Singularity of sparse Bernoulli matrices

Alexander E. Litvak and Konstantin E. Tikhomirov

Abstract

Let M_n be an $n \times n$ random matrix with i.i.d. Bernoulli(p) entries. We show that there is a universal constant $C \geq 1$ such that, whenever p and n satisfy $C \log n/n \leq p \leq C^{-1}$,

$$\begin{aligned}\mathbb{P}\big\{M_n \text{ is singular}\big\} &= (1+o_n(1))\mathbb{P}\big\{M_n \text{ contains a zero row or column}\big\} \\ &= (2+o_n(1))n\,(1-p)^n,\end{aligned}$$

where $o_n(1)$ denotes a quantity which converges to zero as $n \to \infty$. We provide the corresponding upper and lower bounds on the smallest singular value of M_n as well.

AMS 2010 Classification: primary: 60B20, 15B52; secondary: 46B06, 60C05.
Keywords: Littlewood–Offord theory, Bernoulli matrices, sparse matrices, smallest singular value, invertibility

Contents

1 Introduction

Invertibility of discrete random matrices attracts considerable attention in the literature. The classical problem in this direction — estimating the singularity probability of a square random matrix B_n with i.i.d. ± 1 entries — was first addressed by Komlós in the 1960-es. Komlós [21] showed that $\mathbb{P}\{B_n \text{ is singular}\}$ decays to zero as the dimension grows to infinity. A breakthrough result of Kahn–Komlós–Szemerédi [19] confirmed that the singularity probability of B_n is exponentially small in the dimension. Further improvements on the singularity probability were obtained by Tao–Vu [47, 48] and Bourgain–Vu–Wood [7]. An old conjecture states that $\mathbb{P}\{B_n \text{ is singular}\} = \big(\frac{1}{2} + o_n(1)\big)^n$. The conjecture was resolved in [51].

Other models of non-symmetric discrete random matrices considered in the literature include adjacency matrices of d-regular digraphs, as well as the closely related model of sums of independent uniform permutation matrices [22, 9, 10, 25, 26, 27, 28, 29, 2]. In particular, the recent breakthrough works [16, 38, 39] confirmed that the adjacency matrix of a uniform random d–regular digraph of a constant degree $d \geq 3$ is non-singular with probability decaying to zero as the number of vertices of the graph grows to infinity. A closely related line of research deals with the rank of random matrices over finite fields. We refer to [36] for some recent results and further references.

The development of the *Littlewood–Offord theory* and a set of techniques of geometric functional analysis reworked in the random matrix context, produced strong invertibility results for a broad class of distributions. Following works [50, 42] of Tao–Vu and Rudelson, the paper [44] of Rudelson and Vershynin established optimal small ball probability estimates for the smallest singular value in the class of square matrices with i.i.d. subgaussian entries, namely, it was shown that any $n \times n$ matrix A with i.i.d. subgaussian entries of zero mean and unit variance satisfies $\mathbb{P}\{s_{\min}(A) \leq t\, n^{-1/2}\} \leq Ct + 2\exp(-cn)$ for all $t > 0$ and some $C, c > 0$ depending only on the subgaussian moment. The assumptions of identical distribution of entries and of bounded subgaussian moment were removed in subsequent works [40, 33, 34]. This line of research lead to positive solution of the Bernoulli matrix conjecture mentioned in the first paragraph. Let us state the result of [51] for future reference.

Theorem (Invertibility of dense Bernoulli matrices, [51])**.**

- *For each n, let B_n be the $n \times n$ random matrix with i.i.d. ± 1 entries. Then for any $\varepsilon > 0$ there is C depending only on ε such that the smallest singular value $s_{\min}(B_n)$ satisfies*

$$\mathbb{P}\big\{s_{\min}(B_n) \leq tn^{-1/2}\big\} \leq Ct + C(1/2 + \varepsilon)^n, \quad t > 0.$$

In particular, $\mathbb{P}\big\{B_n \text{ is singular}\big\} = (1/2 + o_n(1))^n$, where the quantity $o_n(1)$ tends to zero as n grows to infinity.

- *For each $\varepsilon > 0$ and $p \in (0, 1/2]$ there is $C > 0$ depending on ε and p such that for any n and for random $n \times n$ matrix M_n with i.i.d. Bernoulli(p) entries,*

$$\mathbb{P}\big\{ s_{\min}(M_n) \le tn^{-1/2} \big\} \le Ct + C(1 - p + \varepsilon)^n, \quad t > 0.$$

In particular, for a fixed $p \in (0, 1/2]$, we have $\mathbb{P}\big\{M_n \text{ is singular}\big\} = (1 - p + o_n(1))^n$.

Sparse analogs of the Rudelson–Vershynin invertibility theorem [44] were obtained, in particular, in works [49, 14, 32, 3, 4, 5], with the strongest small ball probability estimates in the i.i.d. subgaussian setting available in [3, 4, 5]. Here, we state a result of Basak–Rudelson [3] for Bernoulli(p_n) random matrices.

Theorem (Invertibility of sparse Bernoulli matrices, [3])**.** *There are universal constants $C, c > 0$ with the following property. Let $n \in \mathbb{N}$ and let $p_n \in (0, 1)$ satisfy $C \log n/n \le p_n \le 1/2$. Further, let M_n be the random $n \times n$ matrix with i.i.d. Bernoulli(p_n) entries (that is, 0/1 random variables with expectation p_n). Then*

$$\mathbb{P}\big\{ s_{\min}(M_n) \le t \exp\big(- C \log(1/p_n)/ \log(np_n)\big) \sqrt{p_n/n} \big\} \le Ct + 2\exp(-cnp_n), \quad t > 0.$$

The singularity probabilities implied by the results [51, 3] may be regarded as suboptimal in a certain respect. Indeed, while [51] produced an asymptotically sharp base of the power in the singularity probability of B_n, the estimate of [51] is off by a factor $(1 + o_n(1))^n$ which may (and in fact does, as analysis of the proof shows) grow to infinity with n superpolynomially fast. Further, the upper bound on the singularity probability of sparse Bernoulli matrices implied by [3] captures an exponential dependence on np_n, but does not recover an asymptotically optimal base of the power.

A folklore conjecture for matrices B_n asserts that $\mathbb{P}\{B_n \text{ is singular}\} = (1 + o_n(1))n^2 2^{1-n}$, where the right hand side of the expression is the probability that two rows or two columns of the matrix B_n are equal up to a sign (see, for example, [19]). This conjecture can be naturally extended to the model with Bernoulli(p_n) (0/1) entries as follows.

Conjecture 1.1 (Stronger singularity conjecture for Bernoulli matrices)**.** *For each n, let $p_n \in (0, 1/2]$, and let M_n be the $n \times n$ matrix with i.i.d. Bernoulli(p_n) entries. Then*

$$\begin{aligned}&\mathbb{P}\{M_n \text{ is singular}\}\\ &\quad = (1 + o_n(1))\mathbb{P}\big\{\text{a row or a column of } M_n \text{ equals zero, or two rows or columns are equal}\big\}.\end{aligned}$$

In particular, if $\limsup p_n < 1/2$ then

$$\mathbb{P}\{M_n \text{ is singular}\} = (1 + o_n(1))\mathbb{P}\big\{\text{either a row or a column of } M_n \text{ equals zero}\big\}.$$

Conceptually, the above conjecture asserts that the main causes for singularity are local in the sense that the linear dependencies typically appear within small subsets of rows or columns. In a special regime $np_n \le \ln n + o_n(\ln \ln n)$, the conjecture was positively resolved in [5] (note that if $np_n \le \ln n$ then the matrix has a zero row with probability at least $1 - 1/e - o_n(1)$). However, the regime $\liminf(np_n/ \log n) > 1$ was not covered in [5].

The main purpose of our paper is to develop methods capable of capturing the singularity probability with a sufficient precision to answer the above question. Interestingly, this appears to be more accessible in the sparse regime, when p_n is bounded above by a small universal constant (we discuss this in the next section in more detail). It is not difficult to show that when $\liminf(np_n/ \ln n) > 1$, the events that a given row or a given column equals zero, almost do not intersect, so that

$$\mathbb{P}\big\{\text{either a row or a column of } M_n \text{ equals zero}\big\} = (2 + o_n(1))n\,(1 - p_n)^n.$$

Our main result can be formulated as follows.

Theorem 1.2. *There is a universal constant $C \geq 1$ with the following property. Let $n \geq 1$ and let M_n be an $n \times n$ random matrix such that*

$$\text{The entries of } M_n \text{ are i.i.d. Bernoulli}(p), \text{ with } p = p_n \text{ satisfying } C \ln n \leq np \leq C^{-1} n. \tag{\textbf{A}}$$

Then

$$\mathbb{P}\big\{M_n \text{ is singular}\big\} = (2 + o_n(1)) n (1-p)^n,$$

where $o_n(1)$ is a quantity which tends to zero as $n \to \infty$. Moreover, for every $t > 0$,

$$\mathbb{P}\big\{s_{\min}(M_n) \leq t \exp(-3\ln^2(2n))\big\} \leq t + (1 + o_n(1))\mathbb{P}\big\{M_n \text{ is singular}\big\} = t + (2 + o_n(1)) n (1-p)^n.$$

In fact, our approach gives much better estimates on $s_{\min}$ in the regime when p_n is constant, see Theorem 7.1 below. At the same time, we note that obtaining small ball probability estimates for $s_{\min}$ was not the main objective of this paper, and the argument was not fully optimized in that respect.

Geometrically, the main result of our work asserts that (under appropriate assumptions on p_n) the probability that a collection of n independent random vectors $X_1^{(n)}, \dots, X_n^{(n)}$ in $\mathbb{R}^n$, with i.i.d Bernoulli(p_n) components is *linearly* dependent, is equal (up to $(1 + o_n(1))$ factor) to probability of the event that either $X_i^{(n)}$ is zero for some $i \leq n$ or $X_1^{(n)}, \dots, X_n^{(n)}$ are contained in the same coordinate hyperplane:

$$\begin{aligned}\mathbb{P}\big\{X_1^{(n)}, \dots, X_n^{(n)} \text{ are linearly dependent}\big\} &= (1 + o_n(1))\, \mathbb{P}\big\{X_i^{(n)} = \mathbf{0} \text{ for some } i \leq n\big\} \\ &+ (1 + o_n(1))\, \mathbb{P}\big\{\exists \text{ a coordinate hyperplane } H \text{ such that } X_i^{(n)} \in H \text{ for all } i \leq n\big\}.\end{aligned}$$

Thus, the linear dependencies between the vectors, when they appear, typically have the prescribed structure, falling into one of the two categories described above with the (conditional) probability $\frac{1}{2} + o_n(1)$.

The paper is organized as follows. In the next section, we give an overview of the proof of the main result. In Section 3, we gather some preliminary facts and important notions to be used later. In Section 4, we consider new anti-concentration inequalities for random 0/1 vectors with prescribed number of non-zero components, and introduce a functional (the *u-degree* of a vector) which enables us to classify vectors on the sphere according to anti-concentration properties of inner products with the random 0/1 vectors. In the same section, we prove a key technical result — Theorem 2.2 — which states, roughly speaking, that with very high probability a random unit vector orthogonal to $n-1$ columns of M_n is either close to being sparse or to being a constant multiple of $(1, 1, \dots, 1)$, or the vector is *very unstructured,* i.e., has a very large u-degree.

In Section 5, we consider a special regime of constant probability of success p. In this regime, estimating the event that M_n has an "almost null" vector which is either close to sparse or almost constant, is relatively simple. The reader who is interested only in the regime of constant p can thus skip the more technical Section 6 and have the proof of the main result as a combination of the theorems in Sections 4 and 5. In Section 6, we consider the entire range for p. Here, the treatment of "almost null" vectors which are either almost constant or close to sparse, is much more challenging and involves a careful analysis of multiple cases. Finally, in Section 7 we establish an *invertibility via distance* lemma and prove the main result of the paper. Some further questions are discussed in Section 8.

2 Overview of the proof

In this section, we provide a high-level overview of the proof; technical details will be discussed further in the text. The proof utilizes some known approaches to the matrix invertibility, which involve, in particular, a decomposition of the space into *structured* and *unstructured* parts, a form of *invertibility via distance* argument, small ball probability estimates based on the Esseen lemma, and various forms of the ε–net

argument. The novel elements of the proof are anti-concentration inequalities for random vectors with a prescribed cardinality of the support, a structural theorem for normals to random hyperplanes spanned by vectors with i.i.d. Bernoulli(p) components, and a sharp analysis of the matrix invertibility over the set of structured vectors. We will start the description with our use of the partitioning trick, followed by a modified invertibility-via-distance lemma, and then consider the anti-concentration inequality and the theorem for normals (Subsection 2.1) as well as invertibility over the structured vectors (Subsection 2.2).

The use of decompositions of the space $\mathbb{R}^n$ into structured and unstructured vectors has become rather standard in the literature. A common idea behind such partitions is to apply the Littlewood–Offord theory to analyse the unstructured vectors and to construct a form of the ε–net argument to treat the structured part. Various definitions of structured and unstructured have been used in works dealing with the matrix invertibility. One of such decomposition was introduced in [31] and further developed in [44]. In this splitting the structured vectors are *compressible*, having a relatively small Euclidean distance to the set of *sparse* vectors, while the vectors in the complement are *incompressible*, having a large distance to sparse vectors and, as a consequence, many components of roughly comparable magnitudes. In our work, the decomposition of $\mathbb{R}^n$ is closer to the one introduced in [27, 30].

Let x^* denote a non-increasing rearrangement of absolute values of components of a vector x, and let $r, \delta, \rho \in (0,1)$ be some parameters. Further, let $\mathbf{g}$ be a non-decreasing function from $[1,\infty)$ into $[1,\infty)$; we shall call it *the growth function*. At this moment, the choice of the growth function is not important; we can assume that $\mathbf{g}(t)$ grows roughly as $t^{\ln t}$. Define the set of *gradual non-constant vectors* as

$$\begin{aligned}\mathcal{V}_n = \mathcal{V}_n(r,\mathbf{g},\delta,\rho) := \big\{ x \in \mathbb{R}^n \,:\, x^*_{\lfloor rn \rfloor} = 1,\ x^*_i \le \mathbf{g}(n/i) \text{ for all } i \le n, \quad \text{and} \\ \exists\, Q_1, Q_2 \subset [n] \quad \text{such that} \quad |Q_1|, |Q_2| \ge \delta n \quad \text{and} \quad \max_{i\in Q_2} x_i \le \min_{i \in Q_1} x_i - \rho \big\}. \end{aligned} \tag{1}$$

In a sense, constant multiples of the gradual non-constant vectors occupy most of the space $\mathbb{R}^n$, they play role of the unstructured vectors in our argument. By negation, the structured vectors,

$$\mathcal{S}_n = \mathcal{S}_n(r,\mathbf{g},\delta,\rho) := \mathbb{R}^n \setminus \bigcup_{\lambda \ge 0} (\lambda\, \mathcal{V}_n(r,\mathbf{g},\delta,\rho)), \tag{2}$$

are either almost constant (with most of components nearly equal) or have a very large ratio of x^*_i and $x^*_{\lfloor rn \rfloor}$ for some $i < rn$.

For simplicity, we only discuss the problem of singularity at this moment. As M_n and $M_n^\top$ are equidistributed, to show that $\mathbb{P}\big\{M_n \text{ is singular}\big\} = (2+o_n(1))n\,(1-p)^n$, it is sufficient to verify that

$$\mathbb{P}\Big(\big\{ M_n x = 0 \text{ for some } x \in \mathcal{V}_n \big\} \cap \Big\{ M_n^\top x \ne 0 \text{ for all } x \in \mathcal{S}_n \Big\} \Big) = o_n(n)\,(1-p)^n, \tag{3}$$

and

$$\mathbb{P}\Big\{ M_n x = 0 \text{ for some } x \in \mathcal{S}_n \Big\} = (1+o_n(1))n\,(1-p)^n.$$

The first relation is dealt with by using a variation of the *invertibility via distance* argument which was introduced in [44] to obtain sharp small ball probability estimates for the smallest singular value. In the form given in [44], the argument reduces the problem of invertibility over unstructured vectors to estimating distances of the form $\operatorname{dist}(\mathbf{C}_i(M_n), H_i(M_n))$, where $\mathbf{C}_i(M_n)$ is the i–th column of M_n, and $H_i(M_n)$ is the linear span of columns of M_n except for the i–th. In our setting, however, the argument needs to be modified to pass to estimating the distance *conditioned* on the size of the support of the column, as this allows using much stronger anti-concentration inequalities (see the following subsection). By the invariance of the distribution of M_n under permutation of columns, it can be shown that in order to prove the relation (3), it is enough to verify that

$$\mathbb{P}\big\{ |\operatorname{supp} \mathbf{C}_1(M_n)| \in [\tfrac{pn}{8}, 8pn] \text{ and } \langle \mathbf{Y}, \mathbf{C}_1(M_n)\rangle = 0 \text{ and } \mathbf{Y}/\mathbf{Y}^*_{\lfloor rn \rfloor} \in \mathcal{V}_n \big\} = o_n(n)\,(1-p)^n, \tag{4}$$

where $\mathbf{Y}$ is a non-zero random vector orthogonal to and measurable with respect to $H_1(M_n)$ (see Lemma 7.4 and the beginning of the proof of Theorem 1.2). In this form, the question can be reduced to studying the anti-concentration of the linear combinations $\sum_{i=1}^n \mathbf{Y}_i b_i$, where the Bernoulli random variables $b_1, \dots, b_n$ are mutually independent with $\mathbf{Y}$ and conditioned to sum up to a fixed number in $[pn/8, 8pn]$. This intermediate problem is discussed in the next subsection.

The approach to the set of structured vectors, $\mathcal{S}_n$, will be discussed in Subsection 2.2.

2.1 New anti-concentration inequalities for random vectors with prescribed support cardinality

The *Littlewood–Offord theory* — the study of anti-concentration properties of random variables — has been a crucial ingredient of many recent results on invertibility of random matrices, starting with the work of Tao–Vu [50]. In particular, the breakthrough result [44] of Rudelson–Vershynin mentioned in the introduction, is largely based on studying the Lévy function $\mathcal{Q}(\langle \mathbf{C}_1(A), \mathbf{Y}\rangle, t)$, with $\mathbf{C}_1(A)$ being the first column of the random matrix A and $\mathbf{Y}$ — a random unit vector orthogonal to the remaining columns of A.

We recall that given a random vector X taking values in $\mathbb{R}^n$, the *Lévy concentration function* $\mathcal{Q}(X, t)$ is defined by

$$\mathcal{Q}(X, t) := \sup_{y \in \mathbb{R}^n} \mathbb{P}\big\{\|X - y\| \le t\big\}, \quad t \ge 0;$$

in particular for a scalar random variable ξ we have $\mathcal{Q}(\xi, t) := \sup_{\lambda \in \mathbb{R}} \mathbb{P}\{|\xi - \lambda| \le t\}$. A common approach is to determine structural properties of a fixed vector which would imply desired upper bounds on the Lévy function of its scalar product with a random vector (say, a matrix' column). The classical result of Erdős–Littlewood–Offord [12, 24] asserts that whenever X is a vector in $\mathbb{R}^n$ with i.i.d. ± 1 components, and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ is such that $|y_i| \ge 1$ for all i, we have

$$\mathcal{Q}(\langle X, y\rangle, t) \le C t\, n^{-1/2} + C n^{-1/2},$$

where $C > 0$ is a universal constant. It can be further deduced from the Lévy–Kolmogorov–Rogozin inequality [41] that the above assertion remains true whenever X is a random vector with independent components X_i satisfying $\mathcal{Q}(X_i, c) \le 1 - c$ for some constant $c > 0$. More delicate structural properties, based on whether components of y can be embedded into a generalized arithmetic progression with prescribed parameters were employed in [50] to prove superpolynomially small upper bounds on the singularity probability of discrete random matrices.

The Least Common Denominator (LCD) of a unit vector introduced in [44] played a central role in establishing the exponential upper bounds on the matrix singularity under more general assumptions on the entries' distributions. We recall that the LCD of a unit vector y in $\mathbb{R}^n$ can be defined as

$$\mathrm{LCD}(y) := \inf\big\{\theta > 0 :\ \mathrm{dist}(\theta y, \mathbb{Z}^n) \le \min(c_1 \|\theta y\|, c_2\sqrt{n})\big\}$$

for some parameters $c_1, c_2 \in (0, 1)$. The small ball probability theorem of Rudelson and Vershynin [44] states that given a vector X with i.i.d. components of zero mean and unit variance satisfying some additional mild assumptions,

$$\mathcal{Q}(\langle X, y\rangle, t) \le Ct + \frac{C'}{\mathrm{LCD}(y)} + 2e^{-c'n} \tag{5}$$

for some constants $C, C', c' > 0$ (see [45] for a generalization of the statement). The LCD, or its relatives, were subsequently used in studying invertibility of non-Hermitian square matrices under broader assumptions [40, 33, 34], and delocalization of eigenvectors of non-Hermitian random matrices [46, 37, 35], among many other works.

Anti-concentration properties of random linear combinations naturally play a central role in the current work, however, the measures of unstructuredness of vectors existing in the literature do not allow to obtain the precise estimates we are aiming for. Here, we develop a new functional for dealing with linear combinations of *dependent* Bernoulli variables.

Given $n \in \mathbb{N}$, $1 \le m \le n/2$, a vector $y \in \mathbb{R}^n$ and parameters $K_1, K_2 \ge 1$, we define the *degree of unstructuredness (u-degree)* of vector y by

$$\mathbf{UD}_n(y, m, K_1, K_2) := \sup\Big\{ t > 0 : \; A_{nm} \sum_{S_1,\dots,S_m} \int_{-t}^{t} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E} \exp\big(2\pi \mathbf{i}\, y_{\eta[S_i]}\, m^{-1/2} s\big)\big|\big)\, ds \le K_1 \Big\}, \quad (6)$$

where the sum is taken over all sequences $(S_i)_{i=1}^m$ of disjoint subsets $S_1, \dots, S_m \subset [n]$, each of cardinality $\lfloor n/m \rfloor$ and

$$A_{nm} = \frac{\big((\lfloor n/m \rfloor)!\big)^m\, (n - m\lfloor n/m \rfloor)!}{n!}. \quad (7)$$

Here $\eta[S_i]$, $i \le m$, denote mutually independent integer random variables uniformly distributed on respective S_i's. The function ψ_{K_2} in the definition acts as a smoothing of $\max(\frac{1}{K_2}, t)$, with $\psi_{K_2}(t) = \frac{1}{K_2}$ for all $t \le \frac{1}{2K_2}$ and $\psi_{K_2}(t) = t$ for all $t \ge \frac{1}{K_2}$ (we prefer to skip discussion of this purely technical element of the proof in this section, and refer to the beginning of Section 4 for the full list of conditions imposed on ψ_{K_2}).

The functional $\mathbf{UD}_n(y, m, K_1, K_2)$ can be understood as follows. The expression inside the supremum is the average value of the integral

$$\int_{-t}^{t} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E} \exp\big(2\pi \mathbf{i}\, y_{\eta[S_i]}\, m^{-1/2} s\big)\big|\big)\, ds,$$

with the average taken over all choices of sequences $(S_i)_{i=1}^m$. The function under the integral, disregarding the smoothing ψ_{K_2}, is the absolute value of the characteristic function of the random variable $\langle y, Z\rangle$, where Z is a random 0/1–vector with exactly m ones, and with the i-th one distributed uniformly on S_i. A relation between the magnitude of the characteristic function and anti-concentration properties of a random variable (the Esseen lemma, see Lemma 3.12 below) has been commonly used in works on the matrix invertibility (see, for example, [43]), and determines the shape of the functional $\mathbf{UD}_n(\cdot)$. The definition of the u-degree is designed specifically to work with random 0/1–vectors having a fixed sum (equal to m). The next statement follows from the definition of $\mathbf{UD}_n(\cdot)$ and the Esseen lemma.

Theorem 2.1 (A Littlewood–Offord-type inequality in terms of the u-degree)**.** *Let m, n be positive integers with $m \le n/2$, and let $K_1, K_2 \ge 1$. Further, let $v \in \mathbb{R}^n$, and let $X = (X_1, \dots, X_n)$ be a random 0/1–vector in $\mathbb{R}^n$ uniformly distributed on the set of vectors with m ones and $n - m$ zeros. Then*

$$\mathcal{Q}\Big(\sum_{i=1}^{n} v_i X_i, \sqrt{m}\,\tau\Big) \le C_{2.1}\left(\tau + \mathbf{UD}_n(v, m, K_1, K_2)^{-1}\right) \quad \text{for all } \tau > 0,$$

where $C_{2.1} > 0$ may only depend on K_1.

The principal difference distinguishing the u-degree and the above theorem from the notion of the LCD and (5) is that the former allows one to obtain stronger anti-concentration inequalities in the same regime of sparsity, assuming that the coefficient vector y is sufficiently unstructured. In fact, under certain conditions, *sparse random 0/1 vectors with prescribed support cardinality admit stronger anti-concentration inequalities compared to the i.i.d. model.*

The last principle can be illustrated by taking the coefficient vector y as a "typical" vector on the sphere S^{n-1}. First, assume that $b_1, \dots, b_n$ are i.i.d. Bernoulli(p) , with $p < 1/2$. Then it is easy to see that for almost all (with respect to normalized Lebesgue measure) vectors $y \in S^{n-1}$,

$$\mathcal{Q}\Big(\sum_{i=1}^n y_i b_i, 0\Big) = (1-p)^n.$$

In words, for a typical coefficient vector y on the sphere, the linear combination $\sum_{i=1}^n y_i b_i$ takes distinct values for any two distinct realizations of $(b_1, \dots, b_n)$, and thus the Lévy function at zero is equal to the probability measure of the largest atom of the distribution of $\sum_{i=1}^n y_i b_i$ which corresponds to all b_i equal to zero. In contrast, if the vector $(b_1, \dots, b_n)$ is uniformly distributed on the set of 0/1–vectors with support of size $d = pn$, then for almost all $y \in S^{n-1}$, the random sum $\sum_{i=1}^n y_i b_i$ takes $\binom{n}{d}$ distinct values. Thus,

$$\mathcal{Q}\Big(\sum_{i=1}^n v_i b_i, 0\Big) = \binom{n}{np}^{-1},$$

where $\binom{n}{np}^{-1} \ll (1-p)^n$ for small p.

The above example provides only qualitative estimates and does not give an information on the location of the atoms of the distribution of $\sum_{i=1}^n y_i b_i$. The notion of the u-degree addresses this problem. The following theorem, which is the main result of Section 4, asserts that with a very large probability the normal vector to the (say, last) $n-1$ columns of our matrix M_n is either very structured or has a very large u-degree, much greater than the critical value $(1-p)^{-n}$.

Theorem 2.2. *Let $r, \delta, \rho \in (0,1)$, $s > 0$, $R \geq 1$, and let $K_3 \geq 1$. Then there are $n_0 \in \mathbb{N}$, $C \geq 1$ and $K_1 \geq 1$, $K_2 \geq 4$ depending on $r, \delta, \rho, R, s, K_3$ such that the following holds. Let $n \geq n_0$, $p \leq C^{-1}$, and $s \ln n \leq pn$. Let $\mathbf{g} : [1, \infty) \to [1, \infty)$ be an increasing (growth) function satisfying*

$$\forall a \geq 2\ \forall t \geq 1: \quad \mathbf{g}(a\,t) \geq \mathbf{g}(t) + a \qquad \text{and} \qquad \prod_{j=1}^{\infty} \mathbf{g}(2^j)^{j\,2^{-j}} \leq K_3. \tag{8}$$

Assume that M_n is an $n \times n$ Bernoulli(p) random matrix. Then with probability at least $1 - \exp(-Rpn)$ one has

$$\{\textit{Set of normal vectors to } \mathbf{C}_2(M_n), \dots, \mathbf{C}_n(M_n)\} \cap \mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \subset$$
$$\{x \in \mathbb{R}^n : \ x^*_{\lfloor rn \rfloor} = 1, \ \ \mathbf{UD}_n(x, m, K_1, K_2) \geq \exp(Rpn) \ \ \textit{for all} \ \ pn/8 \leq m \leq 8pn\}.$$

We would like to emphasize that the parameter s in this theorem can take values less than one, in the regime when the matrix M_n typically has null rows and columns. In this respect, the restriction $p \geq C \ln n/n$ in the main theorem comes from the treatment of structured vectors.

The proof of Theorem 2.2 is rather involved, and is based on a double counting argument and specially constructed lattice approximations of the normal vectors. We refer to Section 4 for details. Here, we only note that, by taking R as a sufficiently large constant, the theorem implies the relation (4), hence, accomplishes the treatment of unstructured vectors.

2.2 Almost constant, steep and $\mathcal{R}$-vectors

In this subsection we discuss our treatment of the set of structured vectors, $\mathcal{S}_n$. In the proof we partition the set $\mathcal{S}_n$ into several subsets and work with them separately. In a simplistic form, the structured vectors are dealt with in two ways: either by constructing discretizations and taking the union bound (variations

of the ε–net argument), or via deterministic estimates in the case when there are only few very large components in the vector. We note here that the discretization procedure has to take into account the non-centeredness of our random matrix model: while in case of centered matrices with i.i.d. components (and under appropriate moment conditions) the norm of the matrix is typically of order $\sqrt{n}$ times the standard deviation of an entry, for our Bernoulli(p) model it has order pn (i.e., roughly $\sqrt{p}n$ times the standard deviation of an entry), which makes a direct application of the ε–net argument impossible. Fortunately, this large norm is attained only in one direction — the direction of the vector $\mathbf{1} = (1, 1, \dots, 1)$ while on the orthogonal complement of $\mathbf{1}$ the typical norm is $\sqrt{pn}$. Therefore it is enough to take a standard net in the Euclidean norm and to make it denser in that one direction, which almost does not affect the cardinality of the net. We refer to Section 3.6 for details.

Let us first describe our approach in the (simpler) case when $p \in (q, c)$, where c is a small enough absolute constant and $q \in (0, c)$ is a fixed parameter (independent of n). We introduce four auxiliary sets and show that the set of unit structured vectors, $\mathcal{S}_n \cap S^{n-1}$, is contained in the closure of their union.

The first set, $\mathcal{B}_1$, consists of unit vectors close to vectors of the canonical basis, specifically, unit vectors x satisfying $x_1^* > 6pnx_2^*$, where x^* denotes the non-inreasing rearrangement of the vector $(|x_i|)_{i\le n}$. For any such vector x the individual bound is rather straightforward — conditioned on the event that there are no zero columns in our matrix M, and that the Euclidean norms of the matrix rows are not too large, we get $Mx \neq 0$. This class is the main contributor to the bound $(1 + o_n(1))n(1-p)^n$ for non-invertibility over the structured vectors $\mathcal{S}_n$.

For the other three sets we use anti-concentration probability estimates and discretizations. An application of Rogozin's lemma (Proposition 3.9) implies that probability of having small inner product of a given row of our matrix with x is small, provided that there is a subset $A \subset [n]$ such that the maximal coordinate of $P_A x$ is bounded above by $c\sqrt{p}\|P_A x\|$, where $\|\cdot\|$ denotes the standard Euclidean norm and P_A is the coordinate projection onto $\mathbb{R}^A$. Combined with the tensorization Lemma 3.8 this implies exponentially (in n) small probability of the event that $\|Mx\|$ is close to zero — see Proposition 3.10 below. Specifically, we define $\mathcal{B}_2$ as the set of unit vectors satisfying the above condition with $A = [n]$, that is, satisfying $x_1^* \le c\sqrt{p}$, and for $\mathcal{B}_3$ we take all unit vectors satisfying the condition with $A = \sigma_x([2, n])$, that is, satisfying $x_2^* \le c\sqrt{p}\|P_{\sigma_x([2,n])}x\|$, where σ_x is a permutation satisfying $x_i^* = |x_{\sigma_x(i)}|$, $i \le n$. For vectors from these two sets we have very good individual probability estimates, but, unfortunately, the complexity of both sets is large — they don't admit nets of small cardinality. To overcome this issue, we have to redefine these sets by intersecting them with specially chosen sets of vectors having many almost equal coordinates. For the precise definition of such sets, denoted by $U(m, \gamma)$, see Subsection 3.6. A set $U(m, \gamma)$ is a variant of the class of *almost constant* vectors, $\mathcal{AC}(\rho)$ (see (9) below), introduced to deal with general p. Having a large part of coordinates of a vector almost equal to each other reduces the complexity of the set making possible to construct a net of small cardinality. This resolves the problem and allows us to deal with these two classes of sets. The remaining class of vectors, $\mathcal{B}_4$, consists of vectors x with $x_1^* \ge x_2^* \ge c\sqrt{p}\|P_{\sigma_x([2,n])}x\|$, i.e., vectors with relatively big two largest components. For such vectors we produce needed anti-concentration estimates for the matrix-vector products by using only these two components, i.e., we consider anti-concentration for the vector $P_A x$, where $A = \sigma_x(\{1, 2\})$. Since the Rogozin lemma is not suitable for this case, we compute the anti-concentration directly in Proposition 3.11. As for the classes $\mathcal{B}_2, \mathcal{B}_3$, we actually intersect the fourth class with appropriately chosen sets of almost constant vectors in order to control cardinalities of the nets. The final step is to show that the set $\mathcal{S}_n$ is contained in the union of four sets described here. Careful analysis of this approach shows that the result can be proved with all constants and parameters r, δ, ρ depending only on q. Thus, it works for p being between the two constants q and c.

The case of small p, that is, the case $C(\ln n)/n \le p \le c$, requires a more sophisticated splitting of $\mathcal{S}_n$ — we split it into *steep vectors* and $\mathcal{R}$*-vectors*. The definition and the treatment of steep vectors essentially follows [27, 30], with corresponding adjustments for our model. The set of *steep* vectors consists

of vectors having a large jump between order statistics measured at certain indices. The first subclass of steep vectors, $\mathcal{T}_0$, is the same as the class $\mathcal{B}_1$ described above — vectors having very large maximal coordinate — and is treated as $\mathcal{B}_1$. Similarly to the case of constant p, this class is the main contributor to the bound $(1+o_n(1))n(1-p)^n$ for non-invertibility over structured vectors. Next we fix certain $m \approx 1/p$ and consider a sequence $n_0 = 2$, $n_{j+1}/n_j = \ell_0$, $j \le s_0 - 1$, $n_{s_0+1} = m$ for some specially chosen parameters ℓ_0 and s_0 depending on p and n. The class $\mathcal{T}_1$ will be defined as the class of vectors such that there exists j with $x^*_{n_{j+1}} > 6pnx^*_{n_j}$. To work with vectors from this class, we first show that for a given j the event that for every choice of two disjoint sets $|J_1| = n_j$ and $|J_2| = n_{j+1} - n_j$, a random Bernoulli(p) matrix has a row with exactly one 1 in components indexed by J_1 and no 1's among components indexed by J_2, holds with a very high probability. Then, conditioned on this event, for every $x \in \mathcal{T}_1$, we choose J_1 corresponding to x^*_i, $i \le n_j$, and J_2 corresponding to x^*_i, $n_j < i \le n_{j+1}$, and the corresponding row. Then the inner product of this row with x will be large in absolute value due to the jump (see Lemma 6.9 for the details). Thus, conditioned on the described event, for every $x \in \mathcal{T}_1$ we have a good lower bound on $\|Mx\|$. Then next two classes of steep vectors, $\mathcal{T}_2$ and $\mathcal{T}_3$, consist of vectors having a jump of order $C\sqrt{pn}$, namely, vectors in $\mathcal{T}_2$ satisfy $x^*_m > C\sqrt{pn}x^*_k$ and vectors in $\mathcal{T}_3$ satisfy $x^*_k > C\sqrt{pn}x^*_\ell$, where $k \approx \sqrt{n/p}$ and $\ell = \lfloor rn \rfloor$ (r is the parameter from the definition of $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$). Trying to apply the same idea for these two subclasses one sees that the size of corresponding sets J_1 and J_2 is too large to have exactly one 1 among a row's components indexed by $J_1 \cup J_2$ with a high probability. Therefore the proof of individual probability bounds is delicate and technical as a construction of corresponding nets for $\mathcal{T}_2, \mathcal{T}_3$. We discuss the details in Subsection 6.6.

The class of $\mathcal{R}$-vectors consists of non-steep vectors to which Rogozin's lemma (Proposition 3.9) can be applied when we project a vector on $n-k$ smallest coordinates with $m < k \le n/\ln^2(pn)$, thus vectors from this class satisfy $\|P_A x\| \le c\sqrt{p}\|P_A x\|_\infty$ for $A = \sigma_x([k,n]$ (we will take union over all choices of integer k in the interval $(m, n/\ln^2(pn)]$). Thus, the individual probability bounds for $\mathcal{R}$-vectors will follow from Rogozin's lemma together with tensorization lemma as for classes $\mathcal{B}_2$, $\mathcal{B}_3$, described above. Thus the remaining part is to construct a good net for $\mathcal{R}$-vectors. For simplicity, dealing with such vectors, we fix the normalization $x^*_{\lfloor rn \rfloor} = 1$. Since vectors are non-steep, we have a certain control of largest coordinates and, thus, on the Euclidean norm of a vector. The upper bound on k is chosen in such a way that the cardinality on a net corresponding to largest coordinates of a vector is relatively small (it lies in $n/\ln^2(pn)$-dimensional subspace). For the purpose of constructing of a net of small cardinality, we need to control the Euclidean norm of $P_A x$ for an $\mathcal{R}$-vector. Therefore we split $\mathcal{R}$-vectors into level sets according to the value of $\|P_A x\|$. There will be two different types of level sets — vectors with relatively large Euclidean norm of $P_A x$ and vectors with small $\|P_A x\|$. A net for level sets with large $\|P_A x\|$ is easier to construct, since we can zero all coordinates starting with $x^*_{\lfloor rn \rfloor} = 1$. If the Euclidean norm is small, we cannot do this, so we intersect this subclass with almost constant vectors (in fact we incorporate this intersection into the definition of $\mathcal{R}$-vectors), defined by

$$\mathcal{AC}(\rho) := \{x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ s. t. } |\lambda| = x^*_{\lfloor rn \rfloor} \text{ and } |\{i \le n : |x_i - \lambda| \le \rho|\lambda|\}| > n - \lfloor rn \rfloor\}. \tag{9}$$

As in the case of constant p, this essentially reduces the dimension corresponding to almost constant part to one and therefore reduce the cardinality of a net. The rather technical construction of nets is presented in Subsection 6.3. In some aspects the construction follows ideas developed in [27].

3 Preliminaries

3.1 General notation

By *universal* or *absolute* constants we always mean numbers independent of all involved parameters, in particular independent of p and n. Given positive integers $\ell < k$ we denote sets $\{1, 2, \dots, \ell\}$ and

$\{\ell, \ell+1, \dots, k\}$ by $[\ell]$ and $[\ell, k]$ correspondingly. Having two functions f and g we write $f \approx g$ if there are two absolute positive constants c and C such that $cf \le g \le Cf$. As usual, Π_n denotes the permutation group on $[n]$.

For every vector $x = (x_i)_{i=1}^n \in \mathbb{R}^n$, by $(x_i^*)_{i=1}^n$ we denote the non-increasing rearrangement of the sequence $(|x_i|)_{i=1}^n$ and we fix one permutation σ_x satisfying $|x_{\sigma_x(i)}| = x_i^*$, $i \le n$. We use $\langle \cdot, \cdot \rangle$ for the standard inner product on $\mathbb{R}^n$, that is $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. Further, we write $\|x\|_\infty = \max_i |x_i|$ and $\|x\| = (\sum_{i=1}^n |x_i|^2)^{1/2}$ for the ℓ_∞- and ℓ_2-norms of x. We also denote $\mathbf{1} = (1, 1, \dots, 1)$.

