indices of the (possibly) non-zero entries of $\mathbf{c}_{i_1}(F), \dots, \mathbf{c}_{i_s}(F)$ must not belong to S. It is possible that these columns do not contain any non-zero entries.

(ii) There exist real coefficients $\alpha_{i_1}, \dots, \alpha_{i_s}$ such that $\alpha_{i_1}\mathbf{c}_{i_1}(M'') + \dots + \alpha_{i_s}\mathbf{c}_{i_s}(M'') = \mathbf{0}$ but $\alpha_{i_1}\mathbf{u}_{i_1} + \dots + \alpha_{i_s}\mathbf{u}_{i_s} \ne \mathbf{0}$, where we recall that $\mathbf{u}_i = (m_{(n_1+1)i}, \dots, m_{ni})$.

We then add the row indices of the possibly non-zero entries of $\mathbf{c}_{i_1}(F), \dots, \mathbf{c}_{i_s}(F)$ to S and move on to the next step. The process terminates if we are not able to proceed further.

By definition, the linear combination $\sum_{j=1}^{s} \alpha_{i_j} \mathbf{c}_{i_j}(M)$ in each step equals $\sum_{j=1}^{s} \alpha_{i_j}(\mathbf{c}_{i_j}(F) \oplus \mathbf{u}_{i_j})$. But as the indices i_j are chosen to be disjoint from the previously-used indices $i_{j'}$, and as the columns of M are orthogonal,

$$(\sum_{j=1}^{s} \alpha_{i_j} \mathbf{c}_{i_j}(M)) \cdot (\sum_{j'=1}^{s'} \alpha_{i_{j'}} \mathbf{c}_{i_{j'}}(M)) = 0.$$

On the other hand, by (i) and by definition that each column of F contains at most one non-zero entry, $(\sum_{j=1}^{s} \alpha_{i_j} \mathbf{c}_{i_j}(F)) \cdot (\sum_{j'=1}^{s'} \alpha_{i_{j'}} \mathbf{c}_{i_{j'}}(F)) = 0$. This implies that

$$(\sum_{j=1}^{s} \alpha_{i_j} \mathbf{u}_{i_j}) \cdot (\sum_{j'=1}^{s'} \alpha_{i_{j'}} \mathbf{u}_{i_{j'}}) = 0. \tag{15}$$

By (15) and by the fact from (ii) that all the vectors $\sum_{j=1}^{s} \alpha_{i_j} \mathbf{u}_{i_j}$ are non-zero in $\mathbf{R}^{n-n_1}$, our process must terminate after at most $n - n_1 = O(n^{1-\varepsilon})$ steps. As such, the final index set S has size at most

$$|S| \leq (r+1)(n-n_1) = O(n^{1-\varepsilon}). \tag{16}$$

Now we consider the collection of columns $\mathbf{c}_i(M')$ of M' where the row index of the possibly non-zero entries of $\mathbf{c}_i(F)$ does not belong to S. Let $T \subset [n_2]$ be the collection of these indices. By (14), (16), and again by the assumption that each column of F contains at most one non-zero entry,

$$|T| = n - O(n^{1-\varepsilon}).$$

For the process cannot be continued, as by (ii), any vanishing linear combination $\alpha_{i_1} \mathbf{c}_{i_1}(M'') + \cdots + \alpha_{i_s} \mathbf{c}_{i_s}(M'') = \mathbf{0}, i_1, \ldots, i_s \in T$, also implies $\alpha_{i_1} \mathbf{u}_{i_1} + \cdots + \alpha_{i_s} \mathbf{u}_{i_s} = \mathbf{0}$. In other words,

$$\mathbf{r}_i^{[T]}(M) \in span\big(\mathbf{r}_j^{[T]}(M''), 1 \leq j \leq n_2\big), \forall n_1 + 1 \leq i \leq n, \tag{17}$$

where we recall that $\mathbf{r}_i^{[T]}(M)$ denote the projection of $\mathbf{r}_i(M)$ onto the components indexing in T, and similarly for $\mathbf{r}_i^{[T]}(M'')$.

