ON THE SINGULARITY OF ADJACENCY MATRICES
FOR RANDOM REGULAR DIGRAPHS

NICHOLAS A. COOK

Abstract. We prove that the (non-symmetric) adjacency matrix of a uniform random d-
regular directed graph on n vertices is asymptotically almost surely invertible, assuming
\( \min(d, n - d) = \omega(\log^2 n) \). This improves an earlier result of the author from [12], which
treated the dense case \( \min(d, n - d) = \Omega(n) \). The proof makes use of a coupling of random
regular digraphs formed by “shuffling” the neighborhood of a pair of vertices, as well as
concentration results for the distribution of edges, proved in [13]. We also apply our general
approach to prove a.a.s. invertibility of Hadamard products \( M \odot \Xi \), where \( \Xi \) is a matrix
of iid uniform ±1 signs, and \( M \) is a 0/1 matrix whose associated digraph satisfies certain
“expansion” properties.

Contents

1. Introduction 2
   1.1. Background 2
   1.2. Main results and conjectures 5
   1.3. The general strategy 9
   1.4. Organization of the paper 11
   1.5. Notation 11
   1.6. Acknowledgements 12

2. Ideas of the proof 12
   2.1. The strategy of Komlós 13
   2.2. Structured and unstructured null vectors 15
   2.3. Injecting a random walk 15
   2.4. Dealing with sparsity 18

3. Preliminaries 18
   3.1. The shuffling coupling 18
   3.2. Discrepancy properties 20
   3.3. Concentration of measure 23
1. Introduction

For $n \geq 1$ and $d \in [n]$, let $\mathcal{M}_{n,d}$ be the set of $n \times n$ matrices with entries taking values in $\{0, 1\}$ satisfying the constraint that all row and column sums are equal to $d$. (For instance, we have that $\mathcal{M}_{n,1}$ is the set of $n \times n$ permutation matrices.) One may interpret the elements of $\mathcal{M}_{n,d}$ as the adjacency matrices of $d$-regular digraphs – that is, directed graphs on $n$ labeled vertices with each vertex having $d$ in-neighbors and $d$ out-neighbors (allowing self-loops). We can also identify $\mathcal{M}_{n,d}$ with the set of $d$-regular bipartite graphs on $n + n$ vertices in the obvious way.

We denote by $M$ a uniform random element of $\mathcal{M}_{n,d}$, and refer to $M$ as an “rrd matrix” (for “random regular digraph”). Our objective in this paper is to determine whether $M$ is invertible with high probability when $n$ is large and for some range of the parameter $d$. Before stating our main result, we give an overview of related work on other random matrix models.

1.1. Background. Much work on the singularity of random matrices has focused on iid sign matrices $\Xi$, which have iid uniform $\pm 1$ entries. It is already a non-trivial problem to prove that $\Xi$ is invertible with probability tending to 1; this was first accomplished by Komlós in the works [19, 20] from the 1960s. His proof was later refined to give the following quantitative bound:

**Theorem 1.1** (Komlós [21]). Let $\Xi$ be an $n \times n$ matrix of iid uniform signs. Then

$$\Pr(\det(\Xi) = 0) = O(n^{-1/2}).$$  \hspace{1cm} (1.1)
The asymptotic notation in (1.1) and throughout this paper is with respect to the large \( n \) limit – see Section 1.5 for our notational conventions.

A key ingredient in the proof of Theorem 1.1 was a bound of Littlewood-Offord type due to Erdős (Theorem 2.1 below) from the seemingly unrelated field of additive combinatorics. This inspired a sequence of works which improved Komlós’ bound to exponential bounds

\[
\mathbb{P} (\det(\Xi) = 0) \ll c^n
\]

for \( c < 1 \) by making heavier use of additive combinatorics machinery. Specifically, the base \( c = .999 \) was obtained by Kahn, Komlós and Szemerédi in [18], and was lowered to \( c = 3/4 + o(1) \) by Tao and Vu [40], and to \( c = 1/\sqrt{2} + o(1) \) by Bourgain, Vu and Wood [7]. The latter two works relied on the inverse Littlewood-Offord theory developed in [40]. These bounds still fall short the conjectured asymptotic \( c = 1/2 + o(1) \). More precisely, we have the following folklore conjecture

\[
\mathbb{P} (\det(\Xi) = 0) = n^2(1 + o(1))2^{1-n}
\]

which has been stated in [18, 20]. The lower bound in (1.3) is easily proved by considering the event that \( \Xi \) has a pair of rows or columns that are parallel.

One source of motivation for controlling the singularity probability is its relation to the problem of proving limit laws for the distribution of eigenvalues. Define the (rescaled) empirical spectral distribution of \( \Xi \) to be the random probability measure

\[
\mu_{\frac{1}{\sqrt{n}}\Xi} := \sum_{i=1}^{n} \delta_{\lambda_i(\frac{1}{\sqrt{n}}\Xi)},
\]

distributed uniformly over the eigenvalues of the normalized matrix \( \frac{1}{\sqrt{n}}\Xi \). In [42] Tao and Vu proved the circular law for \( \Xi \), which states that almost surely, as \( n \to \infty \), \( \mu_{\frac{1}{\sqrt{n}}\Xi} \) converges weakly to the uniform measure on the unit disc in \( \mathbb{C} \). Tao and Vu actually proved a universality principle, which implies that the circular law holds for any matrix with iid entries having mean 0 and variance 1.

The main technical hurdle in the proof of the circular law was to obtain good lower bounds on the least singular value \( \sigma_n(\Xi) \) holding with high probability. (Actually, it was necessary to do this for scalar shifts \( \Xi - zI \) of the matrix \( \Xi \).) Proving lower bounds on \( \sigma_n(\Xi) \) is an extension of the singularity probability problem – indeed, the latter is to bound \( \mathbb{P}(\sigma_n(\Xi) = 0) \). Polynomial lower bounds on the least singular value of general iid matrices were first obtained by Rudelson in [31], and were subsequently improved by Tao and Vu [41] and Rudelson and Vershynin [33].

See [6] for a survey of works on the circular law for iid matrices and other matrix models. The assumption of joint independence of the entries has been relaxed in some directions. For instance, the circular law is established for matrices with log-concave isotropic unconditional laws by Adamczak and Chafaï in [2]. Together with Wolff in [3], the same authors extend the circular law to matrices with exchangeable entries satisfying some moment bounds. (Note that while the rows and columns of the rrd matrix \( M \) are exchangeable, the individual entries are not.)

Apart from iid matrices, a lot of activity has concentrated on random matrix models with constraints on row and column sums. In [5], Bordenave, Caputo and Chafaï proved the circular law for random Markov matrices, obtained by normalizing the rows of an iid matrix
with continuous entry distributions. On the discrete side, in [28] Nguyen proved that a uniform random 0/1 matrix constrained to have all row-sums equal to $n/2$ (say $n$ is even) is invertible with probability $1 - O_C(n^{-C})$. Nguyen and Vu subsequently proved the circular law for a more general class of random discrete matrices with constant row sums [30].