3.2 Lower bound on the singularity probability

Here, we provide a simple argument showing that for the sequence of random Bernoulli(p_n) matrices (M_n), with p_n satisfying $(np_n - \ln n) \longrightarrow \infty$ as $n \to \infty$, we have

$$\mathbb{P}\big\{M_n \text{ contains a zero row or column}\big\} \ge (2 - o_n(1)) n\,(1-p)^n.$$

Our approach is similar to that applied in [5] in a related context.

Fix $n > 1$ and write $p = p_n$. Let $\mathbf{1}_R$ be the indicator of the event that there is a zero row in the matrix M_n, and, similarly, let $\mathbf{1}_C$ be the indicator of the event that M_n has a zero column. Then, obviously,

$$\mathbb{E}\,\mathbf{1}_R = \mathbb{E}\,\mathbf{1}_C = 1 - \big(1 - (1-p)^n\big)^n,$$

hence,

$$\mathbb{E}(\mathbf{1}_R + \mathbf{1}_C)^2 \ge 2 - 2\big(1 - (1-p)^n\big)^n.$$

On the other hand,

$$\mathbb{E}\,\mathbf{1}_R\,\mathbf{1}_C \le \sum_{i=1}^n \sum_{j=1}^n \mathbb{P}\big\{i\text{–th row and } j\text{–th column of } M_n \text{ are zero}\big\} = n^2(1-p)^{2n-1},$$

implying

$$\mathbb{E}(\mathbf{1}_R + \mathbf{1}_C)^2 = \mathbb{P}\big\{\mathbf{1}_R + \mathbf{1}_C = 1\big\} + 4\,\mathbb{P}\big\{\mathbf{1}_R\,\mathbf{1}_C = 1\big\} \le \mathbb{P}\big\{\mathbf{1}_R + \mathbf{1}_C = 1\big\} + 4n^2(1-p)^{2n-1}.$$

Therefore,

$$\begin{aligned}\mathbb{P}\big\{M_n \text{ contains a zero row or column}\big\} &\ge \mathbb{P}\big\{\mathbf{1}_R + \mathbf{1}_C = 1\big\} \\ &\ge \mathbb{E}(\mathbf{1}_R + \mathbf{1}_C)^2 - 4n^2(1-p)^{2n-1} \\ &\ge 2 - 2\big(1 - (1-p)^n\big)^n - 4n^2(1-p)^{2n-1}.\end{aligned}$$

It remains to note that, with our assumption on the growth rate of $p = p_n$, we have $n(1-p)^n \longrightarrow 0$, which implies

$$\frac{1}{n(1-p)^n}\big(2 - 2\big(1 - (1-p)^n\big)^n - 4n^2(1-p)^{2n-1}\big) \longrightarrow 2.$$

3.3 Gradual non-constant vectors

For any $r \in (0,1)$, we define $\Upsilon_n(r)$ as the set of all vectors x in $\mathbb{R}^n$ with $x^*_{\lfloor rn \rfloor} = 1$. We will call these vectors *r-normalized*. By a *growth function* $\mathbf{g}$ we mean any non-decreasing function from $[1, \infty)$ into $[1, \infty)$.

Let $\mathbf{g}$ be an arbitrary growth function. We will say that a vector $x \in \Upsilon_n(r)$ is *gradual* (with respect to the function $\mathbf{g}$) if $x_i^* \le \mathbf{g}(n/i)$ for all $i \le n$. Further, if $x \in \Upsilon_n(r)$ satisfies

$$\exists\, Q_1, Q_2 \subset [n] \quad \text{such that} \quad |Q_1|, |Q_2| \ge \delta n \quad \text{and} \quad \max_{i \in Q_2} x_i \le \min_{i \in Q_1} x_i - \rho \tag{10}$$

then we say that the vector x is *essentially non-constant* or just *non-constant* (with parameters δ, ρ). Recall that the set $\mathcal{V}_n = \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ was defined in (1) as

$$\mathcal{V}_n = \big\{x \in \Upsilon_n(r) : \ x \text{ is gradual with } \mathbf{g} \text{ and satisfies (10)}\big\}.$$

Vectors from this set we call *gradual non-constant vectors*.

Recall that the set $\mathcal{S}_n = \mathcal{S}_n(r, \mathbf{g}, \delta, \rho)$ of structured vectors was defined in (2) as the complement of scalar multiples of $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$. The next simple lemma will allow us to reduce analysis of $\{x/\|x\| : \ x \in \mathcal{S}_n\}$ to the treatment of the set $\{x/\|x\| : \ x \in \Upsilon_n(r) \setminus \mathcal{V}_n\}$.

Lemma 3.1. *For any choice of parameters $r, \mathbf{g}, \delta, \rho$, the set $\{x/\|x\| : \ x \in \mathcal{S}_n\}$ is contained in the closure of the set $\{x/\|x\| : \ x \in \Upsilon_n(r) \setminus \mathcal{V}_n\}$.*

Proof. Let y be a unit vector such that $y = x/\|x\|$ for some $x \in \mathcal{S}_n$. If $x^*_{\lfloor rn \rfloor} \neq 0$ then $y = z/\|z\|$, where $z = x/x^*_{\lfloor rn \rfloor} \in \Upsilon_n(r) \setminus \mathcal{V}_n$. If $x^*_{\lfloor rn \rfloor} = 0$, we can consider a sequence of vectors $(x(j))_{j \geq 1}$ in $\mathbb{R}^n$ defined by

$$x(j)_i = \begin{cases} x_i, & \text{if } \ |x_i| > 1/j, \\ 1/j, & \text{if } \ |x_i| \leq 1/j. \end{cases}$$

Let

$$y(j) := x(j)/x(j)^*_{\lfloor rn \rfloor} = jx(j) \in \Upsilon_n(r), \quad j \geq 1.$$

Clearly, $y(j)^*_1 = jx^*_1 \longrightarrow \infty$, so for all sufficiently large j we have $y(j) \notin \mathcal{V}_n$. Thus, for all large j,

$$y(j)/\|y(j)\| \in \{z/\|z\| : \ z \in \Upsilon_n(r) \setminus \mathcal{V}_n\},$$

whereas $y(j)/\|y(j)\| = x(j)/\|x(j)\| \longrightarrow x/\|x\|$. This implies the desired result. □

We will need two following lemmas. The first one states that vectors which do not satisfy (10) are almost constant (that is, have large part of coordinates nearly equal to each other). The second one is a simple combinatorial estimate, so we omit its proof.

Lemma 3.2. *Let $n \geq 1$, $\delta, \rho, r \in (0,1)$. Denote $k = \lceil \delta n \rceil$ and $m = \lfloor rn \rfloor$ and assume $n \geq 2m > 4k$. Assume $x \in \Upsilon_n(r)$ does not satisfy (10). Then there exist $A \subset [n]$ of cardinality $|A| > n - m$ and λ with $|\lambda| = 1$ such that $|x_i - \lambda| < \rho$ for every $i \in A$.*

Proof. By $(x^\#_i)_{i=1}^n$ denote the non-increasing rearrangement of $(x_i)_{i=1}^n$ (we would like to emphasize that we do not take absolute values). Note that there are two subsets $Q_1, Q_2 \subset [n]$ with $|Q_1|, |Q_2| \geq k$ satisfying $\max_{i \in Q_2} x_i \leq \min_{i \in Q_1} x_i - \rho$ if and only if $x^\#_k - x^\#_{n-k+1} \geq \rho$. Therefore, using that x does not satisfy (10), we observe $x^\#_k - x^\#_{n-k+1} < \rho$. Next consider the set

$$A := \{x^\#_i \ : \ k < i \leq n - k\}.$$

Then $|A| = n - 2k > n - m$. Since $x^*_m = 1$ we obtain that

$$|\{i : |x_i| > 1\}| < m \leq n - m \quad \text{ and } \quad |\{i : |x_i| < 1\}| \leq n - m.$$

Therefore, there exist two indices $j, \ell \in A$ such that either $x_j < -1 < x_\ell < 1$ in which case we take $\lambda = -1$ or $-1 < x_\ell < 1 < x_j$ in which case we take $\lambda = 1$. Then for every $i \in A$, $|x_i - \lambda| < x^\#_k - x^\#_{n-k+1} < \rho$. This completes the proof. □

Lemma 3.3. *For any $\delta \in (0,1]$ there are $n_\delta \in \mathbb{N}$, $c_\delta > 0$ and $C_\delta \geq 1$ depending only on δ with the following property. Let $n \geq n_\delta$ and let $m \in \mathbb{N}$ satisfy $n/m \geq C_\delta$. Denote by $\mathcal{S}$ the collection of sequences $(S_1, \ldots, S_m) \subset [n]$ with $|S_i| = \lfloor n/m \rfloor$ and $S_i \cap S_j = \emptyset$ for all $i \neq j$. Let A_{nm} be as in (7). Then for any pair Q_1, Q_2 of disjoint subsets of $[n]$ of cardinality at least δn each, one has*

$$\Big|\Big\{(S_1, \ldots, S_m) \in \mathcal{S} : \ \min(|S_i \cap Q_1|, |S_i \cap Q_2|) \geq \frac{\delta}{2}\lfloor n/m \rfloor \text{ for at most } c_\delta m \text{ indices } i\Big\}\Big| \leq e^{-c_\delta n} A^{-1}_{nm}.$$

3.4 Auxiliary results for Bernoulli r.v. and random matrices

Let $p \in (0,1)$, δ is Bernoulli random variable taking value 1 with probability p and 0 with probability $1-p$. We say that δ is a Bernoulli(p) random variable. A random matrix with i.i.d. entries distributed as δ will be called *Bernoulli(p) random matrix*.

Here we provide four lemmas needed below. We start with notations for random matrices used throughout the paper. The class of all $n \times n$ matrices having 0/1 entries we denote by $\mathcal{M}_n$. We will consider a probability measure on $\mathcal{M}_n$ induced by the distribution of an $n \times n$ Bernoulli(p) random matrix. We will use the same notation $\mathbb{P}$ for this probability measure; the parameter p will always be clear from the context. Let $M = \{\mu_{ij}\} \in \mathcal{M}_n$. By $\mathbf{R}_i = \mathbf{R}_i(M)$ we denote the i-th row of M, and by $\mathbf{C}_i(M)$ — the i-th column, $i \le n$. By $\|M\|$ we always denote the operator norm of M acting as an operator $\ell_2 \to \ell_2$. This norm is also called spectral norm and equals the largest singular number.

We will need the following form of Bennett's inequality.

Lemma 3.4. *Let $n \ge 1$, $0 < q < 1$, and δ be a Bernoulli(q) random variable. Let δ_i and δ_{ij}, $i, j \le n$, be independent copies of δ. Define the function $h(u) := (1+u)\ln(1+u) - u$, $u \ge 0$. Then for every $t > 0$,*

$$\max\left(\mathbb{P}\left(\sum_{i=1}^n \delta_i > qn + t\right), \mathbb{P}\left(\sum_{i=1}^n \delta_i < qn - t\right)\right) \le \exp\left(-\frac{nq(1-q)}{\max^2(q,1-q)}\, h\left(\frac{t\max(q,1-q)}{nq(1-q)}\right)\right).$$

In particular, for $0 < \varepsilon \le q \le 1/2$,

$$\max\left(\mathbb{P}\left(\sum_{i=1}^n \delta_i > (q+\varepsilon)n\right), \mathbb{P}\left(\sum_{i=1}^n \delta_i < (q-\varepsilon)n\right)\right) \le \exp\left(-\frac{n\varepsilon^2}{2q(1-q)}\,\left(1 - \frac{\varepsilon}{3q}\right)\right),$$

and for $q \le 1/2$, $\tau > e$,

$$\mathbb{P}\left(\sum_{i=1}^n \delta_i > (\tau+1)qn\right) \le \exp\left(-\tau\ln(\tau/e)qn\right).$$

Furthermore, for $50/n \le q \le 0.1$,

$$\mathbb{P}\Big(qn/8 \le \sum_{i=1}^n \delta_i \le 8qn\Big) \ge 1 - (1-q)^{n/2}.$$

Moreover, if $n \ge 30$ and $(4\ln n)/n \le p = q \le 1/2$ then denoting

$$\mathcal{E}_{sum} := \Big\{M = \{\delta_{ij}\}_{i,j\le n} \in \mathcal{M}_n \,:\, \sum_{j=1}^n \delta_{ij} \le 3.5pn \quad \text{for every} \quad i \le n\Big\}$$

we have $\mathbb{P}(\mathcal{E}_{sum}) \ge 1 - \exp(-1.5np)$.

Proof. Recall that Bennett's inequality states that for mean zero independent random variables $\xi_1, \ldots, \xi_n$ satisfying $\xi_i \le \rho$ (for a certain fixed $\rho > 0$) almost surely for $i \le n$, one has for every $t > 0$,

$$\mathbb{P}\left(\sum_{i=1}^n \xi_i > t\right) \le \exp\left(-\frac{\sigma^2}{\rho^2}\, h\left(\frac{\rho t}{\sigma^2}\right)\right),$$

where $\sigma^2 = \sum_{i=1}^n \mathbb{E}\xi_i^2$ (see e.g. Theorem 1.2.1 on p. 28 in [8] or Exercise 2.2 on p. 11 in [11] or Theorem 2.9 in [6]). Take $\xi_i = \delta_i - q$, $\xi_i' = -\xi_i$, $i \le n$. Then for every $i \le n$, ξ_i' and ξ_i are centered,

$|\xi_i'| = |\xi_i| = \max(q, 1-q)$, and $\sigma^2 = nq(1-q)$. Applying the Bennett's inequality with $\rho = \max(q, 1-q)$ twice — to ξ_i and ξ_i', we observe the first inequality. To prove the second inequality, we take $t = \varepsilon n$ and use that $h(\cdot)$ is an increasing function satisfying $h(u) \geq u^2/2 - u^3/6$ on $\mathbb{R}^+$. The third inequality follows by taking $t = \tau qn$ and using $h(u) \geq u \ln(u/e)$.

For the "furthermore" part, we apply the third inequality with $\tau = 7$, to get

$$\mathbb{P}\Big\{ \sum_{i=1}^n \delta_i > 8qn \Big\} \leq \exp(-6qn).$$

On the other hand, using $q \leq 0.1$,

$$\begin{aligned}\mathbb{P}\Big\{ \sum_{i=1}^n \delta_i < qn/8 \Big\} &= \sum_{i=0}^{\lfloor qn/8 \rfloor} \binom{n}{i} q^i (1-q)^{n-i} \leq (1-q)^n + \sum_{i=1}^{\lfloor qn/8 \rfloor} \left(\frac{enq}{i(1-q)} \right)^i (1-q)^n \\ &\leq (1-q)^n + \frac{qn}{8} \left(\frac{8e}{1-q} \right)^{qn/8} (1-q)^n \leq (1-q)^n + \frac{qn}{8} \left(\frac{80e}{9} \right)^{qn/8} (1-q)^n.\end{aligned}$$

Since $(80e/9)^{1/8} \leq e^{0.4}$, $(1-q)^n \leq \exp(-qn)$, $qn \geq 50$, and $\ln x \leq x/e$ on $[0, \infty)$, this implies

$$\mathbb{P}\Big(qn/8 \leq \sum_{i=1}^n \delta_i < qn/8 \Big) \leq \exp(-6qn) + (1 + \exp(0.45qn))(1-q)^n \leq (1-q)^{n/2}.$$

Finally, to get the last inequality, we take $t = 2.5qn = 2.5pn$, then

$$\mathbb{P}\left(\sum_{j=1}^n \delta_{ij} > 3.5pn \right) \leq \exp\left(-\frac{np}{1-p} h(2.5) \right) \leq \exp\left(-np \ (3.5 \ln 3.5 - 2.5) \right) \leq \exp\left(-1.8np \right).$$

Since under our assumptions, $n \exp(-1.8np) \leq \exp(-1.5np)$, the bound on $\mathbb{P}(\mathcal{E}_{sum})$ follows by the union bound. □

We need the following simple corollary of Bennet's inequality.

Lemma 3.5. *For any $R \geq 1$ there is $C_{3.5} = C_{3.5}(R) \geq 1$ with the following property. Let $n \geq 1$ and $p \in (0,1)$ satisfy $C_{3.5}p \leq 1$ and $C_{3.5} \leq pn$. Further, let M be an $n \times n$ be Bernoulli(p) random matrix. Then with probability at least $1 - \exp(-n/C_{3.5})$ one has*

$$8pn \geq |\operatorname{supp} \mathbf{C}_i(M)| \geq pn/8 \quad \text{for all but} \quad \lfloor (pR)^{-1} \rfloor \quad \text{indices } i \in [n] \setminus \{1\}.$$

Proof. For each $i \in [n] \setminus \{1\}$, let ξ_i be the indicator of the event

$$\big\{ 8pn < |\operatorname{supp} \mathbf{C}_i(M)| \qquad \text{or} \qquad |\operatorname{supp} \mathbf{C}_i(M)| < pn/8 \big\}.$$

By Lemma 3.4, $\mathbb{E}\,\xi_i \leq e^{-pn/2}$. Since ξ_i's are independent, by Markov's inequality,

$$\mathbb{P}\Big\{ \sum_{i=2}^n \xi_i \geq \frac{1}{pR} \Big\} \leq \binom{n-1}{\lfloor (pR)^{-1} \rfloor} \big(e^{-pn/2} \big)^{\lfloor (pR)^{-1} \rfloor} \leq \binom{n-1}{\lfloor (pR)^{-1} \rfloor} e^{-n/(4R)}.$$

The result follows. □

The following lemma provides a bound on the norm of a random Bernoulli matrix. It is similar to [5, Theorem 1.14], where the case of symmetric matrices was treated. For the sake of completeness we sketch its proof.

Lemma 3.6. *Let n be large enough and $(4\ln n)/n \le p \le 1/2$. Let $M = (\delta_{ij})_{i,j}$ be a Bernoulli(p) random matrix. Then for every $t \ge 30$ one has*

$$\mathbb{P}\big\{\|M - \mathbb{E}M\| \ge 2t\sqrt{np}\big\} \le 4e^{-t^2pn/4} \qquad \text{and} \qquad \mathbb{P}\big\{\|M\| \ge 2t\sqrt{np} + pn\big\} \le 4e^{-t^2pn/4}.$$

In particular, taking $t = \sqrt{pn}$,

$$\mathbb{P}\left(\|M\mathbf{1}\| \ge 3pn^{3/2}\right) \le 4\exp(-n^2p^2/4). \tag{11}$$

Proof. Given an $n \times n$ random matrix $T = (t_{ij})_{i,j}$ with independent entries taking values in $[0,1]$. We consider it as a vector in $\mathbb{R}^m$ with $m = n^2$. Then the Hilbert–Schmidt norm of T is the standard Euclidean norm on $\mathbb{R}^m$. Let f be any function in $\mathbb{R}^m$ which is convex and is 1-Lipschitz with respect to the standard Euclidean norm. Then Talagrand's inequality (see e.g. Corollary 4.10 and Proposition 1.8 in [23]) gives that for every $s > 0$,

$$\mathbb{P}\left(f(T) \ge \mathbb{E}f(T) + s + 4\sqrt{\pi}\right) \le 4\exp(-s^2/4).$$

We apply this inequality twice, first with the function $f(T) := \|T\|$ to the matrix $T := M - \mathbb{E}M$. At the end of this proof we show that $\mathbb{E}\|M - \mathbb{E}M\| \le 20\sqrt{pn}$. Therefore, taking $s = t\sqrt{pn}$ with $t \ge 30$, we obtain the first bound. For the second bound, note that all entries of $\mathbb{E}M$ equal p, hence $\|\mathbb{E}M\| = pn$. Thus, the second bound follows by the triangle inequality.

It remains to prove that $\mathbb{E}\|M - \mathbb{E}M\| \le 20\sqrt{pn}$. Recall that δ_{ij} are the entries of M. Let δ'_{ij}, $i,j \le n$ be independent copies of δ_{ij} and set $M' := (\delta'_{ij})_{i,j}$. Denote by r_{ij} independent Rademacher random variables and by g_{ij} independent standard Gaussian random variables. We assume that all our variables are mutually independent and set $\xi_{ij} := \delta_{ij} - \delta'_{ij}$. Since for every $i,j \le n$, ξ_{ij} is symmetric, it has the same distribution as $|\xi_{ij}|r_{ij}$ and the same as $\sqrt{2/\pi}|\xi_{ij}|r_{ij}\mathbb{E}|g_{ij}|$. Then we have

$$\mathbb{E}_\delta\|M - \mathbb{E}M\| = \mathbb{E}_\delta\|M - \mathbb{E}_{\delta'}M'\| \le \mathbb{E}_\delta\mathbb{E}_{\delta'}\|M - M'\| = \mathbb{E}_\xi\|(\xi_{ij})_{i,j}\| = \sqrt{2/\pi}\,\mathbb{E}_{\xi,r}\|(\xi_{ij}r_{ij}\mathbb{E}_g|g_{ij}|)_{i,j}\|$$

$$\le \sqrt{2/\pi}\,\mathbb{E}_{\xi,r,g}\|(\xi_{ij}r_{ij}|g_{ij}|)_{i,j}\| = \sqrt{2/\pi}\,\mathbb{E}_\xi\mathbb{E}_g\|(\xi_{ij}|g_{ij}|)_{i,j}\|.$$

Applying a result of Bandeira and Van Handel (see the beginning of Section 3.1 in [1]), we obtain

$$\mathbb{E}_\delta\|M - \mathbb{E}M\| \le \mathbb{E}_\xi(4\max(\sigma_1,\sigma_2) + 15\sigma_*\sqrt{\ln(2n)}),$$

where

$$\sigma_1 = \max_{i\le n}\sqrt{\sum_{j=1}^n \xi_{ij}^2}, \quad \sigma_2 = \max_{j\le n}\sqrt{\sum_{i=1}^n \xi_{ij}^2}, \quad \text{and} \quad \sigma_* = \max_{i,j\le n}|\xi_{ij}| \le 1.$$

Note that ξ_{ij}^2 are Bernoulli(q) random variables with $q = 2p(1-p)$. Since $(4\ln n)/n \le p \le 1/2$ we have $(4\ln n)/n \le p \le q \le 1/2$. Applying the "moreover part" of Lemma 3.4, we obtain that

$$\mathbb{P}\left(\max(\sigma_1,\sigma_2) > \sqrt{7pn}\right) \le 2\exp(-1.5nq) \le 2/n^6.$$

Moreover, since $\xi_{ij}^2 \le 1$, we have also $\max(\sigma_1,\sigma_2) \le \sqrt{n}$. Therefore,

$$\mathbb{E}_\xi(4\max(\sigma_1,\sigma_2) + 15\sigma_*\sqrt{\ln(2n)}) \le 4\sqrt{7pn} + 8/n^5 + 15\sqrt{\ln(2n)} \le 20\sqrt{pn}.$$

□

As an elementary corollary of the above lemma, we have the following statement where the restriction $pn \ge 4\ln n$ is removed.

Corollary 3.7. *For every $s > 0$ and $R \geq 1$ there is $C_{3.7} \geq 1$ depending on s, R with the following property. Let $n \geq 16/s$ be large enough and let $p \in (0, 1/4]$ satisfy $s \ln n \leq pn$. Let M_n be an $n \times n$ Bernoulli(p) random matrix. Then*

$$\mathbb{P}\big\{\|M_n - \mathbb{E}M_n\| \leq C_{3.7}\sqrt{pn}\big\} \geq 1 - \exp(-Rpn).$$

Proof. Let $w := \max(1, \lceil 8/s \rceil)$, $\widetilde{n} := w\, n$, and let $\widetilde{M}_n$ be $\widetilde{n} \times \widetilde{n}$ Bernoulli(p) matrix. Assuming that n is sufficiently large, we get

$$p\,\widetilde{n} = wpn \geq s \max(1, \lceil 8/s \rceil) \ln n \geq 4 \ln \widetilde{n}.$$

Thus, the previous lemma is applicable, and we get

$$\mathbb{P}\big\{\|\widetilde{M}_n - \mathbb{E}\widetilde{M}_n\| \leq C_{3.7}\sqrt{pn}\big\} \geq 1 - \exp(-Rpn),$$

for some $C_{3.7} > 0$ depending only on s, R. Since the norm of a matrix is not less than the norm of any of its submatrices, and because any $n \times n$ submatrix of $\widetilde{M}_n$ is equidistributed with M_n, we get the result. □

3.5 Anti-concentration

In this subsection we combine anti-concentration inequalities with the following tensorization lemma (see Lemma 3.2 in [51], Lemma 2.2 in [44] and Lemma 5.4 in [42]). We also provide Esseen's lemma.

Lemma 3.8 (Tensorization lemma)**.** *Let $\lambda, \gamma > 0$. Let $\xi_1, \xi_2, \ldots, \xi_m$ be independent random variables. Assume that for all $j \leq m$, $\mathbb{P}(|\xi_j| \leq \lambda) \leq \gamma$. Then for every $\varepsilon \in (0, 1)$ one has*

$$\mathbb{P}(\|(\xi_1, \xi_2, ..., \xi_m)\| \leq \lambda\sqrt{\varepsilon m}) \leq (e/\varepsilon)^{\varepsilon m}\gamma^{m(1-\varepsilon)}.$$

Moreover, if there exists $\varepsilon_0 > 0$ and $K > 0$ such that for every $\varepsilon \geq \varepsilon_0$ and for all $j \leq m$ one has $\mathbb{P}(|\xi_j| \leq \varepsilon) \leq K\varepsilon$ then there exists an absolute constant $C_{3.8} > 0$ such that for every $\varepsilon \geq \varepsilon_0$,

$$\mathbb{P}(\|(\xi_1, \xi_2, ..., \xi_m)\| \leq \varepsilon\sqrt{m}) \leq (C_{3.8}K\varepsilon)^m.$$

Recall that for a real-valued random variable ξ its *Lévy concentration function* $\mathcal{Q}(\xi, t)$ is defined as

$$\mathcal{Q}(\xi, t) := \sup_{\lambda \in \mathbb{R}} \mathbb{P}\big\{|\xi - \lambda| \leq t\big\}, \quad t > 0.$$

We will need bounds on the Lévy concentration function of sums of independent random variables. Such inequalities were investigated in many works, starting with Lévi, Doeblin, Kolmogorov, Rogozin. We quote here a result due to Kesten [20], who improved Rogozin's estimate [41].

Proposition 3.9. *There exists an absolute positive constant C such that the following holds. Let $\xi_1, \xi_2, \ldots, \xi_m$ be independent random variables and $\lambda, \lambda_1, ..., \lambda_m > 0$ satisfy $\lambda \geq \max_{i \leq m} \lambda_i$. Then*

$$\mathcal{Q}\Big(\sum_{i=1}^m \xi_i, \lambda\Big) \leq \frac{C\,\lambda\, \max_{i \leq m} \mathcal{Q}(\xi_i, \lambda)}{\sqrt{\sum_{i=1}^m \lambda_i^2 (1 - \mathcal{Q}(\xi_i, \lambda_i))}}.$$

This proposition together with Lemma 3.8 immediately implies the following consequence, in which, given $A \subset [m]$ and $x \in \mathbb{R}^m$, x_A denotes coordinate projection of x on $\mathbb{R}^A$.

Proposition 3.10. *There exists an absolute constant $C_0 \geq 1$ such that the following holds. Let $p \in (0, 1/2]$. Let δ be a Bernoulli(p) random variable. Let δ_j, $j \leq n$, and δ_{ij}, $i, j \leq n$, be independent copies of δ. Let $M = (\delta_{ij})_{ij}$. Let $A \subset [n]$ and $x \in \mathbb{R}^n$ be such that $\|x_A\|_\infty \leq C_0^{-1}\sqrt{p}\,\|x_A\|$. Then*

$$\mathbb{P}\Big(\|Mx\| \leq \frac{\sqrt{pn}}{3\sqrt{2}C_0}\|x_A\|\Big) \leq e^{-3n}.$$

Moreover, if $\lambda := \frac{\sqrt{p}\,\|x_A\|}{3C_0} \leq 1/3$ then $\mathcal{Q}\Big(\sum_{j=1}^n \delta_j x_j, \lambda\Big) \leq e^{-8}$.

Proof. We start with the "moreover" part. Assume $\sqrt{p}\,\|x_A\| \le C_0$. Let $\lambda_j = |x_j|/3$. Clearly, for every $j \le n$, $\mathcal{Q}(x_j\delta_j, |x_j|/3) = \mathcal{Q}(\delta_j, 1/3) = 1-p$. Independence of δ_i's and Proposition 3.9 imply that for every λ satisfying $\max_{j\in A}\lambda_j \le \lambda \le 1/3$ one has

$$\mathcal{Q}\Big(\sum_{j=1}^n x_j\delta_j, \lambda\Big) \le \mathcal{Q}\Big(\sum_{j\in A} x_j\delta_j, \lambda\Big) \le \frac{C\,\lambda}{\sqrt{\sum_{j\in A}\lambda_j^2\, p}} = \frac{3C\,\lambda}{\sqrt{p}\,\|x_A\|}.$$

Choosing $C_0 = Ce^8$ and $\lambda = \sqrt{p}\,\|x_A\|/(3C_0)$ (note that the assumption on $\|x_A\|_\infty$ ensures that $\lambda \ge \lambda_j$ for all $j \in A$) we obtain the "moreover" part.

Now apply Lemma 3.8 with $\xi_i = (Mx)_i = \sum_{j=1}^n x_j\delta_{ij}$, $\varepsilon = 1/2$, $\gamma = e^{-8}$, $m = n$. We have

$$\mathbb{P}\Big(\|Mx\| \le \lambda\sqrt{n/2}\Big) \le (2e)^{n/2}\exp(-4n) \le \exp(-3n).$$

This implies the bound under assumption $\sqrt{p}\,\|x_A\| \le C_0$, which can be removed by normalizing x. □

We also will need the following combination of a simple anti-concentration fact with Lemma 3.8.

Proposition 3.11. *Let $p \in (0, 1/2]$ and $\alpha > 0$. Let δ be a Bernoulli(p) random variable. Let δ_j, $j \le n$, and δ_{ij}, $i, j \le n$, be independent copies of δ. Let $M = (\delta_{ij})_{ij}$. Let $x \in \mathbb{R}^n$ be such that $x_2^* \ge \alpha$. Then*

$$\mathcal{Q}\Big(\sum_{j=1}^n x_j\delta_j, \alpha/2.01\Big) \le 4^{-p} \qquad \text{and} \qquad \mathbb{P}\Big(\|Mx\| \le \frac{\alpha\sqrt{pn}}{10\sqrt{\ln(e/p)}}\Big) \le \exp(-1.2pn).$$

Proof. Without loss of generality we assume that $x_1^* = |x_1|$ and $x_2^* = |x_2|$. Note that $x_1\delta_1 + x_2\delta_2$ takes values in $E_1 := \{0, x_1 + x_2\}$ with probability $(1-p)^2 + p^2$ and in $E_2 := \{x_1, x_2\}$ with probability $2p(1-p)$. Using that $p \le 1/2$, we observe

$$\max\{(1-p)^2 + p^2,\, 2p(1-p)\} \le 4^{-p}.$$

Since the distance between sets E_1 and E_2 equals to $\min\{|x_1|, |x_2|\} = |x_2|$ and since we clearly have $\mathcal{Q}\big(\sum_{j=1}^n x_j\delta_j, \lambda\big) \le \mathcal{Q}\big(\sum_{j=1}^2 x_j\delta_j, \lambda\big)$, the first inequality follows.

We now apply Lemma 3.8 with $\xi_i = (Mx)_i = \sum_{j=1}^n x_j\delta_{ij}$, $\varepsilon = p/(24\ln(e/p))$, $\gamma = 4^{-p}$, $m = n$. Note that $\varepsilon \le 1/(48\ln 2e) \le 0.02$, and $x \ge e\ln x$ for $x \ge 0$, hence

$$\begin{aligned} p(1-\varepsilon)\ln 4 - \varepsilon\ln(e/\varepsilon) &\ge p\,0.98\,\ln 4 - \frac{p}{24\ln(e/p)}\ln\left(\frac{24e\ln(e/p)}{p}\right) \\ &\ge 1.35p - \frac{p}{24}\left(\frac{\ln 24}{\ln 2e} + 1 + \frac{\ln\ln(e/p)}{\ln(e/p)}\right) \ge 1.2p. \end{aligned}$$

Thus Lemma 3.8 yields

$$\mathbb{P}\Big(\|Mx\| \le \frac{\alpha\sqrt{pn}}{2.01\sqrt{24\ln(e/p)}}\Big) \le (e/\varepsilon)^{\varepsilon n}4^{-pn(1-\varepsilon))} \le \exp(-1.2pn).$$

This completes the proof. □

Finally we state Esseen's lemma [13], needed to prove Theorem 2.1.

Lemma 3.12 (Esseen)**.** *There exists an absolute constant $C > 0$ such that the following holds. Let ξ_i, $i \le m$, be independent random variables. Then for every $\tau > 0$,*

$$\mathcal{Q}\Big(\sum_{i=1}^m \xi_i, \tau\Big) \le C\int_{-1}^{1}\prod_{i=1}^m |\mathbb{E}\exp(2\pi\mathbf{i}\xi_i s/\tau)|\,ds.$$

3.6 Net argument

Here we discuss special nets that will be used and corresponding approximations. We fix the following notations. Let $\mathbf{e} = \mathbf{1}/\sqrt{n}$ be the unit vector in the direction of $\mathbf{1}$. Let $P_{\mathbf{e}}$ be the projection on $\mathbf{e}^\perp$ and $P_{\mathbf{e}}^\perp$ be the projection on $\mathbf{e}$, that is $P_{\mathbf{e}}^\perp = \langle \cdot, \mathbf{e}\rangle\, \mathbf{e}$. Similarly, for $j \le n$, let P_j be the projection on $e_j^\perp$ and $P_j^\perp$ be the projection on e_j. Recall that for $x \in \mathbb{R}^n$, the permutation σ_x satisfies $|x_{\sigma_x(i)}| = x_i^*$, $i \le n$. Define a (non-linear) operator $Q : \mathbb{R}^n \to \mathbb{R}^n$ by $Qx = P_{F(x)}x$ — the coordinate projection on $\mathbb{R}^{F(x)}$, where $F(x) = \sigma_x([2,n])$, in other words Q annihilates the largest coordinate of a vector. Consider the triple norm on $\mathbb{R}^n$ defined by

$$|||x|||^2 := \|P_{\mathbf{e}}x\|^2 + pn\|P_{\mathbf{e}}^\perp x\|^2$$

(note that $\|P_{\mathbf{e}}^\perp x\| = |\langle x, \mathbf{e}\rangle|$). We will use the following notion of shifted sparse vectors. Given $m \le n$ and a parameter $\gamma > 0$, define

$$U(m,\gamma) := \Big\{ x \in \mathbb{R}^n \ :\ \exists A \subset [n], |A| = n-m,\ \exists |\lambda| \le \frac{2}{\sqrt{m}}\ \forall i \in A \ \text{ one has } \ |x_i - \lambda| \le \frac{\gamma}{\sqrt{n}} \Big\}.$$

Further, given another parameter $\beta > 0$, define the set

$$V(\beta) := \{ x \in \mathbb{R}^n \ :\ \|x\|_\infty \le 1 \ \text{ and } \ \|Qx\| \le \beta \}.$$

Lemma 3.13. *Let $0 < 8\gamma \le \varepsilon \le \beta$ and $1 \le m \le n$. Then there exists an ε-net in $V(\beta) \cap U(m,\gamma)$ with respect to $|||\cdot|||$ of cardinality at most*

$$\frac{2^{10}\sqrt{p}\, n^2}{\varepsilon^2 \sqrt{m}} \left(\frac{9\beta}{\varepsilon}\right)^m \binom{n}{m}.$$

Proof. Denote $V := V(\beta) \cap U(m,\gamma)$. For each $x \in V$ let $A(x)$ be the set from the definition of $U(m,\gamma)$ (if the choice of $A(x)$ is not unique, we fix one of them).

Fix $E \subset [n]$ of cardinality m. We first consider vectors $x \in V$ satisfying $A(x) = E^c$. Fix $j \le n$ and denote

$$V_j = V_j(E) := \{ x \in V \ :\ j = \sigma_x(1) \ \text{ and } \ A(x) = E^c \}$$

(thus $x_1^* = |x_j|$ on V_j). We now construct a net for V_j. It will be obtained as the sum of four nets, where the first one deals with just one coordinate, j, annihilating the maximal coordinate; the second one deals with the non-constant part of the vector, consisting of at most m coordinates (excluding x_1^*); the third one deals with almost constant coordinates (corresponding to $A(x)$); and the fourth net deals with the direction of the constant vector. This way, three of our four nets are 1-dimensional. Let P_W be the coordinate projection onto $\mathbb{R}^W$, where $W = E \setminus \{j\}$. Note that the definition of $V(\beta)$ implies that $\|P_W(x)\| \le \beta$ for every $x \in V_j$. Let, as before, $P_j^\perp$ be the projection onto e_j.

Let $\mathcal{N}_1$ be an $\varepsilon/4$-net in $P_j^\perp(V_j) \subset [-1,1]e_j$ of cardinality at most $8/\varepsilon$. Let $\mathcal{N}_2$ be an $\varepsilon/4$-net (with respect to the Euclidean metric) in $P_W(V_j)$ of cardinality at most $(1 + 8\beta/\varepsilon)^m$.

Further, let $\mathcal{N}_3'$ be an $\varepsilon/(8\sqrt{n})$-net in the segment $[-2/\sqrt{m}, 2/\sqrt{m}]$ (approximating λ in the definition of $U(m,\gamma)$) with cardinality at most $32\sqrt{n}/(\varepsilon\sqrt{m})$. Let $\mathcal{N}_3$ be the set of all vectors of the type $\lambda_0 \sum_{i \in E^c \setminus \{j\}} e_i$, where $\lambda_0 \in \mathcal{N}_3'$. Then by the construction of the nets and by the definition of $U(m,\gamma)$ for every $x \in V_j$ there exist $y_x^i \in \mathcal{N}_i$, $i \le 3$, such that for $y_x = y_x^1 + y_x^2 + y_x^3$,

$$\|x - y_x\|^2 \le \frac{\varepsilon^2}{16} + \frac{\varepsilon^2}{16} + \sum_{i \in E^c \setminus \{j\}} \left(\frac{\gamma}{\sqrt{n}} + \frac{\varepsilon}{8\sqrt{n}}\right)^2 \le \frac{3\varepsilon^2}{16};$$

in particular, $\|P_{\mathbf{e}}(x-y_x)\| \le \sqrt{3/16}\varepsilon$. Finally, let $\mathcal{N}_4$ be an $\varepsilon/(4\sqrt{pn})$-net in the segment $(\varepsilon/2)[-\mathbf{e},\mathbf{e}]$ with cardinality at most $4\sqrt{pn}$. Then for every $x \in V_j$ there exists y_x as above and $y_x^4 \in \mathcal{N}_4$ with

$$|||x-y_x-y_x^4|||^2 = |||P_{\mathbf{e}}(x-y_x)+P_{\mathbf{e}}^{\perp}(x-y_x)-y_x^4|||^2 = \|P_{\mathbf{e}}(x-y_x)\|^2 + pn\|P_{\mathbf{e}}^{\perp}(x-y_x)-y_x^4\|^2 \le \varepsilon^2/4.$$

Thus the set $\mathcal{N}_{E,j} = \mathcal{N}_1+\mathcal{N}_2+\mathcal{N}_3+\mathcal{N}_4$ is an $(\varepsilon/2)$-net for V_j with respect to $|||\cdot|||$ and its cardinality is bounded by

$$\frac{2^{10}\sqrt{p}\,n}{\varepsilon^2\sqrt{m}}\left(1+\frac{8\beta}{\varepsilon}\right)^m.$$

Taking union of such nets over all choices of $E \subset [n]$ and all $j \le n$ we obtain an $(\varepsilon/2)$-net $\mathcal{N}_0$ in $|||\cdot|||$ for V of desired cardinality. Using standard argument, we pass to an ε-net $\mathcal{N} \subset V$ for V. □

Later we apply Lemma 3.13 with the following proposition.