Step 2. It follows from (17), by the same argument as in the proof of (i) of Lemma 3.2 (and also of Theorem 2.6), that the GAP containing the columns of M'', $\mathbf{c}_i(M''), i \in T$, can be extended to a GAP of the same rank and size that contains the column vectors $\mathbf{c}_i(M)$. We complete the proof by letting the new M' be the restriction of M onto the column index set T (and hence the new F is obtained from the old one by adding another $n - n_1$ zero rows). □

To prove Theorem 2.10, we need further preparations. By permuting the columns if necessary, one can assume that the first r columns of M'' (corresponding to M' of size $n \times n_2$ obtained from Theorem 5.1) span the whole columns of M''. To avoid trivial degeneracy, we will regularize the system further as follows.

Claim 5.2 (Regularization)**.** *With an extra loss of at most $O(n^{1-\varepsilon})$ in the number of columns of M', one can assume that every $n_2 - r$ rows (in $\mathbf{R}^r$) among n rows of the $n \times r$ matrix spanned by $\mathbf{c}_1(M''), \dots, \mathbf{c}_r(M'')$ have full rank.*

Proof. (of Claim 5.2) Assume otherwise that there are $n_2 - r$ rows which span a subspace of dimension at most $r-1$. We next restrict M' onto these row indices. By the assumption that the first r columns of the original M'' span its column space, the columns of the new (after restriction) M'' belongs to the span of its first r columns, and hence a subspace of dimension at most $r-1$.

By applying the steps (1) and (2) of the proof of Theorem 5.1, one obtains a new M' of size $n \times n_2'$, with $n_2' = n_2 - r - O(n^{1-\varepsilon})$, where the columns of the additive part M'' span a subspace of dimension at most $r-1$. In the next step, choose $r-1$ columns that span this subspace. We continue the process if there are $n_2' - (r-1)$ rows (in the submatrix generated by these $r-1$ columns) that span a subspace of dimension at most $r-2$ in $\mathbf{R}^{r-1}$, etc. As the process must terminate after at most r steps, the number of columns remaining at termination is at least $n - r \times O(n^{1-\varepsilon}) = n - O(n^{1-\varepsilon})$. □

With Claim 5.2 in hand, we now show that the additive part M'' can be described as in Theorem 2.10. As usual, we can simplify our matrix further as follows.

- By deleting the columns where F vanishes or has more than one non-zero entries, one can assume that each column of F in $M' = M'' + F$ contains exactly one non-zero entry, which can be assumed to be 1 after an appropriate column sign change.
- Also, by permuting the columns, one can assume that M' consists of the first n_2 columns of M, and the first r columns of M'' span the whole column vectors of M'.
- Finally, by permuting the rows, one can assume that the first n_2 columns of M take the form $\mathbf{u}_i + \mathbf{e}_i$, where $\mathbf{u}_i = (u_{1i}, \dots, u_{ni}) = \mathbf{c}_i(M'') \in \mathbf{R}^n$ belongs to a GAP Q of bounded rank and small size, and $\mathbf{e}_i$ are the standard vectors.

We restate Theorem 2.10 as follows.

Theorem 5.3. *For any $r+1 \le i \le n_2$, one can represent $\mathbf{u}_i$ as*

$$\mathbf{u}_i = d_{i1}\mathbf{u}_1 + \dots + d_{ir}\mathbf{u}_r,$$

where $d_{i1}, \dots, d_{ir}$ are uniquely determined from the first r column vectors $\mathbf{u}_1, \dots, \mathbf{u}_r$ by the formula

$$(u_{i1}, \dots, u_{ir}) = d_{i1}(u_{11}, \dots, u_{1r}) + \dots + d_{ir}(u_{r1}, \dots, u_{rr}).$$

In other words, the matrix M'', when restricting to the first n_2 rows, can be written as

$$\begin{pmatrix} U & UD^T \\ DU & DUD^T \end{pmatrix},$$

where U is a square matrix of size r, and D is an $(n_2 - r) \times r$ matrix.

Proof. (of Theorem 5.3) Because the first r columns span the whole column space of M'', for any $r + 1 \le i \le n_2$, there exist real numbers $x_1, \dots, x_r$ (not necessarily uniquely determined) such that $\mathbf{u}_i = x_1\mathbf{u}_1 + \cdots + x_r\mathbf{u}_r$. For any $1 \le k \le r$, the condition of orthogonality $(\mathbf{u}_i + \mathbf{e}_i) \cdot (\mathbf{u}_k + \mathbf{e}_k) = 0$ implies that