The approach in [28] and [30] was to use a conditioning trick, which we will now briefly outline. As in [28], assume $n$ is even, and let $Q$ be a uniform random $0/1$ matrix with all row sums equal to $n/2$. Suppose we want to control the probability that some property $P$ holds for the first row $R_1$ of $Q$. Now for $Y_1 \in \{0, 1\}^n$ a uniform random 0/1 vector, let $E$ be the event that the components of $Y_1$ sum to $n/2$. One can easily show that

$$P(E) \gg n^{-1/2}. \quad (1.4)$$

Moreover, we have that conditional on $E$, $Y_1$ is identically distributed to $R_1$:

$$Y_1 \mid E \overset{d}{=} R_1. \quad (1.5)$$

It follows that we can bound

$$P(P \text{ holds for } R_1) = P(P \text{ holds for } Y_1 \mid E) \leq \frac{P(P \text{ holds for } Y_1)}{P(E)} \ll n^{1/2}P(P \text{ holds for } Y_1) \quad (1.6)$$

and we proceed to bound the last term using the theory already developed for iid matrices (the loss of a factor $O(n^{1/2})$ turns out to be acceptable).

The results from [5], [28] and [30] still relied on the independence between rows. For the rrd matrix $M$ considered in the present work there is no independence among rows or columns. In particular, an approach by conditioning iid variables as in [28] can not treat each row separately, and instead must condition on the event that the entire iid matrix is in $\mathcal{M}_{n,d}$. Letting $p := d/n$, we draw a random 0/1 matrix $M_p$ with iid $\text{Bernoulli}(p)$ entries, and let

$$E_{n,d} = \{M_p \in \mathcal{M}_{n,d}\}. \quad (1.7)$$

Then $M_p \mid E_{n,d} \overset{d}{=} M$. We have

$$P(E_{n,d}) \sim \sqrt{2\pi d(n-d)} \exp \left( -n \log \left( \frac{2\pi d(n-d)}{n} \right) \right) \quad (1.8)$$

which follows from an asymptotic formula for the cardinality of $\mathcal{M}_{n,d}$ established for the sparse case $d = np = o(\sqrt{n})$ by McKay and Wang in [27] and for the dense range $\min(d, n-d) \gg n/\log n$ by Canfield and McKay in [8].

Although enumeration results for the range $\sqrt{n} \ll d \ll n/\log n$ are unavailable as of this writing (though it is natural to conjecture that the formula (1.8) extends to hold in this range), in [47] Tran used an argument from [34] of Shamir and Upfal to show that for $d = \Omega(\log n)$,

$$P(E_{n,d}) \geq \exp \left( -O(n\sqrt{d}) \right). \quad (1.9)$$

While weaker than (1.8), this lower bound was enough to prove the quarter-circular law for the singular value distribution of $M$ using the conditioning trick (in fact Tran treated the more general case of rectangular 0/1 matrices with constant row and column sums, for which he proved the Marchenko-Pastur law). The semi-circular law was established for undirected random regular graphs with $d \to \infty$ by a similar approach in [48]. It is worth noting that
the Marchenko-Pastur and semi-circular laws were also obtained in [15] and [14] for the sparse regime \( \omega(1) \leq d \leq n^{o(1)} \), using the fact that \( d \)-regular graphs converge locally (in a quantitative Benjamini-Schramm sense) to \( d \)-regular trees.

With (1.9) one is limited to importing properties of the iid matrix \( M_p \) that hold with probability \( 1 - O(\exp(-Cn \sqrt{np})) \) for some sufficiently large \( C \), and this can be slightly relaxed by instead using the formula (1.8) for the appropriate range of \( d \). We note in particular that the results of the present work cannot be obtained by the restriction method.

Back on the continuous side, a similar conditioning approach was used to study uniform random doubly stochastic matrices in [11] and [29]. Specifically, it was noted in [11] by Chatterjee, Diaconis and Sly that this distribution can be obtained as the restriction of the distribution of an iid matrix with exponentially distributed entries. They used this observation to prove the quarter circular law by similar lines to [47], relying on another asymptotic formula of Canfield and McKay from [9] for the volume of the Birkhoff polytope. Nguyen built on this work in [29] to prove the circular law for this model.

1.2. Main results and conjectures. Our main result is an analogue of Komlós’ Theorem 1.1 for rrd matrices, assuming that the matrix is not too sparse or too dense. Specifically, we assume that \( \min(d, n - d) = \omega(\log^2 n) \). This significantly improves an earlier result of the author in [12], which treated the case of linear density \( \min(d, n - d) = \Omega(n) \). As in that work, our approach is by couplings rather than by the conditioning trick described above. We give more detail and motivation for the proof strategy in Section 1.3 below.

**Theorem 1.2** (Main result). Assume \( \min(d, n - d) = \omega(\log^2 n) \), and let \( M \) be a uniform random element of \( \mathcal{M}_{n,d} \). Then

\[
P(\det(M) = 0) = O(d^{-c})
\]

for some absolute constant \( c > 0 \).

**Remark 1.3.** The proof shows that we may take \( c = 1/20 \), though we do not expect this bound to be optimal (see 1.5 below).

**Remark 1.4.** One can easily show that a matrix \( M \in \mathcal{M}_{n,d} \) is invertible if and only if the “complementary” matrix \( M' \) with entries

\[
M'(i, j) = 1 - M(i, j)
\]

is invertible. Hence, we may and will assume that \( p := d/n \leq 1/2 \) throughout the proof.

We believe that when \( d \) is of linear size, the singularity probability is exponentially small, similarly to the bound (1.2) for iid sign matrices.

**Conjecture 1.5.** Assume \( \min(d, n - d) \geq p_0 n \) for some fixed \( p_0 \in (0, \frac{1}{2}) \). Then

\[
P(\det(M) = 0) \leq C e^{-cn}
\]

for constants \( C, c > 0 \) depending only on \( p_0 \).

We also conjecture that rrd matrices are invertible with high probability for much smaller values of \( d \):
Conjecture 1.6. There are absolute constants $C, c > 0$ such that for any $3 \leq d \leq n - 3$ we have

$$\Pr(\det(M) = 0) \leq Cn^{-c}.$$  

This mirrors a similar conjecture of Vu in [49] on the adjacency matrices of undirected $d$-regular graphs, which are the symmetric analogue of $M$. When $d$ is bounded, considering the event that two columns of $M$ are parallel shows that we cannot hope for better than a polynomial bound on the singularity probability. $M$ is obviously invertible when $d = 1$ as it is a permutation matrix in this case. As for the case $d = 2$, the following observation may be surprising at first.

Observation 1.7. Let $M$ be a uniform random element of $\mathcal{M}_{n,2}$. Then $M$ is singular asymptotically almost surely.