Proposition 3.14. *Let n be large enough, $(4\ln n)/n \le p < 1/2$, and $\varepsilon > 0$. Denote*

$$\mathcal{E}_{nrm} := \{M \in \mathcal{M}_n : \|M - p\mathbf{1}\mathbf{1}^\top\| \le 60\sqrt{np} \quad \text{and} \quad \|M\mathbf{1}\| \le 3pn^{3/2}\}.$$

Then for every $x \in \mathbb{R}^n$ satisfying $|||x||| \le \varepsilon$ and every $M \in \mathcal{E}_{nrm}$ one has $\|Mx\| \le 100\sqrt{pn}\varepsilon$.

Proof. Let $w = P_{\mathbf{e}}^{\perp}x$. Then, by the definition of the triple norm, $\|w\| \le |||x|||/\sqrt{pn} \le \varepsilon/\sqrt{pn}$. Clearly,

$$(p\mathbf{1}\mathbf{1}^\top)(x-w) = (p\mathbf{1}\mathbf{1}^\top)P_{\mathbf{e}}x = 0.$$

Therefore, using that $M \in \mathcal{E}_{nrm}$, we get

$$\|M(x-w)\| = \|(M-p\mathbf{1}\mathbf{1}^\top)(x-w)\| \le 60\sqrt{pn}\|x-w\| \le 70\sqrt{pn}\varepsilon.$$

Since $w = \pm\mathbf{1}\|w\|/\sqrt{n}$ and $\|w\| \le \varepsilon/\sqrt{pn}$, using again that $M \in \mathcal{E}_{nrm}$, we observe that

$$\|Mw\| \le \frac{\varepsilon}{\sqrt{p}\,n}\|M\mathbf{1}\| \le 3\sqrt{pn}\varepsilon.$$

The proposition follows by the triangle inequality. □

4 Unstructured vectors

The goal of this section is to prove Theorem 2.2.

Recall that given growth function $\mathbf{g}$ and parameters $r,\delta,\rho \in (0,1)$, the set of vectors $\mathcal{V}_n = \mathcal{V}_n(r,\mathbf{g},\delta,\rho)$ was defined in (1). In the next two sections (dealing with invertibility over structured vectors), we work with two different growth functions; one will be applied to the case of constant p and the other one (giving a worse final estimate) is suitable in the general case. For this reason, and to increase flexibility of our argument, rather than fixing a specific growth function here, we will work with an arbitrary non-decreasing function $\mathbf{g} : [1,\infty) \to [1,\infty)$ satisfying the additional assumption (8) with a "global" parameter $K_3 \ge 1$.

4.1 Degree of unstructuredness: definition and basic properties

Below, for any non-empty finite integer subset S, we denote by $\eta[S]$ a random variable uniformly distributed on S. Additionally, for any $K_2 \ge 1$, we fix a smooth version of $\max(\frac{1}{K_2},t)$. More precisely, let us fix a function $\psi_{K_2} : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying

- The function ψ_{K_2} is twice continuously differentiable, with $\|\psi'_{K_2}\|_\infty = 1$ and $\|\psi''_{K_2}\|_\infty < \infty$;
- $\psi_{K_2}(t) = \frac{1}{K_2}$ for all $t \le \frac{1}{2K_2}$;
- $\frac{1}{K_2} \ge \psi_{K_2}(t) \ge t$ for all $\frac{1}{K_2} \ge t \ge \frac{1}{2K_2}$;
- $\psi_{K_2}(t) = t$ for all $t \ge \frac{1}{K_2}$.

In what follows, we view the maximum of the second derivative of ψ_{K_2} as a function of K_2 (the nature of this function is completely irrelevant as we do not attempt to track magnitudes of constants involved in our arguments).

Fix an integer $n \ge 1$ and an integer $m \le n/2$. Recall that given a vector $v \in \mathbb{R}^n$ and parameters $K_1, K_2 \ge 1$, the *degree of unstructuredness (u-degree)* $\mathbf{UD}_n = \mathbf{UD}_n(v, m, K_1, K_2)$ of v was defined in (6). The quantity $\mathbf{UD}_n$ will serve as a measure of unstructuredness of the vector v and in its spirit is similar to the notion of the essential least common denominator introduced earlier by Rudelson and Vershynin [44]. Here *unstructuredness* refers to the uniformity in the locations of components of v on the real line. The larger the degree is, the better anti-concentration properties of an associated random linear combination are. The functions ψ_{K_2} employed in the definition will be important when discussing certain stability properties of $\mathbf{UD}_n$.

We start with a proof of Theorem 2.1 which connects the definition of the u-degree with anti-concentration properties.

Proof of Theorem 2.1. For any sequence of disjoint subsets $S_1, \dots, S_m$ of $[n]$ of cardinality $\lfloor n/m \rfloor$ each, set

$$\mathcal{E}_{S_1,\dots,S_m} := \big\{ \operatorname{supp} X \cap S_i = 1 \text{ for all } i \le m \big\}.$$

Note that each point ω of the probability space belongs to the same number of events from the collection $\{\mathcal{E}_{S_1,\dots,S_m}\}_{S_1,\dots,S_m}$, therefore, for A_{nm} defined in (7) we have for any $\lambda \in \mathbb{R}$ and $\tau > 0$,

$$\mathbb{P}\Big\{ \Big| \sum_{i=1}^n v_i X_i - \lambda \Big| \le \tau \Big\} = A_{nm} \sum_{S_1,\dots,S_m} \mathbb{P}\Big\{ \Big| \sum_{i=1}^n v_i X_i - \lambda \Big| \le \tau \mid \mathcal{E}_{S_1,\dots,S_m} \Big\}. \tag{12}$$

Further, conditioned on an event $\mathcal{E}_{S_1,\dots,S_m}$, the random sum $\sum_{i=1}^n v_i X_i$ is equidistributed with $\sum_{i=1}^m v_{\eta[S_i]}$ (where we assume that $\eta[S_1], \dots, \eta[S_m]$ are jointly independent with $\mathcal{E}_{S_1,\dots,S_m}$). On the other hand, applying Lemma 3.12, we observe that for every $\tau > 0$,

$$\begin{aligned} \mathcal{Q}\Big(\sum_{i=1}^m v_{\eta[S_i]}, \tau \Big) &\le C' \int_{-1}^{1} \prod_{i=1}^m |\mathbb{E} \exp(2\pi \mathbf{i} v_{\eta[S_i]} s/\tau)| \, ds \\ &= C' \, m^{-1/2} \, \tau \int_{-\sqrt{m}/\tau}^{\sqrt{m}/\tau} \prod_{i=1}^m |\mathbb{E} \exp(2\pi \mathbf{i} v_{\eta[S_i]} \, m^{-1/2} s)| \, ds, \end{aligned}$$

for a universal constant $C' > 0$. Combining this with (12), we get for every $\tau > 0$,

$$\begin{aligned} \mathcal{Q}\Big(\sum_{i=1}^n v_i X_i, \tau \Big) &\le A_{nm} \sum_{S_1,\dots,S_m} \mathcal{Q}\Big(\sum_{i=1}^n v_i X_i, \tau \mid \mathcal{E}_{S_1,\dots,S_m} \Big) \\ &\le \frac{C' \tau A_{nm}}{\sqrt{m}} \sum_{S_1,\dots,S_m} \int_{-\sqrt{m}/\tau}^{\sqrt{m}/\tau} \prod_{i=1}^m |\mathbb{E} \exp(2\pi \mathbf{i} v_{\eta[S_i]} \, m^{-1/2} s)| \, ds. \end{aligned}$$

Setting $\tau := \sqrt{m}/\mathbf{UD}_n$, where $\mathbf{UD}_n = \mathbf{UD}_n(v, m, K_1, K_2)$, we obtain

$$\mathcal{Q}\Big(\sum_{i=1}^n v_i X_i, \sqrt{m}/\mathbf{UD}_n\Big) \le \frac{C' A_{nm}}{\mathbf{UD}_n} \sum_{S_1,\dots,S_m} \int_{-\mathbf{UD}_n}^{\mathbf{UD}_n} \prod_{i=1}^m |\mathbb{E}\exp(2\pi\mathbf{i} v_{\eta[S_i]}\, m^{-1/2} s)|\, ds \le \frac{C' K_1}{\mathbf{UD}_n},$$

in view of the definition of $\mathbf{UD}_n(v, m, K_1, K_2)$. The result follows. □

For the future use we state an immediate consequence of Theorem 2.1 and Lemma 3.8.

Corollary 4.1. *Let $n, \ell \in \mathbb{N}$, let $m_1, \dots, m_\ell$ be integers with $m_i \le n/2$ for all i, and let $K_1, K_2 \ge 1$. Further, let $v \in \mathbb{R}^n$, and let B be an $\ell \times n$ random matrix with independent rows such that the i-th row is uniformly distributed on the set of vectors with m_i ones and $n - m_i$ zeros. Then for any non-random vector $Z \in \mathbb{R}^\ell$ we have*

$$\mathbb{P}\big\{\|Bv - Z\| \le \sqrt{\ell}\, t\big\} \le \Big(2C_{3.8}C_{2.1}t/\sqrt{\min_i m_i}\Big)^\ell \quad \text{for all } t \ge \max_i \frac{\sqrt{m_i}}{\mathbf{UD}_n(v, m_i, K_1, K_2)}.$$

The parameter K_2 which did not participate in any way in the proof of Theorem 2.1 is needed to guarantee a certain stability property of $\mathbf{UD}_n(v, m, K_1, K_2)$. We would like to emphasize that the use of functions ψ_{K_2} is a technical element of the argument.

Proposition 4.2 (Stability of the u-degree)**.** *For any $K_2 \ge 1$ there are $c_{4.2}, c'_{4.2} > 0$ depending only on K_2 with the following property. Let $K_1 \ge 1$, $v \in \mathbb{R}^n$, $k \in \mathbb{N}$, $m \le n/2$, and assume that $\mathbf{UD}_n(v, m, K_1, K_2) \le c'_{4.2} k$. Then there is a vector $y \in \big(\frac{1}{k}\mathbb{Z}\big)^n$ such that $\|v - y\|_\infty \le \frac{1}{k}$, and such that*

$$\mathbf{UD}_n(y, m, c_{4.2}K_1, K_2) \le \mathbf{UD}_n(v, m, K_1, K_2) \le \mathbf{UD}_n(y, m, c_{4.2}^{-1}K_1, K_2)$$

To prove the proposition we need two auxiliary lemmas.

Lemma 4.3. *Let $0 \ne z \in \mathbb{C}$, $\varepsilon \in [0, |z|/2]$ and let W be a random vector in $\mathbb{C}$ with $\mathbb{E}W = 0$ and with $|W| \le \varepsilon$ everywhere on the probability space. Then*

$$\big|\mathbb{E}|z + W| - |z|\big| \le \frac{\varepsilon^2}{|z|}.$$

Proof. We can view both z and W as vectors in $\mathbb{R}^2$, and can assume without loss of generality that $z = (z_1, 0)$, with $z_1 = |z|$. Then $|z_1 + W_1| = z_1 + W_1$ and

$$z_1 + W_1 \le |z + W| = \sqrt{(z_1 + W_1)^2 + W_2^2} \le (z_1 + W_1) + \frac{W_2^2}{2|z_1 + W_1|} \le (z_1 + W_1) + \frac{\varepsilon^2}{2(|z| - \varepsilon)}.$$

Hence,

$$|z| = z_1 = \mathbb{E}(z_1 + W_1) \le \mathbb{E}|z + W| \le \mathbb{E}(z_1 + W_1) + \frac{\varepsilon^2}{|z|} = |z| + \frac{\varepsilon^2}{|z|},$$

which implies the desired estimate. □

Lemma 4.4. *Let $\lambda, \mu \in \mathbb{R}$, and let ξ be a random variable in $\mathbb{R}$ with $\mathbb{E}\xi = \mu$ and with $|\xi - \mu| \le \lambda$ everywhere on the probability space. Then for any $s \in \mathbb{R}$ we have*

$$\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,\xi\, s\big) - \exp\big(2\pi\mathbf{i}\,\mu\, s\big)\big| \le (2\pi\lambda s)^2.$$

Proof. Denote $\xi' = \xi - \mu$. Then $\mathbb{E}\xi' = 0$ and $|\xi'| \le \lambda$. Therefore, using that $|\sin x| \le |x|$ and $|\sin x - x| \le x^2/2$ for every $x \in \mathbb{R}$, we obtain

$$\begin{aligned}\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,\xi\,s\big) - \exp\big(2\pi\mathbf{i}\,\mu\,s\big)\big| &= \big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,\xi' s\big) - 1\big| = \big|\mathbb{E}\cos\big(2\pi\xi' s\big) - 1 + \mathbf{i}\,\mathbb{E}\sin\big(2\pi\xi' s\big)\big| \\ &= \big| - 2\mathbb{E}\sin^2\big(\pi\xi' s\big) + \mathbf{i}\,\mathbb{E}\big(\sin\big(2\pi\xi' s\big) - 2\pi\xi' s\big)\big| \le 2(\pi\lambda s)^2 + (2\pi\lambda s)^2/2 = (2\pi\lambda s)^2.\end{aligned}$$

□

Proof of Proposition 4.2. To prove the proposition, we will use the *randomized rounding* which is a well known notion in computer science, and was recently applied in the random matrix context in [33] (see also [51, 34]). Define a random vector Y in $\big(\frac{1}{k}\mathbb{Z}\big)^n$ with independent components $Y_1, \dots, Y_n$ such that each component Y_i has distribution

$$Y_i = \begin{cases} \frac{1}{k}\lfloor kv_i\rfloor, & \text{with probability } \lfloor kv_i\rfloor - kv_i + 1, \\ \frac{1}{k}\lfloor kv_i\rfloor + \frac{1}{k}, & \text{with probability } kv_i - \lfloor kv_i\rfloor. \end{cases}$$

Then $\mathbb{E}Y_i = v_i$, $i \le n$ and, deterministically, $\|v - Y\|_\infty \le 1/k$.

Fix for a moment a number $s \in (0, k/(14\pi K_2)]$ and a subset $S \subset [n]$ of cardinality $\lfloor n/m\rfloor$. Our intermediate goal is to estimate the quantity

$$\mathbb{E}\,\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big).$$

Denote

$$V = V_S := \left|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,v_j\,s\big)\right| = \big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,v_{\eta[S]}\,s\big)\big|$$

and consider two cases.

Case 1. $V \le \frac{1}{2K_2} - \frac{2\pi s}{k}$. Using that $|e^{\mathbf{i}x} - 1| \le |x|$ for every $x \in \mathbb{R}$, we observe that deterministically

$$|\exp\big(2\pi\mathbf{i}\,v_j\,s\big) - \exp\big(2\pi\mathbf{i}\,Y_j\,s\big)| \le 2\pi s/k. \tag{13}$$

Therefore, by the definition of the function ψ_{K_2}, in this case we have on the entire probability space

$$\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big) = \psi_{K_2}(V) = \frac{1}{K_2}.$$

Case 2. $V > \frac{1}{2K_2} - \frac{2\pi s}{k} \ge \frac{1}{4K_2}$. Set

$$z := \frac{1}{\lfloor n/m\rfloor}\mathbb{E}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big) \quad \text{and} \quad W := \frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big) - z.$$

Then $\mathbb{E}W = 0$ and, using again $|e^{\mathbf{i}x} - 1| \le |x|$, we see that $|W| \le 2\pi s/k$ everywhere. By Lemma 4.4, $|z - V| \le (2\pi s/k)^2$, in particular, $z \ge V - (2\pi s/k)^2 \ge 1/(3K_2) \ge 4\pi s/k \ge |W|/2$. Therefore we may apply Lemma 4.3, to obtain

$$\big|\mathbb{E}|W + z| - |z|\big| \le \frac{4\pi^2 s^2}{|z|k^2} \le \frac{12\pi^2 K_2 s^2}{k^2}.$$

This implies,

$$\Big|\mathbb{E}\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big| - V\Big| = \big|\mathbb{E}|W + z| - |z| + |z| - V\big| \le \frac{16\pi^2 K_2 s^2}{k^2}. \tag{14}$$

To convert the last relation to estimating $\psi_{K_2}(\cdot)$, we will use the assumption that the second derivative of ψ_{K_2} is uniformly bounded. Applying Taylor's expansion around the point V, we get

$$\mathbb{E}\,\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big) = \psi_{K_2}\big(V\big) + \mathbb{E}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big| - V\Big)\,\psi'_{K_2}(V) + C''\Big\|\,\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big| - V\Big\|_\infty^2,$$

for some $C''>0$ which depends only on K_2. Here, $\|\cdot\|_\infty$ denotes the essential supremum of the random variable, and is bounded above by $2\pi s/k$ by 13. Together with (14) and with $\|\psi'_{K_2}\|_\infty\le 1$, this gives

$$\Big|\mathbb{E}\,\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big) - \psi_{K_2}(V)\Big| \le \frac{\bar C\,s^2}{k^2},$$

where $\bar C$ depends only on K_2.

Since $\psi'_{K_2}\ge 1/(2K_2)$, in both cases we obtain for some $\hat C>0$ depending only on K_2,

$$\Big|\mathbb{E}\,\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big) - \psi_{K_2}(V)\Big| \le \frac{\hat C\,s^2}{k^2}\psi_{K_2}(V).$$

Using this inequality together with definition of $V=V_S$, integrating over s, and summing over all choices of disjoint subsets $S_1,\dots,S_m$ of cardinality $\lfloor n/m\rfloor$, for every $t\in(0,k/(14\pi K_2)]$ we get the relation

$$\begin{aligned}
&\sum_{S_1,\dots,S_m}\int_{-t}^{t}\max\Big(0,1-\frac{c_0\,s^2}{k^2}\Big)^m\prod_{i=1}^m\psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,v_{\eta[S_i]}\,s\big)\big|\big)\,ds\\
&\quad\le\sum_{S_1,\dots,S_m}\int_{-t}^{t}\prod_{i=1}^m\mathbb{E}_Y\,\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S_i}\exp\big(2\pi\mathbf{i}\,Y_j\,s\big)\Big|\Big)\,ds\\
&\quad\le\sum_{S_1,\dots,S_m}\int_{-t}^{t}\Big(1+\frac{C_0\,s^2}{k^2}\Big)^m\prod_{i=1}^m\psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,v_{\eta[S_i]}\,s\big)\big|\big)\,ds,
\end{aligned}$$

where $C_0,c_0>7\pi K_2$ are constants that may only depend on K_2. Using independence of the components of Y, we can take the expectation with respect to Y out of the integral.

Given a vector $Q=(q_1,\dots,q_n)\in\mathbb{R}^n$ and $t\in(0,k/(14\pi K_2)]$, denote

$$g_t(Q):=\sum_{S_1,\dots,S_m}\int_{-t}^{t}\prod_{i=1}^m\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S_i}\exp\big(2\pi\mathbf{i}\,q_j\,s\big)\Big|\Big)\,ds.$$

The above relation implies that there are two (non-random) realizations Y' and Y'' of Y such that for

$$g_t(Y')\ge I_1:=\max\Big(0,1-\frac{c_0\,t^2}{k^2}\Big)^m\sum_{S_1,\dots,S_m}\int_{-t}^{t}\prod_{i=1}^m\psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\,v_{\eta[S_i]}\,s\big)\big|\big)\,ds$$

and

$$g_t(Y'') \le I_2 := \Big(1 + \frac{C_0\, t^2}{k^2}\Big)^m \sum_{S_1,\dots,S_m} \int_{-t}^{t} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\, v_{\eta[S_i]}\, s\big)\big|\big)\, ds.$$

Using properties of the function ψ_{K_2}, we note that for any two non-random vectors $\widetilde{Y}$ and $\hat{Y}$ in the range of Y such that they differ on a single coordinate, one has $g_t(\widetilde{Y}) \le 4K_2\, g_t(\hat{Y})$. Consider a path $Y^{(1)} = Y', Y^{(2)}, Y^{(3)}, \dots, Y''$ from Y' to Y'' consisting of a sequence of non-random vectors in the range of Y such that each adjacent pair $Y^{(i)}, Y^{(i+1)}$ differs on a single coordinate and let

$$S := \{i \,:\, g_t(Y^{(i)}) > 4K_2 I_2\} \subset [1, n-1].$$

If $S = \emptyset$, take $\mathbf{Y} = Y^{(1)}$. Otherwise, let $\ell = \max\{i \,:\, g_t(Y^{(i)}) > 4K_2 I_2\}$. Then take $\mathbf{Y} = Y^{(\ell+1)}$ and note $g_t(Y^{(\ell+1)}) \ge g_t(Y^{(\ell)})/(4K_2) \ge I_2 \ge I_1$. Thus the vector $\mathbf{Y}$ is in the range of Y and

$$I_1 \le g_t(\mathbf{Y}) \le 4K_2 I_2.$$

Making substitutions $s' = \sqrt{m}s$, $t' = \sqrt{m}t$ in the integrals in I_1, I_2, and assuming that $t' \le k/\max(2C_0, 2c_0)$ (in this case the condition $t \le k/(14\pi K_2)$ is satisfied), we can rewrite the last inequalities as

$$\begin{aligned}
&\frac{1}{2}\sum_{S_1,\dots,S_m} \int_{-t'}^{t'} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\, v_{\eta[S_i]} m^{-1/2}\, s\big)\big|\big)\, ds \\
&\qquad \le \sum_{S_1,\dots,S_m} \int_{-t'}^{t'} \prod_{i=1}^{m} \psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{j\in S_i} \exp\big(2\pi\mathbf{i}\, \mathbf{Y}_j m^{-1/2}\, s\big)\Big|\Big)\, ds \\
&\qquad\qquad \le 6K_2 \sum_{S_1,\dots,S_m} \int_{-t'}^{t'} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\, v_{\eta[S_i]} m^{-1/2}\, s\big)\big|\big)\, ds.
\end{aligned}$$

The result follows by the definition of $\mathbf{UD}_n(\cdot)$. □

The last statement to be considered in this subsection asserts that the u-degree of any vector from $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ is at least of order $\sqrt{m}$.

Proposition 4.5 (Lower bound on the u-degree)**.** *For any r, δ, ρ there is $C_{4.5} > 0$ depending only on r, δ, ρ with the following property. Let $K_2 \ge 2$, $1 \le m \le n/C_{4.5}$, $K_1 \ge C_{4.5}$ and let $x \in \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$. Then*

$$\mathbf{UD}_n(x, m, K_1, K_2) \ge \sqrt{m}.$$

Lemma 4.6. *For any $\rho > 0$ and $\kappa \in (0, 1/2]$ there is a constant $\widetilde{C} > 0$ depending only on ρ and κ with the following property. Let $S \ne \emptyset$ be a finite subset of $\mathbb{Z}$, and let $(y_w)_{w\in S}$ be a real vector (indexed by S). Assume further that S_1, S_2 are two disjoint subsets of S, each of cardinality at least $\kappa|S|$ such that $\min_{w\in S_1} y_w \ge \max_{w\in S_2} y_w + \rho$. Let $K_2 \ge 2$ and f be a function on $[0,1]$ defined by*

$$f(t) := \psi_{K_2}\Big(\Big|\frac{1}{|S|}\sum_{w\in S} \exp(2\pi\mathbf{i}\, y_w\, t)\Big|\Big), \quad t \in [0,1].$$

Then for every $b > 0$ one has

$$\big|\big\{t \in [0,1] :\ f(t) \ge 1 - b^2\big\}\big| \le \widetilde{C} b.$$

Proof. Clearly we may assume that $b \le 1/\sqrt{2}$. Denote $m = \lceil \kappa |S| \rceil$ and

$$g(t) := \Big| \sum_{w \in S} \exp(2\pi \mathbf{i}\, y_w\, t) \Big|, \quad t \in \mathbb{R}.$$

Let $T \subset S_1 \times S_2$ be of cardinality $T = m$ and such that for all $(q,j), (q',j') \in T$ with $(q,j) \ne (q',j')$ one has $q \ne q'$ and $j \ne j'$. Then for all $t \in \mathbb{R}$,

$$g(t) = \Big| \sum_{w \in S_1 \cup S_2} \exp(2\pi \mathbf{i}\, y_w\, t) + \sum_{w \notin S_1 \cup S_2} \exp(2\pi \mathbf{i}\, y_w\, t) \Big| \le \sum_{(q,j) \in T} \big| 1 + \exp(2\pi \mathbf{i}\, (y_j - y_q)\, t) \big| + |S| - 2m.$$

Further, take any $u \in (0, 1/\sqrt{2\kappa})$ and observe that for each $(q,j) \in T$, since $|y_j - y_q| \ge \rho$, we have

$$\big| \big\{ t \in [0,1] : \ \big| 1 + \exp(2\pi \mathbf{i}\, (y_j - y_q)\, t) \big| \ge 2 - 2u^2 \big\} \big| \le C' u,$$

where $C' > 0$ may only depend on ρ. This implies that

$$\Big| \Big\{ t \in [0,1] : \ \big| 1 + \exp(2\pi \mathbf{i}\, (y_j - y_q)\, t) \big| \ge 2 - 2u^2 \text{ for at least } m/2 \text{ pairs } (q,j) \in T \Big\} \Big| \le 2C' u.$$

On the other hand, whenever $t \in [0,1]$ is such that $\big| 1 + \exp(2\pi \mathbf{i}\, (y_j - y_q)\, t) \big| \ge 2 - 2u^2$ for at most $m/2$ pairs $(q,j) \in T$, we have

$$g(t) \le \frac{m}{2}(2 - 2u^2) + \frac{m}{2} \cdot 2 + |S| - 2m = |S| - mu^2 \le |S|(1 - \kappa u^2),$$

whence $f(t) \le \max\big(\frac{1}{K_2}, 1 - \kappa u^2\big) = 1 - \kappa u^2$. Taking $u = \frac{b}{\sqrt{\kappa}}$ we obtain the desired result with $\widetilde{C} = \frac{2C'}{\sqrt{\kappa}}$. $\square$

Proof of Proposition 4.5. Let A_{nm} be defined as in (7) and n_δ, C_δ, $\mathcal{S}$ be from Lemma 3.3. We assume that $n \ge n_\delta$ and $n/m \ge C_\delta$. For every $i \le m$ denote

$$f_i(s) = \psi_{K_2}\big(\big| \mathbb{E} \exp\big(2\pi \mathbf{i}\, x_{\eta[S_i]}\, m^{-1/2} s \big) \big| \big).$$

Further, let subsets Q_1 and Q_2 be taken from the definition of non-constant vectors applied to x. Then by Lemma 3.3 and since $\psi_{K_2}(1) \le 1$,

$$A_{nm} \sum_{(S_1,\dots,S_m) \in \mathcal{S}} \int_{-\sqrt{m}}^{\sqrt{m}} \prod_{i=1}^{m} f_i \, ds \le \ e^{-c_\delta n}\, 2\sqrt{m} + A_{nm} \sum_{(S_1,\dots,S_m) \in \mathcal{S}'} \int_{-\sqrt{m}}^{\sqrt{m}} \prod_{i=1}^{m} f_i \, ds,$$

where $\mathcal{S}'$ is the set of all sequences $(S_1, \dots, S_m) \in \mathcal{S}$ such that is the subset of S such that

$$\min(|S_i \cap Q_1|, |S_i \cap Q_2|) \ge \frac{\delta}{2} \lfloor n/m \rfloor \ \text{ for at least } \ c_\delta m \ \text{ indices } \ i. \tag{15}$$

Take any $(S_1, \dots, S_m) \in \mathcal{S}'$ and denote $m_0 := \lceil c_\delta m \rceil$. Without loss of generality we assume that (15) holds for all $i \le m_0$. Applying Lemma 4.6 with $\kappa := \delta/2$ and $b = \sqrt{1-u}$, we get for all $u \in (0,1]$ and $i \le m_0$,

$$\mu(u) := \Big| \Big\{ s \in [-\sqrt{m}, \sqrt{m}] : \ f_i \ge u \Big\} \Big| \le \widetilde{C} \sqrt{m} \sqrt{1-u},$$

where $\widetilde{C} > 0$ depends only on δ and ρ. This estimate implies that for $i \le m_0$,

$$\int_{-\sqrt{m}}^{\sqrt{m}} (f_i(s))^{m_0} \, ds = \int_0^1 m_0\, u^{m_0 - 1}\, \mu_u \, ds \le \widetilde{C} \sqrt{m}\, m_0\, B(3/2, m_0) \le C_2,$$

where B denotes the Beta-function and $C_2 > 0$ is a constant depending only on ρ and δ. Applying Hölder's inequality, we obtain

$$\int_{-\sqrt{m}}^{\sqrt{m}} \prod_{i=1}^{m} \psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\, x_{\eta[S_i]}\, m^{-1/2} s\big)\big|\big)\, ds \le \int_{-\sqrt{m}}^{\sqrt{m}} \prod_{i=1}^{m_0} \psi_{K_2}\big(\big|\mathbb{E}\exp\big(2\pi\mathbf{i}\, x_{\eta[S_i]}\, m^{-1/2} s\big)\big|\big)\, ds \le C_2,$$

which impies the desired result. □

4.2 No moderately unstructured normal vectors

Let M_n be an $n \times n$ Bernoulli(p) random matrix.. For each $i \le n$, denote by $H_i = H_i(M_n)$ the span of columns $\mathbf{C}_j(M_n)$, $j \ne i$. The goal of this subsection is to prove Theorem 2.2, which asserts that under appropriate restrictions on n and p with a very large probability (say, at least $1 - 2e^{-2pn}$), the subspace $H_i^{\perp}$ is either structured or very unstructured. The main ingredient of the proof — Proposition 4.9 — will be considered in the next subsection. Here, we will only state the proposition to be used as a black box and for this we need to introduce an additional product structure, which, in a sense, replaces the set $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$.

Fix a permutation $\sigma \in \Pi_n$, two disjoint subsets Q_1, Q_2 of cardinality $\lceil \delta n \rceil$ each, and a number $h \in \mathbb{R}$ such that

$$\forall i \in Q_1 :\ h + 2 \le \mathbf{g}(n/\sigma^{-1}(i)) \qquad \text{and} \qquad \forall i \in Q_2 :\ -\mathbf{g}(n/\sigma^{-1}(i)) \le h - \rho - 2. \tag{16}$$

Define the sets $\Lambda_n = \Lambda_n(k, \mathbf{g}, Q_1, Q_2, \rho, \sigma, h)$ by

$$\Lambda_n := \left\{ x \in \frac{1}{k}\mathbb{Z}^n : |x_{\sigma(i)}| \le \mathbf{g}(n/i) \ \text{ for all } i \le n, \ \ \min_{i \in Q_1} x_i \ge h, \ \ \text{and} \ \ \max_{i \in Q_2} x_i \le h - \rho \right\}. \tag{17}$$

In what follows, we adopt the convention that $\Lambda_n = \emptyset$ whenever h does not satisfy (16).

Lemma 4.7. *There exists an absolute constant $C_{4.7} \ge 1$ such that for every $n \ge 1$ there is a subset $\bar{\Pi}_n \subset \Pi_n$ of cardinality at most $\exp(C_{4.7} n)$ with the following property. For any two partitions $(S_i)_{i=1}^m$ and $(S_i')_{i=1}^m$ of $[n]$ with $2^{-i+1} n \ge |S_i| = |S_i'|$, $i \le m$, there is $\sigma \in \bar{\Pi}_n$ such that $\sigma(S_i) = S_i'$, $i \le m$.*

This lemma immediately follows from the fact that the total number of partitions $(S_i)_{i=1}^m$ of $[n]$ satisfying $2^{-i+1} n \ge |S_i|$, $i \le m$, is exponential in n (one can take $C_{4.7} = 23$). Using Lemma 4.7, we provide an efficient approximation of $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$.

Lemma 4.8. *For any $x \in \mathcal{V}_n = \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$, $k \ge 4/\rho$, and any $y \in \frac{1}{k}\mathbb{Z}^n$ with $\|x - y\|_\infty \le 1/k$ one has*

$$y \in \bigcup_{q = \lfloor -4\mathbf{g}(6n)/\rho \rfloor}^{\lceil 4\mathbf{g}(6n)/\rho \rceil} \ \bigcup_{\bar{\sigma} \in \bar{\Pi}_n} \ \bigcup_{|Q_1|, |Q_2| = \lceil \delta n \rceil} \Lambda_n(k, \mathbf{g}(6\,\cdot), Q_1, Q_2, \rho/4, \bar{\sigma}, \rho q/4),$$

where the set of permutations $\bar{\Pi}_n$ is taken from Lemma 4.7.

Proof. Let $x \in \mathcal{V}_n$, and assume that $y \in \frac{1}{k}\mathbb{Z}^n$ satisfies $\|x - y\|_\infty \le 1/k$. Then, by the definition of $\mathcal{V}_n$, there exist sets $Q_1, Q_2 \subset [n]$, each of cardinality $\lceil \delta n \rceil$, satisfying

$$\max_{i \in Q_2} y_i - \frac{1}{k} \le \max_{i \in Q_2} x_i \le \min_{i \in Q_1} x_i - \rho \le \min_{i \in Q_1} y_i - \rho + \frac{1}{k}.$$

Then $\max\limits_{i\in Q_2} y_i \le \min\limits_{i\in Q_1} y_i - \frac{\rho}{2}$, hence we can find a number $h \in \frac{\rho}{4}\mathbb{Z}$ such that

$$\min_{i\in Q_1} y_i \ge h \qquad \text{and} \qquad \max_{i\in Q_2} y_i \le h - \frac{\rho}{4}.$$

By the definition of $\mathcal{V}_n$ we also have $|x_{\sigma_x(i)}| \le \mathbf{g}(n/i)$ for all $i \in [n]$. By the definition of $\bar\Pi_n$, we can find a permutation $\bar\sigma \in \bar\Pi_n$ such that

$$\sigma_x\big(\{\lfloor n/2^\ell\rfloor + 1, \dots, \lfloor n/2^{\ell-1}\rfloor\}\big) = \bar\sigma\big(\{\lfloor n/2^\ell\rfloor + 1, \dots, \lfloor n/2^{\ell-1}\rfloor\}\big) \quad \text{for all } \ell \ge 1.$$

Clearly for such a permutation we have $|x_{\bar\sigma(i)}| \le \mathbf{g}(2n/i)$ for every $i \le n$. Using (8), we obtain

$$|y_{\bar\sigma(i)}| \le |x_{\bar\sigma(i)}| + \frac{1}{k} \le \mathbf{g}(2n/i) + \frac{1}{k} \le \mathbf{g}(6n/i) - 2.$$

Thus

$$\forall i \in \bar\sigma^{-1}(Q_1):\ h \le \min_{i\in Q_1} y_i \le \mathbf{g}(6n/i) - 2 \quad \text{and} \quad \forall i \in \bar\sigma^{-1}(Q_2):\ h - \frac{\rho}{4} \ge \max_{i\in Q_2} y_i \ge 2 - \mathbf{g}(6n/i).$$

Since $h = \rho q/4$ for some $q \in \mathbb{Z}$, this implies the desired result. □

The following statment, together with Theorem 2.1 and Proposition 4.2, is the main ingredient of the proof of Theorem 2.2.

Proposition 4.9. *Let $\varepsilon \in (0, 1/8]$, $\rho, \delta \in (0, 1/4]$ and let the growth function $\mathbf{g}$ satisfies (8). There exist $K_{4.9} = K_{4.9}(\delta, \rho) \ge 1$, $n_{4.9} = n_{4.9}(\varepsilon, \delta, \rho, K_3)$, and $C_{4.9} = C_{4.9}(\varepsilon, \delta, \rho, K_3) \in \mathbb{N}$ with the following property. Let $\sigma \in \Pi_n$, $h \in \mathbb{R}$, and let $Q_1, Q_2 \subset [n]$ be such that $|Q_1|, |Q_2| = \lceil \delta n \rceil$. Let $8 \le K_2 \le 1/\varepsilon$, $n \ge n_{4.9}$, $m \ge C_{4.9}$ with $n/m \ge C_{4.9}$, $1 \le k \le \min\big((K_2/8)^{m/2}, 2^{n/C_{4.9}}\big)$, and let $X = (X_1, \dots, X_n)$ be a random vector uniformly distributed on $\Lambda_n(k, \mathbf{g}, Q_1, Q_2, \rho, \sigma, h)$. Then*

$$\mathbb{P}\big\{\mathbf{UD}_n(X, m, K_{4.9}, K_2) < km^{1/2}/C_{4.9}\big\} \le \varepsilon^n.$$

Let us describe the proof of Theorem 2.2 informally. Assume that the hyperplane H_1 admits a normal vector X which belongs to $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$. We need to show that with a large probability the u-degree $\mathbf{UD}_n(X, m, K_1, K_2)$ of X is very large, say, at least ε^{-m} for a small $\varepsilon > 0$. The idea is to split the collection $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ into about $\log_2(\varepsilon^{-m})$ subsets according to the magnitude of the u-degree (that is, each subset $\mathcal{T}_N$ will have a form $\mathcal{T}_N = \big\{x \in \mathcal{V}_n(r, \mathbf{g}, \delta, \rho):\ \mathbf{UD}_n(x, m, K_1, K_2) \in [N, 2N)\big\}$ for an appropriate N). To show that for each $N \ll \varepsilon^{-m}$ the probability of $X \in \mathcal{T}_N$ is very small, we define a discrete approximation $\mathcal{A}_N$ of $\mathcal{T}_N$ consisting of all vectors $y \in \frac{1}{k}\mathbb{Z}^n$ such that $\|y - x\|_\infty \le 1/k$ for some $x \in \mathcal{T}_N$ and additionally, in view of Proposition 4.2, $\mathbf{UD}_n(y, m, c_{4.2}K_1, K_2) \le 2N$ and $\mathbf{UD}_n(y, m, c_{4.2}^{-1}K_1, K_2) \ge N$. We can bound the cardinality of such set $\mathcal{A}_N$ by $(\tilde\varepsilon\, k)^n$, for a small $\tilde\varepsilon > 0$, by combining Proposition 4.9 with Lemma 4.8 and with the following simple fact.

Lemma 4.10. *Let $k \ge 1$, $h \in \mathbb{R}$, $\rho, \delta \in (0, 1)$, $Q_1, Q_2 \subset [n]$ with $|Q_1|, |Q_2| = \lceil \delta n \rceil$, and $\mathbf{g}$ satisfies (8) with some $K_3 \ge 1$. Then $|\Lambda_n(k, \mathbf{g}, Q_1, Q_2, \rho, \sigma, h)| \le \big(C_{4.10}k\big)^n$, where $C_{4.10} \ge 1$ depends only on K_3.*

On the other hand, for each fixed vector y in the set $\mathcal{A}_N$ we can estimate the probability that it "approximates" a normal vector to H_1 by using Corollary 4.1:

$$\mathbb{P}\big\{y \text{ is an "approximate" normal vector to } H_1\big\} \le (C'/k)^n \quad \text{for every } y \in \mathcal{A}_N,$$

for some constant $C' \ll \tilde\varepsilon^{-1}$. Taking the union bound, we obtain

$$\mathbb{P}\big\{X \in \mathcal{T}_N\big\} \le \mathbb{P}\big\{\mathcal{A}_N \text{ contains an "approximate" normal vector to } H_1\big\} \le (C'/k)^n\,(\tilde\varepsilon\, k)^n \ll 1.$$

Below, we make this argument rigorous.