$$\begin{aligned} 0 &= \sum_{1 \le j \le r} x_j(\mathbf{u}_j \cdot \mathbf{u}_k + \mathbf{u}_j \cdot \mathbf{e}_k) + \mathbf{e}_i \cdot \mathbf{u}_k \\ &= -\sum_{1 \le j \le r} x_j \mathbf{e}_j \cdot \mathbf{u}_k + \mathbf{e}_i \cdot \mathbf{u}_k = (\mathbf{e}_i - \sum_{1 \le j \le r} x_j \mathbf{e}_j) \cdot \mathbf{u}_k, \end{aligned}$$

where we used the initial assumption that $(\mathbf{u}_j + \mathbf{e}_j) \cdot (\mathbf{u}_k + \mathbf{e}_k) = \delta_{jk}$ for $1 \le j \le r$.

On the other hand, we can rewrite $(\mathbf{e}_i - \sum_{1 \le j \le r} x_j \mathbf{e}_j) \cdot \mathbf{u}_k = 0, 1 \le k \le r$, as

$$(u_{i1}, \dots, u_{ir}) = x_1(u_{11}, \dots, u_{1r}) + \cdots + x_r(u_{r1}, \dots, u_{rr}). \tag{18}$$

In particular, the first r rows $(u_{11}, \dots, u_{1r}), \dots, (u_{r1}, \dots, u_{rr})$ span the whole row space of the $n_2 \times r$ submatrix spanned by $\mathbf{u}_1^{[1,\dots,n_2]}, \dots, \mathbf{u}_r^{[1,\dots,n_2]}$. By the assumption of Claim 5.2, these n_2 row vectors have full rank in $\mathbf{R}^r$, and so the representation in (18) is unique: the coefficients $(x_1, \dots, x_r)$ must equal $(d_{i1}, \dots, d_{ir})$ introduced in the statement.

□

6. Application: general matrices

As an application, we deduce from Theorem 2.6 that general matrices of sufficiently small entries cannot be near-invariant with respect to the hypercube.

Theorem 6.1. *For any $C > 0$ and $0 < \varepsilon < 1/2$, there exist $n_0 = n_0(C, \varepsilon)$ and $c = c(C, \varepsilon) > 0$ such that the following holds for all $n \ge n_0$. Let $M = (m_{ij})_{1 \le i,j \le n}$ be a matrix with $\mathrm{rank}(M) \ge (1/2 + \varepsilon)n$ and $|m_{ij}| \le c, \forall 1 \le i, j \le n$. Then $s_0(M) \le n^{-C}$.*[1]

We have not tried to sharpen the requirement of $\mathrm{rank}(M) \ge (1/2 + o(1))n$ here, but it can be easily seen that for Theorem 6.1, the rank of M must be sufficiently large. We also invite

[1]It seems possible that Theorem 6.1 could also be proved by using a result of Alon [4]. We thank an anonymous reviewer for this observation.

the reader to consult Appendix A for a related result with explicit constants for stochastic matrices of large permanent.

To prove Theorem 6.1, we need a simple claim stated below.

Claim 6.2. *Assume that $F \in \mathcal{F}_{n_1 n_2}$ with $n_1, n_2 = (1 - o(1))n$ and $\mathrm{rank}(F) \geq (1/2 + \varepsilon)n$. Then F contains a block of size $\varepsilon n \times \varepsilon n$ which contains exactly one non-zero entry in each row and column.*

Proof. (of Claim 6.2) Recall that each row of F is either zero or contains exactly one non-zero entry. Thus the total number of non-zero entries of F is at most n_1. As such, the number of columns of F that contain at least two non-zero entries is at most $n_1/2$. Consider the submatrix F' of F obtained by deleting these columns; then

$$\mathrm{rank}(F') \geq \mathrm{rank}(F) - n_1/2 \geq \varepsilon n.$$

By definition, each row and column of F' has at most one non-zero entry; thus F' contains a block of size $\varepsilon n \times \varepsilon n$ which contains exactly one non-zero entry in each row and column as desired. □

Proof. (of Theorem 6.1) Assume otherwise. After appropriate row and column permutations and sign changes, by Theorem 2.6 and Claim 6.2, one can assume that the top-left $n' \times n'$ corner of M, where $n' = \varepsilon n$, can be written as $U + F$, with F being the identity matrix $I_{n'}$ and the column vectors $\mathbf{u}_1, \dots, \mathbf{u}_{n'}$ of U belonging to a GAP P of small size and bounded rank r. In what follows we will not use this information in its full strength, but only a weaker fact, that the space generated by $\mathbf{u}_i$ has bounded dimension r (hence, Theorem 4.2 would suffice). Without loss of generality, assume that the first r columns $\mathbf{u}_1, \dots, \mathbf{u}_r$ span this subspace in $\mathbf{R}^{n'}$.