Proof Sketch. One first observes that $M$ is identically distributed to

$$P(I + P_0)$$

where $P$ and $P_0$ are independent permutation matrices, with $P$ uniform random and $P_0$ uniform among permutation matrices with 0 diagonal (i.e. $P_0$ is associated to a uniform random derangement). Hence, the probability that $M$ is invertible is equal to the probability that

$$I + P_0$$

is invertible. Now we conjugate by a permutation matrix $Q$ to put $P_0$ in block diagonal form according to its cycle structure. The resulting block matrix

$$I + Q^T P_0 Q$$

has blocks of the form

$$\begin{pmatrix}
1 & 1 & 0 & \cdots & 0 \\
0 & 1 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 1
\end{pmatrix}$$

A matrix of this form is invertible if and only if it is of odd dimension. Hence, the probability that $M$ is invertible is equal to the probability that a uniform random derangement decomposes into only odd cycles. The reader may verify that for $\sigma \in \text{Sym}(n)$ a uniform random permutation, we have

$$\Pr(\sigma \text{ contains only odd cycles}) = o(1)$$

(1.11)

(for the precise asymptotics of this probability see exercise 5.10 in [35]). The result then follows from the fact that a uniform random permutation is a derangement with probability $\Omega(1)$ – see for instance [51].

Remark 1.8. One may similarly show that the sum of two independent and uniformly distributed permutation matrices $P_1 + P_2$ is singular asymptotically almost surely.

Next we give a consequence of Theorem 1.2 for random sign matrices. Note that if we draw an iid matrix of signs $\Xi$ as in Theorem 1.1 and condition on the event that all rows and columns sum to 0, the resulting matrix $\Xi_0$ will be singular with null vector $1 = (1, 1, \ldots, 1) \in \mathbb{R}^n$. The following result states that this is the only obstruction for invertibility.
Corollary 1.9. Assume $n$ is even, and let $\Xi_0$ be an $n \times n$ matrix of Bernoulli($1/2$) signs conditioned to have each row and column sum to 0. Then with high probability, $\ker(\Xi_0) = \langle 1 \rangle$. In particular, $\text{corank}(\Xi_0) = 1$ with high probability.

Proof. Let

$$M = \frac{1}{2}(\Xi_0 + 1 1^T).$$

Then $M$ is an rrd matrix with $d = n/2$. For $x \in \mathbb{R}^n$, write

$$x = \overline{x} 1 + x_0$$

with $\overline{x} = \frac{1}{n} \sum_i x(i)$ and $x_0$ the orthogonal projection of $x$ to $(1)^\perp$, the space of mean-zero vectors. One then verifies that $x \in \ker(\Xi_0)$ if and only if $x_0 \in \ker(M)$, and the result follows from Theorem 1.2. \qed

Our next result concerns signed rrd matrices. Let $M_{\pm}^{n,d}$ denote the set of $n \times n$ matrices $M_{\pm}$ with entries taking values in $\{\pm 1, 0\}$ satisfying the constraints

$$d = \sum_{i=1}^n |M_{\pm}(i,k)| = \sum_{j=1}^n |M_{\pm}(k,j)|$$

for all $k \in [n]$. We have the following analogue of Theorem 1.2 for signed rrd matrices:

Theorem 1.10 (Signed rrd matrices are invertible a.a.s.). Assume $\min(d, n-d) = \omega(\log^2 n)$, and let $M_{\pm}$ be a uniform random element of $M_{\pm}^{n,d}$. Then

$$P(\det(M_{\pm}) = 0) = O(d^{-1/4}).$$

While we consider the above result to be interesting in its own right, we include it mainly for expository reasons: each piece of its proof will preview a more complicated argument for the proof of Theorem 1.2.

The signed rrd matrix $M_{\pm}$ will be easier to work with than the unsigned rrd matrix $M$ due to the following alternative description. Let $M$ denote the unsigned version of $M_{\pm}$, i.e. for each $i, j \in [n]$,

$$M(i,j) := |M_{\pm}(i,j)|.$$

Then $M$ is an rrd matrix. We may hence view $M_{\pm}$ as being generated in the following way:

1. draw an rrd matrix $M$,
2. draw an $n \times n$ matrix $\Xi$ of iid uniform signs, independent of $M$, and
3. let $M_{\pm} = M \odot \Xi$.

Here $\odot$ denotes the Hadamard (or Schur) product, so that

$$M_{\pm}(i,j) = \Xi(i,j)M(i,j)$$

for each $i, j \in [n]$. We refer to $M$ as the “base” or “support” of the signed rrd matrix $M_{\pm}$. Roughly speaking, our approach to proving Theorem 1.10 will be to condition on a “good” realization of the base rrd matrix $M$ and proceed using only the randomness of the iid sign matrix $\Xi$. We will then have to show that such good configurations $M$ occur with high probability.
We pause to discuss what we mean by “good configurations” in some more detail. The conditions are most naturally stated in terms of the $d$-regular digraph $\Gamma = (V, E)$ which has $M$ as its adjacency matrix: we identify $V$ with $[n]$, and $E \subseteq [n]^2$ is such that for all $i, j \in [n],$

$$M(i, j) = 1 \iff i \rightarrow j.$$  

We associate row and column indices with vertices of $\Gamma$. For $i \in [n]$, let

$$N_M(i) = \{j \in [n] : M(i, j) = 1\}$$  \hfill (1.15)

so that $N_M(i)$ and $N_M^T(i)$ are the out- and in-neighborhoods of the vertex $i$, respectively, in $\Gamma$. For $S \subseteq [n]$ we denote

$$N_M(S) = \bigcup_{i \in S} N_M(i).$$  \hfill (1.16)

For $A, B \subseteq [n]$ we let

$$e_M(A, B) := |(A \times B) \cap E| = \sum_{i \in A} \sum_{j \in B} M(i, j)$$  \hfill (1.17)

the number of directed edges passing from $A$ to $B$.

Roughly speaking, the base matrix $M$ is a good configuration if the associated digraph satisfies certain expansion properties. In Section 3.2 we prove that the rrd matrix $M$ satisfies all of the necessary properties with overwhelming probability. The proofs rely on sharp tail bounds for the edge counts $e_M(A, B)$, which were proved in [13]. The proof of Theorem 1.2 for the unsigned rrd matrix $M$ will make heavier use of the discrepancy and expansion properties proved in Section 3.2.