Proof of Theorem 2.2. We start by defining parameters. We always assume that n is large enough, so all statements used below work for our n. Fix any $R \geq 1$, $r > 0$ and $s > 0$, and set $b := \lfloor (2pR)^{-1} \rfloor$. Let $K_2 = 32\exp(16R)$. Note that the function $\mathbf{g}(6\,\cdot)$ is a growth function that satisfies condition (8) with parameter $K_3' = (K_3)^8$. In particular, choosing j so that $2^{j-1} \leq 6n \leq 2^j$, we have

$$\mathbf{g}(6n) \leq \mathbf{g}(2^j) \leq (K_3')^{2^j/j} \leq (K_3')^{12n/\log_2(6n)} \leq K_3^n.$$

For brevity, we denote

$$C_{3.7} := C_{3.7}(s, 2R),\ C_{3.5} := C_{3.5}(2R),\ c_{4.2}' := c_{4.2}'(K_2),\ c_{4.2} := c_{4.2}(K_2),\ C_{4.10} = C_{4.10}(K_3').$$

Set

$$K_1 := \max\big(K_{4.9}(\delta, \rho/4)/c_{4.2}, C_{4.5}(r, \delta, \rho)\big),$$

and

$$\varepsilon := \min\Big(K_2^{-1},\, c_{4.2}'\big(384eK_3\,\exp(C_{4.7})\,C_{4.10}C_{3.8}C_{2.1}C_{3.7}\big)^{-1}\exp(-3R)\Big)$$

We will assume that pn is sufficiently large so that

$$5\exp(-2Rpn) \leq \exp(-Rpn) \quad \text{and} \quad \exp(-3Rpn) \leq \frac{1}{2Rpn}\exp(-2Rpn).$$

Moreover, we will assume that

$$2RC_{3.5}p \leq 1 \quad \text{and} \quad C_{3.5} \leq pn \tag{18}$$

and

$$\frac{1}{8p} \geq \max(C_{4.9}(\varepsilon, \delta, \rho/4, K_3'), C_{4.5}(r, \delta, \rho)); \quad pn \geq 16C_{4.9}(\varepsilon, \delta, \rho/4, K_3')^2;$$
$$e^{2Rp} \leq 2^{1/C_{4.9}(\varepsilon,\delta,\rho/4,K_3')}; \quad c_{4.2}'/3 \geq \exp(-Rpn); \quad \lfloor \exp(Rpn)/c_{4.2}' \rfloor\, n \leq 2^n.$$

Define two auxiliary random objects as follows. Set

$$Z := \{x \in \mathbb{R}^n :\ x^*_{\lfloor rn \rfloor} = 1,\ \ \mathbf{UD}_n(x, m, K_1, K_2) \geq \exp(Rpn) \ \text{ for all } \ pn/8 \leq m \leq 8pn\},$$

and let X be a random vector measurable with respect to H_1 and such that

- $X \in \big(\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp\big) \setminus Z$ whenever $\big(\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp\big) \setminus Z \neq \emptyset$;
- $X \in \big(\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp\big) \cap Z$ whenever $\big(\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp\big) \setminus Z = \emptyset$ and $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp \neq \emptyset$;
- $X = \mathbf{0}$ whenever $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cap H_1^\perp = \emptyset$.

(Note that $H_1^\perp$ may have dimension larger than one with non-zero probability, and thus $\pm X$ is not uniquely defined). Note that to prove the theorem, it is sufficient to show that with probability at least $1 - \exp(-Rpn)$ one has either $X = \mathbf{0}$ or $X \in Z$.

Next, we denote

$$\xi := \begin{cases} \min\limits_{8pn \geq m \geq pn/8} \mathbf{UD}_n(X, m, K_1, K_2), & \text{whenever } \ X \neq \mathbf{0}; \\ +\infty, & \text{otherwise.} \end{cases}$$

Then, proving the theorem amounts to showing that $\xi < \exp(Rpn)$ with probability at most $\exp(-Rpn)$.

We say that a collection of indices $I \subset [n]$ is admissible if $1 \notin I$ and $|I| \geq n - b - 1$. For admissible sets I consider disjoint collection of events $\{\mathcal{E}_I\}_I$ defined by

$$\mathcal{E}_I := \big\{\forall i \in I : \ |\operatorname{supp} \mathbf{C}_i(M_n)| \in [pn/8, 8pn] \quad \text{and} \quad \forall i \notin I : \ |\operatorname{supp} \mathbf{C}_i(M_n)| \notin [pn/8, 8pn]\big\}.$$

Further, denote

$$\widetilde{\mathcal{E}} := \big\{\|M_n - \mathbb{E}M_n\| \leq C_{3.7}\sqrt{pn}\big\}.$$

According to Corollary 3.7, $\mathbb{P}(\widetilde{\mathcal{E}}) \geq 1 - \exp(-2Rpn)$, while by Lemma 3.5 and (18),

$$\mathbb{P}\Big(\bigcup_I \mathcal{E}_I\Big) \geq 1 - \exp(-n/C_{3.5}) \geq 1 - \exp(-2Rpn).$$

Denote by $\mathcal{I}$ the collection of all admissible I satisfying $2\mathbb{P}(\mathcal{E}_I \cap \widetilde{\mathcal{E}}) \geq \mathbb{P}(\mathcal{E}_I)$. Then for $I \in \mathcal{I}$, we have $\mathbb{P}(\mathcal{E}_I) \geq 2\mathbb{P}(\mathcal{E}_I \cap \widetilde{\mathcal{E}}^c)$, and, using that events $\mathcal{E}_I$ are disjoint,

$$\mathbb{P}\Big(\bigcup_{I\in\mathcal{I}} \mathcal{E}_I\Big) \geq 1 - \exp(-2Rpn) - 2\mathbb{P}(\widetilde{\mathcal{E}}^c) \geq 1 - 3\exp(-2Rpn).$$

Hence,

$$\begin{aligned}\mathbb{P}\big\{\xi < \exp(Rpn)\big\} &\leq \sum_{I\in\mathcal{I}} \mathbb{P}\big(\big\{\xi < \exp(Rpn)\big\} \cap \mathcal{E}_I \cap \widetilde{\mathcal{E}}\big) + \mathbb{P}\Big(\bigcap_{I\in\mathcal{I}} \mathcal{E}_I^c\Big) + \mathbb{P}(\widetilde{\mathcal{E}}^c) \\ &\leq \sum_{I\in\mathcal{I}} \mathbb{P}\big(\big\{\xi < \exp(Rpn)\big\} \mid \mathcal{E}_I \cap \widetilde{\mathcal{E}}\big)\mathbb{P}(\mathcal{E}_I \cap \widetilde{\mathcal{E}}) + 4\exp(-2Rpn).\end{aligned}$$

Therefore, to prove the theorem it is sufficient to show that for any $I \in \mathcal{I}$,

$$\mathbb{P}\big(\big\{\xi < \exp(Rpn)\big\} \mid \mathcal{E}_I \cap \widetilde{\mathcal{E}}\big) \leq \exp(-2Rpn).$$

Fix an admissible $I \in \mathcal{I}$, denote by B_I the $|I| \times n$ matrix obtained by transposing columns $\mathbf{C}_i(M_n)$, $i \in I$, and let $\widetilde{B}_I$ be the non-random $|I| \times n$ matrix with all elements equal to p. Note that, in view of our definition of K_1, the assumptions on p and Proposition 4.5, we have a *deterministic* relation

$$\xi \geq \sqrt{pn/8}$$

everywhere on the probability space. For each real number $N \in J_p := [\sqrt{pn/8}, \exp(Rpn)/2]$, denote by $\mathcal{E}_{N,I}$ the event

$$\mathcal{E}_{N,I} := \big\{\xi \in [N, 2N)\big\} \cap \mathcal{E}_I \cap \widetilde{\mathcal{E}}.$$

Splitting the interval J_p into subintervals, we observe that it is sufficient to show that for every $N \in J_p$ we have

$$\mathbb{P}\big(\mathcal{E}_{N,I} \mid \mathcal{E}_I \cap \widetilde{\mathcal{E}}\big) \leq \exp(-3Rpn) \leq \frac{1}{2Rpn}\exp(-2Rpn).$$

The rest of the argument is devoted to estimating probability of $\mathcal{E}_{N,I}$ for fixed $N \in J_p$ and fixed $I \in \mathcal{I}$. Set $k := \lceil 2N/c'_{4.2} \rceil$. Let $\mathbf{m} : \mathcal{E}_{N,I} \to [pn/8, 8pn]$ be a (random) integer such that

$$\mathbf{UD}_n(X, \mathbf{m}, K_1, K_2) \in [N, 2N) \quad \text{everywhere on} \quad \mathcal{E}_{N,I}.$$

Since on $\mathcal{E}_{N,I}$ we have $\mathbf{UD}_n(X, \mathbf{m}, K_1, K_2) \leq 2N \leq c'_{4.2}k$, applying Proposition 4.2, we can construct a random vector $\mathbf{Y} : \mathcal{E}_{N,I} \to \frac{1}{k}\mathbb{Z}^n$ having the following properties:

- $\|\mathbf{Y} - X\|_\infty \le 1/k$ everywhere on $\mathcal{E}_{N,I}$,
- $\mathbf{UD}_n(\mathbf{Y}, \mathbf{m}, c_{4.2}K_1, K_2) \le 2N$ everywhere on $\mathcal{E}_{N,I}$,
- $\mathbf{UD}_n(\mathbf{Y}, m, c_{4.2}^{-1}K_1, K_2) \ge N$ for all $m \in [pn/8, 8pn]$ and everywhere on $\mathcal{E}_{N,I}$.

The first condition together with the inclusion $\mathcal{E}_{N,I} \subset \widetilde{\mathcal{E}}$ implies that

$$\|(B_I - \widetilde{B}_I)(\mathbf{Y} - X)\| \le C_{3.7}\sqrt{p}n/k.$$

Using that $B_I X = 0$ and that $\widetilde{B}_I(\mathbf{Y} - X) = p(\sum_{i=1}^n (\mathbf{Y}_i - X_i))\,\mathbf{1}_I$, we observe that there is a random number $\mathbf{z} : \mathcal{E}_{N,I} \to [-pn/k, pn/k] \cap \frac{\sqrt{pn}}{k}\mathbb{Z}$ such that everywhere on $\mathcal{E}_{N,I}$ one has

$$\|B_I\mathbf{Y} - \mathbf{z}\,\mathbf{1}_I\| \le 2C_{3.7}\sqrt{p}n/k.$$

Let Λ be a subset of

$$\bigcup_{q=\lfloor -4\mathbf{g}(6n)/\rho\rfloor}^{\lceil 4\mathbf{g}(6n)/\rho\rceil} \bigcup_{\bar\sigma\in\bar\Pi_n} \bigcup_{|Q_1|,|Q_2|=\lceil \delta n\rceil} \Lambda_n(k, \mathbf{g}(6\,\cdot), Q_1, Q_2, \rho/4, \bar\sigma, \rho q/4),$$

consisting of all vectors y such that

- $\mathbf{UD}_n(y, m, c_{4.2}K_1, K_2) \le 2N$ for *some* $m \in [pn/8, 8pn]$;
- $\mathbf{UD}_n(y, m, c_{4.2}^{-1}K_1, K_2) \ge N$ for all $m \in [pn/8, 8pn]$.

Note that by Lemma 4.8 the entire range of $\mathbf{Y}$ on $\mathcal{E}_{N,I}$ falls into Λ.

Combining the above observations,

$$\mathcal{E}_{N,I} \subset \big\{\|B_I y - z\mathbf{1}_I\| \le 2C_{3.7}\sqrt{p}n/k \ \text{ for some } y \in \Lambda,\ z \in [-pn/k, pn/k] \cap \tfrac{\sqrt{pn}}{k}\mathbb{Z}\big\},$$

whence, using that $2\mathbb{P}(\mathcal{E}_I \cap \widetilde{\mathcal{E}}) \ge \mathbb{P}(\mathcal{E}_I)$ by the definition of $\mathcal{I}$,

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{N,I} \,|\, \mathcal{E}_I \cap \widetilde{\mathcal{E}}) &\le 2\mathbb{P}\big\{\|B_I y - z\mathbf{1}_I\| \le 2C_{3.7}\sqrt{p}n/k \ \text{ for some } y \in \Lambda,\ z \in [-pn/k, pn/k] \cap \tfrac{\sqrt{pn}}{k}\mathbb{Z} \mid \mathcal{E}_I\big\}\\ &\le 6|\Lambda|\sqrt{pn}\max_{z\in\frac{\sqrt{pn}}{k}\mathbb{Z}}\max_{y\in\Lambda}\mathbb{P}\big\{\|B_I y - z\mathbf{1}_I\| \le 2C_{3.7}\sqrt{p}n/k \mid \mathcal{E}_I\big\}.\end{aligned}$$

To estimate the last probability, we apply Corollary 4.1 with $t := C_{3.7}\sqrt{8pn}/N$ (note that $k \ge 2N$, $2|I| \ge n$, and that t satisfies the assumption of the corollary). We obtain that for all admissible y and z,

$$\begin{aligned}\mathbb{P}\big\{\|B_I y - z\mathbf{1}_I\| \le 2C_{3.7}\sqrt{p}n/k \mid \mathcal{E}_I\big\} &\le \mathbb{P}\bigg\{\|B_I y - z\mathbf{1}_I\| \le \frac{C_{3.7}\sqrt{8pn}}{N}\sqrt{|I|} \mid \mathcal{E}_I\bigg\}\\ &\le (16C_{3.8}C_{2.1}C_{3.7}/N)^{|I|}.\end{aligned}$$

On the other hand, the cardinality of Λ can be estimated by combining Lemma 4.10, Lemma 4.7 and Proposition 4.9 (note that our choice of parameters guarantees applicability of these statements):

$$|\Lambda| \le 8pn\varepsilon^n\,(9\mathbf{g}(6n)/\rho)\exp(C_{4.7}n)\,2^{2n}(C_{4.10}k)^n \le (72pn/\rho)\varepsilon^n\,K_3^n\,\exp(C_{4.7}n)\,2^{2n}(C_{4.10}k)^n,$$

where $C_{4.10} = C_{4.10}(K_3')$. Thus, using our choice of parameters and assuming in addition that $2^n \ge 72pn/\rho$

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{N,I} \,|\, \mathcal{E}_I \cap \widetilde{\mathcal{E}}) &\le \varepsilon^n\,(8K_3\,\exp(C_{4.7})\,C_{4.10}k)^n\,(16C_{3.8}C_{2.1}C_{3.7}/N)^{|I|}\\ &\le \varepsilon^n\,(8K_3\,\exp(C_{4.7})\,C_{4.10}k)^n\,(48C_{3.8}C_{2.1}C_{3.7}/(c_{4.2}'k))^n N^{1+\lfloor(2pR)^{-1}\rfloor}\\ &\le \varepsilon^n\,(384K_3\,\exp(C_{4.7})\,C_{4.10}C_{3.8}C_{2.1}C_{3.7}/(c_{4.2}'))^n\,e^n\\ &\le \exp(-3Rn),\end{aligned}$$

by our choice of parameters. The result follows. □

4.3 Anti-concentration on a lattice

The goal of this subsection is to prove Proposition 4.9. Thus, in this subsection, we fix $\rho, \delta \in (0, 1/4]$, a growth function $\mathbf{g}$ satisfying (8), which in particular means that $\mathbf{g}(n) \le K_3^{2n/\log_2 n}$, a permutation $\sigma \in \Pi_n$, a number $h \in \mathbb{R}$, two sets $Q_1, Q_2 \subset [n]$ such that $|Q_1|, |Q_2| = \lceil \delta n \rceil$, and we do not repeat these assumptions in lemmas below. We also always use short notation Λ_n for the set $\Lambda_n(k, \mathbf{g}, Q_1, Q_2, \rho, \sigma, h)$ defined in (17).

We start with auxiliary probabilistic statements which are just special forms of Markov's inequality.

Lemma 4.11 (Integral form of Markov's inequality, I)**.** *For each $s \in [a, b]$, let $\xi(s)$ be a non-negative random variable with $\xi(s) \le 1$ a.e. Assume that the random function $\xi(s)$ is integrable on $[a, b]$ with probability one. Assume further that for some integrable function $\phi(s) : [a, b] \to \mathbb{R}_+$ and some $\varepsilon > 0$ we have*

$$\mathbb{P}\big\{\xi(s) \le \phi(s)\big\} \ge 1 - \varepsilon$$

for all $s \in [a, b]$. Then for all $t > 0$,

$$\mathbb{P}\bigg\{ \int_a^b \xi(s)\, ds \ge \int_a^b \phi(s)\, ds + t(b-a) \bigg\} \le \varepsilon/t.$$

Proof. Consider a random set

$$I := \big\{ s \in [a, b] :\ \xi(s) > \phi(s) \big\}.$$

Since $\mathbb{P}\{s \in I\} \le \varepsilon$ for any $s \in [a, b]$, we have $\mathbb{E}|I| \le \varepsilon(b-a)$. Therefore, by Markov's inequality, $\mathbb{P}\big\{|I| \ge t(b-a)\big\} \le \varepsilon/t$ for all $t > 0$. The result follows by noting that

$$\int_a^b \xi(s)\, ds \le |I| + \int_a^b \phi(s)\, ds.$$

□

Lemma 4.12 (Integral form of Markov's inequality, II)**.** *Let I be a finite set, and for each $i \in I$, let ξ_i be a non-negative random variable with $\xi_i \le 1$ a.e. Assume further that for some $\phi(i) : I \to \mathbb{R}_+$ and some $\varepsilon > 0$ we have*

$$\mathbb{P}\big\{\xi_i \le \phi(i)\big\} \ge 1 - \varepsilon$$

for all $i \in I$. Then for all $t > 0$,

$$\mathbb{P}\bigg\{ \frac{1}{|I|} \sum_{i \in I} \xi_i \ge \frac{1}{|I|} \sum_{i \in I} \phi(i) + t \bigg\} \le \varepsilon/t.$$

The proof of Lemma 4.12 is almost identical to that of Lemma 4.11, and we omit it.

Our next statement will be important in an approximation (discretization) argument used later in the proof.

Lemma 4.13 (Lipschitzness of the product $\prod \psi_{K_2}(\cdot)$)**.** *Let $y_1, \dots, y_n \in \mathbb{R}$ and set $y := \max_{w \le n} |y_w|$. Further, let $S_1, \dots, S_m$ be some non-empty subsets of $[n]$. For $i \le m$ denote*

$$f_i(s) := \psi_{K_2}\bigg(\Big| \frac{1}{|S_i|} \sum_{w \in S_i} \exp(2\pi \mathbf{i}\, y_w s) \Big| \bigg) \quad \text{and let} \quad f(s) := \prod_{i=1}^m f_i(s).$$

Then f (viewed as a function of s) is $(8K_2 \pi y\, m)$-Lipschitz.

Proof. By our definition, ψ_{K_2} is 1-Lipschitz for any $K_2 \geq 1$, hence f_i (viewed as a function of s) is $2\pi y$-Lipschitz. Since $\left|\sum_{w\in S_i} \exp(2\pi \mathbf{i}\, y_w s)\right| \leq |S_i|$, by the definition of the function ψ_{K_2}, we have $1/(2K_2) \leq f_i \leq 1$, hence, for all $s, \Delta s \in \mathbb{R}$,

$$\frac{f_i(s)}{f_i(s+\Delta s)} = 1 + \frac{f_i(s) - f_i(s+\Delta s)}{f_i(s+\Delta s)} \leq 1 + 4K_2\pi y\,|\Delta s|.$$

Taking the product, we obtain that

$$\frac{f(s)}{f(s+\Delta s)} \leq \big(1 + 4K_2\pi y\,|\Delta s|\big)^m \leq 1 + 8K_2\pi y\, m\,|\Delta s|$$

whenever $8K_2\pi y\, m\,|\Delta s| \leq 1/2$. This, together with the bound $f \leq 1$ implies for all $s, \Delta s \in \mathbb{R}$,

$$f(s) - f(s+\Delta s) \leq 8K_2\pi y\, m\,|\Delta s|,$$

which completes the proof. □

In the next two lemmas we initiate the study of random variables $\exp(2\pi \mathbf{i}\,\eta[I_w]\, s_j/k)$, more specifically, we will be interested in the property that, under appropriate assumptions on s_j's, the sum of such variables is close to zero in average.

Lemma 4.14. *Let $\varepsilon \in (0,1]$, $k \geq 1$, $\ell \geq 2/\varepsilon$. Let I be an integer interval and recall that $\eta[I]$ denotes a random variable uniformly distributed on I. Assume that $s_1, \dots, s_\ell$ are real numbers such that for all $j \neq u$,*

$$\frac{k}{\varepsilon|I|} \leq |s_j - s_u| \leq \frac{k}{2}.$$

Then

$$\mathbb{E}\left|\sum_{j=1}^{\ell} \exp(2\pi \mathbf{i}\,\eta[I]\, s_j/k)\right|^2 \leq \varepsilon \ell^2.$$

Proof. We will determine the restrictions on parameter R at the end of the proof. We have

$$\begin{aligned}\mathbb{E}\left|\sum_{j=1}^{\ell} \exp(2\pi \mathbf{i}\,\eta[I]\, s_j/k)\right|^2 &= \sum_{j=1}^{\ell}\sum_{u=1}^{\ell} \mathbb{E}\exp\big(2\pi \mathbf{i}\,\eta[I]\,(s_j - s_u)/k\big) \\ &\leq \ell + \Big|\sum_{j\neq u} \mathbb{E}\exp\big(2\pi \mathbf{i}\,\eta[I]\,(s_j - s_u)/k\big)\Big|.\end{aligned} \tag{19}$$

Further, denoting $a = \min I$ and $b = \max I$, we observe for any $j \neq u$,

$$\begin{aligned}&\mathbb{E}\exp\big(2\pi \mathbf{i}\,\eta[I]\,(s_j - s_u)/k\big) \\ &= \frac{1}{|I|}\sum_{v=a}^{b} \exp\big(2\pi \mathbf{i}\, v\,(s_j - s_u)/k\big) \\ &= \frac{1}{|I|}\exp\big(2\pi \mathbf{i}\, a\,(s_j - s_u)/k\big)\cdot \frac{1 - \exp\big(2\pi \mathbf{i}\,(b-a+1)\,(s_j - s_u)/k\big)}{1 - \exp\big(2\pi \mathbf{i}\,(s_j - s_u)/k\big)}.\end{aligned}$$

In view of assumptions on $|s_j - s_u|$,

$$\big|1 - \exp\big(2\pi \mathbf{i}\,(s_j - s_u)/k\big)\big| = \big|2\sin(\pi\,(s_j - s_u)/k)\big| \geq \frac{4|s_j - s_u|}{k} \geq \frac{4}{\varepsilon|I|}.$$

Therefore,

$$\left|\mathbb{E}\exp\left(2\pi\mathbf{i}\,\eta[I]\,(s_j-s_u)/k\right)\right|\le\frac{\varepsilon}{2}.$$

Using (19), we complete the proof. □

Lemma 4.15. *For every $\varepsilon\in(0,1/2]$ there are $R_{4.15}=R_{4.15}(\varepsilon)>0$ and $\ell:=\ell_{4.15}(\varepsilon)\in\mathbb{N}$, $\ell\ge 1000$, with the following property. Let $k\ge 1$, $u\ge\ell$, let I_w ($w=1,2,\dots,u$) be integer intervals, and let $s_1,\dots,s_\ell$ be real numbers such that $|I_w|\,|s_j-s_q|\ge R_{4.15}k$, and $|s_j-s_q|\le k/2$ for all $j\ne q$ and $w\le u$. Then, assuming that random variables $\eta[I_w]$, $w\le u$, are mutually independent, one has*

$$\mathbb{P}\Big\{\Big|\frac{1}{u}\sum_{w=1}^{u}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge\varepsilon\ \text{ for at least }\varepsilon\ell\text{ indices }j\Big\}\le\varepsilon^u.$$

Proof. Fix any $\varepsilon\in(0,1/2]$, and set $\varepsilon_1:=2^{-10}e^{-6}\varepsilon^{4+9/\varepsilon}$. Set $R:=1/\varepsilon_1$ and $\ell:=\lceil 2/\varepsilon_1\rceil$. Assume that $u\ge\ell$, and let numbers s_j and integer intervals I_w satisfy the assumptions of the lemma. Denote the event

$$\Big\{\Big|\frac{1}{u}\sum_{w=1}^{u}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge\varepsilon\ \text{ for at least }\varepsilon\ell\text{ indices }j\Big\}$$

by $\mathcal{E}$, and additionally, for any subset $Q\subset[\ell]$ of cardinality $\lfloor\varepsilon\ell/4\rfloor$ and any vector $z\in\{-1,1\}^2$, set

$$\mathcal{E}_{Q,z}:=\Big\{\Big\langle\Big(\frac{1}{u}\sum_{w=1}^{u}\cos(2\pi\,\eta[I_w]\,s_j/k),\frac{1}{u}\sum_{w=1}^{u}\sin(2\pi\,\eta[I_w]\,s_j/k)\Big),z\Big\rangle\ge\varepsilon\ \text{ for all }j\in Q\Big\}.$$

It is not difficult to see that

$$\mathcal{E}\subset\bigcup_{Q,z}\mathcal{E}_{Q,z},$$

whence it is sufficient to show that for any admissible Q,z,

$$\mathbb{P}(\mathcal{E}_{Q,z})\le\frac{1}{4}\binom{\ell}{\lfloor\varepsilon\ell/4\rfloor}^{-1}\varepsilon^u.\tag{20}$$

Without loss of generality, we can consider $Q=Q_0:=\big[\lfloor\varepsilon\ell/4\rfloor\big]$. Event $\mathcal{E}_{Q_0,z}$ is contained inside the event

$$\Big\{\Big|\sum_{j\in Q_0}\sum_{w=1}^{u}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge 2^{-1/2}\varepsilon u\,\lfloor\varepsilon\ell/4\rfloor\Big\},$$

while the latter is contained inside the event

$$\Big\{\Big|\sum_{j\in Q_0}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge\frac{\varepsilon}{4}\,\lfloor\varepsilon\ell/4\rfloor\text{ for at least }\varepsilon u/4\text{ indices }w\Big\}.$$

Thus, taking the union over all admissible choices of $\lceil\varepsilon u/4\rceil$ indices $w\in[u]$, we get

$$\mathbb{P}(\mathcal{E}_{Q_0,z})\le\binom{u}{\lceil\varepsilon u/4\rceil}\max_{F\subset[u],\,|F|=\lceil\varepsilon u/4\rceil}\mathbb{P}\Big\{\Big|\sum_{j\in Q_0}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge\frac{\varepsilon}{4}\,\lfloor\varepsilon\ell/4\rfloor\text{ for all }w\in F\Big\}.$$

To estimate the last probability, we apply Markov's inequality, together with the bound for the second moment from Lemma 4.14 (applied with ε_1), and using independence of $\eta[I_w]$, $w\le u$. We then get

$$\max_{\substack{F\subset[u]\\|F|=\lceil\varepsilon u/4\rceil}}\mathbb{P}\Big\{\Big|\sum_{j\in Q_0}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|\ge\frac{\varepsilon}{4}\,\lfloor\varepsilon\ell/4\rfloor\text{ for all }w\in F\Big\}\le\left(\frac{\varepsilon_1\ell^2}{(\varepsilon^2\ell/32)^2}\right)^{\lceil\varepsilon u/4\rceil}\le e^{-3\varepsilon u/2}\varepsilon^{2u}.$$

In view of (20) this implies the result, since using $8 \le \ell \le u$ and $\varepsilon < 1/2$, we have

$$4\binom{\ell}{\lfloor \varepsilon \ell/4 \rfloor} \varepsilon^{-u} \binom{u}{\lceil \varepsilon u/4 \rceil} e^{-3\varepsilon u/2} \varepsilon^{2u} \le 4e^{-3\varepsilon u/2} \left(\frac{4e}{\varepsilon}\right)^{\varepsilon \ell/4} \left(\frac{2e}{\varepsilon}\right)^{\varepsilon u/2} \varepsilon^u \le 4(16e^{-3})^{\varepsilon u/4}\, \varepsilon^{u/4} \le 1.$$

□

Our next step is to show that for the vector $X = (X_1, \dots, X_n)$ uniformly distributed on Λ_n the random product $\prod_{i=1}^{m} \psi_{K_2}\big(\big|\frac{1}{\lfloor n/m \rfloor} \sum_{w \in S_i} \exp(2\pi \mathbf{i} X_w s)\big|\big)$ is, in a certain sense, typically small (for most choices of s). To do this we first show that given a collection of distinct numbers $s_1, \dots, s_\ell$ which are pairwise well separated, the above product is small for at least one s_j with very high probability.

Lemma 4.16. *For any $\varepsilon \in (0, 1/2]$ there are $R_{4.16} = R_{4.16}(\varepsilon) \ge 1$ and $\ell := \ell_{4.16}(\varepsilon) \in \mathbb{N}$ with the following property. Let $k, m, n \in \mathbb{N}$ be with $n/m \ge \ell$. Let $1 \le K_2 \le 2/\varepsilon$, $X = (X_1, \dots, X_n)$ be a random vector uniformly distributed on Λ_n, and let $s_1, \dots, s_\ell$ be real numbers in $[0, k/2]$ such that $|s_j - s_q| \ge R_{4.16}$ for all $j \ne q$. Fix disjoint subsets $S_1, \dots, S_m$ of $[n]$ of cardinality $\lfloor n/m \rfloor$ each. Then*

$$\mathbb{P}\Big\{\forall j \le \ell \,:\, \prod_{i=1}^{m} \psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m \rfloor} \sum_{w \in S_i} \exp(2\pi \mathbf{i} X_w s_j)\Big|\Big) \ge (K_2/2)^{-m/2}\Big\} \le \varepsilon^n.$$

Proof. Fix any $\varepsilon \in (0, 1/2]$ and set $\ell := \ell_{4.15}(\varepsilon^5) \ge 1000$ and $R := R_{4.15}(\varepsilon^5)$. Assume that $n/m \ge \ell$. Note that, by our definition of Λ_n, the coordinates of X are independent and, moreover, each variable kX_w is distributed on an integer interval of cardinality at least k. Thus, it is sufficient to prove that for any collection of integer intervals I_j, $j \le n$, satisfying $|I_j| \ge k$, the event

$$\mathcal{E} := \Big\{\forall j \le \ell \,:\, \prod_{i=1}^{m} \psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m \rfloor} \sum_{w \in S_i} \exp(2\pi \mathbf{i}\, \eta[I_w]\, s_j/k)\Big|\Big) \ge (K_2/2)^{-m/2}\Big\}.$$

has probability at most ε^n, where, as usual, we assume that the variables $\eta[I_w]$, $w \in S_i$, $i \le m$, are jointly independent. Observe that, as $\psi_{K_2}(t) \le 1$ for all $t \le 1$, the event $\mathcal{E}$ is contained inside the event

$$\mathcal{E}' := \Big\{\forall j \le \ell \,:\, a_{ij} \ge 2/K_2 \text{ for at least } m/2 \text{ indices } i\Big\},$$

where $a_{ij} := \Big|\frac{1}{\lfloor n/m \rfloor} \sum_{w \in S_i} \exp(2\pi \mathbf{i}\, \eta[I_w]\, s_j/k)\Big|$, $i \le m$, $j \le \ell$. Denoting $b_{ij} = 1$ if $a_{ij} \ge 2/K_2$ and $b_{ij} = 0$ otherwise and using a simple counting argument for the matrix $\{b_{ij}\}_{ij}$, we obtain that

$$\mathcal{E} \subset \mathcal{E}' \subset \mathcal{E}'' := \Big\{\Big|\Big\{i \,:\, \ a_{ij} \ge 2/K_2 \ \text{ for at least } \ell/4 \text{ indices } j\Big\}\Big| \ge m/4\Big\}.$$

To estimate $\mathbb{P}(\mathcal{E}'')$ we use Lemma 4.15 with ε^5. Note that $\varepsilon^5 \le \min(2/K_2, 1/2)$, and that by our choice of R, for any $j \ne q$ we have $|I_w|\,|s_j - s_q| \ge k\,|s_j - s_q| \ge R_{4.15}(\varepsilon^5)k$, while $|s_j - s_q| \le k/2$. Thus,

$$\forall i \le m \,:\quad \mathbb{P}\Big\{a_{ij} \ge 2/K_2 \ \text{ for at least } \ell/4 \text{ indices } j\Big\} \le \varepsilon^{5\lfloor n/m \rfloor}.$$

Hence,

$$\mathbb{P}(\mathcal{E}'') \le \binom{m}{\lceil m/4 \rceil} \varepsilon^{5\lfloor n/m \rfloor\, m/4} \le 2^m \varepsilon^{5\lfloor n/m \rfloor\, m/4} \le \varepsilon^n,$$

which completes the proof. □

Lemma 4.17 (Very small product everywhere except for a set of measure $O(1)$)**.** *For any $\varepsilon \in (0,1/2]$ there are $R_{4.17} = R_{4.17}(\varepsilon) \geq 1$, $\ell = \ell_{4.17}(\varepsilon) \in \mathbb{N}$ and $n_{4.17} = n_{4.17}(\varepsilon, K_3) \in \mathbb{N}$ with the following property. Let $k, m, n \in \mathbb{N}$, $n \geq n_{4.17}$, $k \leq 2^{n/\ell}$, $n/m \geq \ell$, and $4 \leq K_2 \leq 2/\varepsilon$. Let $X = (X_1, \ldots, X_n)$ be a random vector uniformly distributed on Λ_n. Fix disjoint subsets $S_1, \ldots, S_m$ of $[n]$, each of cardinality $\lfloor n/m \rfloor$. Then*

$$\mathbb{P}\Big\{\Big|\Big\{s \in [0,k/2]:\ \prod_{i=1}^{m} \psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_w s)\Big|\Big) \geq (K_2/4)^{-m/2}\Big\}\Big| \leq R_{4.17}\Big\} \geq 1-(\varepsilon/2)^n.$$

Proof. Fix any $\varepsilon \in (0,1/2]$, and define $\widetilde{\varepsilon} := \varepsilon^{3/2}/32$, $\widetilde{\ell} := \ell_{4.16}(\widetilde{\varepsilon})$, $\ell := 2\widetilde{\ell}$, and $R := 4R_{4.16}(\widetilde{\varepsilon})\ell_{4.16}(\widetilde{\varepsilon}) > 1$.

Assume that the parameters k, m, n and $S_1, \ldots, S_m$ satisfy the assumptions of the lemma. In particular, we assume that n is large enough so that $(8K_2\pi n)^{\widetilde{\ell}} \leq 2^n$ and $\mathbf{g}(n)^{\widetilde{\ell}} \leq 2^n$. Denote

$$\beta := (8K_2\pi m\mathbf{g}(n))^{-1}(2K_2)^{-m/2} \quad \text{and} \quad a_{ij} := \Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}\,\eta[I_w]\,s_j/k)\Big|,\ i \leq m,\ j \leq \widetilde{\ell}.$$

Let $T := [0,k/2] \cap \beta\mathbb{Z}$. By Lemma 4.16 for any collection $s_1, \ldots, s_{\widetilde{\ell}}$ of points from T satisfying $|s_j - s_q| \geq R_{4.16}(\widetilde{\varepsilon})$ for all $j \neq q$, we have

$$\mathbb{P}\Big\{\forall j \leq \widetilde{\ell}:\ \prod_{i=1}^{m}\psi_{K_2}(a_{ij}) \geq (K_2/2)^{-m/2}\Big\} \leq \widetilde{\varepsilon}^n.$$

Taking the union bound over all possible choices of $s_1, \ldots, s_{\widetilde{\ell}}$ from T, we get

$$\begin{aligned}\mathbb{P}\Big\{ \prod_{i=1}^{m}\psi_{K_2}(a_{ij}) \geq (K_2/2)^{-m/2} \text{ for all } j \leq \widetilde{\ell} \text{ and for some } s_1, \ldots, s_{\widetilde{\ell}} \in T \\ \text{with } |s_p - s_q| \geq R_{4.16}(\widetilde{\varepsilon}) \text{ for all } p \neq q\Big\} \leq \widetilde{\varepsilon}^n |T|^{\widetilde{\ell}}.\end{aligned} \tag{21}$$

Further, by of Lemma 4.13, for any realization of X_w's the product

$$f(s) := \prod_{i=1}^{m}\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_w s)\Big|\Big),$$

viewed as a function of s, is $(8K_2\pi\mathbf{g}(n)m)$-Lipschitz. This implies that for any pair $(s,s') \in \mathbb{R}^2_+$, satisfying $|s-s'| \leq \beta$, we have

$$f(s) \geq (K_2/2)^{-m/2} \qquad \text{whenever} \quad f(s') \geq (K_2/4)^{-m/2}.$$

Moreover, for any collection $s'_1, \ldots, s'_{\widetilde{\ell}}$ of numbers from $[0,k/2]$ satisfying $|s'_p - s'_q| \geq 2R_{4.16}(\widetilde{\varepsilon})$ for all $p \neq q$ there are numbers $s_1, \ldots, s_{\widetilde{\ell}} \in T$ with $|s_q - s'_q| \leq \beta$ $|s_p - s_q| \geq R_{4.16}(\widetilde{\varepsilon})$ for all $p \neq q$ (we used also $2\beta \leq 1 \leq R_{4.16}(\widetilde{\varepsilon})$). This, together with (21), yields

$$\begin{aligned}&\mathbb{P}\Big\{ \prod_{i=1}^{m}\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_w s'_j)\Big|\Big) \geq (K_2/4)^{-m/2} \text{ for all } j \leq \widetilde{\ell} \text{ and some } s'_1, \ldots, s'_{\widetilde{\ell}} \in [0,k/2] \\ &\quad \text{with } |s'_p - s'_q| \geq 2R_{4.16}(\widetilde{\varepsilon}) \text{ for all } p \neq q\Big\} \\ &\qquad \leq \widetilde{\varepsilon}^n|T|^{\widetilde{\ell}} \leq \widetilde{\varepsilon}^n\,(k/\beta)^{\widetilde{\ell}} \leq \widetilde{\varepsilon}^n\,2^n\,(8K_2\pi m\mathbf{g}(n))^{\widetilde{\ell}}(2K_2)^{m\widetilde{\ell}/2} \\ &\qquad \leq \widetilde{\varepsilon}^n\,8^n\,(4/\varepsilon)^{m\widetilde{\ell}/2} \leq \widetilde{\varepsilon}^n\,\varepsilon^{-n/2}\,16^n \leq (\varepsilon/2)^n.\end{aligned}$$

The event whose probability is estimated above, clearly contains the event in the question —

$$\Big\{\Big|\Big\{s\in[0,k/2]:\ \prod_{i=1}^m\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_ws)\Big|\Big)\ge(K_2/4)^{-m/2}\Big\}\Big|\ge 4R_{4.16}(\widetilde{\varepsilon})\widetilde{\ell}\Big\}.$$

This, and our choice of parameters, implies the result. □

Lemma 4.18 (Moderately small product for almost all s)**.** *For any $\varepsilon\in(0,1]$ and $z\in(0,1)$ there are $\varepsilon'=\varepsilon'(\varepsilon)\in(0,1/2]$, $n_{4.18}=n_{4.18}(\varepsilon,z)\ge 10$, and $C_{4.18}=C_{4.18}(\varepsilon,z)\ge 1$ with the following property. Let $n\ge n_{4.18}$, $2^n\ge k\ge 1$, $C_{4.18}\le m\le n/4$, and $4\le K_2\le 1/\varepsilon$. Let $X=(X_1,\dots,X_n)$ be a random vector uniformly distributed on Λ_n. Fix disjoint subsets $S_1,\dots,S_m$ of $[n]$ of cardinality $\lfloor n/m\rfloor$ each. Then*

$$\mathbb{P}\Big\{\forall s\in[z,\varepsilon' k]\ :\ \prod_{i=1}^m\psi_{K_2}\Big(\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_ws)\Big|\Big)\le e^{-\sqrt{m}}\Big\}\ge 1-(\varepsilon/2)^n.$$

Proof. Let $\varepsilon'>0$ will be chosen later. Fix any $s\in[z,\varepsilon' k]$. Assume $m\ge(\varepsilon' z)^{-4}\ge 10$. For $i\le m$ denote

$$\gamma_i(s):=\Big|\frac{1}{\lfloor n/m\rfloor}\sum_{w\in S_i}\exp(2\pi\mathbf{i}X_ws)\Big|,\qquad f_i(s):=\psi_{K_2}\big(\gamma_i(s)\big),\quad\text{ and }\quad f(s):=\prod_{i=1}^m f_i(s)$$