For any $r + 1 \leq i \leq n'$, there exist $c_{i1}, \dots, c_{ir}$ such that $\mathbf{u}_i = c_{i1}\mathbf{u}_1 + \dots + c_{ir}\mathbf{u}_r$. In particular, for $i = r + 1$,

$$\mathbf{u}_{r+1} = c_{(r+1)1}\mathbf{u}_1 + \dots + c_{(r+1)r}\mathbf{u}_r. \tag{19}$$

Thus, in the $(r+1)^{th}$ component,

$$u_{(r+1)(r+1)} = c_{(r+1)1}u_{(r+1)1} + \dots + c_{(r+1)r}u_{(r+1)r}. \tag{20}$$

By definition, as $|u_{(r+1)(r+1)} + 1| = |m_{(r+1)(r+1)}| \leq c$,

$$|u_{(r+1)(r+1)}| \geq 1 - c. \tag{21}$$

As the off-diagonal terms, $u_{(r+1)1} = m_{(r+1)1}, \dots, u_{((r+1)r} = m_{(r+1)r}$, all have absolute value at most c by assumption, we have

$$1 - c \leq c(|c_{(r+1)1}| + \cdots + |c_{(r+1)r}|) \leq rc \max\{|c_{(r+1)1}|, \ldots, |c_{(r+1)r}|\}.$$

Assume that the maximum above is achieved at some $1 \leq i_0 \leq r$,

$$|c_{(r+1)i_0}| = \max\{|c_{(r+1)1}|, \ldots, |c_{(r+1)r}|\} \geq (1-c)/rc. \tag{22}$$

Next, consider (19) in the i_0^{th} component,

$$u_{i_0(r+1)} = c_{((r+1)1} u_{i_0 1} + \cdots + c_{(r+1)r} u_{i_0 r}.$$

Equivalently,

$$\begin{aligned} -c_{(r+1)i_0} u_{i_0 i_0} = &- u_{i_0(r+1)} + c_{((r+1)1} u_{i_0 1} + \cdots + c_{(r+1)(i_0-1)} u_{i_0(i_0-1)} \\ &+ c_{(r+1)(i_0+1)} u_{i_0(i_0+1)} + \cdots + c_{(r+1)r} u_{i_0 r}. \end{aligned} \tag{23}$$

On the other hand, for the same reason as in (21), the diagonal term $u_{i_0 i_0}$ has large absolute value, $|u_{i_0 i_0}| \geq 1 - c$. Similarly, the off-diagonal terms, $u_{i_0 1} = m_{i_0 1}, \ldots, u_{i_0 r} = m_{i_0 r}, u_{i_0(r+1)} = m_{i_0(r+1)}$, all have absolute value at most c. It thus follows from (23) that

$$|c_{(r+1)i_0}|(1-c) \leq c + rc|c_{(r+1)i_0}|.$$

If c is chosen to be strictly smaller $1/(1+r)$, then this gives

$$|c_{(r+1)i_0}| \leq c/(1-(r+1)c).$$

However, this contradicts (22) when c is selected sufficiently small depending on r. □

7. Application: proof of Theorem 1.4

In this section we prove Theorem 1.4 by invoking Theorem 2.10. Assume that the matrix of the first r columns of M'' has the form (U, DU), where U is a non-singular square matrix of size r, and by Claim 5.2, one can also assume that $\mathrm{rank}(D) = r$.

The fact that the first r columns are orthogonal yields the following

$$(U + I_r)^T (U + I_r) + U^T D^T D U = I_r, \text{ or equivalently, } U + U^T + U^T U + U^T D^T D U = 0. \tag{24}$$

Thus $U + U^T$ is a negative semidefinite matrix of real entries, and the diagonal terms of DUD^T are non-positive because

$$\begin{aligned} -(DUD^T)_{ii} = -(DU^TD^T)_{ii} &= -(D(U+U^T)D^T)_{ii}/2 \\ &= [D(U^TU + U^TD^TDU)D^T]_{ii}/2, \end{aligned} \tag{25}$$

The latter is a sum of two positive semidefinite matrices of real entries.