It turns out that in the proof of Theorem 1.10 we will not need all of the discrepancy properties enjoyed by $M$. Hence, the proof actually gives the following more general result – the hypotheses distill the minimal set of facts we use about the rrd matrix $M$ in the proof of Theorem 1.10. In particular, Theorem 1.11 is independent of the results in [13]. Here and throughout we make use of the notation

$$a \wedge b = \min(a, b), \quad a \vee b = \max(a, b).$$

**Theorem 1.11 (0/±1 matrices with expanding support are invertible a.a.s.).** Let $M$ be a random 0/1 matrix with exchangeable rows and columns. Let $G_d$ be the event that $M$ enjoys the following discrepancy properties with respect to a parameter $d \in [n]$ with $n \geq d \geq \omega(\log^2 n)$:

1. (Minimum degree) Every row and column of $M$ has at least $d$ nonzero entries. That is, for all $i \in [n],$

$$|N_M(i)| \wedge |N_M^T(i)| \geq d.$$

2. (Expansion of small sets) There is some constant $c_0 > 0$ such that for all $\gamma \in (0, c_0],$

for all $S \subseteq [n]$ such that $|S| \leq \frac{n \log n}{2 \gamma^2}$, we have

$$|N_M(S)| \wedge |N_M^T(S)| \geq \frac{\gamma}{\log n}d|S|$$

(that is, the sizes of neighborhoods $N_M(S)$ are within a logarithmic factor of their upper bound $d|S|$).
(3) (No large sparse minors) There are constants $C, c > 0$ such that for all $A, B \subseteq [n]$ satisfying
$$|A| \wedge |B| \geq C \frac{n}{d} \log n,$$
we have
$$e_M(A, B) \geq c \frac{d}{n}|A||B|.$$

(4) (No thin dense minors) For any $S, B \subseteq [n]$,
$$e_M(S, B) \vee e_M(B, S) \leq d|S|.$$

Let $\Xi$ be an iid sign matrix independent of $M$. Then
$$\mathbb{P}(\det(M \odot \Xi) = 0) \leq \mathbb{P}(G_n) + O(d^{-1/4}).$$

In particular, if properties (1)-(4) hold a.a.s. for $M$ with $d = \omega(\log^2 n)$, then $M \odot \Xi$ is invertible a.a.s.

**Remark 1.12.** Note that the iid sign matrix $\Xi$ can be expressed in the above notation as $M \odot \Xi$ with $M = 11^T$, the all-1s matrix. The associated digraph is then the complete digraph, which satisfies properties (1)-(4) with $d = n$. Hence, Theorem 1.11 is a common generalization of Theorem 1.10 and Komlós’ Theorem 1.1 (though the exponent $1/2$ in Theorem 1.1 has been reduced to $1/4$).

1.3. The general strategy. Now we give a high level discussion our couplings approach to proving invertibility of an rrd matrix $M$. The strategy is similar in spirit to the one used by Rudelson and Vershynin in the recent work [33] on the least singular value of perturbations of deterministic matrices by Haar unitary or orthogonal matrices. As the rrd matrix $M$ has discrete distribution, the couplings we define will be of a very different nature from the ones considered in that work. Nevertheless, on a conceptual level at least, we cannot overvalue the influence [33] has had on our approach for dealing with dependent random variables.

In order to improve on the strategy of conditioning on an iid matrix Bernoulli($p$) matrix $M_p$ as in (1.7), we would like to show that the events
$$\{\det(M_p) = 0\} \text{ and } \{M_p \in \mathcal{M}_{n,d}\}$$
are approximately independent in some sense. Indeed, proceeding as in (1.6) gives
$$\mathbb{P}(\det(M) = 0) = \mathbb{P}(\det(M_p) = 0 | M_p \in \mathcal{M}_{n,d}) \leq \frac{\mathbb{P}(\det(M_p) = 0)}{\mathbb{P}(M_p \in \mathcal{M}_{n,d})}$$
(1.19)
which is only sharp for the worst case that we actually have the containment
$$\{\det(M_p) = 0\} \subset \{M_p \in \mathcal{M}_{n,d}\}.$$  

Of course, (1.20) is likely far from the truth. This motivates us to better understand the structure of the set $\mathcal{M}_{n,d}$; specifically, we try to identify symmetries of this set. If we can identify a large class of operations $\Phi : \mathcal{M}_{n,d} \rightarrow \mathcal{M}_{n,d}$ which leave the distribution of a uniform random element $M \in \mathcal{M}_{n,d}$ invariant, then we could select such an operation $\Phi$ at random from this class and form a new rrd matrix
$$\tilde{M} = \Phi(M).$$
Now to bound the event that some property $P$ holds for $M$, we may replace $M$ with $\tilde{M}$:

$$\mathbb{P}(P \text{ holds for } M) = \mathbb{P}(P \text{ holds for } \tilde{M}) = \mathbb{E} \mathbb{P}(P \text{ holds for } \tilde{M} | M)$$

and proceed to bound the inner probability using only the randomness we have “injected” via the map $\Phi$. This approach can be very powerful if we can design the map $\Phi$ to involve a large number of independent random variables.

This strategy was used in [33] to obtain bounds of the form

$$\mathbb{P}(\sigma_n(D + U) \leq t) \ll t^C n^C$$

for some absolute constants $C, c > 0$, where $D$ is a deterministic matrix (satisfying some additional hypotheses) and $U$ is a Haar-distributed unitary or orthogonal matrix. Since the random matrices in this case are drawn from a group, there is no shortage of symmetries to consider for injecting independence. Furthermore, the availability of continuous symmetries allowed for the injection of random variables possessing smooth bounded density (such as iid Gaussians). This gave quick access to anti-concentration estimates, which play the same role as Erdős’ Theorem 2.1 in the proof of Komlos’ Theorem 1.1 (see Section 2 for a sketch of that proof).

The bound (1.21) had implications for the Single Ring Theorem, proved by Guionnet, Krishnapur and Zeitouni in [17], for the limiting spectral distribution of certain random matrices with prescribed singular values; specifically, it was shown that a hypothesis in [17] could be disposed of. (1.21) was also used in the proof by Basak and Dembo in [4] of the limiting spectral distribution for the sum of a fixed number of independent Haar unitary or orthogonal matrices. It was conjectured in [6] that the same law should hold for the sum of a fixed number of independent uniform random permutation matrices, which can be viewed as a sparse version of the rrd matrix $M$.

The present setting of rrd matrices is a little more complicated as the distribution is discrete, and $\mathcal{M}_{n,d}$ is not a group (except when $d = 1$ of course, but then the problem is trivial). The basic building block for our coupled pairs $(M, \tilde{M})$ will be the well-known “simple switching” operation: letting

$$I_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

we can replace a $2 \times 2$ minor of $M$ by $I_2$ if it is $J_2$ and $J_2$ if it is $I_2$ – indeed, note that this preserves the row and column sums. If $i_1, i_2$ and $j_1, j_2$ are the row and column indices, respectively, of such a minor, then in the associated digraph $\Gamma = (V, E)$ we are alternating between the following edge configurations at vertices $i_1, i_2, j_1, j_2$:

![Diagram showing edge configurations](image)

where we use solid arrows to depict directed edges, and dashed arrows to indicate the absence of a directed edge (i.e. “non-edges”). This forms the basis for (a simple instance of) what
is known as the *switchings method*, which has been a successful tool in the study of random regular graphs since its introduction by McKay in [26]; see also section 2.4 of the survey [50].

We will want to form $\tilde{M}$ by applying several switchings independently at non-overlapping $2 \times 2$ minors. Each minor is replaced with $I_2$ or $J_2$ uniformly at random, independently of all other switchings. We can encode the outcomes of the random switchings with iid uniform signs – this will give us access to the anti-concentration estimate of Theorem 2.1.

In the prior work [12] on dense rrd matrices, we selected the $2 \times 2$ minors for switchings by randomly sampling pairs of row and column indices; however, sampling loses its effectiveness when working with sparse matrices. Here we will use a more efficient coupling which we call “shuffling” – see Section 3.1 for details.