Observe that by the definition of ψ_{K_2} for each $i\le m$ we have $f_i(s)=\gamma_i(s)$, provided $\gamma_i(s)\ge 1/K_2$. Next note that if for some complex unit numbers $z_1,...,z_N$ their average $v:=\sum_{i=1}^N z_i/N$ has length $1-\alpha>0$ then, taking the unit complex number z_0 satisfying $\langle z_0,v\rangle=|v|$ we have

$$N(1-\alpha)\le\sum_{i=1}^N Re\,\langle z_i,v\rangle\le N,$$

therefore there are at least $N/2+1$ indices i such that $Re\,\langle z_i,v\rangle\ge 1-4\alpha$. This in turn implies that there exists an index j such that there are at least $N/2$ indices i with $Re\,\langle z_i,\bar{z}_j\rangle\ge 1-16\alpha$. Thus, the event $\big\{f_i(s)\ge 1-\frac{2}{\sqrt{m}}\big\}$ is contained in the event

$$\Big\{\exists\ w'\in S_i:\quad\cos(2\pi s(X_w-X_{w'}))\ge 1-\frac{32}{\sqrt{m}}\ \text{ for at least }\ \frac{n}{2m}\ \text{ indices }\ \ w\in S_i\setminus\{w'\}\Big\}.$$

To estimate the probability of the later event, we take the union bound over all choices of $n/(2m)$ indices from S_i, and over all choices of w'. We then get

$$\mathbb{P}\Big\{f_i(s)\ge 1-\frac{2}{\sqrt{m}}\Big\}\le\frac{n}{m}\,2^{\lfloor n/m\rfloor}\max_{\substack{w'\in S_i,\,F\subset S_i\setminus\{w'\},\\ |F|\ge n/(2m)}}\mathbb{P}\Big\{\forall w\in F:\ \operatorname{dist}(s(X_w-X_{w'}),\mathbb{Z})\le\frac{2}{m^{1/4}}\Big\}.$$

To estimate the latter probability (the probability following maximum in the previous line) we use the definition of Λ_n and independence of coordinates of the vector X. Note that for each fixed w there is an integer interval I_w of the length at least $2k$ such that X_w is uniformly distributed on I_w/k. Therefore, fixing a realization $X_{w'}=b/k$, $b\in\mathbb{Z}$, we need to count how many $a\in I_w$ are such that $s(a-b)/k$ is close to an integer. This can be done by splitting I_w into subintervals of length k and considering cases $z\le s\le 1$, $1<s\le C'k/m^{1/4}$ (this case can be empty), and $C'k/m^{1/4}<s\le\varepsilon' k$. This leads to the following bound with an absolute constant $C''>0$,

$$\mathbb{P}\Big\{f_i(s)\ge 1-\frac{2}{\sqrt{m}}\Big\}\le\frac{n}{m}\,2^{n/m}\left(\max\Big(\frac{C''}{z\,m^{1/4}},C''\varepsilon'\Big)\right)^{n/(2m)}\le\frac{n}{m}\left(4C''\varepsilon'\right)^{n/(2m)}.$$

Using this estimate and the fact that $\psi_{K_2}(t) \le 1$ for $t \le 1$ (so, each $f_i(s) \le 1$), we obtain

$$\mathbb{P}\Big\{ f(s) \ge \Big(1 - \frac{2}{\sqrt{m}}\Big)^{3m/4} \Big\} \le \mathbb{P}\Big\{ f_i(s) \ge 1 - \frac{2}{\sqrt{m}} \ \text{ for at least } m/4 \text{ indices } i \Big\}$$
$$\le 2^m \left(\frac{n}{m} \big(4C''\varepsilon'\big)^{n/(2m)} \right)^{m/4} = \left(\frac{16n}{m} \right)^{m/4} \big(4C''\varepsilon'\big)^{n/8}.$$

The last step of the proof is somewhat similar to the one used in the proof of Lemma 4.17 — we discretize the interval $[z, \varepsilon' k]$ and use the the fact that f is Lipschitz. Recall that $\mathbf{g}(n) \le 2^n$ and thus, by Lemma 4.13, $f(s)$ is $(8K_2\pi 2^n\, m)$-Lipschitz. Let

$$\beta := \big(1 - 2/\sqrt{m}\big)^{3m/4} \big(8K_2\pi\, 2^n m\big)^{-1} \qquad \text{and} \qquad T := [z, \varepsilon' k] \cap \beta\mathbb{Z}.$$

Then for any $s, s' \in [z, \varepsilon' k]$ satisfying $|s - s'| \le \beta$ we have $|f(s) - f(s')| \le \big(1 - 2/\sqrt{m}\big)^{3m/4}$ deterministically. This implies that

$$\mathbb{P}\Big\{ \forall s \in [z, \varepsilon' k] \ : \ \ f(s) \le 2\Big(1 - \frac{2}{\sqrt{m}}\Big)^{3m/4} \Big\} \ge \mathbb{P}\Big\{ \forall s \in T \ : \ \ f(s) \le \Big(1 - \frac{2}{\sqrt{m}}\Big)^{3m/4} \Big\}$$
$$\ge 1 - \frac{k}{\beta} \Big(\frac{16n}{m} \Big)^{m/4} \big(4C''\varepsilon'\big)^{n/8} \ge 1 - (\varepsilon/2)^n,$$

provided that $\varepsilon' := c''\varepsilon^8$ for a sufficiently small universal constant $c'' > 0$. □

Lemma 4.19. *Let $\rho, \varepsilon \in (0, 1]$, $k \ge 1$, $h \in \mathbb{R}$, $a_1 \ge h+1$, $a_2 \le h - \rho - 1$. Let Y_1, Y_2 be independent random variables, with Y_1 uniformly distributed on $[h, a_1] \cap \frac{1}{k}\mathbb{Z}$ and Y_2 uniformly distributed on $[a_2, h - \rho] \cap \frac{1}{k}\mathbb{Z}$. Then for every $s \in [-\varepsilon/8, \varepsilon/8]$ one has*

$$\mathbb{P}\big\{ \big| \exp\big(2\pi\mathbf{i} Y_1 s\big) + \exp\big(2\pi\mathbf{i} Y_2 s\big) \big| > 2 - 2\pi\rho^2 s^2 \big\} \le \varepsilon.$$

Proof. Clearly, it is enough to consider $0 < s < \varepsilon/8$ only. Note that

$$\big| \exp\big(2\pi\mathbf{i} Y_1 s\big) + \exp\big(2\pi\mathbf{i} Y_2 s\big) \big| = \big| 1 + \exp\big(2\pi\mathbf{i}(Y_1 - Y_2)s\big) \big| = 2 \big| \cos\big(\pi\mathbf{i}(Y_1 - Y_2)s\big) \big|.$$

We consider two cases.

Case 1. Assume that $a_1 \le h + 2\varepsilon^{-1}$ and $a_2 \ge h - 2\varepsilon^{-1}$. In this case, deterministically, $\rho \le Y_1 - Y_2 \le 4/\varepsilon$, therefore, using that $\cos t \le 1 - t^2/\pi$ on $[-\pi/2, \pi/2]$, we have for every $s \in (0, \varepsilon/8]$,

$$\big| \exp\big(2\pi\mathbf{i} Y_1 s\big) + \exp\big(2\pi\mathbf{i} Y_2 s\big) \big| \le 2 - 2\pi\rho^2 s^2.$$

Case 2. Assume that either $a_1 > h + 2\varepsilon^{-1}$ or $a_2 < h - 2\varepsilon^{-1}$. Without loss of generality, we will assume the first inequality holds. We condition on a realization $\widetilde{Y}_2$ of Y_2 (further in the proof, we compute conditional probabilities given $Y_2 = \widetilde{Y}_2$). For any $s \le \varepsilon/8$, the event

$$\big\{ \big| 1 + \exp\big(2\pi\mathbf{i}(Y_1 - \widetilde{Y}_2)s\big) \big| \ge 2 - s^2 \big\}$$

is contained inside the event

$$\big\{ \mathrm{dist}\big((Y_1 - \widetilde{Y}_2)s, \mathbb{Z}\big) \le s \big\}.$$

On the other hand, since $(Y_1 - \widetilde{Y}_2)s$ is uniformly distributed on a set $[b_1, b_2] \cap \frac{s}{k}\mathbb{Z}$, for some $b_2 \ge b_1 + 2\varepsilon^{-1}s$, the probability of the last event is less than ε. The result follows.

□

Lemma 4.20 (Integration for small s)**.** *For any $\widetilde{\varepsilon} \in (0,1]$, $\rho \in (0,1/4]$ and $\delta \in (0,1/2]$ there are $n_{4.20} = n_{4.20}(\widetilde{\varepsilon},\delta,\rho)$, $C_{4.20} = C_{4.20}(\widetilde{\varepsilon},\delta,\rho) \ge 1$, and $K_{4.20} = K_{4.20}(\delta,\rho) \ge 1$ with the following property. Let A_{nm} be defined as in (7), $n \ge n_{4.18}$, $k \ge 1$, $m \in \mathbb{N}$ with $n/m \ge C_{4.20}$ and $m \ge 2$, and let $X = (X_1,\dots,X_n)$ be a random vector uniformly distributed on Λ_n. Then for every $K_2 \ge 4$,*

$$\mathbb{P}\Big\{ A_{nm} \sum_{S_1,\dots,S_m} \int\limits_{-\sqrt{m}/C_{4.20}}^{\sqrt{m}/C_{4.20}} \prod_{i=1}^{m} \psi_{K_2}\Big(\Big| \frac{1}{\lfloor n/m \rfloor} \sum_{w\in S_i} \exp\big(2\pi \mathbf{i} X_w m^{-1/2} s\big)\Big|\Big)\, ds \ge K_{4.20} \Big\} \le (\widetilde{\varepsilon}/2)^n,$$

where the sum is taken over all disjoint subsets $S_1,\dots,S_m \subset [n]$ of cardinality $\lfloor n/m \rfloor$ each.

Proof. Let $n_\delta, C_\delta, c_\delta$, and $\mathcal{S}$ be as in Lemma 3.3). For a given choice of subsets $(S_1,\dots,S_m) \in \mathcal{S}$ denote

$$\gamma_i(s) := \Big| \frac{1}{\lfloor n/m \rfloor} \sum_{w\in S_i} \exp(2\pi \mathbf{i} X_w s)\Big|, \qquad f_i(s) := \psi_{K_2}\big(\gamma_i(s)\big), \quad \text{and} \quad f(s) := \prod_{i=1}^m f_i(s)$$

(note that functions $\gamma_i(s)$, $f_i(s)$, $f(s)$ depend on the choice of subsets S_i).

First, we study the distribution of the variable $f(s)$ for a given choice of subsets S_i. We assume that $n \ge n_\delta$ and $n/m \ge C_\delta$. We also denote $\varepsilon := 2^{-10/\delta}\, \widetilde{\varepsilon}^{16/\delta c_\delta}$ and

$$\mathcal{S}' := \Big\{ (S_1,\dots,S_m) \in \mathcal{S} :\ \min(|S_i \cap Q_1|, |S_i \cap Q_2|) \ge \delta \lfloor n/m \rfloor/2 \text{ for at least } c_\delta m \text{ indices } i \Big\}.$$

Fix a sequence $(S_1,\dots,S_m) \in \mathcal{S}'$, and $J \subset [m]$ be a subset of cardinality $\lceil c_\delta m \rceil$ such that

$$\forall i \in J :\ \min(|S_i \cap Q_1|, |S_i \cap Q_2|) \ge \delta \lfloor n/m \rfloor/2.$$

For any $i \in J$, $w_1 \in S_i \cap Q_1$, and $w_2 \in S_i \cap Q_2$ by Lemma 4.19 we have for $s \in [-\varepsilon/8, \varepsilon/8]$,

$$\mathbb{P}\big\{ \big| \exp\big(2\pi \mathbf{i} X_{w_1} s\big) + \exp\big(2\pi \mathbf{i} X_{w_2} s\big)\big| \ge 2 - 2\pi\rho^2 s^2 \big\} \le \varepsilon.$$

Within S_i, we can find at least $\frac{\delta}{2}\lfloor n/m \rfloor$ disjoint pairs of indices $(w_1,w_2) \in Q_1 \times Q_2$ satisfying the above condition. Let T be a set of such pairs with $|T| = \frac{\delta}{2}\lfloor n/m \rfloor$. Using the independence of coordinates of X, and denoting $z := \min\big(\sqrt{1/(\pi\rho^2\delta)}, \varepsilon/8\big)$, we obtain for every $s \in [-z,z]$,

$$\begin{aligned}
&\mathbb{P}\Big\{ \gamma_i(s) \ge 1 - \frac{\pi\rho^2\delta s^2}{2} \Big\} \\
&\quad \le \mathbb{P}\big\{ \big| \exp\big(2\pi \mathbf{i} X_{w_1} s\big) + \exp\big(2\pi \mathbf{i} X_{w_2} s\big)\big| \ge 2 - 2\pi\rho^2 s^2 \text{ for at least } \tfrac{\delta}{4}\lfloor n/m \rfloor \text{ pairs } (w_1,w_2) \in T \big\} \\
&\quad \le 2^{\delta\lfloor n/m \rfloor/2}\, \varepsilon^{\delta\lfloor n/m \rfloor/4} \le (4\varepsilon)^{\delta n/(4m)}.
\end{aligned}$$

Applying this for all $i \in J$ together with observations $f(s) \le 1$ and $f_i(s) = \gamma_i(s)$ (when $\gamma_i(s) \ge 1/K_2$), we conclude that for every $s \in [-z,z]$,

$$\begin{aligned}
\mathbb{P}\Big\{ f(s) \ge \big(1 - \pi\rho^2\delta s^2/2\big)^{|J|/2} \Big\} &\le \mathbb{P}\Big\{ f_i(s) \ge 1 - \pi\rho^2\delta s^2/2 \ \text{ for at least } |J|/2 \text{ indices } i \in J \Big\} \\
&\le 2^{|J|}\, (4\varepsilon)^{\delta |J| n/(8m)}
\end{aligned}$$

At the next step, we apply the Lemma 4.11 with $\xi(s) = f(s)$ to obtain from the previous relation

$$\mathbb{P}\Big\{ \int\limits_{-z}^{z} f(s)\, ds \le \int\limits_{-z}^{z} \Big(1 - \frac{\pi\rho^2\delta s^2}{2}\Big)^{|J|/2} ds + m^{-1/2} \Big\} \ge 1 - 2z m^{1/2}\, 2^{|J|}\, (4\varepsilon)^{\delta|J|n/(8m)}.$$

Next we apply Lemma 4.12) with $I = \mathcal{S}'$, $\xi_i = f(s)$ (recall that $f(s)$ depends also on the choice of $(S_1, \dots, S_m) \in \mathcal{S}$). We obtain

$$\mathbb{P}\Big\{ A_{nm} \sum_{(S_1,\dots,S_m)\in\mathcal{S}'} \int_{-z}^{z} f(s)\,ds \le \int_{-z}^{z} \Big(1 - \frac{\pi\rho^2\delta s^2}{2}\Big)^{|J|/2} ds + 2m^{-1/2} \Big\} \ge 1 - 2zm\,2^{|J|}\,(4\varepsilon)^{\delta|J|n/(8m)}.$$

Further, since by Lemma 3.3 we have $|\mathcal{S}'| \ge (1 - e^{-c_\delta n})|\mathcal{S}|$ and since $f(s) \le 1$, we observe that

$$A_{nm} \sum_{(S_1,\dots,S_m)\in\mathcal{S}\setminus\mathcal{S}'} \int_{-z}^{z} f(s)\,ds \le 2z\,e^{-c_\delta n}$$

deterministically. Recalling that $|J| = \lceil c_\delta m \rceil$, we obtain

$$\mathbb{P}\Big\{ A_{nm} \sum_{(S_1,\dots,S_m)\in\mathcal{S}} \int_{-z}^{z} f(s)\,ds \le C'' m^{-1/2} \Big\} \ge 1 - 2zm\,2^{|J|}\,(4\varepsilon)^{\delta|J|n/(8m)} \ge 1 - (\widetilde{\varepsilon}/2)^n,$$

for some $C'' \ge 1$ depending only on δ and ρ, provided that $n \ge n_0(\widetilde{\varepsilon}, \delta, \rho)$. The result follows by the substitution $s = m^{-1/2}u$ in the integral. □

As a combination of Lemmas 4.17, 4.18, and 4.20, we obtain Proposition 4.9.

Proof of Proposition 4.9. As we mentioned at the beginning of this subsection, we fix $\rho, \delta \in (0, 1/4]$, a growth function $\mathbf{g}$ satisfying (8), a permutation $\sigma \in \Pi_n$, a number $h \in \mathbb{R}$, two sets $Q_1, Q_2 \subset [n]$ such that $|Q_1|, |Q_2| = \lceil \delta n \rceil$, and we use Λ_n for the set $\Lambda_n(k, \mathbf{g}, Q_1, Q_2, \rho, \sigma, h)$ defined in (17). We also fix $\varepsilon \in (0, 1/4]$.

We start by selecting the parameters. Assume that n is large enough. Set $\ell := \ell_{4.17}(\varepsilon)$. Let $\varepsilon' = \varepsilon'(\varepsilon)$ be taken from Lemma 4.18. Set $z := 1/C_{4.20}(\varepsilon, \delta, \rho)$. Fix an integer $m \in [C_{4.18}(\varepsilon, z), n/\max(\ell, C_{4.20})]$ satisfying the condition $R_{4.17}\sqrt{m}\,e^{-\sqrt{m}} \le 1$, and take $1 \le k \le \min\big(2^{n/\ell}, (K_2/8)^{m/2}\big)$. Let A_{nm} be defined as in (7). We assume that h is chosen in such a way that the set Λ_n is non-empty. As before X denotes the random vector uniformly distributed on Λ_n. Let $\mathcal{S}$ be as in Lemma 3.3). A given choice of subsets $(S_1, \dots, S_m) \in \mathcal{S}$ denote

$$f(s) = f_{S_1,\dots,S_m}(s) := \prod_{i=1}^{m} \psi_{K_2}\Big(\Big| \frac{1}{\lfloor n/m \rfloor} \sum_{w\in S_i} \exp\big(2\pi\mathbf{i} X_w m^{-1/2}\, s\big) \Big| \Big).$$

We have

$$A_{nm} \sum_{S_1,\dots,S_m} \int_{-\varepsilon' m^{1/2} k}^{\varepsilon' m^{1/2} k} f(s)\,ds = A_{nm} \sum_{S_1,\dots,S_m} \int_{-z\sqrt{m}}^{z\sqrt{m}} f(s)\,ds + 2A_{nm} \sum_{S_1,\dots,S_m} \int_{z\sqrt{m}}^{\varepsilon' k\sqrt{m}} f(s)\,ds.$$

In view of Lemma 4.20, with probability at least $1 - (\varepsilon/2)^n$ the first summand is bounded above by $K_{4.20}$. To estimate the second summand, we combine Lemmas 4.17 and 4.18 (we assume that $z \le \varepsilon' k$ as otherwise there is no second summand). Fix for a moment a collection $(S_1, \dots, S_m) \in \mathcal{S}$. By Lemma 4.17, with probability at least $1 - (\varepsilon/2)^n$ the function f on $[0, k\sqrt{m}/2]$ is bounded above by $(K_2/4)^{-m/2}$ for all points s outside of some set of measure at most $R_{4.17}\sqrt{m}$ (note that we apply variable transformation $s \to m^{-1/2}s$ to use the lemma here). Further, by Lemma 4.18, with probability at least $1 - (\varepsilon/2)^n$ we have that f is bounded above by $e^{-\sqrt{m}}$ for all $s \in [z\sqrt{m}, \varepsilon' k\sqrt{m}]$. Thus, with probability at least $1 - 2(\varepsilon/2)^n$,

$$\int_{z\sqrt{m}}^{\varepsilon' k\sqrt{m}} f(s)\,ds \le \sqrt{m}k\,\Big(\frac{K_2}{4}\Big)^{-m/2} + R_{4.17}\sqrt{m}\,e^{-\sqrt{m}}.$$

Applying Lemma 4.12 with $I = \mathcal{S}$ and $\xi_i = f(s)$, we obtain that

$$A_{nm} \sum_{S_1,\dots,S_m} \int_{z\sqrt{m}}^{\varepsilon' k\sqrt{m}} f(s)\, ds \le \sqrt{m}k \left(\frac{K_2}{4}\right)^{-m/2} + R_{4.17}\sqrt{m}\, e^{-\sqrt{m}} + 1 \le 3$$

with probability at least $1 - 2(\varepsilon/2)^n$. Thus, taking $K_1 := K_{4.20} + 3$, we obtain

$$\mathbb{P}\{\mathbf{UD}_n(X, m, K_1, K_2) \ge \varepsilon' m^{1/2} k\} \ge 1 - 3(\varepsilon/2)^n \ge 1 - 3\varepsilon^n.$$

□

5 Complement of gradual non-constant vectors: constant p

In this section, we study the problem of invertibility of the Bernoulli(p) matrix M over the set $\mathcal{S}_n$ defined by (2) in the case when the parameter p is a small constant. This setting turns out to be much simpler than treatment of the general case $C \ln n/n \le p \le c$ given in the next section. Although the results of Section 6 essentially absorb the statements of this section, we prefer to include analysis of the constant p in our work, first, because it provides a short and relatively simple illustration of our method and, second, because the estimates obtained here allow to derive better quantitative bounds for the smallest singular value of M.

5.1 Spliting of $\mathbb{R}^n$ and main statements

We define the following four classes of vectors $\mathcal{B}_1, \dots, \mathcal{B}_4$. For simplicity, we normalize vectors with respect to the Euclidean norm. The first class is the set of vectors with one coordinate much larger than the others, namely,

$$\mathcal{B}_1 = \mathcal{B}_1(p) := \{x \in S^{n-1} \,:\, x_1^* > 6pn\, x_2^*\}.$$

For the next sets we fix a parameter $\beta_p = \sqrt{p}/C_0$, where C_0 is the absolute constant from Proposition 3.10. Recall also that the operator Q (which annihilates the maximal coordinate of a given vector) and the set $U(m, \gamma)$ were introduced in Subsection 3.6. We also fix a small enough absolute positive constant c_0. We don't try to compute the actual value of c_0, the conditions on how small c_0 is can be obtained from the proofs. We further fix an integer $1 \le m \le n$.

The second class of vectors consist of those vectors for which the Euclidean norm dominates the maximal coordinate. To control cardinalities of nets (discretizations) we intersect this class with $U(m, c_0)$, specifically, we set

$$\mathcal{B}_2 = \mathcal{B}_2(p, m) := \mathcal{B}_2' \cap U(m, c_0), \quad \text{where} \quad \mathcal{B}_2' := \left\{x \in S^{n-1} \,:\, x \notin \mathcal{B}_1 \ \text{ and } \ x_1^* \le \beta_p\right\}.$$

The next set is similar to $\mathcal{B}_2$, but instead of comparing x_1^* with the Euclidean norm of the entire vector, we compare x_2^* with $\|Qx\|$. For a technical reason, we need to control the magnitude of $\|Qx\|$ precisely; thus we partition the third set into subsets. Let numbers λ_k, $k \le \ell$, be defined by

$$\lambda_1 = \frac{1}{6pn}, \quad \lambda_{k+1} = 3\lambda_k,\ k < \ell - 1, \quad 1/3 \le \lambda_{\ell-1} < 1 \quad \text{ and } \quad \lambda_\ell = 1. \tag{22}$$

Clearly, $\ell \le \ln n$. Then for each $k \le \ell - 1$ we define

$$\mathcal{B}_{3,k} = \mathcal{B}_{3,k}(p, m) := \left\{x \in S^{n-1} \,:\, x \notin \mathcal{B}_1 \cup \mathcal{B}_2',\ x_2^* \le \beta_p \|Qx\|, \ \text{ and } \ \lambda_k \le \|Qx\| < \lambda_{k+1}\right\} \cap U(m, c_0\lambda_k).$$

To explain the choice of λ_1, note that if $x \notin \mathcal{B}_1 \cup \mathcal{B}_2'$ and $\|x\| = 1$, then $x_2^* \geq x_1^*/(6pn) \geq \beta_p/(6pn)$. Thus, if in addition $\beta_p\|Qx\| \geq x_2^*$, then $\|Qx\| \geq 1/(6pn) = \lambda_1$. We set

$$\mathcal{B}_3 = \mathcal{B}_3(p,m) := \bigcup_{k=1}^{\ell-1} \mathcal{B}_{3,k}.$$

The fourth set covers the remaining options for vectors having a large almost constant part. Let numbers μ_k, $k \leq s$, be defined by

$$\mu_1 = \frac{\beta_p}{6pn}, \quad \mu_{k+1} = 3\mu_k, \ k < s-1, \quad 1/3 \leq \mu_{s-1} < 1 \quad \text{and} \quad \mu_s = 1. \tag{23}$$

Clearly, $s \leq \ln n$. Then for each $k \leq s-1$ define the set $\mathcal{B}_{4,k} = \mathcal{B}_{4,k}(p,m)$ as

$$\big\{x \in S^{n-1} \,:\, x \notin \mathcal{B}_1 \cup \mathcal{B}_2', \ x_2^* > \beta_p\|Qx\|, \ \text{ and } \ \mu_k \leq x_2^* < \mu_{k+1}\big\} \cap U(m, c_0\mu_k/\sqrt{\ln(e/p)}).$$

Note that if $x \notin \mathcal{B}_1 \cup \mathcal{B}_2'$ and $\|x\| = 1$, then $x_2^* \geq x_1^*/(6pn) \geq \beta_p/(6pn)$, justifying the choice of μ_1. We set

$$\mathcal{B}_4 = \mathcal{B}_4(p,m) = \bigcup_{k=1}^{\ell-1} \mathcal{B}_{4,k}.$$

Finally define $\mathcal{B}$ as the union of these four classes, $\mathcal{B} = \mathcal{B}(p,m) := \bigcup_{j=1}^4 \mathcal{B}_j$.

In this section we prove two following theorems.

Theorem 5.1. *There exists positive absolute constants c, C such that the following holds. Let n be large enough, let $m \leq cpn/\ln(e/p)$, and $(30\ln n)/n \leq p \leq 1/2$. Let M be an $n \times n$ Bernoulli(p) random matrix. Then*

$$\mathbb{P}\Big\{\exists\, x \in \mathcal{B} \quad \text{such that} \quad \|Mx\| < \frac{1}{C\sqrt{n\ln(e/p)}}\,\|x\|\Big\} \leq n(1-p)^n + e^{-1.1np},$$

where the set $\mathcal{B} = \mathcal{B}(p,m)$ is defined above.

Recall that the set $\mathcal{V}_n$ was introduced in Subsection 3.3. The next theorem shows that, after a proper normalization, the complement of $\mathcal{V}_n$ (taken in $\Upsilon_n(r)$) is contained in $\mathcal{B}$ for some choice of r, δ, ρ and for the growth function $\mathbf{g}(t) = (2t)^{3/2}$ (clearly, satisfying (8)).

Theorem 5.2. *There exists an absolute (small) positive constant c_1 such that the following holds. Let $q \in (0, c_1)$ be a parameter. Then there exist $n_q \geq 1$, $r = r(q), \rho = \rho(q) \in (0,1)$ such that for $n \geq n_q$, $p \in (q, c_1)$, $\delta = r/3$, $\mathbf{g}(t) = (2t)^{3/2}$, and $m = \lfloor rn \rfloor$ one has*

$$\Big\{x/\|x\| \ :\ x \in \Upsilon_n(r) \setminus \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)\Big\} \subset \mathcal{B}(p,m).$$

5.2 Proof of Theorem 5.1

Theorem 5.1 is a consequence of four lemmas that we prove in this section. Each lemma treats one of the classes $\mathcal{B}_i$, $i \leq 4$, and Theorem 5.1 follows by the union bound. Recall that $U(m,\gamma)$ was introduced in Subsection 3.6 and that given x, we fixed one permutation, σ_x, such that $x_i^* = |x_{\sigma_x(i)}|$ for $i \leq n$. Recall also that the event $\mathcal{E}_{nrm}$ was introduced in Proposition 3.14.

Lemma 5.3. *Let $n \geq 1$ and $p \in (0, 1/2]$. Let $\mathcal{E}_{sum}$ (with $q = p$) be the event introduced in Lemma 3.4 and by $\mathcal{E}_{col} \subset \mathcal{M}_n$ denote the subset of 0/1 matrices with no zero columns. Then for every $M \in \mathcal{E}_{sum} \cap \mathcal{E}_{col}$ and every $x \in \mathcal{B}_1$,*

$$\|Mx\| \geq \frac{1}{3\sqrt{n}} \|x\|.$$

In particular,

$$\mathbb{P}\Big\{ M \in \mathcal{M}_n : \exists x \in \mathcal{B}_1 \quad with \quad \|Mx\| \leq \frac{1}{3\sqrt{n}} \Big\} \leq n(1-p)^n + e^{-1.5np}.$$

Proof. Let δ_{ij}, $i, j \leq n$ be entries of $M \in \mathcal{E}_{sum} \cap \mathcal{E}_{col}$. Let $\sigma = \sigma_x$. Denote, $\ell = \sigma(1)$. Since $M \in \mathcal{E}_{col}$, there exists $s \leq n$ such that $\delta_{s\ell} = 1$. Then

$$|\langle R_s(M),\, x\rangle| = \Big| x_\ell + \sum_{j \neq \ell} \delta_{sj} x_j \Big| \geq |x_\ell| - \sum_{j \neq \ell} \delta_{sj}\, x_j \geq |x_\ell| - \sum_{j=1}^{n} \delta_{sj}\, x^*_{n_2}.$$

Using that $M \in \mathcal{E}_{sum}$ we observe that $\sum_{j=1}^n \delta_{sj} \leq 3.5pn$. Thus,

$$\|Mx\| \geq |\langle R_s(M),\, x\rangle| \geq x^*_1 - 3.5pn x^*_{n_2} \geq x^*_1/3.$$

The trivial bound $\|x\| \leq \sqrt{n}\, x^*_1$ completes the first estimate. The "in particular" part follows by the "moreover" part of Lemma 3.4 and since $\mathbb{P}(\mathcal{E}_{col}) \leq n(1-p)^n$. □

Lemma 5.4. *There exists a (small) absolute positive constant c such that the following holds. Let n be large enough and $m \leq cn$. Let $(4 \ln n)/n \leq p < 1/2$ and M be a Bernoulli(p) random matrix. Then*

$$\mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_2 \quad with \quad \|Mx\| \leq \frac{\sqrt{pn}}{5C_0} \Big) \leq e^{-2n}.$$

Proof. By Lemma 3.13 for $\varepsilon \in [8c_0, 1)$ there exists an $(\varepsilon/2)$–net in $V(1) \cap U(m, c_0)$ with respect to the triple norm $|||\cdot|||$, with cardinality at most

$$\frac{Cn^2}{\varepsilon^2} \left(\frac{18en}{\varepsilon m} \right)^m.$$

Since $\mathcal{B}_2 \subset V(1) \cap U(m, c_0)$, by a standard "projection" trick, we can obtain from it an ε–net $\mathcal{N}$ in $\mathcal{B}_2$ of the same cardinality. Let $x \in \mathcal{B}_2$. Let $z \in \mathcal{N}$ be such that $|||x - z||| \leq \varepsilon$. Since on $\mathcal{B}_2$ we have $z^*_1 \leq \beta_p \|z\| = \beta_p$, Proposition 3.10 implies that with probability at least $1 - e^{-3n}$,

$$\|Mz\| \geq \frac{\sqrt{pn}}{3\sqrt{2}C_0}. \tag{24}$$

Further, in view of Proposition 3.14, conditioned on (24) and on $\{M \in \mathcal{E}_{nrm}\}$, we have

$$\|Mx\| \geq \|Mz\| - \|M(x-z)\| \geq \frac{\sqrt{pn}}{3\sqrt{2}C_0} - 100\sqrt{pn}\varepsilon \geq \frac{\sqrt{pn}}{5C_0},$$

where we have chosen $\varepsilon = 1/(5000C_0)$. Using the union bound and our choice of ε, we obtain that

$$\mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_2 \ \text{ with } \ \|Mx\| \leq \frac{\sqrt{pn}}{5C_0} \Big) \leq e^{-3n} |\mathcal{N}| \leq e^{-2n}$$

for sufficiently large n and provided that $c_0 \leq 1/(40000C_0)$ and $m \leq cn$ for small enough absolute positive constant c. This completes the proof. □

Remark 5.5. *Note that we used Proposition 3.10 with the set $A = [n]$. In this case we could use slightly easier construction for nets than the one in Lemma 3.13 — we don't need to distinguish the first coordinate in the net construction, in other words we could have only one special direction, not two. However this would not lead to a better estimate and in the remaining lemmas we will need the full strength of our construction.*

Next we treat the case of vectors in $\mathcal{B}_3$. The proof is similar to the proof of Lemma 5.4, but we need to remove the maximal coordinate and to deal with the remaining part of the vector. Recall that the operator Q serves this purpose.

Lemma 5.6. *There exists a (small) absolute positive constant c such that the following holds. Let n be large enough, and $m \le cpn/\ln(e/p)$, $(4\ln n)/n \le p < 1/2$. Let M be a Bernoulli(p) random matrix. Then*

$$\mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_3 \text{ with } \|Mx\| \le \frac{1}{30C_0\sqrt{pn}}\Big) \le e^{-2n}.$$

Proof. Fix $1 \le k \le \ell - 1$. By Lemma 3.13 for $\varepsilon \in [8c_0\lambda_k, \lambda_{k+1})$ there exists an $(\varepsilon/2)$–net in $V(\lambda_{k+1}) \cap U(m, c_0\lambda_k)$ with respect to $|||\cdot|||$, with cardinality at most

$$\frac{Cn^2}{\varepsilon^2}\left(\frac{18e\lambda_{k+1}n}{\varepsilon m}\right)^m \le \frac{Cn^2}{\varepsilon^2}\left(\frac{54e\lambda_k n}{\varepsilon m}\right)^m.$$

Again using a "projection" trick, we can construct an ε–net $\mathcal{N}_k$ in $\mathcal{B}_{3,k}$ of the same cardinality. Let $x \in \mathcal{B}_{3,k}$. Let $z \in \mathcal{N}_k$ be such that $|||x - z||| \le \varepsilon$. Since on $\mathcal{B}_{3,k}$ we have $z_2^* \le \beta_p\|Qz\|$, Proposition 3.10 applied with $A = \sigma_z([2, n])$ implies that with probability at least $1 - e^{-3n}$,

$$\|Mz\| \ge \frac{\sqrt{pn}\,\|Qz\|}{3\sqrt{2}C_0} \ge \frac{\sqrt{pn}\,\lambda_k}{3\sqrt{2}C_0}.$$

Conditioned on the above inequality and on the event $\{M \in \mathcal{E}_{nrm}\}$, Proposition 3.14 implies that

$$\|Mx\| \ge \|Mz\| - \|M(x - z)\| \ge \frac{\sqrt{pn}\,\lambda_k}{3\sqrt{2}C_0} - 100\sqrt{pn}\varepsilon \ge \frac{\sqrt{pn}\,\lambda_k}{5C_0},$$

where we have chosen $\varepsilon = \lambda_k/(5000C_0)$. Using the union bound, our choice of ε and $\lambda_k \ge 1/(6pn)$, we obtain that

$$P_k := \mathbb{P}\Big(\exists x \in \mathcal{B}_{3,k} \text{ with } \|Mx\| \le \frac{\sqrt{pn}\,\lambda_k}{5C_0}\Big) \le e^{-3n}|\mathcal{N}_k| \le e^{-2.5n}$$

for large enough n and for $m \le cn$, where $c > 0$ is a small enough absolute constant (we also assume $c_0 \le 1/(40000C_0)$). Since $\ell \le \ln n$ and $\lambda_k \ge \lambda_1 \ge 1/(6pn)$, we obtain

$$\mathbb{P}\Big(\exists x \in \mathcal{B}_3 \text{ with } \|Mx\| \le \frac{1}{30C_0\sqrt{pn}}\Big) \le \sum_{k=1}^{\ell-1} P_k \le e^{-2pn}.$$

This completes the proof. □

Finally we threat the case of vectors in $\mathcal{B}_4$.

Lemma 5.7. *There exists a (small) absolute positive constant c such that the following holds. Let n be large enough and let $m \le cpn/\ln(e/p)$, $(30\ln n)/n \le p \le 1/2$. Let M be a Bernoulli(p) random matrix. Then*

$$\mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_4 \text{ with } \|Mx\| \le \frac{1}{78C_0\sqrt{n\ln(e/p)}}\Big) \le e^{-1.14pn}.$$

Proof. Fix $1 \le k \le s-1$. By Lemma 3.13 for $\varepsilon \in [8c_0\mu_k/\sqrt{\ln(e/p)}, \mu_{k+1})$ there exists an $(\varepsilon/2)$–net in

$$V(\mu_{k+1}/\beta_p) \cap U(m, c_0\mu_k/\sqrt{\ln(e/p)})$$

with respect to $|||\cdot|||$ with cardinality at most

$$\frac{Cn^2}{\varepsilon^2}\left(\frac{18e\mu_{k+1}n}{\varepsilon m\beta_p}\right)^m \le \frac{Cn^2}{\varepsilon^2}\left(\frac{54e\mu_k n}{\varepsilon m\beta_p}\right)^m.$$

By the projection trick, we get an ε–net $\mathcal{N}_k$ in $\mathcal{B}_{4,k} \subset V(\mu_{k+1}/\beta_p) \cap U(m, c_0\mu_k/\sqrt{\ln(e/p)})$.

Let $x \in \mathcal{B}_{4,k}$. Let $z \in \mathcal{N}_k$ be such that $|||x-z||| \le \varepsilon$. Since on $\mathcal{B}_4$ we have $z_1^* \ge z_2^* \ge \mu_k$, Proposition 3.11 implies that with probability at least $1 - e^{-1.2np}$,

$$\|Mz\| \ge \frac{\mu_k\sqrt{pn}}{10\sqrt{\ln(e/p)}}.$$

Conditioned on the above and on $\{M \in \mathcal{E}_{nrm}\}$, Proposition 3.14 implies that

$$\|Mx\| \ge \|Mz\| - \|M(x-z)\| \ge \frac{\mu_k\sqrt{pn}}{10\sqrt{\ln(e/p)}} - C_1\sqrt{pn}\varepsilon \ge \frac{\mu_k\sqrt{pn}}{13\sqrt{\ln(e/p)}},$$

where we have chosen

$$\varepsilon = \mu_k/(50C_1\sqrt{\ln(e/p)}) \ge 8c_0\mu_k/\sqrt{\ln(e/p)},$$

provided that $c_0 \le 1/40000$. Using the union bound and our choice of ε we obtain that

$$P_k := \mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_{4,k} \ \text{ with } \ \|Mx\| \le \frac{\mu_k\sqrt{pn}}{13\sqrt{\ln(e/p)}}\Big) \le e^{-1.2pn}|\mathcal{N}_k| \le e^{-1.15pn}$$

for large enough n and for $m \le cpn/\ln(e/p)$, where $c > 0$ is a small enough absolute constant. Since $s \le \ln n$ and $\mu_k \ge \mu_1 \ge \beta_p/(6pn) = 1/(6C_0n\sqrt{p})$, we obtain

$$\mathbb{P}\Big(M \in \mathcal{E}_{nrm} \quad \text{and} \quad \exists x \in \mathcal{B}_4 \ \text{ with } \ \|Mx\| \le \frac{1}{78C_0\sqrt{n\ln(e/p)}}\Big) \le \sum_{k=1}^{s-1} P_k \le e^{-1.14pn}.$$

This completes the proof. □

Proof of Theorem 5.1. Lemmas 5.3, 5.4, 5.6, and 5.7 imply that

$$\mathbb{P}(\mathcal{E}) \le n(1-p)^n + 3e^{-1.14np} + \mathbb{P}(\mathcal{E}_{nrm}^c),$$

where $\mathcal{E}$ denotes the event from Theorem 5.1. Lemma 3.6 applied with $t = 30$ and (11) imply that $\mathbb{P}(\mathcal{E}_{nrm}^c) \le e^{-10pn}$, provided that pn is large enough. This completes the proof. □

5.3 Proof of Theorem 5.2

Proof. We prove the statement with $r = r(q) = cq/\ln(e/q)$, where c is the constant from Theorem 5.1, and $\rho = \rho(q) = c_0\sqrt{r}\beta_q/(6\sqrt{\ln(e/q)})$. Note that under our choice of parameters (and assuming c_1 is small), $9\delta/2 \le c_0\beta_q/\sqrt{\ln(e/q)} \le c_0\beta_p/\sqrt{\ln(e/p)}$.