Furthermore, (24) also implies that $U^{-1} + (U^T)^{-1} + I_r + D^TD = 0$. Let $A := U^{-1} - (U^T)^{-1}$ be the matrix difference. Then A is real asymmetric and $U^{-1} = [A - (I_r + D^TD)]/2$. Thus

$$U = -2(I_r + D^TD - A)^{-1}.$$

Using this formula for U, we show the following key trace-estimate.

Lemma 7.1. *We have*

$$|\mathrm{tr}(DUD^T)| \leq 2r.$$

Proof. (of Lemma 7.1) Using the formula $\mathrm{tr}(XY) = \mathrm{tr}(YX)$, write

$$\mathrm{tr}(DUD^T) = \mathrm{tr}(D^TDU) = -2\mathrm{tr}(D^TD(I_r + D^TD - A)^{-1}).$$

Let $V \in \mathbf{O}(r)$ be an orthogonal matrix which diagonalizes D^TD,

$$V(D^TD)V^T = E,$$

where $E = (\lambda_i)_{1 \leq i \leq r}$ is the diagonal matrix of eigenvalues of D^TD (where we recall that as $\mathrm{rank}(D) = r$, all the eigenvalues λ_i of D^TD are positive).

Plugging into the trace identity above,

$$\begin{aligned} \mathrm{tr}(D^TD(I_r + D^TD - A)^{-1}) &= \mathrm{tr}((V^TEV)(I_r + D^TD - A)^{-1}) = \mathrm{tr}(EV(I_r + D^TD - A)^{-1}V^T) \\ &= \mathrm{tr}(E(I_r + VD^TDV^T - VAV^T)^{-1}) = \mathrm{tr}(E(I_r + E - VAV^T)^{-1}) \\ &= \mathrm{tr}(E(I_r + E - B)^{-1}), \end{aligned}$$

where $B := VAV^T$ is a real asymmetric matrix of order r.

Next, let F be the diagonal matrix of positive entries $F = (\sqrt{\lambda_i})_{1 \leq i \leq r}$. Thus $E = FF^T$, and so

$$\begin{aligned}\mathrm{tr}(D^TD(I_r+D^TD-A)^{-1}) &= \mathrm{tr}(E(I_r+E-B)^{-1}) = \mathrm{tr}(FF^T(I_r+E-B)^{-1}) \\ &= \mathrm{tr}(F^T(I_r+E-B)^{-1}F) = \mathrm{tr}(I_r+E^{-1}-F^{-1}B(F^T)^{-1})^{-1}.\end{aligned}$$

Again, notice that the matrix $B' := F^{-1}B(F^T)^{-1}$ is another real asymmetric matrix, and the diagonal terms of the inverse matrix $E' = E^{-1}$ are positive. To this end, we introduce the following estimate.

Claim 7.2. *Assume that $E' = (e_i)$ is a diagonal matrix with positive entries, and $B' = (b_{ij})$ is a real asymmetric matrix, all of order r. Then*

$$0 \leq \mathrm{tr}(I_r + E' - B')^{-1} \leq r.$$

Proof. (of Claim 7.2) Let $\mu_i, 1 \leq i \leq r$, be the eigenvalues of $I_r + E' - B'$. Then

$$\mathrm{tr}(I_r + E' - B')^{-1} = \sum_i \mu_i^{-1} = \sum_i \mathrm{Re}\mu_i^{-1}.$$

On the other hand, for any eigenvalue μ with unit eigenvector $\mathbf{x} = (x_1, \ldots, x_r) \in \mathbf{C}^r$, the identity $(I_r + E' - B')\mathbf{x} = \mu\mathbf{x}$ implies that

$$\begin{aligned}\mu &= (I_r + E' - B')\mathbf{x}\cdot\bar{\mathbf{x}} = (I_r+E')\mathbf{x}\cdot\bar{\mathbf{x}} - B'\mathbf{x}\cdot\bar{\mathbf{x}} = 1 + \sum_i e_i|x_i|^2 - \sum_{i\neq j} b_{ij}x_i\bar{x}_j \\ &= 1 + \sum_i e_i|x_i|^2 - \sum_{1\leq i<j\leq r} b_{ij}(x_i\bar{x}_j - x_j\bar{x}_i).\end{aligned}$$

Because $b_{ij} \in \mathbf{R}$, the second summand is purely imaginary, and so $\mathrm{Re}\mu = 1 + \sum_i e_i|x_i|^2 \geq 1$. It thus follows that

$$0 \leq \mathrm{Re}\mu^{-1} \leq 1.$$

Summing over all μ_i, we hence obtain the claim.