1.4. Organization of the paper. The rest of the paper is organized as follows. In Section 2 we describe the ideas of the proof in more detail, reviewing the approach introduced by Komlós to classify potential null vectors as structured and unstructured, and illustrating our use of couplings by considering a toy problem. Section 3 gives the formal statements and proofs for the tools that were motivated in Section 2, namely the “shuffling” coupling of rrd matrices, discrepancy properties for random regular digraphs (including results from [13]), and a concentration inequality for the symmetric group. In Section 4 we bound the event that $M$ or $M_\pm$ have “unstructured” null vectors, and in Section 5 we take care of “structured” null vectors. In both sections, we first treat the signed rrd matrix $M_\pm$ as a warmup to the more complicated arguments for $M$. The appendix contains the proof of a key technical lemma – Lemma 5.3 – for the proofs in Section 5.

1.5. Notation. We make use of the following asymptotic notation with respect to the limit $n \to \infty$ (though all statements have content for $n$ fixed sufficiently large). $f \ll g$, $g \gg f$, $f = O(g)$, and $g = \Omega(f)$ are all synonymous to the statement that $|f| \leq Cg$ for all $n \geq C$ for some absolute constant $C$. $f \asymp g$ and $f = \Theta(g)$ mean $f \ll g$ and $f \gg g$. $f = o(g)$ and $g = \omega(f)$ mean that $f/g \to 0$ as $n$ tends to infinity. For a parameter $\alpha \in \mathbb{R}$, $f \ll_\alpha g$, $f = O_\alpha(g)$ etc. mean that $|f| \leq C_\alpha g$ for all $n \geq C_\alpha$, with $C_\alpha$ a constant depending only on $\alpha$. $C, c, c', c_1, \ldots, c_k$ etc. denote absolute constants whose value may change from line to line. We stress that while many constants are left unspecified, the proofs are completely effective.

Events will be denoted by the letters $\mathcal{E}, \mathcal{B}$, and $\mathcal{G}$, where the latter two denote “bad” and “good” events, respectively. Their meaning may vary from proof to proof, but will remain fixed for the duration of each proof. $1_\mathcal{E}$ denotes the indicator random variable corresponding to the event $\mathcal{E}$. $\mathbb{E}_X$ and $\mathbb{P}_X$ denote expectation and probability, respectively, conditional on all random variables but $X$.

We make use of the following terminology for sequences of events.

**Definition 1.1 (Frequent events).** An event $\mathcal{E}$ depending on $n$ holds

- asymptotically almost surely (a.a.s.) if $\mathbb{P}(\mathcal{E}) = 1 - o(1)$,
- with high probability (w.h.p.) if $\mathbb{P}(\mathcal{E}) = 1 - O(n^{-c})$ for some absolute constant $c > 0$,
- with overwhelming probability (w.o.p.) if $\mathbb{P}(\mathcal{E}) = 1 - O_C(n^{-C})$ for any constant $C > 0$.

Given ordered tuples of row and column indices $(i_1, \ldots, i_a)$ and $(j_1, \ldots, j_b)$, we denote by

$$M_{(i_1,\ldots,i_a) \times (j_1,\ldots,j_b)}$$

(1.23)
the $a \times b$ matrix with $(k, l)$ entry equal to the $(i_k, j_l)$ entry of $M$. (Note for instance that the sequence $(i_1, \ldots, i_a)$ need not be increasing.) For $A, B \subset [n]$, by $M_{A \times B}$ we mean the $|A| \times |B|$ minor of $M$ as in (1.23), with the increasing ordering of the elements of $A, B$ implied. The rows and columns of $M$ will simply be denoted by $R_i, X_j$, respectively.

We view a vector $x \in \mathbb{R}^n$ as a function from $[n]$ to $\mathbb{R}$. In particular, the $i$th component of $x$ is denoted $x(i)$. We also define the support of $x$ as

$$\text{spt}(x) := \{i \in [n] : x(i) \neq 0\} = [n] \setminus x^{-1}(0).$$

We let $1$ denote the all-ones vector $(1, \ldots, 1) \in \mathbb{R}^n$. “Null vector” will mean “right null vector” unless otherwise stated.

As noted in Section 1.2, it will be convenient to use some terminology reflecting the association of $M$ with a $d$-regular digraph $\Gamma = (V, E)$. In addition to the notation (1.15), (1.16), (1.17), for distinct row indices $i_1, i_2 \in [n]$ we define the sets of column indices

$$\text{Co}_M(i_1, i_2) = \mathcal{N}_M(i_1) \cap \mathcal{N}_M(i_2)$$

$$= \left\{ j \in [n] : M_{(i_1, i_2)\times j} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

$$\text{Ex}_M(i_1, i_2) = \mathcal{N}_M(i_1) \setminus \mathcal{N}_M(i_2)$$

$$= \left\{ j \in [n] : M_{(i_1, i_2)\times j} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

so that

$$\text{Ex}_M(i_2, i_1) = \left\{ j \in [n] : M_{(i_1, i_2)\times j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}. \quad (1.24)$$

These three sets partition the vertex-pair neighborhood $\mathcal{N}_M\{i_1, i_2\}$. We denote their cardinalities with lower case:

$$\text{co}_M(i_1, i_2) = |\text{Co}_M(i_1, i_2)|, \quad \text{ex}_M(i_1, i_2) = |\text{Ex}_M(i_1, i_2)|$$

the first of these being the usual (out-)codegree of vertices $i_1, i_2$.

1.6. Acknowledgements. The author thanks Terence Tao for invaluable discussions on this problem and on random matrix theory in general, as well as for helpful feedback on preliminary versions of the manuscript. Thanks also go to Ioanna Dumitriu and Jamal Najim for the suggestion to consider signed rrd matrices, in large part because the proof of Theorem 1.10 inspired arguments to improve the main theorem, allowing the degree $d$ to lower from $n^{1/2+\varepsilon}$ to $\omega(\log^2 n)$.

2. Ideas of the proof

Our general approach to Theorem 1.2 is inspired by Komlós’ proof of the analogous theorem for iid sign matrices. After briefly reviewing Komlós’ argument below, we will discuss the new ideas that are necessary to treat rrd matrices.
2.1. The strategy of Komlós. A key ingredient of Komlós’ proof is the following anti-concentration bound for random walks due to Erdős.

**Theorem 2.1** (Anti-concentration for random walks [16]). Let \( \xi = (\xi_1, \ldots, \xi_n) \in \{+1, -1\}^n \) be a vector of iid uniform signs. Then

\[
\sup_{a \in \mathbb{R}} \mathbb{P} \left( x : \xi = a \right) \ll |\text{spt}(x)|^{-1/2}
\]

(2.1)

where we recall the notation \( \text{spt}(x) = \{ i \in [n] : x(i) \neq 0 \} \).