Assume that $x \in \Upsilon_n(r) \setminus \mathcal{V}_n$. By $(x_i^\#)_i$ denote the non-increasing rearrangement of $(x_i)_i$ (we would like to emphasize that we do not take absolute values). Note that for any $t > 0$ there are two subsets

$Q_1, Q_2 \subset [n]$ with $|Q_1|, |Q_2| \geq \lceil \delta n \rceil$ satisfying $\max_{i \in Q_2} x_i \leq \min_{i \in Q_1} x_i - t$ if and only if $x^{\#}_{\lceil \delta n \rceil} - x^{\#}_{n - \lceil \delta n \rceil + 1} \geq t$. This leads to the two following cases.

Case 1. $x^{\#}_{\lceil \delta n \rceil} - x^{\#}_{n-\lceil \delta n \rceil + 1} \geq \rho$. Since $x \notin \mathcal{V}_n$, in this case there exists an index $j \leq n$ with $x^*_j > (2n/j)^{3/2}$. Note that since $x^*_{\lfloor rn \rfloor} = 1$, we have $j < rn = 3\delta n$.

Subcase 1a. $1 < j < 3\delta n$. Since $x^*_j > (2n/j)^{3/2}$ we get

$$\|Qx\|^2 \geq \sum_{i=2}^{j} (x^*_i)^2 \geq \sum_{i=2}^{j} (2n/i)^3 \geq \frac{j}{2} (2n/j)^3 = n(2n/j)^2.$$

Therefore,

$$\frac{x^*_{\lfloor rn \rfloor + 1}}{\|Qx\|} \leq \frac{1}{\sqrt{n}} \frac{j}{2n} \leq \frac{(3\delta/2)}{\sqrt{n}}.$$

Now let $y = x/\|x\|$. Then

$$y^*_{\lfloor rn \rfloor + 1} = \frac{x^*_{\lfloor rn \rfloor + 1}}{\|x\|} \leq \frac{3\delta/2}{\sqrt{n}} \frac{\|Qx\|}{\|x\|} = \frac{3\delta/2}{\sqrt{n}} \|Qy\|. \tag{25}$$

Our goal is to show that $y \in \mathcal{B}(p, m)$ (with $m = \lfloor rn \rfloor$).

If $y \in \mathcal{B}_1(p)$, we are done.

Otherwise, if $y \in \mathcal{B}'_2$, then (25) implies that $y^*_{\lfloor rn \rfloor + 1} \leq c_0/\sqrt{n}$, that is, there are at least $n - m$ coordinates at the distance at most $c_0/\sqrt{n}$ from zero. Thus $y \in U(m, c_0)$ and hence $y \in \mathcal{B}_2$.

If $y \notin \mathcal{B}_1 \cup \mathcal{B}'_2$ and $y^*_2 \leq \beta_p \|Qy\|$, then necessarily $\lambda_k \leq \|Qy\| < \lambda_{k+1} \leq 3\lambda_k$ for some k, where λ_k, λ_{k+1} are defined according to (22). Then (25) implies that $y^*_{\lfloor rn \rfloor + 1} \leq c_0 \lambda_k/\sqrt{n}$, that is, there are at least $n - m$ coordinates at the distance at most $c_0 \lambda_k/\sqrt{n}$ from zero. Thus $y \in U(m, c_0\lambda_k)$ and hence $y \in \mathcal{B}_{3,k}$.

If $y \notin \mathcal{B}_1 \cup \mathcal{B}'_2$ and $y^*_2 > \beta_p \|Qy\|$ then necessarily $\mu_k \leq y^*_2 < \mu_{k+1} \leq 3\mu_k$, where μ_k, μ_{k+1} are given by (23). Then, similarly,

$$y^*_{\lfloor rn \rfloor + 1} \leq \frac{3\delta/2}{\sqrt{n}} \|Qy\| \leq \frac{3\delta/2}{\sqrt{n}} \frac{y^*_2}{\beta_p} \leq \frac{9\delta/2}{\beta_p \sqrt{n}} \mu_k \leq \frac{c_0 \mu_k}{\sqrt{\ln(e/p)} \sqrt{n}}.$$

This implies that $y \in U(m, c_0\mu_k/\sqrt{\ln(e/p)})$ and, thus, $y \in \mathcal{B}_{4,k}$.

Subcase 1b. $j = 1$. In this case $x^*_1 \geq (2n)^{3/2}$. Assume $x \notin \mathcal{B}_1$, that is $x^*_1 < 6pnx^*_2$. Then

$$\frac{x^*_{\lfloor rn \rfloor + 1}}{\|Qx\|} \leq \frac{1}{x^*_2} \leq \frac{6pn}{(2n)^{3/2}} = \frac{6p}{2^{3/2}\sqrt{n}}.$$

We can now define $y := x/\|x\|$ and, having noted that $y^*_{\lfloor rn \rfloor + 1} \leq \frac{6p}{2^{3/2}\sqrt{n}} \|Qy\|$, proceed similarly to the Subcase 1a. We will need to use the condition $18p \leq 2^{3/2} c_0 \beta_p / \sqrt{\ln(e/p)}$, which holds for small enough p.

Case 2. $x^{\#}_{\lceil \delta n \rceil} - x^{\#}_{n - \lceil \delta n \rceil + 1} < \rho$. Set σ be a permutation of $[n]$ such that $x^{\#}_i = x_{\sigma(i)}$, $i \leq n$ (note that σ is in general different from the permutation σ_x defined in connection with the non-increasing rearrangement of the absolute values $|x_i|$). Define the following set, which will play the role of the set in the definition of $U(m, \gamma)$ (see Subsection 3.6),

$$A := \{\sigma(i) : \lceil \delta n \rceil < i \leq n - \lceil \delta n \rceil\}.$$

Then $|A| = n - 2\lceil \delta n \rceil$, and $m > 2\lceil \delta n \rceil = 2\lceil rn/3 \rceil$. Since $x^*_m = 1$, we observe that either $x^{\#}_{\lceil \delta n \rceil + 1} \geq 1$ or $x^{\#}_{n - \lceil \delta n \rceil} \leq -1$ (or both). Moreover, since $r < 1/2$, we necessarily have that $|x^{\#}_i| \leq 1$ for some $\lceil \delta n \rceil < i \leq$

$n - \lceil \delta n \rceil$. Therefore, there exists an index $j \in A$ such that $|x_j| = 1$. Taking $b = x_j$, we observe that for every $i \in A$, $|x_i - b| < \rho$. On the other hand we have

$$\|x\|^2 \geq \|Qx\|^2 \geq \sum_{i=2}^{m} x_i^* \geq m - 1 \geq m/2 \quad \text{and} \quad \forall i \in A : \frac{|x_i - b|}{\|Qx\|} \leq \frac{\sqrt{2}\,\rho}{\sqrt{m}} \leq \frac{1}{\sqrt{n}} \frac{2\rho}{\sqrt{r}}.$$

Now let $y = x/\|x\|$. Then

$$\forall i \in A : \left| y_i - \frac{b}{\|x\|} \right| = \frac{|x_i - b|}{\|Qx\|} \frac{\|Qx\|}{\|x\|} \leq \frac{1}{\sqrt{n}} \frac{2\rho}{\sqrt{r}} \|Qy\|. \tag{26}$$

The end of the proof is similar to the end of the proof of Case 1. If $y \in \mathcal{B}_1$, we are done. If $y \in \mathcal{B}_2'$, then using (26), $\|Qy\| \leq \|y\| = 1$, and $6\rho/\sqrt{r} \leq c_0$ we obtain that $y \in U(m, c_0)$ and, thus, $y \in \mathcal{B}_2$. If $y \notin \mathcal{B}_1 \cup \mathcal{B}_2'$, $y_2^* \leq \beta_p \|Qy\|$, and $\lambda_k \leq \|Qy\| < \lambda_{k+1} \leq 3\lambda_k$ then, using (26) and $6\rho/\sqrt{r} \leq c_0$ we obtain that $y \in U(m, c_0\lambda_k)$ and, thus, $y \in \mathcal{B}_{3,k}$. If $y \notin \mathcal{B}_1 \cup \mathcal{B}_2'$, $y_2^* \geq \beta_p \|Qy\|$, and $\mu_k \leq y_2^* < \mu_{k+1} \leq 3\mu_k$ then, similarly, using (26) and $6\rho/\sqrt{r} \leq c_0\beta_p/\sqrt{\ln(e/p)}$, we obtain that $y \in U(m, c_0\mu_k/\sqrt{\ln(e/p)})$ and, thus, $y \in \mathcal{B}_{4,k}$. This completes the proof. □

6 Complement of gradual non-constant vectors: general case

We split $\mathbb{R}^n$ into two classes of vectors. The first class, the class of *steep* vectors $\mathcal{T}$, is constructed in essentially the same way as in [27] and [30]. The proof of the bound for this class resembles corresponding proofs in [27] and [30], however, due to the differences of the models of randomness, there are important modifications. The second class $\mathcal{R}$, which we call $\mathcal{R}$-vectors, will consist of vectors to which Proposition 3.10 can be applied, therefore dealing with this class is simpler. To control the cardinality of nets, part of this class will be intersected with the almost constant vectors. Then we show that the complement of $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ in $\Upsilon_n(r)$ is contained in $\mathcal{T} \cup \mathcal{R}$.

We now introduce the following parameters, which will be used throughout this section. It will be convenient to denote $d = pn$. We always assume that $p \leq 0.0001$ and n is large enough (that is, larger than a certain positive absolute constant). We also always assume that the "average degree" $d = pn \geq 200 \ln n$. Fix a sufficiently small absolute positive constant r and sufficiently large absolute positive constant C_τ (we do not try to estimate the actual values of r and C_τ, the conditions on how large $1/r$ and C_τ can be extracted from the proofs, in particular, the condition on C_τ comes from (38)). We also fix two positive integers ℓ_0 and s_0 such that

$$\ell_0 = \left\lfloor \frac{pn}{4\ln(1/p)} \right\rfloor \quad \text{and} \quad \ell_0^{s_0 - 1} \leq \frac{1}{64p} = \frac{n}{64d} < \ell_0^{s_0}. \tag{27}$$

Note that $\ell_0 \geq 50$ and that $s_0 > 1$ implies $p \leq c\sqrt{(\ln n)/n}$.

For $1 \leq j \leq s_0$ we set

$$n_0 := 2, \quad n_j := 30\ell_0^{j-1}, \qquad n_{s_0+2} := \left\lfloor \sqrt{n/p} \right\rfloor = \left\lfloor \frac{n}{\sqrt{d}} \right\rfloor, \quad \text{and} \quad n_{s_0+3} := \lfloor rn \rfloor.$$

Then, in the case $\lfloor 1/(64p) \rfloor \geq 15 n_{s_0}$ we set $n_{s_0+1} = \lfloor 1/(64p) \rfloor$. Otherwise, let $n_{s_0+1} = n_{s_0}$. Note that with this definition we always have $n_{s_0+2} > n_{s_0+1}$. **The indices n_j, $j \leq s_0 + 3$, are global parameters which will be used throughout the section.** Below we provide the proof only for the case $\lfloor 1/(64p) \rfloor = n_{s_0+1} \geq 15 n_{s_0}$, the other case is treated similarly (in particular, in that other case the set $\mathcal{T}_{1(s_0+1)}$ defined below, will be empty).

We also will use another parameter,

$$\kappa = \kappa(p) := \frac{\ln(6pn)}{\ln \ell_0}. \tag{28}$$

Note that the function $f(p) = \ln(6pn)/(4\ln(1/p))$ is a decreasing function on $(0,1)$, therefore for $p \geq (100\ln n)/n$ and sufficiently large n we have $1 < \kappa \leq \ln\ln n$. Moreover, it is easy to see that if $p \geq (100\ln^2 n)/n$, then $\kappa \leq 2$. We also notice that if $pn \geq 6(5\ln n)^{1+\gamma}$ for some $\gamma \in (0,1)$ then $\kappa \leq 1 + 1/\gamma$ and, using the definition of ℓ_0 and s_0,

$$(6d)^{s_0-1} = \ell_0^{(s_0-1)\kappa} \leq 1/(64p)^{\kappa}. \tag{29}$$

6.1 Two classes of vectors and main results

We first introduce the class of steep vectors. It will be constructed as a union of four subclasses. Recall that the notation x^* was introduced in Subsection 3.1. Set

$$\mathcal{T}_0 := \{x \in \mathbb{R}^n \,:\, x_1^* > 6d\, x_2^*\} \quad \text{and} \quad \mathcal{T}_{11} := \{x \in \mathbb{R}^n \,:\, x \notin \mathcal{T}_0 \ \text{ and } \ x_2^* > 6d\, x_{n_1}^*\}.$$

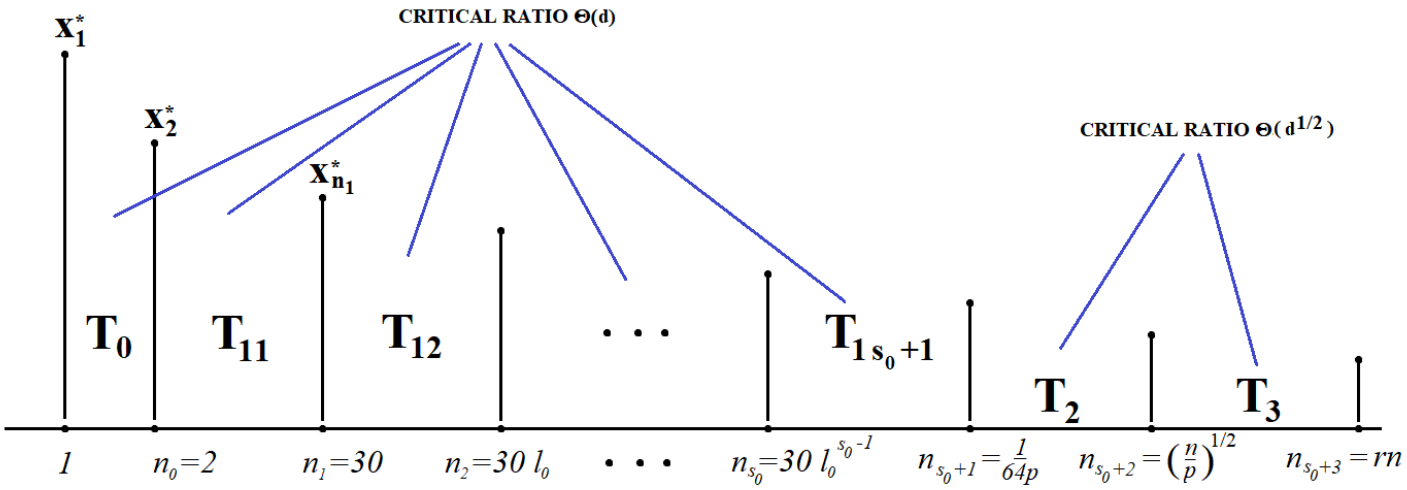
Then for $2 \leq j \leq s_0 + 1$,

$$\mathcal{T}_{1j} := \left\{ x \in \mathbb{R}^n \,:\, x \notin \mathcal{T}_0 \cup \bigcup_{i=1}^{j-1} \mathcal{T}_{1i} \ \text{ and } \ x_{n_{j-1}}^* > 6d\, x_{n_j}^* \right\} \quad \text{and} \quad \mathcal{T}_1 := \bigcup_{i=1}^{s_0+1} \mathcal{T}_{1i}.$$

Finally, for $k = 2, 3$ set $j = j(k) = s_0 + k$ and define

$$\mathcal{T}_k := \left\{ x \in \mathbb{R}^n \,:\, x \notin \bigcup_{i=0}^{k-1} \mathcal{T}_i \ \text{ and } \ x_{n_{j-1}}^* > C_\tau \sqrt{d}\, x_{n_j}^* \right\}.$$

The set of steep vectors is $\mathcal{T} := \mathcal{T}_0 \cup \mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3$. The "rules" of the partition are summarized in the diagram.



For this class we prove the following bound.

Theorem 6.1. *There exist positive absolute constants c and C such that the following holds. Let $n \geq C$, and let $0 < p < c$ satisfy $pn \geq C\ln n$. Let M be a Bernoulli(p) random matrix and denote*

$$\mathcal{E}_{steep} := \left\{ \exists\, x \in \mathcal{T} \quad \textit{such that} \quad \|Mx\| < \frac{c(64p)^{\kappa}}{(pn)^2} \min\left(1, \frac{1}{p^{1.5}n}\right) \|x\| \right\},$$

where as before $\kappa = \kappa(p) := (\ln(6pn))/\ln \ell_0$. *Then*

$$\mathbb{P}(\mathcal{E}_{steep}) \le n(1-p)^n + 2e^{-1.4pn}.$$

Next we introduce the class of $\mathcal{R}$-vectors, denoted by $\mathcal{R}$. Let C_0 be the constant from Proposition 3.10 and recall that the class $\mathcal{AC}(\rho)$ of almost constant vectors was defined by (9) in Subsection 2.2. Given $n_{s_0+1} < k \le n/\ln^2 d$ denote $A = A(k) := [k, n]$ and consider the sets

$$\mathcal{R}_k^1 := \left\{ x \in \big(\Upsilon_n(r) \setminus \mathcal{T}\big) \cap \mathcal{AC}(\rho) \ : \ \frac{\|x_{\sigma_x(A)}\|}{\|x_{\sigma_x(A)}\|_\infty} \ge \frac{C_0}{\sqrt{p}} \quad \text{and} \quad \sqrt{n/2} \le \|x_{\sigma_x(A)}\| \le C_\tau \sqrt{dn} \right\},$$

and

$$\mathcal{R}_k^2 := \left\{ x \in \Upsilon_n(r) \setminus \mathcal{T} \ : \ \frac{\|x_{\sigma_x(A)}\|}{\|x_{\sigma_x(A)}\|_\infty} \ge \frac{C_0}{\sqrt{p}} \quad \text{and} \quad \frac{2\sqrt{n}}{r} \le \|x_{\sigma_x(A)}\| \le C_\tau^2 d\sqrt{n} \right\}.$$

Define $\mathcal{R} := \bigcup_{n_{s_0+1} < k \le n/\ln^2 d} (\mathcal{R}_k^1 \cup \mathcal{R}_k^2)$.

The class $\mathcal{R}$ should be thought of as the class of *sufficiently spread* vectors, not steep, but possibly without having two subsets of coordinates of size proportional to n, which are separated by ρ (which would allow us to treat those vectors as part of the set $\mathcal{V}_n$). Crucially, the sets $\mathcal{R}_k^1$ and $\mathcal{R}_k^2$ are "low complexity" sets because they admit ε–nets of relatively small cardinalities (see Subsection 6.3). For the class $\mathcal{R}$ we prove the following bound.

Theorem 6.2. *There are absolute constants* r_0, ρ_0, C *with the following property. Let* $0 < r \le r_0$, $0 < \rho \le \rho_0$, *let* $n \ge 1$ *and* $p \in (0, 0.001]$ *be such that* $d = pn \ge C \ln n$. *Then*

$$\mathbb{P}\left(\left\{ \exists x \in \mathcal{R} \ : \ \|Mx\| \le \frac{\sqrt{p}n}{12C_0} \right\} \right) \le e^{-2n} + e^{-200pn}.$$

Finally we show that together with $\mathcal{V}_n$, the classes $\mathcal{T}$ and $\mathcal{R}$ cover all (properly normalized) vectors for the growth function defined by

$$\mathbf{g}(t) = (2t)^{3/2} \quad \text{for } 1 \le t < 64pn \qquad \text{and} \qquad \mathbf{g}(t) = \exp(\ln^2(2t)) \quad \text{for } t \ge 64pn. \tag{30}$$

It is straightforward to check that $\mathbf{g}$ satisfies (8) with some absolute constant K_3.

Theorem 6.3. *There are universal constants* $c, C > 0$ *with the following property. Let* $n \ge C$, $p \in (0, c)$, *and assume that* $d = pn \ge 100 \ln n$. *Let* $r \in (0, 1/2)$, $\delta \in (0, r/3)$, $\rho \in (0, 1)$, *and let* $\mathbf{g}$ *be as in (30). Then*

$$\Upsilon_n(r) \setminus \mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \subset \mathcal{R} \cup \mathcal{T}.$$

6.2 Auxiliary lemmas

In the following lemma we provide a simple bound on the Euclidean norms of vectors in the class $\mathcal{T}$ and its complement in terms of their order statistics.

Lemma 6.4. *Let* n *be large enough and* $(200 \ln n)/n < p < 0.001$. *Consider the vectors* $x \in \mathcal{T}_{1j}$ *for some* $1 \le j \le s_0 + 1$, $y \in \mathcal{T}_2$, $z \in \mathcal{T}_3$ *and* $w \in \mathcal{T}^c$. *Then*

$$\frac{\|x\|}{x^*_{n_{j-1}}} \le \frac{64(pn)^2}{(64p)^\kappa}, \quad \frac{\|y\|}{y^*_{n_{s_0+1}}} \le \frac{384(pn)^3}{(64p)^\kappa}, \quad \frac{\|z\|}{z^*_{n_{s_0+2}}} \le \frac{384C_\tau(pn)^{3.5}}{(64p)^\kappa}, \quad \text{and} \quad \frac{\|w\|}{w^*_{n_{s_0+3}}} \le \frac{384C_\tau^2(pn)^4}{(64p)^\kappa}.$$

Proof. Let $d = pn$. Since $x \in \mathcal{T}_{1j}$, denoting $m = n_{j-1}$, we have

$$x_1^* \le (6d)x_2^* \le (6d)^2 x_{n_1}^* \le \ldots \le (6d)^j x_{n_{j-1}}^* = (6d)^j x_m^*.$$

Since $n_i = 30\ell_0^{i-1} \le 30d^{i-1}$, $i \le s_0$, since $\kappa > 1$, and in view of (29), we obtain

$$\begin{aligned}
\|x\|^2 &= (x_1^*)^2 + (x_2^* + \ldots + (x_{n_1}^*)^2) + ((x_{n_1+1}^*)^2 + \cdots + (x_{n_2}^*)^2) + \ldots \\
&\le ((6d)^{2j} + n_1(6d)^{2(j-1)} + n_2(6d)^{2(j-2)} \ldots + n_{j-1}(6d)^2 + n)(x_m^*)^2 \\
&\le \Big((6d)^{2j} + 5(6d)^{2j-2}\sum_{i\ge 0}(6d)^{-i} + n\Big)(x_m^*)^2 \le \left(2(6d)^{2(s_0+1)} + n\right)(x_m^*)^2 \\
&\le \left(2(6d)^4/(64p)^{2\kappa} + n\right)(x_m^*)^2 \le \left(3(6d)^4/(64p)^{2\kappa}\right)(x_m^*)^2.
\end{aligned}$$

This implies the first bound. The bounds for y, z, w are obtained similarly. □

The next two Lemmas 6.5 and 6.6 will be used to bound from below the norm of the matrix-vector product Mx for vectors x with a "too large" almost constant part which does not allow to directly apply the Lévy–Kolmogorov–Rogozin anti-concentration inequality together with the tensorization argument. Lemma 6.5 will be used to bound $\|Mx\|$ by a single inner product $|\langle \mathbf{R}_i(M), x\rangle|$ for a specially chosen index i, while Lemma 6.6 will allow to extract a subset of "good" rows having large inner products with x.

Lemma 6.5. *Let $n \ge 30$ and $0 < p < 0.001$ satisfy $pn \ge 200 \ln n$. Let $m, \ell = \ell(m) \ge 2$ be such that either*

$$m = 2 \text{ and } \ell = 15,$$

or

$$m \ge 30, \quad \ell m \le \frac{1}{64p} \qquad \text{and} \qquad \ell \le \frac{np}{4\ln\frac{1}{pm}}.$$

Let M be an $n \times n$ Bernoulli(p) random matrix. By $\mathcal{E}_{col} = \mathcal{E}_{col}(\ell, m)$ denote the event that for any choice of two disjoint sets $J_1, J_2 \subset [n]$ of cardinality $|J_1| = m$, $|J_2| = \ell m - m$ there exists a row of M with exactly one 1 among components indexed by J_1 and no 1s among components indexed by J_2. Then $\mathbb{P}(\mathcal{E}_{col}) \ge 1 - \exp(-1.5pn)$.

Proof. We first treat the case $m \ge 30$. Fix two disjoint sets $J_1, J_2 \subset [n]$ of required cardinality. The probability that a fixed row has exactly one 1 among components indexed by J_1 and no 1s among components indexed by J_2 equals

$$q := mp(1-p)^{\ell m - 1} \ge mp\exp(-2p\ell m) \ge 29mp/30,$$

where we used $\ell mp \le 1/64$. Since the rows are independent, the probability that M does not have such a row is

$$(1-q)^n \le \exp(-nq) \le \exp(-29mpn/30).$$

Note that the number of all choices of J_1 and J_2 satisfying the conditions of the lemma is

$$\binom{n}{\ell m - m}\binom{n - \ell m + m}{m} \le \left(\frac{en}{(\ell-1)m}\right)^{\ell m - m}\left(\frac{en}{m}\right)^m \le \left(\frac{3n}{\ell m}\right)^{\ell m}(2\ell)^m.$$

Thus union bound over all choices of J_1 and J_2 implies

$$\mathbb{P}((\mathcal{E}_{col})^c) \le \left(\frac{3n}{\ell m}\right)^{\ell m}(2\ell)^m \exp(-29mpn/30).$$

Using that $m \le 1/(64p)$ and $\ell \le \frac{np}{4\ln(1/(pm))}$, we observe $\left(\frac{3n}{\ell m}\right)^{\ell m} \le \exp(mpn/2)$. Since $np \ge 200\ln n$, we have $(2\ell)^m \le \exp(2mpn/5)$. Thus,

$$\mathbb{P}((\mathcal{E}_{col})^c) \le \exp(-mpn/15) \le \exp(-2pn),$$

which proves this case.

The case $m = 2$, $\ell = 15$ is similar. Fixing two disjoint sets $J_1, J_2 \subset [n]$ of the required cardinality, the probability that a fixed row has exactly one 1 among components indexed by J_1 and no 1s among components indexed by J_2 equals

$$q := 2p(1-p)^{29} \ge 2p\exp(-29p).$$

Since rows are independent, the probability that M does not have such a row is

$$(1-q)^n \le (1-2p\exp(-29p))^n \le \exp(-2pn\exp(-29p)) \le \exp(-1.8pn).$$

Using union bound over all choices of J_1 and J_2 we obtain

$$\mathbb{P}(\mathcal{E}^c_{sum}) \le \frac{n^{30}}{2\cdot 28!}\exp(-1.8pn) \le \exp(-1.5pn),$$

which proves the lemma. □

In the next lemma we restrict a matrix to a certain set of columns and estimate the cardinality of a set of rows having exactly one 1. To be more precise, for any $J \subset [n]$ and a 0/1 matrix M denote

$$I_J = I(J,M) := \{i \le n : |\operatorname{supp}\mathbf{R}_i(M) \cap J| = 1\}.$$

The following statement is similar to Lemma 2.7 from [27] and Lemma 3.6 in [30].

Lemma 6.6. *Let $\ell \ge 1$ be an integer and $p \in (0, 1/2]$ be such that $p\ell \le 1/32$. Let M be a Bernoulli(p) random matrix. Then with probability at least*

$$1 - 2\binom{n}{\ell}\exp\left(-n\ell p/4\right)$$

for every $J \subset [n]$ of cardinality ℓ one has

$$\ell pn/16 \le |I(J,M)| \le 2\ell np.$$

In particular, if $\ell = 2\lfloor 1/(64p)\rfloor \le n$, $n \ge 10^5$, and $p \in [100/n, 0.001]$ then, denoting

$$\mathcal{E}_{card} = \mathcal{E}_{card}(\ell) := \{M \in \mathcal{M}_n : \forall J \subset [n] \text{ with } |J| = \ell \text{ one has } |I(J,M)| \in [\ell pn/16, 2\ell pn]\},$$

we have

$$\mathbb{P}\left(\mathcal{E}_{card}\right) \ge 1 - 2\exp\left(-n/500\right).$$

Proof. Fix $J \subset [n]$ of cardinality ℓ. Denote $q = \ell p(1-p)^{\ell-1}$. Since $\ell p \le 1/32$,

$$15\ell p/16 \le \ell p(1-2p\ell) \le \ell p\exp(-2p\ell) \le q \le \ell p \le 1/2.$$

For every $i \le n$, let ξ_i be the indicator of the event $\{i \in I(J,M)\}$. Clearly, ξ_i's are independent Bernoulli(q) random variables and $|I(J,M)| = \sum_{i=1}^n \xi_i$. Applying Lemma 3.4, we observe that for every $0 < \varepsilon < q$

$$\mathbb{P}\left(|I(J,M)| \in [(q-\varepsilon)n, (q+\varepsilon)n]\right) \ge 1 - 2\exp\left(-\frac{n\varepsilon^2}{2q(1-q)}\left(1-\frac{\varepsilon}{3q}\right)\right).$$

Taking $\varepsilon = 14q/15$ we obtain that

$$(q-\varepsilon)n = qn/15 \ge \ell pn/16 \qquad \text{and} \qquad (q+\varepsilon)n \le 2qn \le 2\ell pn,$$

and

$$\frac{n\varepsilon^2}{2q(1-q)}\left(1-\frac{\varepsilon}{3q}\right) \ge \frac{98\cdot 31nq}{225\cdot 45} \ge 0.3n\ell p(1-2\ell p) \ge n\ell p/4.$$

This implies the bound for a fixed J. The lemma follows by the union bound. □

6.3 Cardinality estimates for ε–nets

In this subsection we provide bounds on cardinality of certain discretizations of the sets of vectors introduced earlier. Recall that $\mathbf{e}$ denotes the vector $\mathbf{1}/\sqrt{n}$, $P_{\mathbf{e}}$ denotes the projection on $\mathbf{e}^\perp$, and $P_{\mathbf{e}}^\perp$ is the projection on $\mathbf{e}$, that is $P_{\mathbf{e}}^\perp = \langle \cdot, \mathbf{e}\rangle \mathbf{e}$. We recall also that given $A \subset [n]$, x_A denotes coordinate projection of x on $\mathbb{R}^A$, and that given $x \in \mathbb{R}^n$, σ_x is a (fixed) permutation corresponding to non-increasing rearrangement of $\{|x_i|\}_{i=1}^n$.

Our first lemma deals with nets for $\mathcal{T}_2$ and $\mathcal{T}_3$. We will consider the following normalization:

$$\mathcal{T}_2' = \{x \in \mathcal{T}_2 : x^*_{n_{s_0+1}} = 1\} \quad \text{and} \quad \mathcal{T}_3' = \{x \in \mathcal{T}_3 : x^*_{n_{s_0+2}} = 1\}.$$

The triple norm is defined by the equation $|||x|||^2 = \|P_{\mathbf{e}}x\|^2 + pn\|P_{\mathbf{e}}^\perp x\|^2$.

Lemma 6.7. *Let $n \ge 1$, $p \in (0, 0.001]$, and assume that $d = pn$ is sufficiently large. Let $i \in \{2,3\}$. Then there exists a set $\mathcal{N}_i = \mathcal{N}_i' + \mathcal{N}_i''$, $\mathcal{N}_i' \subset \mathbb{R}^n$, $\mathcal{N}_i'' \subset \operatorname{span}\{\mathbf{1}\}$, with the following properties:*

- $|\mathcal{N}_i| \le \exp\left(2n_{s_0+i}\ln d\right)$.
- *For every $u \in \mathcal{N}_i'$ one has $u_j^* = 0$ for all $j \ge n_{s_0+i}$.*
- *For every $x \in \mathcal{T}_i'$ there are $u \in \mathcal{N}_i'$ and $w \in \mathcal{N}_i''$ satisfying*

$$\|x-u\|_\infty \le \frac{1}{C_\tau\sqrt{d}}, \quad \|w\|_\infty \le \frac{1}{C_\tau\sqrt{d}}, \quad \text{and} \quad |||x-u-w||| \le \frac{\sqrt{2n}}{C_\tau\sqrt{d}}.$$

Since the proof of this lemma in many parts repeats the proofs of Lemma 3.8 from [27] and of Lemma 6.8, we only sketch it below.

Proof. Fix $\mu = 1/(C_\tau\sqrt{d})$ and $i \in \{2,3\}$. We first repeat the proof of Lemma 3.8 from [27] with our choice of parameters. See also the beginning of the proof of Lemma 6.8 below — many definitions, constructions, and calculations are exactly the same, however note that the normalization is slightly different. In particular, the definitions of sets $B_1(x)$, $B_2(x)$ (with $k-1 = n_{s_0+i-1}$), $B_3(x)$ are the same (we do not need the sets $B_0(x)$ and $B_4(x)$). This will show (for large enough d) the existence of a μ-net $\mathcal{N}_i'$ (in the ℓ_∞ metric) for $\mathcal{T}_i'$ such that for every $u \in \mathcal{N}_i'$ one has $u_j^* = 0$ for all $j \ge n_{s_0+i}$ and $|\mathcal{N}_i'| \le \exp\left(1.1n_{s_0+i}\ln d\right)$.

Next given $x \in \mathcal{T}_i'$ let $u = u(x) \in \mathcal{N}_i'$ be such that $\|x-u\|_\infty \le \mu$. Then $\|P_{\mathbf{e}}^\perp(x-u)\| \le \mu\sqrt{n}$. Let $\mathcal{N}_i''$ be a $(\mu\sqrt{n/d})$-net in the segment $\mu\sqrt{n}\,[-\mathbf{e},\mathbf{e}]$ of cardinality at most $2\sqrt{d}$ (note, we are in the one-dimensional setting). Note that every $w \in \mathcal{N}_i''$ is of the form $w = a\,\mathbf{e} = a\,\mathbf{1}/\sqrt{n}$, $|a| \le \mu\sqrt{n}$, in particular, $\|w\|_\infty \le \mu$. Then for x (and the corresponding $u = u(x)$), there exists $w \in \mathcal{N}_i''$ such that

$$|||x-u-w|||^2 = \|P_{\mathbf{e}}(x-u-w)\|^2 + d\|P_{\mathbf{e}}^\perp(x-u-w)\|^2 = \|P_{\mathbf{e}}(x-u)\|^2 + d\|P_{\mathbf{e}}^\perp(x-u)-w\|^2 \le 2\mu^2 n.$$

Finally, note that $|\mathcal{N}_i' + \mathcal{N}_i''| \le 2\sqrt{d}\exp\left(1.1 n_{s_0+i}\ln d\right) \le \exp\left(2 n_{s_0+i}\ln d\right)$. This completes the proof. □

Let $\mathcal{R}_k^1$, $\mathcal{R}_k^2$ be the vector subsets introduced in Subsection 6.1. Consider the increasing sequence $\lambda_1 < \lambda_2 < \ldots < \lambda_m$, $m \ge 1$, defined by

$$\lambda_1 = 1/\sqrt{2}, \quad \lambda_{i+1} = 3\lambda_i \quad \text{for } 1 < i < m, \quad \text{and} \quad \lambda_{m-1} < \lambda_m = C_\tau^2 d \le 3\lambda_{m-1}. \tag{31}$$

Clearly $m \le n$. For $s \in \{1,2\}$, $n_{s_0+1} < k \le n/\ln^2 d$ and $i \le m$ set

$$\mathcal{R}_{ki}^s := \left\{x \in \mathcal{R}_k^s \ :\ \lambda_i\sqrt{n} \le \|x_{\sigma_x([k,n])}\| \le \lambda_{i+1}\sqrt{n}\right\}.$$

It is not difficult to see that the union of $\mathcal{R}_{ki}^s$'s over admissible i gives $\mathcal{R}_k^s$. The sets $\mathcal{R}_{ki}^s$ are "low complexity" sets in the sense that they admit efficient ε-nets. For $s = 1$, the low complexity is a consequence of the condition that $\mathcal{R}_{ki}^1 \subset \mathcal{AC}(\rho)$, i.e., the vectors have a very large almost constant part. For the sets $\mathcal{R}_{ki}^2$, we do not assume the almost constant behavior, but instead rely on the assumption that $\|x_{\sigma_x([k,n])}\|$ is large (much larger than $\sqrt{n}$). This will allow us to pick ε much larger than $\sqrt{n}$, and thus construct a net of small cardinality.

Lemma 6.8. *Let $R \ge 40$ be a (large) constant. Then there is $r_0 > 0$ depending on R with the following property. Let $0 < r \le r_0$, $0 < \rho \le 1/(2R)$, let $n \ge 1$ and $p \in (0, 0.001]$ so that $d = pn$ is sufficiently large (larger than a constant depending on R, r). Let $s \in \{1,2\}$, $n_{s_0+1} < k \le n/\ln^2 d$, $t \le m$, and $40\lambda_t\sqrt{n}/R \le \varepsilon \le \lambda_t\sqrt{n}$, where λ_t and m are defined according to relation* (31). *Then there exists an ε-net $\mathcal{N}_{kt}^s \subset \mathcal{R}_{kt}^s$ for $\mathcal{R}_{kt}^s$ with respect to $|||\cdot|||$ of cardinality at most $(e/r)^{3rn}$.*

Proof. Note that in case of $s = 2$ the set $\mathcal{R}_{kt}^2$ is empty whenever $3\lambda_t < \frac{2}{r}$. So, in the course of the proof we will implicitly assume that $3\lambda_t \ge \frac{2}{r}$ whenever $s = 2$.

We follow ideas of the proof of Lemma 3.8 from [27]. We split a given vector from $\mathcal{R}_{kt}^s$ into few parts according to magnitudes of its coordinates and approximate each part separately. Then we construct nets for vectors with the same splitting and take the union over all nets. We now discuss the splitting. For each $x \in \mathcal{R}_{kt}^s$ consider the following (depending on x) partition of $[n]$. If $s = 2$, set $B_0'(x) = \emptyset$. If $s = 1$ then $x \in \mathcal{AC}(\rho)$ and we set

$$B_0'(x) := \sigma_x(\{j \le n \ :\ |x_j - \lambda_x| \le \rho\}),$$

where $\lambda_x = \pm 1$ is from the definition of $\mathcal{AC}(\rho)$ (note that under the normalization in $\Upsilon_n(r)$ we have $x^*_{n_{s_0+3}} = 1$). Then $|B_0'(x)| > n - n_{s_0+3}$ for $s = 1$. Next, we set

$$\begin{aligned}
B_1(x) &= \sigma_x([n_{s_0+1}]);\\
B_2(x) &= \sigma_x([k-1]) \setminus B_1(x);\\
B_3(x) &= \sigma_x([n_{s_0+3}]) \setminus (B_1(x) \cup B_2(x));\\
B_0(x) &= B_0'(x) \setminus (B_1(x) \cup B_2(x) \cup B_3(x));\\
B_4(x) &= [n] \setminus (B_0(x) \cup B_1(x) \cup B_2(x) \cup B_3(x))
\end{aligned}$$

(one of the sets $B_0(x)$, $B_4(x)$ could be empty). Denote $\ell_x := |B_0(x)|$. Note that the definition of $B_3(x)$ and $B_4(x)$ imply that $\ell_x \le n - n_{s_0+3}$, while the condition $k - 1 \le n_{s_0+3}$ and the above observation for $B_0'(x)$ give $n - 2n_{s_0+3} < \ell_x$ for $s = 1$. Clearly, $\ell_x = 0$ for $s = 2$.