□

It is clear that Claim 7.2 implies Lemma 7.1. The proof of this lemma is therefore complete. □

We now conclude the main result of this section.

Proof. (of Theorem 1.4) Let $k = n^{\varepsilon}$, and let K be the number of indices $i_0, 1 \leq i_0 \leq n_2 - r$, where $(DUD^T)_{i_0 i_0} \leq -k/n$. By (25) and Lemma 7.1,

$$K \leq \frac{2rn}{k} = 2rn^{1-\varepsilon}.$$

Thus there are at least $n_2 - r - O(n^{1-\varepsilon}) = n - O(n^{1-\varepsilon})$ indices i_0 such that

$$-k/n \leq (DUD^T)_{i_0 i_0} \leq 0.$$

With such i_0, by Theorem 2.10,

$$|(M'' + F)_{i_0 i_0}| = |(DUD^T)_{i_0 i_0} + (I)_{i_0 i_0}| \geq 1 - n^{-1+\varepsilon}.$$

□

8. Open problems

One obvious problem is to improve our result, to show that if M maps many points in $\mathcal{C}_n$ to points in $\mathcal{C}_n$, then M is close to a permuted diagonal matrix on all but $O(\log n)$ rows (rather than all but $O\left(n^{1-\varepsilon}\right)$ rows, as the current result gives).

A second problem is to generalize our result (regarding $s_0(M)$ [2]) from orthogonal matrices and real matrices to unitary matrices and general complex matrices.

A third problem is to prove analogous results constraining the form of near-isometries, but for objects other than hypercubes.

However, perhaps the most interesting problem is to generalize this paper's treatment from the "exact" score function $s_0(M)$ to the original score function $s(M)$, thereby answering questions 1.2 and 1.1 by the second named author and Hance (at least for the case of real and orthogonal matrices).

Acknowledgments. The authors are grateful to T. Tao, A. Arkhipov, S. Garg, and an anonymous reviewer for invaluable comments and suggestions.

Appendix A. Permaments of stochastic matrices

The goal of this section is to show a strong variant of Theorem 1.4 and Theorem 6.1 for (column) stochastic matrices of large permanent.

[2] In the complex setting, the exact score function $s_0(M)$ is the probability that $M\mathbf{x}$ lies exactly on the product of n unit circles, that is $s_0(M) = \mathbf{P}_{\mathbf{x} \in \mathcal{C}_n}(|(M\mathbf{x})_1| = \cdots = |(M\mathbf{x})_n| = 1)$.

Theorem A.1. *Let $A = (a_{ij})$ be an $n \times n$ stochastic matrix, and suppose $\mathrm{per}(A) \geq n^{-C}$. Then all but $O_C(\log n)$ of the rows of A contain an entry that is at least 0.8, with the remaining entries in that row summing to at most 0.1. (Of course, by stochasticity, these 0.8 entries must all lie in separate columns.)*

Proof. (of Theorem A.1) Let E be the event that, if we throw n balls independently into n bins, with the j^{th} ball thrown according to the probability distribution $\mathbf{P}(\text{bin } i) = a_{ij}$, then all n balls land in separate bins (i.e., there are no collisions). Then observe that $\mathrm{per}(A)$ is simply $\mathbf{P}(E)$.

Let $\mathbf{r}_i = (a_{ij})_j$ be the i^{th} row vector in A. Also let $\|\mathbf{r}_i\|_1 := \sum_j a_{ij}$. Observe that $\sum_i \|\mathbf{r}_i\|_1 = n$. In the balls-in-bins experiment, let B_i be the number of balls that land in the i^{th} bin. By definition, $\sum_i B_i = n$ and $\mathbf{E}(B_i) = \|\mathbf{r}_i\|_1$. Also, the event E holds only when $B_i = 1\ \forall i$.