**Proof of Theorem 1.1.** We want to bound the bad event

\( \mathcal{B} := \{ \det(\Xi) = 0 \} = \{ \exists \text{ nonzero } x \in \mathbb{R}^n : \Xi x = 0 \} \). (2.2)

The idea is to separately consider the possibility of “structured” and “unstructured” null vectors \( x \). Here the right notion of structure is sparsity. Say that \( x \in \mathbb{R}^n \) is \( k \)-sparse if \( |\text{spt}(x)| \leq k \).

**Proposition 2.2** (No structured null vectors for \( \Xi \)). For any fixed \( \eta \in (0, 1) \), with overwhelming probability \( \Xi \) has no nontrivial \((1 - \eta)n\)-sparse null vectors.

**Remark 2.3.** The proof will show that we can take \( \eta \) as small as \( C_0 n^{-1/4} \) for some absolute constant \( C_0 > 0 \). This is not important for the proof of Theorem 1.1, but a similar observation will be used in the proofs of Theorems 1.2 and 1.10 (and is responsible for the exponent \( 1/4 \) in (1.13)).

We defer the proof of this proposition to the end. Fix \( \eta \in (0, 1) \). We say that \( x \in \mathbb{R}^n \) is “structured” if \( x \) is \((1 - \eta)n\)-sparse, and “unstructured” otherwise. Since \( \Xi \) is identically distributed to its transpose, we may now restrict to the event \( \mathcal{G} \) on which \( \Xi \) has no structured left or right null vectors.

For each \( i \in [n] \), let \( R_i \) denote the \( i \)th row of \( \Xi \), and denote

\[ V_i = \text{span}(R_{i'} : i' \neq i) \]

Define the events

\[ \mathcal{B}_i := \mathcal{G} \land \{ R_i \in V_i \} \]

On \( \mathcal{B} \land \mathcal{G} \), \( \Xi \) must have an unstructured left null vector, which implies that \( \mathcal{B}_i \) holds for at least \((1 - \eta)n\) values of \( i \in [n] \). By double counting we then have that

\[ \sum_{i=1}^{n} \mathbb{P}(\mathcal{B}_i) \geq (1 - \eta)n \mathbb{P}(\mathcal{B} \land \mathcal{G}). \]

(2.3)

Since the rows of \( \Xi \) are exchangeable, all of the summands on the left hand side are equal to \( \mathbb{P}(\mathcal{B}_1) \), say, and so

\[ \mathbb{P}(\mathcal{B} \land \mathcal{G}) \leq \frac{1}{1 - \eta} \mathbb{P}(\mathcal{B}_1). \]

(2.4)

By our bound on \( \mathbb{P}(\mathcal{G}^c) \) from Proposition 2.2, it only remains to show that \( \mathbb{P}(\mathcal{B}_1) \ll n^{-1/2} \).

We condition on the rows \( R_2, \ldots, R_n \) of \( \Xi \), which fixes their span \( V_1 \). Conditional also on a unit normal vector \( u \in V_1^\perp \), drawn independently of \( R_1 \). We have

\[ \{ R_1 \in V_1 \} \subset \{ R_1 \cdot u = 0 \} \].

(2.5)
On $\mathcal{B}_1$ we have that $u$ is perpendicular to every row of $\Xi$, and is hence a left null vector. By our restriction to $\mathcal{G}$ we may hence assume that $u$ is unstructured. By Theorem 2.1 we may now use the randomness of $R_1$ to conclude the desired bound

$$\mathbb{P}(\mathcal{B}_1) \leq \mathbb{P}(\mathcal{G} \land \{R_1 \cdot u = 0\}) \ll \eta n^{-1/2}.$$ 

We turn to the proof of Proposition 2.2. Define the events

$$\mathcal{E}_k = \{\exists x \in \mathbb{R}^n : 0 < |spt(x)| \leq k, \Xi x = 0\}.$$ 

Our aim is to bound

$$\mathbb{P}(\mathcal{E}_{(1-\eta)n}) = \sum_{k=2}^{[(1-\eta)n]} \mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1})$$

(2.6)

(noting that $\mathcal{E}_1$ is empty). It suffices to show that $\mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1})$ is exponentially small for arbitrary fixed $2 \leq k \leq (1 - \eta)n$.

Fix $k$ in this range. On $\mathcal{E}_k \setminus \mathcal{E}_{k-1}$ there is a right null vector $x$ with exactly $k$ nonzero components. We may spend a factor $\binom{n}{k}$ to assume that $x$ is supported on $[k]$ (using column exchangeability). Now on the complement of $\mathcal{E}_{k-1}$, the first $k$ columns of $\Xi$ must span a space of dimension $k - 1$. It follows that there are $k - 1$ linearly independent rows of the left $n \times k$ minor of $\Xi$. By row exchangeability we may spend another factor $\binom{n}{k-1}$ to assume the first $k - 1$ rows are linearly independent. To summarize,

$$\mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \leq \binom{n}{k} \binom{n}{k-1} \mathbb{P}(\mathcal{E}'_k)$$

(2.7)

where

$$\mathcal{E}'_k := \{\exists x \in \mathbb{R}^n : \Xi x = 0, \text{spt}(x) = [k], R_1, \ldots, R_{k-1} \text{ are linearly independent}\}.$$ 

Now note that by linear independence, on $\mathcal{E}'_k$ $x$ is determined by the first $k - 1$. Conditioning on these rows fixes $x$. Then by the independence of the rows of $\Xi$ we have

$$\mathbb{P}(\mathcal{E}'_k \mid R_1, \ldots, R_{k-1}) \leq \mathbb{P}(R_i \cdot x = 0 \forall i \in [k, n])$$

$$= \mathbb{P}(R_n \cdot x = 0)^{n-k+1}.$$ 

(2.8)

Since $|spt(x)| = k$, by Theorem 2.1 we can bound

$$\mathbb{P}(R_n \cdot x = 0) \leq \min \left[ 1/2, Ck^{-1/2} \right]$$

(2.9)

for some $C > 0$ absolute. Combining this bound with (2.8), (2.7) and the inequality $\binom{n}{k} \leq (en/k)^k$ we conclude

$$\mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \ll \exp \left\{ (n - k) \left[ C + 2 \log \frac{n}{n-k} - \log \sqrt{k} \right] \right\}$$

(2.10)

which is more than sufficiently small provided $n-k \geq C_0 n^{-1/4}$ for a sufficiently large absolute constant $C_0$. □
2.2. Structured and unstructured null vectors. It turns out that Proposition 2.2 is robust under some zeroing out of the entries of $M$. Specifically, we can show an analogous result for matrices of the form $M \odot \Xi$ as in Theorem 1.11. We do this for the special case of signed rrd matrices $M_{\pm}$ in Section 5.1, where we prove the following:

**Proposition 2.4** (No structured null vectors for $M_{\pm}$). For $\eta \in (0, 1]$, let $\mathcal{G}_{\pm}^\text{sp}(\eta)$ denote the event that $M_{\pm}$ has no nontrivial $(1 - \eta)n$-sparse null vectors. With hypotheses as in Theorem 1.10, we have that $\mathcal{G}_{\pm}^\text{sp}(\eta)$ holds with overwhelming probability if $\eta \in [C_0 d^{-1/4}, 1]$ for $C_0$ a sufficiently large absolute constant.