Moreover, we have both for $s = 1$ and $s = 2$:

$$|B_1(x)| = n_{s_0+1}, \quad |B_2(x)| = k - 1 - n_{s_0+1}, \quad |B_3(x)| = n_{s_0+3} - k + 1, \quad |B_4(x)| = n - \ell_x - n_{s_0+3}. \quad (32)$$

Thus, given $\ell \in \{0\} \cup [n - n_{s_0+3} - k + 1, n - k + 1]$ and a partition of $[n]$ into five sets B_i, $0 \le i \le 4$, with cardinalities as in (32), it is enough to construct a net for vectors $x \in \mathcal{R}^s_{kt}$ with $B_i(x) = B_i$, $0 \le i \le 4$, $\ell_x = \ell$, and then to take the union of nets over all possible realizations of ℓ and all such partitions $\{B_0, B_1, B_2, B_3, B_4\}$ of $[n]$.

Now we describe our construction. Fix ℓ as above and fix two parameters $\mu = 1/(C_\tau\sqrt{d})$, and $\nu = 9\lambda_t\sqrt{n}/R$. We would like to emphasize that for the actual calculations in this lemma, taking μ to be a small constant multiple of R^{-1} would be sufficient, however, we would like to run the proof with the above choice of μ because this corresponds to the parameter choice in the previous Lemma 6.7 whose proof we only sketched. Note that for $x \in \mathcal{R}^s_{kt}$ we have $x \notin \mathcal{T}$, hence $x^*_{n_{s_0+1}} \le C_\tau\sqrt{d}x^*_{n_{s_0+2}} \le C_\tau^2 d$ and

$$x^*_1 \le (6d)x^*_2 \le (6d)^2 x^*_{n_1} \le \ldots \le (6d)^{s_0+2} x^*_{n_{s_0+1}} \le C_\tau^2 d(6d)^{s_0+2}. \quad (33)$$

Fix $I_0 \subset [n]$ with $|I_0| = n_{s_0+1}$ (which will play the role of B_1). We will construct a μ-net $\mathcal{N}_{I_0}$ (in the ℓ_∞-metric) for the set

$$\mathcal{T}_{I_0} := \big\{P_{B_1(x)}x : \ x \in \mathcal{R}^s_{kt}, \ B_1(x) = I_0\big\}.$$

Clearly, the nets $\mathcal{N}_{I_0}$ for various I_0's can be related by appropriate permutations, so without loss of generality we can assume for now that $I_0 = [n_{s_0+1}]$. First, consider the partition of I_0 into sets $I_1, \ldots, I_{s_0+2}$ defined by

$$I_1 = [2] \quad \text{and} \quad I_j = [n_{j-1}] \setminus [n_{j-2}], \ \text{ for } \ 2 \le j \le s_0 + 2.$$

Consider the set

$$\mathcal{T}^* := \big\{x \in \mathcal{T}_{[n_{s_0+1}]} : \ \sigma_x(I_j) = I_j, \ \ j = 1, 2, \ldots, s_0 + 2\big\}.$$

By the definition of $\mathcal{T}_{I_0}$, for every $x \in \mathcal{T}^*$, one has $\|P_{I_j}x\|_\infty \le b_j := C_\tau^2 d(6d)^{s_0+3-j}$ for every $j \le s_0 + 2$ (where as usual P_I denotes the coordinate projection onto $\mathbb{R}^I$). Define a μ–net (in the ℓ_∞-metric) for $\mathcal{T}^*$ by setting

$$\mathcal{N}^* := \mathcal{N}_1 \oplus \mathcal{N}_2 \oplus \cdots \oplus \mathcal{N}_{s_0+2},$$

where $\mathcal{N}_j$ is a μ-net (in the ℓ_∞-metric) of cardinality at most

$$(3b_j/\mu)^{|I_j|} \le (C_\tau^3 d^{3/2}(6d)^{s_0+3-j})^{n_{j-1}} \le (C_\tau^3(6d)^{s_0+5-j})^{n_{j-1}}$$

in the coordinate projection of the cube $P_{I_j}(b_j B^n_\infty)$. Recall that $n_0 = 2$, $n_j = 30\ell_0^{j-1}$, $1 \le j \le s_0$, where ℓ_0 and s_0 are given by (27). Since d is large enough,

$$2s_0 + 8 + 30\sum_{j=2}^{s_0+1}(s_0 + 5 - j)\ell_0^{j-2} = 2s_0 + 8 + 30\sum_{m=1}^{s_0-1}(m+3)\ell_0^{s_0-m} \le 121\ell_0^{s_0-1} \le 4.1n_{s_0+1},$$

which implies

$$|\mathcal{N}^*| \le \prod_{j=1}^{s_0+2}|\mathcal{N}_j| \le \exp(7.1n_{s_0+1}\ln(6C_\tau^2 d)).$$

To pass from the net for $\mathcal{T}^*$ to the net for $\mathcal{T}_{[n_{s_0+1}]}$, let $\mathcal{N}_{[n_{s_0+1}]}$ be the union of nets constructed as $\mathcal{N}^*$ but for arbitrary partitions $I'_1, \ldots, I'_{s_0+2}$ of $[n_{s_0+1}]$ with $|I'_j| = |I_j|$. Using that

$$\sum_{j=1}^{s_0+1} n_{j-1} \le 2 + 30\sum_{j=0}^{s_0-1}\ell_0^j \le 2 + 30\ell_0^{s_0-1}/(1 - 1/\ell_0) \le 2n_{s_0+1}$$

and $e\ell_0 \le d$, we obtain that the cardinality of $\mathcal{N}_{[n_{s_0+1}]}$ is at most

$$|\mathcal{N}^*| \prod_{j=1}^{s_0+1} \binom{n_j}{n_{j-1}} \le |\mathcal{N}^*| \prod_{j=1}^{s_0+1} \Big(\frac{en_j}{n_{j-1}}\Big)^{n_{j-1}} \le |\mathcal{N}^*| \prod_{j=1}^{s_0+1} (e\ell_0)^{n_{j-1}} \le \exp(9.1 n_{s_0+1} \ln(6C_\tau^2 d)).$$

Next we construct a net for the parts of the vectors corresponding to B_2. Fix $J_0 \subset [n]$ with $|J_0| = k-1-n_{s_0+1}$ (it will play the role of B_2). We construct a μ-net (in the ℓ_∞-metric) for the set

$$\mathcal{T}^2_{J_0} := \{P_{B_2(x)}x \,:\, x \in \Upsilon_n(r) \setminus \mathcal{T},\ B_2(x) = J_0\}.$$

Since by (33), we have $x^*_{n_{s_0+1}} \le C_\tau^2 d$ for every $x \in \Upsilon_n(r) \setminus \mathcal{T}$, it is enough to take a μ-net $\mathcal{K}_{J_0}$ of cardinality at most

$$|\mathcal{K}_{J_0}| \le (3C_\tau^2 d/\mu)^{|J_0|} \le (3C_\tau^3 d^{3/2})^k$$

in the coordinate projection of the cube $P_{J_0}(C_\tau^2 d B_\infty^n)$.

Now we turn to the part of the vectors corresponding to B_3. Fix $D_0 \subset [n]$ with $|D_0| = n_{s_0+3} - k + 1$ (it will play the role of B_3). For this part we use ℓ_2-metric and construct a ν-net (in the *Euclidean metric* this time) for the set

$$\mathcal{T}^3_{D_0} := \{P_{B_3(x)}x \,:\, x \in \mathcal{R}^s_{kt},\ B_3(x) = D_0\}.$$

Since for $x \in \mathcal{R}^s_{kt}$ we have $\|x_{B_3(x)}\| \le \|x_{\sigma_x([k,n])}\| \le 3\lambda_t\sqrt{n}$, there exists a corresponding ν-net $\mathcal{L}_{D_0}$ in the coordinate projection of the Euclidean ball $P_{D_0}(3\lambda_t\sqrt{n}B_2^n)$ of cardinality at most

$$|\mathcal{L}_{D_0}| \le (9\lambda_t\sqrt{n}/\nu)^{|D_0|} \le R^{n_{s_0+3}} \le R^{rn}.$$

Next we approximate the almost constant part of a vector (corresponding to B_0), provided that it is not empty (otherwise we skip this step). Fix $A_0 \subset [n]$ with $|A_0| = \ell$ (it will play the role of B_0) and denote

$$\mathcal{T}^0_{A_0} := \{P_{B_0(x)}x \,:\, x \in \big(\Upsilon_n(r) \setminus \mathcal{T}\big) \cap \mathcal{AC}(\rho),\ B_0(x) = A_0\}.$$

Let $\mathcal{K}^0_{A_0} := \{\pm P_{A_0}\mathbf{1}\}$. Since for every $x \in \Upsilon_n(r)$ we have either $\lambda_x = 1$ or $\lambda_x = -1$, by the definition of $B_0(x)$, every $z \in \mathcal{T}^0_{A_0}$ is approximated by one of $\pm P_{A_0}\mathbf{1}$ within error ρ in the ℓ_∞-metric.

We use 0 to approximate the last part of the vector, which corresponds to B_4. Note that for any $x \in \mathcal{R}^1_{kt}$ we have $\|P_{B_4(x)}x\| \le \sqrt{rn} \le \sqrt{2r}\lambda_t\sqrt{n}$, in view of the condition $x \in \mathcal{AC}(\rho)$. On the other hand, for $x \in \mathcal{R}^2_{kt}$ we have $\|P_{B_4(x)}x\| \le \sqrt{n} \le \frac{3r}{2}\lambda_t\sqrt{n}$.

Now we combine our nets. Consider the net

$$\mathcal{N}_0 := \bigcup_{\ell, I_0, J_0, D_0, A_0} \big\{y = y_1 + y_2 + y_3 + y_0 :\, y_1 \in \mathcal{N}_{I_0},\, y_2 \in \mathcal{K}_{J_0},\, y_3 \in \mathcal{L}_{D_0}, y_0 \in \mathcal{K}^0_{A_0}\big\},$$

where the union is taken over all $\ell \in \{0\} \cup [n-2n_{s_0+3}, n-n_{s_0+3}]$ and all partitions of $[n]$ into I_0, J_0, D_0, A_0, B with $|I_0| = n_{s_0+1}$, $|J_0| = k-1-n_{s_0+1}$, $|D_0| = n_{s_0+3}-k+1$, $|A_0| = \ell$, and $B = [n] \setminus (I_0 \cup J_0 \cup D_0 \cup A_0)$. Then the cardinality of $\mathcal{N}_0$,

$$|\mathcal{N}_0| \le n\binom{n}{n_{s_0+1}}\binom{n-n_{s_0+1}}{k-1-n_{s_0+1}}\binom{n-k+1}{n_{s_0+3}-k+1}\binom{n-n_{s_0+3}}{\ell} \max_{I_0}|\mathcal{N}_{I_0}| \max_{J_0}|\mathcal{K}_{J_0}| \max_{D_0}|\mathcal{L}_{D_0}| \max_{A_0}|\mathcal{K}^0_{A_0}|.$$

Using that $n_{s_0+1} \le n/(64d)$, $k \le n/\ln^2 d$, $n_{s_0+3} \le rn$, $\ell = 0$ or $\ell \ge n - 2n_{s_0+3}$, the obtained bounds on nets, as well as that d is large enough and r is small enough (smaller than a constant depending on R), we observe that the cardinality of $\mathcal{N}_0$ is bounded by

$$n\,(ed)^{n/d}\,\big(2e\ln^2 d\big)^{n/\ln^2 d}\,(2e/r)^{rn}\,(2e/r)^{rn}\,\exp(9.1n\ln(6C_\tau^2 d)/(64d))\,(3C_\tau^3 d^{3/2})^{n/\ln^2 d} R^{rn}\cdot 2 \le (e/r)^{2.5rn}.$$

By construction, for every $x \in \mathcal{R}^s_{kt}$ there exists $y = y_1 + y_2 + y_3 + y_0 \in \mathcal{N}_0$ such that

$$\begin{aligned}\|x-y\| &\le \|P_{B_1(x)}x - y_1\| + \|P_{B_2(x)}x - y_2\| + \|P_{B_3(x)}x - y_3\| + \|P_{B_4(x)}x\| + \|P_{B_0(x)}x - y_0\| \\ &\le \mu\sqrt{n_{s_0+1}} + \mu\sqrt{k-1-n_{s_0+1}} + \nu + \sqrt{2r}\lambda_t\sqrt{n} + \rho\sqrt{n} \le \frac{2\sqrt{n}}{C_\tau\sqrt{d}} + \rho\sqrt{n} + \frac{9\lambda_t\sqrt{n}}{R} \le \frac{10\lambda_t\sqrt{n}}{R},\end{aligned}$$

where we used that $\rho \le 1/(2R) \le \lambda_1/(\sqrt{2}R) \le \lambda_t/(\sqrt{2}R)$ and that r is sufficiently small.

Finally we adjust our net to $|||\cdot|||$. Note that by Lemma 6.4 for every $x \in \Upsilon_n(r) \setminus \mathcal{T}$,

$$|\langle x, \mathbf{e}\rangle| = \left|\sum_{i=1}^n \frac{x_i}{\sqrt{n}}\right| \le \|x\| \le \frac{384 C_\tau^2 d^4}{(64p)^{\ln(6d)}} \le e^{rn}.$$

Therefore, there exists an $\varepsilon/(4\sqrt{pn})$-net $\mathcal{N}_*$ in $P_{\mathbf{e}}^\perp \mathcal{R}^s_{kt}$ of cardinality $8\sqrt{pn}e^{rn}/\varepsilon$ (note, the rank of $P_{\mathbf{e}}^\perp$ is one). Then, by the constructions of nets, for every $x \in \mathcal{R}^s_{kt}$ there exist $y \in \mathcal{N}_0$ and $y_* \in \mathcal{N}_*$ such that

$$|||x - P_{\mathbf{e}}y - y_*|||^2 = \|P_{\mathbf{e}}(x-y)\|^2 + pn\|P_{\mathbf{e}}^\perp x - y_*\|^2 \le \frac{100\lambda_t^2 n}{R^2} + \varepsilon^2/16 \le \varepsilon^2/8.$$

Thus the set $\mathcal{N} = P_{\mathbf{e}}(\mathcal{N}_0) + \mathcal{N}_*$ is an $(\varepsilon/2)$-net for $\mathcal{R}^s_{kt}$ with respect to $|||\cdot|||$ and its cardinality is bounded by $(e/r)^{3rn}$. Using standard argument we pass to an ε-net $\mathcal{N}^s_{kt} \subset \mathcal{R}^s_{kt}$ for $\mathcal{R}^s_{kt}$. □

6.4 Proof of Theorem 6.2

Proof. Recall that the sets $\mathcal{R}^s_{ki}$ were introduced just before Lemma 6.8 and the event $\mathcal{E}_{nrm}$ was defined in Proposition 3.14.

Fix $s \in \{1,2\}$, $k \le n/\ln^2 d$, $A := [k,n]$, $i \le m$. Set $\varepsilon := \lambda_i\sqrt{n}/(600\sqrt{2}C_0)$, where λ_i and m are defined according to (31). Applying Lemma 6.8 with $R = 24000\sqrt{2}C_0$, we find an ε-net (in the $|||\cdot|||$–norm) $\mathcal{N}^s_{ki} \subset \mathcal{R}^s_{ki}$ for $\mathcal{R}^s_{ki}$ of cardinality at most $(e/r)^{3rn}$. Take for a moment any $y \in \mathcal{N}^s_{ki}$. Note that $\|y_{\sigma(A)}\| \ge C_0\|y_{\sigma(A)}\|_\infty/\sqrt{p}$, $\|y_\sigma(A)\| \ge \lambda_i\sqrt{n}$ (where $\sigma = \sigma_y$). Then Proposition 3.10 implies $\mathbb{P}(\mathcal{E}_y^c) \le e^{-3n}$, where

$$\mathcal{E}_y = \left\{\|My\| > \frac{\sqrt{pn}}{3\sqrt{2}C_0}\,\|y_{\sigma(A)}\|\right\}.$$

Condition on the event

$$\mathcal{E}_{nrm} \cap \bigcap_{y \in \mathcal{N}^s_{ki}} \mathcal{E}_y.$$

Using the definition of $\mathcal{N}^s_{ki}$ and $\mathcal{R}^s_{ki}$, the triangle inequality, and the definition of $\mathcal{E}_{nrm}$ from Proposition 3.14, we get that for any $x \in \mathcal{R}^s_{ki}$ there is $y \in \mathcal{N}^s_{ki}$ such that $|||x-y||| \le \varepsilon$, and hence

$$\|Mx\| \ge \|My\| - \|M(x-y)\| > \frac{\sqrt{pn}}{3\sqrt{2}C_0}\,\|y_{\sigma(A)}\| - 100\sqrt{pn}\varepsilon \ge \frac{\sqrt{p}\lambda_i n}{6\sqrt{2}C_0}.$$

Using that $|\mathcal{N}^s_{ki}| \le (e/r)^{3rn}$, that $\lambda_i \ge 1/\sqrt{2}$, and the union bound, we obtain

$$\mathbb{P}\left(\mathcal{E}_{nrm} \cap \left\{\exists x \in \mathcal{R}^s_{ki} \,:\, \|Mx\| \le \frac{\sqrt{p}n}{12C_0}\right\}\right) \le \mathbb{P}\Big(\mathcal{E}_{nrm} \cap \bigcup_{y \in \mathcal{N}^s_{ki}} \mathcal{E}_y^c\Big) \le e^{-3(1-r\ln(e/r))n}.$$

Since $\mathcal{R} = \bigcup_{k,i}(\mathcal{R}^1_{ki} \cup \mathcal{R}^2_{ki})$ and r is small enough, the result follows by the union bound together with (11) and Lemma 3.6 applied with $t = 30$ in order to estimate $\mathbb{P}(\mathcal{E}_{nrm})$. □

6.5 Lower bounds on $\|Mx\|$ for vectors from $\mathcal{T}_0 \cup \mathcal{T}_1$

The following lemma provides a lower bound on the ratio $\|Mx\|/\|x\|_2$ for vectors x from $\mathcal{T}_0 \cup \mathcal{T}_1$.

Lemma 6.9. *Let $n \geq 1$, $0 < p < 0.001$, and assume that $d = pn \geq 200 \ln n$. Then*

$$\mathbb{P}\left(\left\{ \exists\, x \in \mathcal{T}_0 \cup \mathcal{T}_1 \quad \textit{such that} \quad \|Mx\| \leq \frac{(64p)^\kappa}{192(pn)^2}\, \|x\| \right\} \right) \leq n(1-p)^n + e^{-1.4np},$$

where κ is defined by (28).

Proof. Let δ_{ij}, $i, j \leq n$ be entries of M. Let $\mathcal{E}$ be the event that there are no zero columns in M. Clearly, $\mathbb{P}(\mathcal{E}) \geq 1 - n(1-p)^n$.

Also, for each $1 \leq j \leq s_0 + 1$, let $\mathcal{E}_j = \mathcal{E}_{col}(\ell_0, n_{j-1})$ be the event introduced in Lemma 6.5 (with s_0, ℓ_0 defined in (27)), and observe that, according to Lemma 6.5, $\mathbb{P}(\mathcal{E}_j) \geq 1 - e^{-1.5np}$ for every j.

Recall that σ_x denotes a permutation $[n]$ such that $x^*_i = |x_{\sigma(i)}|$ for $i \leq n$. Pick any $x \in \mathcal{T}_0 \cup \mathcal{T}_1$. In the case $x \in \mathcal{T}_0$ set $m = m_1 = 1$ and $m_2 = 2$. In the case $x \in \mathcal{T}_{1j}$ for some $1 \leq j \leq s_0 + 1$ set $m = m_1 = n_{j-1}$ and $m_2 = n_j$. Then by the definition of sets $\mathcal{T}_0, \mathcal{T}_1$ we have $x^*_m > 6d x^*_{m_2}$. Let

$$J^\ell = J^\ell(x) = \sigma_x([m]), \quad J^r = J^r(x) = \sigma_x([m_2 - 1] \setminus [m]), \quad \text{and} \quad J(x) = (J^\ell \cup J^r)^c$$

(if $x \in \mathcal{T}_0$ then $J^r = \emptyset$). Note that by our definition we have $|x_i| > 6d|x_u|$ for any $i \in J^\ell(x)$ and $u \in J(x)$, and that $\max_{i \in J(x)} |x_i| \leq x^*_{m_2}$. Denote by $I^\ell(x)$ the (random) set of rows of M having exactly one 1 in $J^\ell(x)$ and no 1's in $J^r(x)$. Now we recall that the event $\mathcal{E}_{sum}$ was introduced in Lemma 3.4 (we use it with $q = p$) and set

$$\mathcal{E}' := \mathcal{E} \cap \mathcal{E}_{sum} \cap \bigcap_{j=1}^{s_0+1} \mathcal{E}_j.$$

Clearly, conditioned on $\mathcal{E}'$, the set $I^\ell(x)$ is not empty for any $x \in \mathcal{T}_0 \cup \mathcal{T}_1$. By definition, for every $s \in I^\ell(x)$ there exists $j(s) \in J^\ell(x)$ such that

$$\operatorname{supp} R_s(M) \cap J^\ell(x) = \{j(s)\} \qquad \text{and} \qquad \operatorname{supp} R_s(M) \cap J^r(x) = \emptyset.$$

Since $j(s) \in J^\ell(x)$ (which implies $|x_{j(s)}| \geq x^*_m > 6d x^*_{m_2}$), we obtain

$$|\langle R_s(M),\, x\rangle| = \Big| x_{j(s)} + \sum_{j \in J(x)} \delta_{sj} x_j \Big| \geq |x_{j(s)}| - x^*_{m_2} \sum_{j \in J(x)} \delta_{sj} \geq x^*_m - \frac{x^*_m}{6d} \sum_{j \in J(x)} \delta_{sj}.$$

Observe that conditioned on $\mathcal{E}_{sum}$ we have $\sum_{j \in J(x)} \delta_{sj} \leq \sum_{j=1}^n \delta_{sj} \leq 3.5pn = 3.5d$. Thus, everywhere on $\mathcal{E}'$ we have for all $x \in \mathcal{T}_0 \cup \mathcal{T}_1$,

$$\|Mx\| \geq |\langle R_s(M),\, x\rangle| \geq x^*_m/3, \quad s \in I^\ell(x).$$

Finally, in the case $x \in \mathcal{T}_0$ we have $m = 1$ and $\|x\| \leq \sqrt{n} x^*_1$. In the case $x \in \mathcal{T}_{1j}$ by Lemma 6.4 we have

$$\|x\| \leq \frac{64(pn)^2}{(64p)^\kappa}\, x^*_m,$$

This proves the lower bound on $\|Mx\|/\|x\|$ conditioned on $\mathcal{E}'$. The probability bound follows by the union bound, Lemmas 3.4 and 6.5, and since $s_0 \leq \ln n$, indeed

$$\mathbb{P}\left(\mathcal{E} \cap \mathcal{E}_{sum} \cap \bigcap_{j=1}^{s_0+1} \mathcal{E}_j \right) \geq 1 - n(1-p)^n - (s_0+2)e^{-1.5np} \geq 1 - n(1-p)^n - e^{-1.4np}.$$

□

6.6 Individual bounds for vectors from $\mathcal{T}_2 \cup \mathcal{T}_3$

In this section we provide individual probability bounds for vectors from the nets constructed in Lemma 6.7. To obtain the lower bounds on $\|Mx\|$, we consider the behavior of the inner products $\langle \mathbf{R}_i(M), x\rangle$, more specifically, of the Lévy concentration function for $\langle \mathbf{R}_i(M), x\rangle$. To estimate this function, we will consider $2m$ columns of M corresponding to the m biggest and m smallest (in absolute value) coordinates of x, where $m = n_{s_0+1}$ or $m = n_{s_0+2}$. In a sense, our anti-concentration estimates will appear in the process of swapping 1's and 0's within a specially chosen subset of the matrix rows. A crucial element in this process is to extract a pair of subsets of indices on which the chosen matrix rows have only one non-zero component. This will allow to get anti-concentration bounds by "sending" the non-zero component into the other index subset from the pair. The main difficulty in this scheme comes from the restriction $2mp \le 1/32$ from Lemma 6.6, which guarantees existence of sufficiently many required subsets (and rows) but which cannot be directly applied to $m = n_{s_0+2}$. To resolve this problem we use idea from [30]. We split the initially fixed set of $2m$ columns into smaller subsets of columns of size at most $1/(64p)$ each, and create independent random variables corresponding to this splitting. Then we apply Proposition 3.9, allowing to deal with the Lévy concentration function for sums of independent random variables.

We first describe subdivisions of $\mathcal{M}_n$ used in [30]. Recall that $\mathcal{M}_n$ denotes the class of all $n \times n$ matrices with 0/1 entries. We recall also that the probability measure $\mathbb{P}$ on $\mathcal{M}_n$ is always assumed to be induced by a Bernoulli(p) random matrix. Given $J \subset [n]$ and $M \in \mathcal{M}_n$ denote

$$I(J, M) = \{i \le n \,:\, |\operatorname{supp} \mathbf{R}_i(M) \cap J| = 1\}.$$

By $\mathcal{M}_J$ we denote the set of $n \times |J|$ matrices with 0/1 entries and with columns indexed by J. Fix $q_0 \le n$ and a partition J_0, J_1, ..., J_{q_0} of $[n]$. Given subsets $I_1, \dots, I_{q_0}$ of $[n]$ and $V = (v_{ij}) \in \mathcal{M}_{J_0}$, denote $\mathcal{I} = (I_1, \dots, I_{q_0})$ and consider the class

$$\mathcal{F}(\mathcal{I}, V) = \{M = (\mu_{ij}) \in \mathcal{M}_n \,:\, \forall q \in [q_0] \quad I(J_q, M) = I_q \ \text{ and } \ \forall i \le n \, \forall j \in J_0 \ \ \mu_{ij} = v_{ij}\}\,.$$

In words, we fix the columns indexed by J_0 and for each $q \in [q_0]$ we fix the row indices having exactly one 1 in columns indexed by J_q. Then, for any fixed partition J_0, J_1, ..., J_{q_0}, $\mathcal{M}_n$ is the disjoint union of classes $\mathcal{F}(\mathcal{I}, V)$ over all $V \in \mathcal{M}_{J_0}$ and all $\mathcal{I} \in (\mathcal{P}([n]))^{q_0}$, where $\mathcal{P}(\cdot)$ denotes the power set.

The following is an important, but simple observation.

Lemma 6.10. *Let $\mathcal{F}(\mathcal{I}, V)$ be a non-empty class (defined as above), and denote by $\mathbb{P}_{\mathcal{F}}$ the induced probability measure on $\mathcal{F}(\mathcal{I}, V)$, i.e., let*

$$\mathbb{P}_{\mathcal{F}}(B) := \frac{\mathbb{P}(B)}{\mathbb{P}(\mathcal{F}(\mathcal{I}, V))}, \quad B \subset \mathcal{F}(\mathcal{I}, V).$$

Then the matrix rows for matrices in $\mathcal{F}(\mathcal{I}, V)$ are mutually independent with respect to $\mathbb{P}_{\mathcal{F}}$, in other words, a random matrix distributed according to $\mathbb{P}_{\mathcal{F}}$ has mutually independent rows.

Finally, given a vector $v \in \mathbb{R}^n$, a class $\mathcal{F}(\mathcal{I}, V)$, indices $i \le n$, $q \le q_0$, define

$$\xi_q(i) = \xi_q(M, v, i) := \sum_{j \in J_q} \delta_{ij} v_j, \quad M = (\delta_{ij}) \in \mathcal{F}(\mathcal{I}, V). \tag{34}$$

We will view $\xi_q(i)$ as random variables on $\mathcal{F}(\mathcal{I}, V)$ (with respect to the measure $\mathbb{P}_{\mathcal{F}}$). It is not difficult to see that for every fixed i, the variables $\xi_1(i), \dots \xi_{q_0}(i)$ are mutually independent, and, moreover, whenever $i \in I_q$, the variable $\xi_q(i)$ is uniformly distributed on the multiset $\{v_j\}_{j \in J_q}$. Thus, we may apply Proposition 3.9 to

$$|\langle \mathbf{R}_i(M), v\rangle| = \Big| \sum_{q=0}^{q_0} \xi_q(i) \Big|$$

with some $\alpha > 0$ satisfying $\mathcal{Q}(\xi_q(i), 1/3) \le \alpha$ for every $i \in I_q$. This gives

$$\mathbb{P}_{\mathcal{F}}\{|\langle \mathbf{R}_i(M), x+y\rangle| \le 1/3\} \le \frac{C_0\alpha}{\sqrt{(1-\alpha)|\{q \ge 1 :\, i \in I_q\}|}}, \tag{35}$$

where C_0 is a positive absolute constant.

We are ready now to estimate individual probabilities.

Lemma 6.11 (Individual probabilities)**.** *There exist absolute constants $C, C' > 1 > c_1 > 0$ such that the following holds. Let $p \in (0, 1/64]$, $d = pn \ge 2$, Set $m_0 = \lfloor 1/(64p)\rfloor$ and let m_1 and m_2 be such that*

$$1 \le m_1 < m_2 \le n - m_1.$$

Let $y \in \operatorname{span}\{\mathbf{1}\}$ and assume that $x \in \mathbb{R}^n$ satisfies

$$x^*_{m_1} > 2/3 \quad \text{and} \quad x^*_i = 0 \ \text{ for every } \ i > m_2.$$

Denote $m = \min(m_0, m_1)$ and consider the event

$$E(x,y) = \left\{M \in \mathcal{M}_n \,:\, \|M(x+y)\| \le \sqrt{c_1 md}\right\}.$$

Then in the case $m_1 \le m_0$ one has

$$\mathbb{P}(E(x,y) \cap \mathcal{E}_{card}) \le 2^{-md/20},$$

and in the case $m_1 > C'm_0$ one has

$$\mathbb{P}(E(x,y) \cap \mathcal{E}_{card}) \le \left(\frac{Cn}{m_1 d}\right)^{md/20},$$

where $\mathcal{E}_{card}$ is the event introduced in Lemma 6.6 with $\ell = 2m$.

Remark 6.12. *Below we apply Lemma 6.11 for sets $\mathcal{T}_i$ with the following choice of parameters. For $i = 2$ we set*

$$m_1 = m_0 = n_{s_0+1} = \max(30\ell_0^{s_0-1}, \lfloor 1/(64p)\rfloor), \quad m_2 = n_{s_0+2}, \quad \text{and} \quad p \le 0.001,$$

obtaining

$$\mathbb{P}(E(x,y) \cap \mathcal{E}_{card}) \le 2^{-n_{s_0+1}d/20}.$$

For $i = 3$, we set

$$m_1 = n_{s_0+2} = \lfloor n/\sqrt{d}\rfloor > m_0 = n_{s_0+1}, \quad m_2 = n_{s_0+3}, \quad \text{and} \quad p \le 0.001,$$

obtaining for large enough d,

$$\mathbb{P}(E(x,y) \cap \mathcal{E}_{card}) \le \left(\frac{Cn}{n_{s_0+2}d}\right)^{n_{s_0+1}d/20} \le \left(\sqrt{d}/(2C)\right)^{-n_{s_0+1}d/20}.$$

To prove Lemma 6.11 it will be convenient to use the same notation as in Lemma 6.9. Given two disjoint subsets J^ℓ, $J^r \subset [n]$ and a matrix $M \in \mathcal{M}_n$, denote

$$I^\ell = I^\ell(M) := \{i \le n :\, |\operatorname{supp}\mathbf{R}_i(M) \cap J^\ell| = 1 \ \text{ and } \ \operatorname{supp}\mathbf{R}_i(M) \cap J^r = \emptyset\},$$

and

$$I^r = I^r(M) := \{i \le n :\, \operatorname{supp}\mathbf{R}_i(M) \cap J^\ell = \emptyset \ \text{ and } \ |\operatorname{supp}\mathbf{R}_i(M) \cap J^r| = 1\}.$$

Here the upper indices ℓ and r refer to *left* and *right*.

Proof. Let $d = pn$ and fix $\gamma = mp/72 = md/(72n)$.

Fix $x \in \mathbb{R}^n$ and $y \in \operatorname{span}\{\mathbf{1}\}$ satisfying the conditions of the lemma. Let $\sigma = \sigma_x$, that is, a permutation of $[n]$ such that $x_i^* = |x_{\sigma(i)}|$ for all $i \le n$. Denote $q_0 = m_1/m$ and without loss of generality assume that either $q_0 = 1$ or that q_0 is a large enough integer. Let $J_1^\ell, J_2^\ell, \dots, J_{q_0}^\ell$ be a partition of $\sigma([m_1])$ into sets of cardinality m each, and let $J_1^r, J_2^r, \dots, J_{q_0}^r$ be a partition of $\sigma([n - m_1 + 1, n])$ into sets of cardinality m each. Denote

$$J_q := J_q^\ell \cup J_q^r \quad \text{for} \quad q \in [q_0] \qquad \text{and} \qquad J_0 := [n] \setminus \bigcup_{q=1}^{q_0} J_q.$$

Then J_0, J_1, ..., J_{q_0} is a partition of $[n]$, which we fix in this proof. Let M be a $0/1$ $n \times n$ matrix. For every pair J_q^ℓ, J_q^r, let the sets $I_q^\ell(M)$ and $I_q^r(M)$ be defined as after Remark 6.12 and let $I_q(M) = I_q^\ell(M) \cup I_q^r(M)$. Since

$$|J_q| = 2m \le 2m_0 \le 1/(32p),$$

and by the definition of the event $\mathcal{E}_{card}$ (see Lemma 6.6 with $\ell = 2m$), we have

$$|I_q(M)| \in [md/8,\, 4md] \tag{36}$$

everywhere on $\mathcal{E}_{card}$. Now we represent $\mathcal{M}_n$ as a disjoint union of classes $\mathcal{F}(\mathcal{I}, V)$ defined at the beginning of this subsection with $V \in \mathcal{M}_{J_0}$ and $\mathcal{I} = (I_1, \dots, I_q)$. Since it is enough to prove a uniform upper bound for classes $\mathcal{F}(\mathcal{I}, V) \cap \mathcal{E}_{card}$ and since for every such non-empty class $\mathcal{I}$ must satisfy (36) for every $q \le q_0$, we have

$$\mathbb{P}(E(x,y) \cap \mathcal{E}_{card}) \le \max \mathbb{P}(E(x,y) \cap \mathcal{E}_{card} \,|\, \mathcal{F}(\mathcal{I}, V)) \le \max \mathbb{P}(E(x,y) |\, \mathcal{F}(\mathcal{I}, V)),$$

where the first maximum is taken over all $\mathcal{F}(\mathcal{I}, V)$ with $\mathcal{F}(\mathcal{I}, V) \cap \mathcal{E}_{card} \ne \emptyset$ and the second maximum is taken over all $\mathcal{F}(\mathcal{I}, V)$ with I_q's satisfying condition (36).

Fix any class $\mathcal{F}(\mathcal{I}, V)$, where $\mathcal{I}$ satisfies (36), and denote the corresponding induced probability measure on the class by $\mathbb{P}_{\mathcal{F}}$, that is

$$\mathbb{P}_{\mathcal{F}}(\cdot) = \mathbb{P}(\cdot \,|\, \mathcal{F}(\mathcal{I}, V)).$$

Let

$$I := \bigcup_{q=1}^{q_0} I_q.$$

Note that $|I| \le 4q_0 md$. We first show that the set of i's which belongs to many I_q's is large. More precisely, denote

$$A_i = \{q \in [q_0] \,:\, i \in I_q\}, \;\; i \in [n], \qquad \text{and} \qquad I_0 = \{i \le n \,:\, |A_i| \ge \gamma q_0\}.$$

Then, using bounds on cardinalities of I_q's, one has

$$md q_0/8 \le \sum_{q=1}^{q_0} |I_q| = \sum_{i=1}^{n} |A_i| \le |I_0| q_0 + (n - |I_0|)\gamma q_0 \le |I_0| q_0 + n\gamma q_0.$$

Thus,

$$|I_0| \ge md/8 - n\gamma \ge md/9.$$

Without loss of generality we assume that $I_0 = \{1, 2, \dots |I_0|\}$ and only consider the first $k := \lceil md/9 \rceil$ indices from it. Then $[k] \subset I_0$.

Now, by definition, for matrices $M \in E(x,y)$ we have

$$\|M(x+y)\|^2 = \sum_{i=1}^{n} |\, \langle \mathbf{R}_i(M), x+y \rangle \,|^2 \le c_1\, md.$$

Therefore there are at most $9c_1md$ rows with $|\langle \mathbf{R}_i(M), x+y)\rangle| \geq 1/3$. Hence,

$$|\{i \leq k \,:\, |\langle \mathbf{R}_i(M), x+y\rangle| < 1/3\}| \geq md/9 - 9c_1md \geq (1/9 - 9c_1)md.$$

Let $k_0 := \lceil (1/9 - 9c_1)md \rceil$ and for every $i \leq k$ denote

$$\Omega_i := \{M \in \mathcal{F}(\mathcal{I}, V) \,:\, |\,\langle \mathbf{R}_i(M), x+y\rangle\,| < 1/3\} \quad \text{and} \quad \Omega_0 = \mathcal{F}(\mathcal{I}, V).$$

Then

$$\mathbb{P}_{\mathcal{F}}(E(x,y)) \leq \sum_{\substack{B\subset[k]\\|B|=k_0}} \mathbb{P}_{\mathcal{F}}\Big(\bigcap_{i\in B}\Omega_i\Big) \leq \binom{k}{k_0} \max_{\substack{B\subset[k]\\|B|=k_0}} \mathbb{P}_{\mathcal{F}}\Big(\bigcap_{i\in B}\Omega_i\Big).$$

Without loss of generality we assume that the maximum above is attained at $B = [k_0]$. Then

$$\mathbb{P}_{\mathcal{F}}(E(x,y)) \leq (e/(81c_1))^{9c_1md} \prod_{i=1}^{k_0} \mathbb{P}_{\mathcal{F}}(\Omega_i|\,\Omega_1\cap\ldots\cap\Omega_{i-1}) = (e/(81c_1))^{9c_1md} \prod_{i=1}^{k_0} \mathbb{P}_{\mathcal{F}}(\Omega_i), \qquad (37)$$

where at the last step we used mutual independence of the events Ω_i (with respect to measure $\mathbb{P}_{\mathcal{F}}$), see Lemma 6.10.