Call the i^{th} row *little* if $\|\mathbf{r}_i\|_1 \leq 0.9$, and *splittable* if one can partition its entries into two parts, both of which sum to at least 0.1. Let $L, S \subseteq [n]$ be the sets of little and splittable rows respectively. Our strategy will be to show that

$$\mathbf{P}(E) \leq \exp\left(-\Omega\left(\max\left\{|L|, |S|\right\}\right)\right).$$

To begin with the little rows: for each $i \in L$, Markov's inequality implies that $\mathbf{P}(B_i \geq 1) \leq 0.9$. Furthermore, consider the events $B_i \geq 1$ for all i. We claim that, if we condition on any subset of these events occurring, then we can only *decrease*, not increase, the probability that other such events occur. To see this, note that each time we condition on a $B_i \geq 1$ event, what we are doing is eliminating the possible states in which no balls land in the i^{th} bin. However, the probability distribution remains symmetric among the bins j for which we have *not* conditioned on $B_j \geq 1$–we have merely eliminated a part of the distribution that had more balls in circulation among those bins. This can only decrease the probability of $B_j \geq 1$ for any such $j \neq i$. Another way of saying this is that the events $B_i \geq 1$ behave as a submartingale. So by Azuma's inequality, we have

$$\mathbf{P}(E) \leq \mathbf{P}(B_i \geq 1\ \forall i \in L) \leq \exp\left(-\frac{(0.1|L|)^2}{2|L|}\right) = \exp\left(-\frac{|L|}{200}\right),$$

where we also used the fact that

$$\mathbf{E}\left(\sum_{i \in L} B_i\right) = \sum_{i \in L} \|\mathbf{r}_i\|_1 \leq 0.9|L|.$$

For the splittable rows: for each $i \in S$, we claim that

$$\mathbf{P}(B_i \geq 2) \geq \left(1 - e^{-0.1}\right)^2 > 0.009.$$

The reason is that we can partition the n balls into two sets P and Q, both of which have at least 0.1 balls landing in the i^{th} bin in expectation. Because the balls are thrown independently, this implies that P (and likewise, Q) must have at least one ball landing in the i^{th} bin with probability at least $1 - e^{-0.1}$. Moreover, these events are independent between P and Q.

Now, by an analogous argument to the one above, the events $B_i \leq 1$ behave as a submartingale: conditioned on some of them occurring, we can only *decrease* the probability that others occur, by increasing the expected number of balls that are available for other bins. So by Azuma's inequality,

$$\mathbf{P}(E) \leq \mathbf{P}(B_i \leq 1 \ \forall i \in S) \leq \exp\left(-\frac{(0.009|S|)^2}{2|S|}\right) < \exp\left(-\frac{|S|}{25000}\right).$$

So, in conclusion, if $\mathrm{per}(A) = \mathbf{P}(E)$ is n^{-C}, then $|L|$ and $|S|$ must both be $O(C \log n)$. Now consider a row i that is neither little nor splittable. We have $\|\mathbf{r}_i\|_1 > 0.9$. Moreover, $\mathbf{r}_i$ must contain an entry j that is at least 0.8, since otherwise we could split $\mathbf{r}_i$, by setting $P = \{j\}$ and $Q = [n] \setminus \{j\}$.

□

References

[1] S. Aaronson, `http://mathoverflow.net/questions/172723/`.

[2] S. Aaronson and A. Arkhipov, The Computational Complexity of Linear Optics, Theory of Computing, Volume 9 (4), 2013, pp. 143-252.

[3] S. Aaronson and T. Hance, Generalizing and Derandomizing Gurvits's Approximation Algorithm for the Permanent, Quantum Information and Computation, 14 (7-8), 541-559, 2014.

[4] N. Alon, Perturbed identity matrices have high rank: proof and applications, Combinatorics, Probability and Computing 18 (2009), 3-15.

[5] P. Erdős, On a lemma of Littlewood and Offord, Bull. Amer. Math. Soc. 51 (1945), 898-902.

[6] D. Glynn, The permanent of a square matrix. European Journal of Combinatorics, 31(7):1887-1891, 2010.

[7] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, Advances in Mathematics, 226 (2011), no. 6, 5298-5319.

[8] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, Journal of American Mathematical Society, 20 (2007), 603-628.

[9] T. Tao and V. Vu, Inverse Littlewood-Offord theorems and the condition number of random matrices, Annals of Mathematics (2) 169 (2009), no. 2, 595-632.

[10] T. Tao and V. Vu, A sharp inverse Littlewood-Offord theorem, Random Structures Algorithms 37 (2010), no. 4, 525-539.

E-mail address: `aaronson@csail.mit.edu`

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139

E-mail address: `nguyen.1261@math.osu.edu`

Department of Mathematics, The Ohio State University, Columbus, OH 43210