Examination of the proof shows that the above statement holds for the more general case that $M$ is a 0/1 matrix as in Theorem 1.11 and lies in the event $\mathcal{G}_d$.

To prove Theorem 1.2 for the rrd matrix $M$, we will also treat structured null vectors separately. It turns out that sparsity is no longer the right notion of structure in this setting – instead, we will need to show that null vectors of $M$ have small level sets:

**Proposition 2.5** (No structured null vectors for $M$). For $\eta \in (0, 1]$, let $\mathcal{G}_{\text{sls}}(\eta)$ be the event that for any nontrivial left or right null vector $x$ of $M$ and for any $\lambda \in \mathbb{R}$,

$$|x^{-1}(\lambda)| \leq \eta n. \quad (2.11)$$

With hypotheses as in Theorem 1.2, we have that $\mathcal{G}_{\text{sls}}(\eta)$ holds with overwhelming probability if $\eta \gg d^{-c}$ for a sufficiently small absolute constant $c > 0$ (we can actually take $c = 1/20$).

We prove this proposition in Section 5. The reason for ruling out null vectors with large level sets will become apparent in the next section.

2.3. Injecting a random walk. The proof in Section 2.1 proceeded by reducing to the event that $R_1 \cdot u = 0$, where $R_1$ is the first row of $\Xi$ and $u$ is a unit vector in $V_1^\perp = \text{span}(R_2, \ldots, R_n)^\perp$. Then we used independence of the entries of $\Xi$ in two ways:

1. Independence of the rows of $\Xi$ allowed us to condition on $R_2, \ldots, R_n$ to fix $u$, without affecting the distribution of $R_1$.

2. Independence of the components of $R_1$ allowed us to view the dot product $R_1 \cdot u$ as a random walk, to which we could apply the anti-concentration result Theorem 2.1.

The rrd matrix $M$ enjoys neither of these properties. However, we will be able to accomplish something like (2) above by defining an appropriate coupling of rrd matrices using switchings. It will take some care to implement this without having the independence between rows (1).

To illustrate our couplings approach, let us consider a toy problem: to control the event that the first two rows lie in the span of the remaining rows, i.e. to show

$$\mathbb{P} \left\{ R_1, R_2 \in V_{1,2}^\perp \right\} = o(1) \quad (2.12)$$

where $V_{1,2} := \text{span}(R_3, \ldots, R_n)$. We will see later that this can be used to control the event that $M$ has corank at least 2 (see Lemma 4.3). For now we will operate under the following

**Assumption 2.6.** $n/2 \geq d \gg n$.

Thus, we are assuming $M$ is a dense rrd matrix, as in the work [12]. In the next section we will discuss some of the new ideas necessary to treat sparse matrices.
Blindly following the proof from Section 2.1, we condition on the rows $R_3, \ldots, R_n$ to fix the space $V_{1,2}$, and pick a unit vector $u \in V^\perp_{1,2}$, say uniformly and independently of $R_1, R_2$ under the conditioning. Now it suffices to show

$$\mathbb{P}\{R_1 \cdot u = 0 \mid R_3, \ldots, R_n\} = o(1). \quad (2.13)$$

We need to understand how $R_1$ and $R_2$ are distributed under the conditioning on $R_3, \ldots, R_n$. Recall our notation

$$\text{Co}_M(1,2) = \left\{ j \in [n] : M_{(1,2) \times j} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

$$\text{Ex}_M(1,2) = \left\{ j \in [n] : M_{(1,2) \times j} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

$$\text{Ex}_M(2,1) = \left\{ j \in [n] : M_{(1,2) \times j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

which partition the vertex-pair neighborhood $\mathcal{N}_M(\{1,2\})$. Having fixed $R_3, \ldots, R_n$, we have fixed which columns of $M$ need 2, 1, or 0 more entries equal to 1 in the first two rows to meet the column sum constraint of $d$. This fixes the sets $\text{Co}_M(1,2)$ and $\text{Ex}_M(1,2) \cup \text{Ex}_M(2,1)$. Furthermore, by the row sums constraint, we must have

$$|\text{Ex}_M(1,2)| = d - |\text{Co}_M(1,2)| \quad (2.14)$$

$$= |\text{Ex}_M(2,1)|.$$ 

It follows that with $R_3, \ldots, R_n$ fixed, the only remaining randomness is in the uniform random equipartition of the deterministic set $\text{Ex}_M(1,2) \cup \text{Ex}_M(2,1)$ into the sets $\text{Ex}_M(1,2)$, $\text{Ex}_M(2,1)$. See Figure 1.

We re-randomize the sets $\text{Ex}_M(1,2)$, $\text{Ex}_M(2,1)$ in the following way. Under this conditioning, pick a bijection

$$\pi : \text{Ex}_M(1,2) \rightarrow \text{Ex}_M(2,1)$$

uniformly at random. Now for each $j \in \text{Ex}_M(1,2)$ we have

$$M_{(1,2) \times (j, \pi(j))} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =: \mathbf{I}_2.$$
Having obtained a sequence of “switchable” 2×2 minors, we can now apply random switchings (with terminology as in Section 1.3). Let \( \tilde{\xi} = (\xi_j)_{j=1}^n \) be a sequence of iid uniform signs, independent of all other variables. For each \( j \in \text{Ex}_M(1, 2) \), we replace the minor \( M_{(1,2) \times (j,\pi(j))} \) with the random minor
\[
\begin{align*}
\mathbf{1}_2 \mathbf{1}_{\xi_j=+1} + \mathbf{J}_2 \mathbf{1}_{\xi_j=-1}.
\end{align*}
\]
Call the resulting matrix \( \tilde{M} \). It is not hard to show that \( \tilde{M} \) is also an rrd matrix after undoing all of the conditioning (see the proof of Lemma 3.1 below).

We have hence obtained a coupled pair \((M, \tilde{M})\) of rrd matrices. Let \( \tilde{R}_i \) denote the rows of \( \tilde{M} \). Replacing \( M \) with \( \tilde{M} \) in (2.13), it now suffices to show
\[
\mathbb{P}\left\{ \tilde{R}_1 \cdot u = 0 \mid M \right\} = o(1). \tag{2.15}
\]
Now in the randomness of the iid signs \( \xi_j \), one can see that the dot product \( \tilde{R}_1 \cdot u \) is a random walk:
\[
\tilde{R}_1 \cdot u = W_0 + \sum_{j \in \text{Ex}_M(1, 2)} \xi_j \frac{u(j) - u(\pi(j))}{2}, \tag{2.16}
\]
where \( W_0 \) is a term that does not depend on \( \pi \) or \( \tilde{\xi} \). Applying Theorem 2.1 we have
\[
\mathbb{P}\{ \tilde{R}_1 \cdot u = 0 \} \ll \left| \left\{ j \in \text{Ex}_M(1, 2) : u(j) \neq u(\pi(j)) \right\} \right|^{-1/2}.
\]
It remains to get a lower bound on the number of \( j \in \text{Ex}_M(1, 2) \) for which \( u(j) \neq u(\pi(j)) \).