Next we estimate the factors in the product. Fix $i \leq k_0$ and $A_i = \{q \,:\, i \in I_q\}$. Since, by our assumptions, $i \in I_0$, we have $|A_i| \geq \gamma q_0$. Consider the random variables $\xi_q(i) = \xi_q(M, x+y, i)$, $q \in A_i$, defined in (34). Then by (35) we have

$$\begin{aligned}\mathbb{P}_{\mathcal{F}}(\Omega_i) = \mathbb{P}_{\mathcal{F}}\big\{|\,\langle \mathbf{R}_i(M), x+y\rangle\,| < 1/3\big\} &\leq \mathcal{Q}_{\mathcal{F}}\Big(\sum_{q=0}^{q_0}\xi_q(i), 1/3\Big)\\ &\leq \mathcal{Q}_{\mathcal{F}}\Big(\sum_{q\in A_i}\xi_q(i), 1/3\Big) \leq \frac{C_0\alpha}{\sqrt{(1-\alpha)|A_i|}} \leq \frac{C_0\alpha}{\sqrt{(1-\alpha)\gamma q_0}},\end{aligned}$$

where $\alpha = \max_{q\in A_i}\mathcal{Q}_{\mathcal{F}}(\xi_q(i), 1/3)$. Moreover, in the case $q_0 = 1$ we just have

$$\mathbb{P}_{\mathcal{F}}(\Omega_i) \leq \alpha = \mathcal{Q}(\xi_1(i), 1/3).$$

Thus it remains to estimate $\mathcal{Q}_{\mathcal{F}}(\xi_q(i), 1/3)$ for $q \in A_i$. Fix $q \in A_i$, so that $i \in I_q$. Recall that, by construction, the intersection of the support of $\mathbf{R}_i(M)$ with J_q is a singleton everywhere on $\mathcal{F}(\mathcal{I}, V)$. Denote the corresponding index by $j(q, M) = j(q, M, i)$. Then

$$\xi_q(i) = \xi_q(M, x+y, i) = \sum_{j\in J_q}\delta_{ij}(x_j + y_1) = x_{j(q,M)} + y_1,$$

and note that $|x_{j(q,M)}| > 2/3$ whenever $j(q,M) \in J_q^\ell$ and $x_{j(q,M)} = 0$ whenever $j(q,M) \in J_q^r$. Observe further that $\mathbb{P}_{\mathcal{F}}\big\{j(q,M) \in J_q^r\big\} = \mathbb{P}_{\mathcal{F}}\big\{j(q,M) \in J_q^\ell\big\} = 1/2$. Hence, we obtain

$$\mathcal{Q}_{\mathcal{F}}(\xi_q(i), 1/3) \leq 1/2 =: \alpha.$$

Combining the probability estimates starting with (37) and using that $\gamma = md/(72n)$, we obtain in the case $q_0 = m_1/m \geq C'$,

$$\begin{aligned}\mathbb{P}_{\mathcal{F}}(E(x,y)) &\leq \left(\frac{e}{81c_1}\right)^{9c_1md} \left(\frac{C_0}{\sqrt{2\gamma q_0}}\right)^{(1/9-9c_1)md}\\ &= \left(\frac{e}{81c_1}\right)^{9c_1md} \left(\frac{6C_0\sqrt{n}}{\sqrt{m_1 d}}\right)^{(1/9-9c_1)md} \leq \left(\frac{C_1 n}{m_1 d}\right)^{md/20},\end{aligned}$$

provided that c_1 is small enough and $C_1 = 36C_0^2$. Note that the bound is meaningful only if C' is large enough. In the case $q_0 = 1$ we have

$$\mathbb{P}_{\mathcal{F}}(E(x,y)) \le \left(\frac{e}{81c_1}\right)^{9c_1md} \left(\frac{1}{2}\right)^{(1/9-9c_1)md} \le \left(\frac{1}{2}\right)^{md/20},$$

provided that c_1 is small enough. This completes the proof. □

6.7 Proof of Theorem 6.1

We are ready to complete the proof. Denote

$$m = m_0 = n_{s_0+1} := \max(30\ell_0^{s_0-1}, \lfloor 1/(64p) \rfloor) \in [n/(64d), n/(2d)].$$

Lemma 6.9 implies that

$$\mathbb{P}\left(\left\{\exists\, x \in \mathcal{T}_0 \cup \mathcal{T}_1 \quad \text{such that} \quad \|Mx\| \le \frac{(64p)^\kappa}{192(pn)^2}\,\|x\|\right\}\right) \le n(1-p)^n + e^{-1.4np}.$$

We now turn to the remaining cases. Fix $j \in \{2,3\}$. Let

$$\mathcal{E}_j := \left\{M \in \mathcal{M}_n \,:\, \exists\, x \in \mathcal{T}_j \text{ such that } \|Mx\| \le \frac{\sqrt{c_1md}}{2\,b_j}\,\|x\|\right\},$$

where c_1 is the constant from Lemma 6.11, and $b_2 = 384(pn)^3/(64p)^\kappa$, $b_3 = 384C_\tau(pn)^{3.5}/(64p)^\kappa$.

Recall that $\mathcal{E}_{nrm}$ was defined in Proposition 3.14. For any matrix $M \in \mathcal{E}_j \cap \mathcal{E}_{nrm}$ there exists $x = x(M) \in \mathcal{T}_j$ satisfying

$$\|Mx\| \le \frac{\sqrt{c_1md}}{2\,b_j}\,\|x\|.$$

Normalize x so that $x^*_{n_{s_0+j-1}} = 1$, that is, $x \in \mathcal{T}_j'$. By Lemma 6.4 we have $\|x\| \le b_j$.

Let $\mathcal{N}_j = \mathcal{N}_j' + \mathcal{N}_j''$ be the net constructed in Lemma 6.7. Then there exist $u \in \mathcal{N}_j'$ with

$$u^*_{s_0+j-1} \ge 1 - 1/(C_\tau\sqrt{d}) > 2/3$$

and $u^*_\ell = 0$ for $\ell > n_{s_0+j}$, and $w \in \mathcal{N}_j'' \subset \operatorname{span}\{\mathbf{1}\}$, such that $|||x - (u+w)||| \le \sqrt{2n}/(C_\tau\sqrt{d})$. Applying Proposition 3.14 (where $\mathcal{E}_{nrm}$ was introduced), and using that C_τ is large enough, we obtain that for every matrix $M \in \mathcal{E}_j \cap \mathcal{E}_{nrm}$ there exist $u = u(M) \in \mathcal{N}_j'$ and $w = w(M) \in \mathcal{N}_j'' \subset \operatorname{span}\{\mathbf{1}\}$ with

$$\|M(u+w)\| \le \|Mx\| + \|M(x-u-w)\| \le \sqrt{c_1md}/2 + 200\sqrt{2n}/C_\tau \le \sqrt{c_1md}. \tag{38}$$

Using our choice of n_{s_0+1}, n_{s_0+2}, n_{s_0+3}, Lemma 6.7, and Lemma 6.11 twice — first with $m_1 = m_0 = n_{s_0+1}$, $m_2 = n_{s_0+2}$, then with $m_1 = n_{s_0+2} > m_0 = n_{s_0+1}$, $m_2 = n_{s_0+3}$ (see Remark 6.12), we obtain that for small enough r and large enough d the probability $\mathbb{P}\left(\mathcal{E}_2 \cap \mathcal{E}_{nrm} \cap \mathcal{E}_{card}\right)$ is bounded by

$$\exp\left(2n_{s_0+2}\ln d\right) 2^{-n_{s_0+1}d/20} \le \exp\left(-n_{s_0+1}d/30\right) \le \exp\left(-n/2000\right)$$

and that the probability $\mathbb{P}\left(\mathcal{E}_3 \cap \mathcal{E}_{nrm} \cap \mathcal{E}_{card}\right)$ is bounded by

$$\exp\left(2n_{s_0+3}\ln d\right)\left(\sqrt{d}/(2C)\right)^{-n_{s_0+1}d/20} \le \exp\left(-n\ln d/10000\right),$$

where $\mathcal{E}_{card}$ is the event introduced in Lemma 6.6 with $\ell = 2m$.

Combining all three cases we obtain that the desired bound holds for all $x \in \mathcal{T}$ with probability at most

$$2\exp\left(-n/2000\right) + \mathbb{P}\left(\mathcal{E}_{norm}^c\right) + \mathbb{P}\left(\mathcal{E}_{card}^c\right).$$

It remains to note that since np is large, by Lemma 3.6 (applied with $t = 30$) and by Lemma 6.6,

$$\mathbb{P}\left(\mathcal{E}_{nrm}^c\right) + \mathbb{P}\left(\mathcal{E}_{card}^c\right) \le 4e^{-225np} + 2\exp(-n/500) \le \exp(-10pn).$$

□

6.8 Proof of Theorem 6.3

Proof. Clearly, it is enough to show that $\Upsilon_n(r) \setminus (\mathcal{V}_n(r, \mathbf{g}, \delta, \rho) \cup \mathcal{T}) \subset \mathcal{R}$. Let $x \in \Upsilon_n(r) \setminus \mathcal{T}$ and set $\sigma := \sigma_x$. Note that $|x_{n_{s_0+2}}| \le C_\tau\sqrt{d}$, where s_0 was defined in (27). Denote $m_0 = \lfloor n/\ln^2 d \rfloor > 2n_{s_0+2}$.

Assume first that x does not satisfy (10). Then by Lemma 3.2, $x \in \mathcal{AC}(\rho)$. If $x^*_{m_0} \le \ln^2 d$ then denoting $k = m_0$, $A = [k, n]$, and using the definition of $\mathcal{AC}(\rho)$, we observe

$$\|x_{\sigma(A)}\| \ge \sqrt{(n - n_{s_0+3} - k)(1-\rho)} \ge \sqrt{n/2},$$

whence

$$\frac{\|x_{\sigma(A)}\|}{\|x_{\sigma(A)}\|_\infty} \ge \frac{\sqrt{n/2}}{\ln^2 d} \ge \frac{C_0}{\sqrt{p}}.$$

On the other hand, $x^*_{m_0} \le |x_{n_{s_0+2}}| \le C_\tau\sqrt{d}$, hence $\|x_{\sigma(A)}\| \le C_\tau\sqrt{dn}$. This implies that $x \in \mathcal{R}^1_k \subset \mathcal{R}$.

Now, if $x^*_{m_0} > \ln^2 d$ then denoting $k = n_{s_0+2}$, $A = [k, n]$, we get

$$\|x_{\sigma(A)}\|^2 \ge \sum_{i=n_{s_0+2}}^{m_0} (x_i^*)^2 \ge (m_0/2)\ln^4 d \ge (n/4)\ln^2 d,$$

whence

$$\frac{\|x_{\sigma(A)}\|}{\|x_{\sigma(A)}\|_\infty} \ge \frac{\sqrt{n}\ln d}{2C_\tau\sqrt{d}} \ge \frac{C_0}{\sqrt{p}}.$$

As in the previous case we have $\|x_{\sigma(A)}\| \le C_\tau\sqrt{dn}$, which implies that $x \in \mathcal{R}^1_k \subset \mathcal{R}$.

Next we assume that x does satisfy (10). Then, by the definition of the set $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ and our function $\mathbf{g}$, x does not satisfy the following condition:

$$\forall i \le \frac{1}{64p} : \;\; x_i^* \le \exp(\ln^2(2n/i)) \quad \text{and} \quad \forall \frac{1}{64p} < i \le n : \;\; x_i^* \le (2n/i)^{3/2}.$$

We fix the smallest value of $j \ge 1$ which breaks this condition and consider several cases. Note that since $x \in \Upsilon_n(r)$, we must have $j \le rn$.

Case 1. $2m_0 \le j \le rn$. In this case by the conditions and by minimality of j, we have $x^*_{m_0} \le (2n/m_0)^{3/2}$ and $x_j^* \ge (2n/j)^{3/2}$. Take $k = m_0$ and $A = [k, n]$. Then we have

$$\|x_{\sigma(A)}\| \ge \sqrt{j - m_0 + 1}\, x_j^* \ge \sqrt{j/2}\,(2n/j)^{3/2} \ge \sqrt{rn/2}\,(2/r)^{3/2} = 2\sqrt{n}/r,$$

hence

$$\frac{\|x_{\sigma(A)}\|}{\|x_{\sigma(A)}\|_\infty} \ge \left(\frac{2}{r}\right)\frac{\sqrt{n}}{(2n/m_0)^{3/2}} \ge \left(\frac{2}{r}\right)\frac{\sqrt{n}}{(2\ln d)^3} \ge \frac{C_0}{\sqrt{p}}.$$

As above we have $\|x_{\sigma(A)}\| \le C_\tau\sqrt{dn}$, which implies that $x \in \mathcal{R}_k^2 \subset \mathcal{R}$.

Case 2. $16C_0^2 n/d \le j \le 2m_0$. Take $k = \lceil j/2 \rceil$ and $A = [k, n]$. Then we have $x_k^* \le (2n/k)^{3/2} \le (4n/j)^{3/2}$, $x_j \ge (2n/j)^{3/2}$, and

$$\|x_{\sigma(A)}\| \ge \sqrt{j-k+1}\, x_j^* \ge \sqrt{j/2}\,(2n/j)^{3/2} \ge (2/r)\sqrt{n}.$$

Therefore,

$$\frac{\|x_{\sigma(A)}\|}{\|x_{\sigma(A)}\|_\infty} \ge \left(\frac{j}{2}\right)^{1/2} \frac{(2n/j)^{3/2}}{(4n/j)^{3/2}} \ge \frac{C_0}{\sqrt{p}}.$$

Since $x \notin \mathcal{T}$, we observe $x_k^* \le C_\tau^2 d$, hence $\|x_{\sigma(A)}\| \le C_\tau^2 d\sqrt{n}$ and $x \in \mathcal{R}_k^2 \subset \mathcal{R}$.

In the rest of the proof we show that we must necessarily have $j \ge 16C_0^2 n/d$.

Case 3. $n_{s_0+1} \le j < C_1 n/d$, where $C_1 = 16C_0^2$. Using that $x \notin \mathcal{T}$, in this case we have

$$C_\tau^2 d \ge x_j^* \ge \left(\frac{2n}{j}\right)^{3/2} \ge \left(\frac{2d}{C_1}\right)^{3/2},$$

which is impossible for large enough d.

Case 4. $n_{s_0} \le j < n_{s_0+1}$. Using that $x \notin \mathcal{T}$ and that $n_{s_0+1} = \lfloor 1/(64p) \rfloor = \lfloor n/(64d) \rfloor$, in this case we have

$$(6d)C_\tau^2 d \ge x_j^* \ge \exp(\ln^2(2n/j)) \ge \exp(\ln^2(2n/n_{s_0+1})) \ge \exp(\ln^2(128d))$$

which is impossible for large enough d.

Case 5. $n_k \le j < n_{k+1}$ for some $1 \le k \le s_0 - 1$. Recall that $n_k = 30\ell_0^{k-1}$ and recall also that if $s_0 > 1$ (as in this case) then $p \le c\sqrt{n \ln n}$. Using that $x \notin \mathcal{T}$, in this case we have

$$(C_\tau^2 d)(6d)^{s_0-k+1} \ge x_j^* \ge \exp(\ln^2(2n/j)) \ge \exp(\ln^2(2n/(30\ell_0^k))),$$

hence

$$(C_\tau^2 d)(6d)^{s_0+1} \ge (6d)^k \exp(\ln^2(2n/(30\ell_0^k))). \tag{39}$$

Considering the function $f(k) := k\ln(6d) + \ln^2(2n/(30\ell_0^k)$, we observe that its derivative is linear in k, therefore f attains its maximum either at $k = 1$ or at $k = s_0 - 1$. Thus, to show that (39) is impossible, it is enough to consider $k = 1, s_0 - 1$ only. Let $k = 1$. By (29), $(6d)^{s_0} \le (6d)\,1/(64p)^\kappa$, where $\kappa = \frac{\ln(6d)}{\ln \ell_0}$. Therefore, the logarithm of the left hand side of (39) is

$$\ln((C_\tau^2 d)(6d)^{s_0+1}) \le 4\ln d + \frac{\ln(6d)}{\ln \ell_0}\,\ln(1/64p). \tag{40}$$

On the other hand, $n/\ell_0 \ge (4\ln(1/p))/p$, therefore the logarithm of the left hand side of (39) is larger than $\ln^2(\ln(1/p)/(4p))$. Thus, it is enough to check that

$$(1/2)\ln^2(\ln(1/p)/(4p)) \ge 4\ln d \quad \text{and} \quad (1/2)\ln^2(\ln(1/p)/(4p))\ln \ell_0 \ge \ln(6d)\,\ln(1/64p).$$

Both inequalities follows since $p \le c\sqrt{n\ln n}$, $d = pn$, d and n are large enough, and since $\ell_0 \ge 25$. Next assume that $k = s_0 - 1$. Note that in this case $\ell_0^k \le n/(64d)$. Thus, to disprove (39), it is enough to show that

$$\ln^2(64d/15) \ge \ln(36C_\tau^2 d^3),$$

which clearly holds for large enough d.

Case 6. $2 \le j < 30$. In this case we have

$$(C_\tau^2 d)(6d)^{s_0+1} \ge x_j^* \ge \exp(\ln^2(2n/j)) \ge \exp(\ln^2(2n/30)),$$

By (40) this implies

$$4\ln d + \frac{\ln(6d)}{\ln \ell_0}\,\ln(1/64p) \geq \ln^2(2n/30),$$

which is impossible.

Case 7. $j = 1$. In this case we have $(C_\tau^2 d)(6d)^{s_0+2} \geq x_1^* \geq \exp(\ln^2(2n))$ and we proceed as in Case 6. □

7 Proof of the main theorem

In this section, we combine the results of Sections 4, 5, and 6, as well as Subsection 3.2 to prove the main theorems, Theorems 1.2 and the following improvement for the case of constant p.

Theorem 7.1. *There exists an absolute positive constant c with the following property. Let $q \in (0, c)$ be a parameter (independent of n). Then there exist C_q and $n_q \geq 1$ (both depend only on q), such that for every $n \geq n_q$ and every $p \in (q, c)$ a Bernoulli(p) $n \times n$ random matrix M_n satisfies*

$$\mathbb{P}\big\{M_n \text{ is singular}\big\} = (2 + o_n(1))n\,(1-p)^n,$$

and, moreover, for every $t > 0$,

$$\mathbb{P}\big\{s_{\min}(M_n) \leq C_q\, n^{-2.5}\, t\big\} \leq t + (1 + o_n(1))\mathbb{P}\big\{M_n \text{ is singular}\big\} = t + (2 + o_n(1))n\,(1-p)^n.$$

At this stage, the scheme of the proof to a large extent follows the approach of Rudelson and Vershynin developed in [44]. However, a crucial part of their argument — "invertibility via distance" (see [44, Lemma 3.5]) — will be reworked in order to keep sharp probability estimates for the matrix singularity and to be able to bind this part of the argument with the previous sections, where we essentially condition on row- and column-sums of our matrix.

We start by restating main results of Sections 5 and 6 using the vector class $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ defined by (1), together with Lemma 3.1.

Corollary 7.2. *There are universal constants $C \geq 1$, $\delta, \rho \in (0, 1)$ and $r \in (0, 1)$ with the following property. Let M_n be a random matrix satisfying $(\mathbf{A})$ with C and let the growth function $\mathbf{g}$ be given by (30). Then*

$$\mathbb{P}\Big\{\|M_n x\| \leq a_n^{-1}\|x\| \quad \text{for some} \quad x \notin \bigcup_{\lambda \geq 0} \big(\lambda\, \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)\big)\Big\} = (1 + o_n(1))n\,(1-p)^n, \tag{41}$$

where

$$a_n = \frac{(pn)^2}{c(64p)^\kappa}\,\max\big(1, p^{1.5}n\big) \qquad \text{and} \qquad \kappa = \kappa(p) := (\ln(6pn))/\ln\left\lfloor \frac{pn}{4\ln(1/p)} \right\rfloor.$$

Further, Theorems 5.1, 5.2 and Lemma 3.1 are combined as follows.

Corollary 7.3. *There are universal positive constants c, C with the following property. Let $q \in (0, c)$ be a parameter. Then there exist $n_0 = n_0(q) \geq 1$, $r = r(q), \rho = \rho(q) \in (0, 1)$ such that for $n \geq n_0$, $p \in (q, c)$, $\delta = r/3$, $\mathbf{g}(t) = (2t)^{3/2}$, a random Bernoulli($p$) $n \times n$ matrix M_n satisfies (41) with $a_n = C\sqrt{n\ln(e/p)}$.*

Below is our version of "invertibility via distance," which deals with *pairs* of columns.

Lemma 7.4 (Invertibility via distance)**.** *Let $r, \delta, \rho \in (0,1)$, and let $\mathbf{g}$ be a growth function. Further, let $n \geq 6/r$ and let A be an $n \times n$ random matrix. Then for any $t > 0$ we have*

$$\begin{aligned}\mathbb{P}\big\{\|Ax\| \leq t\,\|x\| \quad \text{for some} \quad x \in \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)\big\} \\ \leq \frac{2}{(rn)^2} \sum_{i \neq j} \mathbb{P}\big\{\operatorname{dist}(H_i(A), \mathbf{C}_i(A)) \leq t\, b_n \quad \text{and} \quad \operatorname{dist}(H_j(A), \mathbf{C}_j(A)) \leq t\, b_n\big\},\end{aligned}$$

where the sum is taken over all ordered pairs (i,j) with $i \neq j$ and $b_n = \sum_{i=1}^n \mathbf{g}(i)$.

Proof. For every $i \neq j$, denote by $\mathbf{1}_{ij}$ the indicator of the event

$$\mathcal{E}_{ij} := \big\{\operatorname{dist}(H_i(A), \mathbf{C}_i(A)) \leq t\, b_n \quad \text{and} \quad \operatorname{dist}(H_j(A), \mathbf{C}_j(A)) \leq t\, b_n\big\}.$$

The condition

$$\|Ax\| \leq t\,\|x\|$$

for some $x \in \mathcal{V}_n = \mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$ implies that for every $i \leq n$,

$$|x_i|\operatorname{dist}(H_i(A), \mathbf{C}_i(A)) \leq \|Ax\| \leq t\, b_n,$$

where the last inequality follows from the definition of $\mathcal{V}_n$. Since $x^*_{\lfloor rn \rfloor} = 1$, we get that everywhere on the event $\{\|Ax\| \leq t\,\|x\| \text{ for some } x \in \mathcal{V}_n\}$ there are at least $\lfloor rn \rfloor (\lfloor rn \rfloor - 1) \geq (rn)^2/2$ ordered pairs of indices (i,j) such that for each pair the event $\mathcal{E}_{ij}$ occurs. Rewriting this assertion in terms of indicators, we observe

$$\{\|Ax\| \leq t\,\|x\| \text{ for some } x \in \mathcal{V}_n\} \subset \Big\{\sum_{i \neq j} \mathbf{1}_{ij} \geq (rn)^2/2\Big\}.$$

Applying Markov's inequality in order to estimate probability of the event on the right hand side, we obtain the desired result. □

Proof of Theorems 1.2 and 7.1. The proofs of both theorems are almost the same, the only difference is that Theorem 1.2 uses Corollary 7.3 while Theorem 1.2 uses Corollary 7.2. Let parameters $\delta, \rho, r, \mathbf{g}, a_n$ be taken from Corollary 7.2 or from Corollary 7.3 correspondingly. We always write $\mathcal{V}_n$ for $\mathcal{V}_n(r, \mathbf{g}, \delta, \rho)$. Let $b_n = \sum_{i=1}^n \mathbf{g}(i)$. Without loss of generality, we can assume that $n \geq 6/r$. Fix $t \in (0,1]$, and denote by $\mathcal{E}$ the complement of the event

$$\Big\{\|M_n x\| \leq a_n^{-1}\|x\| \;\text{ or }\; \|M_n^\top x\| \leq a_n^{-1}\|x\| \quad \text{for some} \quad x \notin \bigcup_{\lambda \geq 0} (\lambda\, \mathcal{V}_n)\Big\}.$$

For $i = 1, 2$ denote

$$\mathcal{E}_i := \big\{\operatorname{dist}(H_i(M_n), \mathbf{C}_i(M_n)) \leq a_n^{-1}\, t\big\}.$$

Applying Corollary 7.2 (or Corollary 7.3), Lemma 7.4 and the invariance of the conditional distribution of M_n given $\mathcal{E}$ under permutation of columns, we obtain

$$\begin{aligned}&\mathbb{P}\big\{s_{\min}(M_n) \leq (a_n b_n)^{-1} t\big\} \\ &\quad \leq (2 + o_n(1)) n\,(1-p)^n + \mathbb{P}\big(\big\{\|M_n x\| \leq (a_n b_n)^{-1}\, t \|x\| \quad \text{for some} \quad x \in \mathcal{V}_n\big\} \cap \mathcal{E}\big) \\ &\quad \leq (2 + o_n(1)) n\,(1-p)^n + \frac{2}{r^2}\,\mathbb{P}\big(\mathcal{E} \cap \mathcal{E}_1 \cap \mathcal{E}_2\big).\end{aligned}$$

At the next step, we consider events

$$\Omega_i := \big\{|\operatorname{supp} \mathbf{C}_i(M_n)| \in [pn/8, 8pn]\big\},\ i = 1, 2, \quad \text{and} \quad \Omega := \Omega_1 \cup \Omega_2.$$

Since columns of M are independent and consist of i.i.d. Bernoulli(p) variables, applying Lemma 3.4, we observe

$$\mathbb{P}\big(\Omega^c\big) = \mathbb{P}\big(\Omega_1^c\big)\mathbb{P}\big(\Omega_2^c\big) \le (1-p)^n.$$

Therefore, in view of equidistribution of the first two columns, we get

$$\mathbb{P}\big(\mathcal{E}\cap\mathcal{E}_1\cap\mathcal{E}_2\big) \le (1-p)^n + \mathbb{P}\big(\mathcal{E}\cap\mathcal{E}_1\cap\mathcal{E}_2\cap\Omega\big) \le (1-p)^n + 2\mathbb{P}\big(\mathcal{E}\cap\mathcal{E}_1\cap\Omega_1\big).$$

Denote by $\mathbf{Y}$ a random unit vector orthogonal to (and measurable with respect to) $H_1(M_n)$. Note that on the event $\mathcal{E}_1$ the vector $\mathbf{Y}$ satisfies

$$|\langle \mathbf{Y}, \mathbf{C}_1(M_n)\rangle| = \|M_n^\top \mathbf{Y}\| \le a_n^{-1}\, t\, \|\mathbf{Y}\|,$$

which implies that on the event $\mathcal{E}\cap\mathcal{E}_1$ we also have $\mathbf{Y}^*_{\lfloor rn\rfloor} \ne 0$, and $\mathbf{Z} := \mathbf{Y}/\mathbf{Y}^*_{\lfloor rn\rfloor} \in \mathcal{V}_n$. By the definition of $\mathcal{V}_n$, we have $\|\mathbf{Z}\| \le b_n$, therefore,

$$P_0 := \mathbb{P}\big(\mathcal{E}\cap\mathcal{E}_1\cap\Omega_1\big) \le \mathbb{P}\big(\Omega_1 \cap \big\{\exists\, Z \in H_1(M_n)^\perp \cap \mathcal{V}_n :\ |\langle Z, \mathbf{C}_1(M_n)\rangle| \le a_n^{-1}\, b_n\, t\big\}\big).$$

On the other hand, applying Theorem 2.2 with $R = 2$, we get that for some constants $K_1 \ge 1$ and $K_2 \ge 4$, with probability at least $1 - \exp(-2pn)$,

$$H_1(M_n)^\perp \cap \mathcal{V}_n \subset \big\{x \in \Upsilon_n(r) :\ \mathbf{UD}_n(x, m, K_1, K_2) \ge \exp(2pn) \quad \text{for any} \quad m \in [pn/8, 8pn]\big\}.$$

Combining the last two assertions and applying Theorem 2.1, we observe

$$\begin{aligned}
P_0 &\le \exp(-2pn) + \mathbb{P}\big(\Omega_1 \cap \big\{\exists\, Z \in H_1(M_n)^\perp \cap \mathcal{V}_n :\ |\langle Z, \mathbf{C}_1(M_n)\rangle| \le a_n^{-1}\, b_n\, t \ \text{ and} \\
&\qquad\qquad \forall m \in [pn/8, 8pn] :\ \mathbf{UD}_n(Z, m, K_1, K_2) \ge \exp(2pn)\big\}\big) \\
&\le \exp(-2pn) + \sup_{\substack{m\in[pn/8,8pn],\, y\in\Upsilon_n(r),\\ \mathbf{UD}_n(y,m,K_1,K_2)\ge\exp(2pn)}} \mathbb{P}\big\{|\langle y, \mathbf{C}_1(M_n)\rangle| \le a_n^{-1} b_n\, t \ \big|\ |\operatorname{supp}\mathbf{C}_1(M_n)| = m\big\} \\
&\le (1 + C_{2.1})\exp(-2pn) + \frac{C_{2.1} b_n}{a_n\sqrt{pn/8}}\, t.
\end{aligned}$$

Thus

$$\mathbb{P}\big\{s_{\min}(M_n) \le (a_n b_n)^{-1} t\big\} \le (2 + o_n(1)) n\,(1-p)^n + \frac{8C_{2.1} b_n}{r^2\, a_n\sqrt{pn}}\, t.$$

By rescaling t we obtain

$$\mathbb{P}\Big\{s_{\min}(M_n) \le \frac{r^2\sqrt{pn}}{(8C_{2.1}b_n^2)}\, t\Big\} \le (2 + o_n(1)) n\,(1-p)^n + t, \quad 0 \le t \le \frac{8C_{2.1} b_n}{r^2\, a_n\sqrt{pn}}.$$

In the case of constant p (applying Corollary 7.3) we have $a_n = C\sqrt{n\ln(e/p)}$ and $b_n \le 2\sqrt{3}n^{3/2}$, and we get the small ball probability estimate of Theorem 7.1.

In the general case (applying Corollary 7.2) we have $a_n = \frac{(pn)^2}{c(64p)^\kappa}\, \max(1, p^{1.5}n)$ and $b_n \le \exp(1.5\ln^2(2n))$. Therefore,

$$\frac{r^2\sqrt{pn}}{(8C_{2.1}b_n^2)} \ge \exp(-3\ln^2(2n))$$

for large enough n, and the $s_{\min}$ estimate follows.

In both cases the upper bound on t, $\frac{8C_{2.1}b_n}{r^2\, a_n\sqrt{pn}}$, is greater than 1, so we may omit it.

Finally, applying the argument of Subsection 3.2, we get the matching lower bound for the singularity probability. This completes the proof. □

8 Further questions

The result of this paper leaves open the problem of estimating the singularity probability for Bernoulli matrices in two regimes: when np_n is logarithmic in n and when p_n is larger than the constant C^{-1} from Theorem 1.2.

For the first regime, we recall that the singularity probability of M_n, with np_n in a (small) neighborhood of $\ln n$, was determined up to the $1 + o_n(1)$ multiple in the work of Basak–Rudelson [5].

Problem 8.1 (A bridge: Theorem 1.2 to Basak–Rudelson). *Let p_n satisfy*

$$1 \le \liminf np_n / \ln n \le \limsup np_n / \ln n < \infty,$$

and for each n let M_n be the $n \times n$ matrix with i.i.d. Bernoulli(p_n) entries. Show that

$$\mathbb{P}\big\{M_n \text{ is singular}\big\} = (1 + o_n(1))\mathbb{P}\big\{M_n \text{ has a zero row or a zero column}\big\}.$$

A few months after our paper was posted on arXiv, a positive solution to the above problem was given by Huang in [15], thus completing the research program of sharp singularity probability estimates for sparse Bernoulli matrices.

The second problem is the singularity of random Bernoulli matrices with large values of p_n.

Problem 8.2 (Optimal singularity probability for dense Bernoulli matrices below the 1/2 threshold). *Let the sequence p_n satisfy*

$$0 < \liminf p_n \le \limsup p_n < 1/2.$$

Show that

$$\mathbb{P}\big\{M_n \text{ is singular}\big\} = (1 + o_n(1))\mathbb{P}\big\{M_n \text{ has a zero row or a zero column}\big\} = (2 + o_n(1))n\,(1 - p_n)^n.$$

As with the first problem, a few months after our paper was posted on arXiv, a positive solution was obtained by Jain, Sah, and Sawhney in [17, 18] (their result in fact covers a more general model of randomness). Conditioning on the sums of Bernoulli random vectors exploited in the present paper, is also one of crucial elements of [17].

Acknowledgments

K.T. was partially supported by the Sloan Research Fellowship.

References

[1] A.S. Bandeira, R. van Handel, Sharp nonasymptotic bounds on the norm of random matrices with independent entries. Ann. Probab. 44 (2016), 2479–2506.

[2] A. Basak, N. Cook, and O. Zeitouni, Circular law for the sum of random permutation matrices, Electronic Journal of Probability, **23** (2018), Paper No. 33, 51 pp.

[3] A. Basak and M. Rudelson, Invertibility of sparse non-Hermitian matrices, Adv. Math. **310** (2017), 426–483. MR3620692

[4] A. Basak and M. Rudelson, The circular law for sparse non-Hermitian matrices, Ann. Probab. **47** (2019), 2359–2416. MR3980923

[5] A. Basak and M. Rudelson, Sharp transition of the invertibility of the adjacency matrices of random graphs, Probab. Theory Relat. Fields, to appear.

[6] S. Boucheron, G. Lugosi, P. Massart, Concentration inequalities. A nonasymptotic theory of independence. With a foreword by Michel Ledoux. Oxford University Press, Oxford, 2013.

[7] J. Bourgain, V. H. Vu and P. M. Wood, On the singularity probability of discrete random matrices, J. Funct. Anal. **258** (2010), no. 2, 559–603. MR2557947

[8] D. Chafaï, O. Guédon, G. Lecué, A. Pajor, Interactions between compressed sensing random matrices and high dimensional geometry, Panoramas et Synthèses [Panoramas and Syntheses], **37**. Soc. Math. de France, Paris, 2012.

[9] N. A. Cook, On the singularity of adjacency matrices for random regular digraphs, Probab. Theory Related Fields **167** (2017), no. 1-2, 143–200. MR3602844

[10] N. Cook, The circular law for random regular digraphs, Ann. Inst. Henri Poincare Probab. Stat., **55** (2019), 2111–2167. MR4029149

[11] L. Devroye; G. Lugosi, Combinatorial methods in density estimation. Springer Series in Statistics. Springer-Verlag, New York, 2001.

[12] P. Erdös, On a lemma of Littlewood and Offord, Bull. Amer. Math. Soc. **51** (1945), 898–902. MR0014608

[13] C. G. Esseen, On the Kolmogorov-Rogozin inequality for the concentration function, Z. Wahrsch. Verw. Gebiete **5** (1966), 210–216. MR0205297

[14] F. Götze and A. Tikhomirov, The circular law for random matrices, Ann. Probab. **38** (2010), no. 4, 1444–1491. MR2663633

[15] H. Huang, Rank of Sparse Bernoulli Matrices, preprint, arXiv:2009.13726, 2020.

[16] J. Huang, Invertibility of adjacency matrices for random d-regular graphs, preprint, arXiv:1807.06465, 2018.

[17] V. Jain, A. Sah, M. Sawhney, Singularity of discrete random matrices I, preprint, arXiv:2010.06553, 2020.

[18] V. Jain, A. Sah, M. Sawhney, Singularity of discrete random matrices II, preprint, arXiv:2010.06554, 2020.

[19] J. Kahn, J. Komlós and E. Szemerédi, On the probability that a random ± 1-matrix is singular, J. Amer. Math. Soc. **8** (1995), no. 1, 223–240. MR1260107

[20] H. Kesten, A sharper form of the Doeblin-Lévy-Kolmogorov-Rogozin inequality for concentration functions, Math. Scand. **25** (1969), 133–144. MR0258095

[21] J. Komlós, On the determinant of (0, 1) matrices, Studia Sci. Math. Hungar **2** (1967), 7–21. MR0221962

[22] B. Landon, P. Sosoe, H. Yau, Fixed energy universality of Dyson Brownian motion, Adv. Math. **346** (2019), 1137–1332.

[23] M. Ledoux, The concentration of measure phenomenon. Mathematical Surveys and Monographs, **89**. American Mathematical Society, Providence, RI, 2001.

[24] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation. III, Rec. Math. [Mat. Sbornik] N.S. **12(54)** (1943), 277–286. MR0009656

[25] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, Anti-concentration property for random digraphs and invertibility of their adjacency matrices, C. R. Math. Acad. Sci. Paris **354** (2016), no. 2, 121–124. MR3456885

[26] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, Adjacency matrices of random digraphs: singularity and anti-concentration, J. Math. Anal. Appl. **445** (2017), no. 2, 1447–1491. MR3545253

[27] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, The smallest singular value of a shifted d-regular random square matrix, Probab. Theory Related Fields, **173** (2019), 1301–1347.

[28] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, The circular law for sparse random regular digraphs, J. European Math. Soc., **23** (2021), 467–501.

[29] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, The rank of random regular digraphs of constant degree, J. of Complexity, **48** (2018), 103–110.

[30] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann and P. Youssef, Structure of eigenvectors of random regular digraphs, Trans. Amer. Math. Soc., **371** (2019), 8097–8172.

[31] A. E. Litvak, A. Pajor, M. Rudelson and N. Tomczak-Jaegermann, Smallest singular value of random matrices and geometry of random polytopes, Adv. Math. **195** (2005), no. 2, 491–523. MR2146352

[32] A. E. Litvak and O. Rivasplata, Smallest singular value of sparse random matrices, Studia Math. **212** (2012), no. 3, 195–218. MR3009072

[33] G. V. Livshyts, The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding, Journal d'Analyse Mathematique, to appear, arXiv:1811.07038.

[34] G. V. Livshyts, K. Tikhomirov, R. Vershynin, The smallest singular value of inhomogeneous square random matrices, Ann. Probab., to appear, arXiv:1909.04219

[35] A. Lytova, K. Tikhomirov, On delocalization of eigenvectors of random non-Hermitian matrices, Probab. Theor. Rel. Fields, **177** (2020), 465–524. MR4095020

[36] K. Luh, S. Meehan, H.H. Nguyen, Some new results in random matrices over finite fields, J. of London Math. Soc., to appear.

[37] K. Luh, S. O'Rourke, Eigenvector Delocalization for Non-Hermitian Random Matrices and Applications, Random Structures Algorithms **57** (2020), 169–210. MR4120597

[38] A. Mészáros, The distribution of sandpile groups of random regular graphs, Trans. Amer. Math. Soc. **373** (2020), 6529–6594. MR4155185

[39] H.H. Nguyen and M.M. Wood, Cokernels of adjacency matrices of random r-regular graphs, preprint, arXiv: 1806.10068, 2018.

[40] E. Rebrova, K. Tikhomirov, Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries, Israel J. Math., **227** (2018), 507–544.

[41] B. A. Rogozin, On the increase of dispersion of sums of independent random variables, Teor. Verojatnost. i Primenen **6** (1961), 106–108. MR0131894

[42] M. Rudelson, Invertibility of random matrices: norm of the inverse, Ann. of Math. **168** (2008), 575–600.

[43] M. Rudelson, Recent developments in non-asymptotic theory of random matrices, in *Modern aspects of random matrix theory*, 83–120, Proc. Sympos. Appl. Math., 72, Amer. Math. Soc., Providence, RI. MR3288229

[44] M. Rudelson and R. Vershynin, The Littlewood–Offord problem and invertibility of random matrices, Adv. Math. **218** (2008), no. 2, 600–633. MR2407948

[45] M. Rudelson and R. Vershynin, Smallest singular value of a random rectangular matrix, Comm. Pure Appl. Math. **62** (2009), no. 12, 1707–1739. MR2569075

[46] M. Rudelson and R. Vershynin, No-gaps delocalization for general random matrices, Geom. Funct. Anal. **26** (2016), no. 6, 1716–1776. MR3579707

[47] T. Tao and V. Vu, On random ± 1 matrices: singularity and determinant, Random Structures Algorithms **28** (2006), no. 1, 1–23. MR2187480

[48] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, J. Amer. Math. Soc. **20** (2007), no. 3, 603–628. MR2291914

[49] T. Tao and V. Vu, Random matrices: the circular law, Commun. Contemp. Math. **10** (2008), 261–307. MR2409368

[50] T. Tao and V. H. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices, Ann. of Math. **169** (2009), 595–632. MR2480613

[51] K. Tikhomirov, Singularity of random Bernoulli matrices, Annals of Math., **191** (2020), 593–634.

Alexander E. Litvak
Dept. of Math. and Stat. Sciences,
University of Alberta,
Edmonton, AB, Canada, T6G 2G1.
e-mail: aelitvak@gmail.com

Konstantin E. Tikhomirov
School of Math., GeorgiaTech,
686 Cherry street,
Atlanta, GA 30332, USA.
e-mail: ktikhomirov6@gatech.edu