First we deal with the possibility that \( \text{Ex}_M(1, 2) \) is a very small set. On average, we expect \( \text{Co}_M(1, 2) \) to be of size roughly \( p^2 n = d^2/n \). By (2.14) and our assumption \( p \leq 1/2 \) (from Remark 1.4) we have
\[
\mathbb{E}\left| \text{Ex}_M(1, 2) \right| \gg d.
\]
It was shown in [13] that codegrees in random regular digraphs are sharply concentrated (see Theorem 3.3) from which we can show
\[
\left| \text{Ex}_M(1, 2) \right| \gg d \tag{2.17}
\]
on a negligibly small event.

Now we apply Proposition 2.5 and the randomness of \( \pi \) to argue that for most \( j \in \text{Ex}_M(1, 2) \) we have \( u(j) \neq u(\pi(j)) \). Since \( u \in V_{1,2}^+ \) we have that \( u \) is a (right) null vector of the \((n-2) \times n \) matrix \( M_{[3,\pi]\times [n]} \). By a small extension of Proposition 2.5 we may assume that \( u \) is unstructured, i.e. that all of its level sets are of size at most \( \eta m \), with \( \eta \) of size \( \Theta(d^{-c}) \) for some \( c > 0 \) absolute. (In the actual proof we will argue that \( u \) is unstructured in a slightly different way, but in any case it comes down to an application of Proposition 2.5.) Now by (2.17) and 2.6, the sets \( \text{Ex}_M(1, 2) \), \( \text{Ex}_M(2, 1) \) are much larger than the level sets of \( u \). Hence, in the randomness of \( \pi \), it is very unlikely that we have \( \pi(j) \in u^{-1}(u(j)) \) for a large number of indices \( j \in \text{Ex}_M(1, 2) \). Thus, off a negligibly small event we have shown that most of the steps taken by the random walk (2.16) are nonzero, and hence
\[
\mathbb{P}_{\xi, \pi}\{ \tilde{R}_1 \cdot u = 0 \} \ll \left| \text{Ex}_M(1, 2) \right|^{-1/2} \ll d^{-1/2}. \tag{2.18}
\]
Since we are assuming \( d = \omega(1) \), we have completed the proof of (2.12).

To summarize, we bounded \( \mathbb{P}(R_1, R_2 \in V_{1,2}^+) \) by replacing \( M \) with a coupled rrd matrix \( \tilde{M} \), formed by introducing the new variables \( \pi \) and \( \tilde{\xi} \). The variables \( M, \pi \) and \( \tilde{\xi} \) each played a special role:
(1) In the randomness of $M$, we simply restricted to a couple of “good events”: the event $G_{\text{dis}}(\eta)$, and the event that $|\text{Ex}_M(1, 2)| \gg d$.

(2) Conditional on $M$ satisfying the good events, $\pi$ was used to pair indices in $\text{Ex}_M(1, 2)$ with indices in $\text{Ex}_M(2, 1)$ to show that, off a small event, the random walk $R_1 \cdot u$ takes many nonzero steps $\frac{1}{2}(u(j) - u(\pi(j)))$.

(3) Conditional on good realizations of $M$ and $\pi$, the randomness of $\xi$ was used with Theorem 2.1 to finish the proof.

As remarked above, (2.12) can be used to deduce that $M$ has corank at most 1 a.a.s. It then remains to deal with the event that $\text{corank}(M) = 1$. This task is a little more complicated, and involves expressing a certain $2 \times 2$ determinant involving two randomly sampled rows of $M$ as a random walk. See Section 4.5 for details.

2.4. Dealing with sparsity. In the previous section, we used 2.6 to guarantee that the level sets of the normal vector $u$ were small in comparison to the sets $\text{Ex}_M(1, 2)$ and $\text{Ex}_M(2, 1)$. Indeed, since the level sets are of size at most $\eta m = \Theta(nd^{-c})$ by Proposition 2.5, and since $|\text{Ex}_M(1, 2)| \gg d$ with high probability, we see upon rearranging that in the above argument we must assume $d \geq Cn^{1/(1+\epsilon)}$ for a sufficiently large constant $C > 0$. It turns out that the best value for the constant $c$ in Proposition 2.5 that can be obtained by our approach is $1/8$ (see [12]). Hence, the argument of the previous section is limited to $d \geq Cn^{8/9}$.

In the present work we are able to take $d$ as small as $\omega(\log^2 n)$ using some new ideas. Rather than consider the event that $R_1, R_2 \in V^\perp_{1,2}$, we will draw row indices $I_1, I_2$ at random and seek to bound

$$\mathbb{P}\{ R_{I_1}, R_{I_2} \in V^\perp_{I_1,I_2}\}.$$  \hfill (2.19)

Since the rows of $M$ are exchangeable, a bound on (2.19) can still be used to control the event that $\text{corank}(M) \geq 2$ (see Lemma 4.3). Conditional on $I_1, I_2$ and the remaining rows $(R_i)_{i \notin \{I_1, I_2\}}$, we will again select a unit normal vector $u$ uniformly at random.

Whereas in Section 2.3 the distribution of $u$ played no special role, here we will use it along with the randomness of $I_1, I_2$ to argue that it is very unlikely that a level set of $u$ has large overlap with the sets $\text{Ex}_M(I_1, I_2), \text{Ex}_M(I_2, I_1)$. Under conditioning on $I_1, I_2$, one can see that the “bad” realizations of $u$ form an algebraic set. We will then use the simple fact that a proper algebraic subset of the sphere has surface measure zero. (This is perhaps the only part of the proof that is not strictly combinatorial in nature.) The argument requires some care as the vector $u$ and the sets $\text{Ex}_M(I_1, I_2), \text{Ex}_M(I_2, I_1)$ are both dependent on $I_1, I_2$. See Section 4.4 for the detailed proof.

3. Preliminaries

3.1. The shuffling coupling. In this section we formally define the coupling $(M, \tilde{M})$ of rrd matrices described in the previous section, where $\tilde{M}$ is obtained by re-randomizing the neighborhood $\mathcal{N}_M(i_1, i_2)$ in a certain way. On an intuitive level, the shuffling operation is somewhat similar to performing a “riffle shuffling” of the “deck” $\text{Ex}_M(i_1, i_2) \cup \text{Ex}_M(i_2, i_1)$, then cutting the deck into two equal parts. The set of common neighbors $\text{Co}_M(i_1, i_2) = \mathcal{N}_M(i_1) \cap \mathcal{N}_M(i_2)$ is preserved by the shuffling.
Definition 3.1 (Shuffling). Let $M \in \mathcal{M}_{n,d}$ and $i_1, i_2 \in [n]$ distinct. For a bijection\[ \pi : \text{Ex}_M(i_1, i_2) \to \text{Ex}_M(i_2, i_1) \]
and a sequence of signs $\vec{\xi} = (\xi_j)_{j=1}^n \in \{+1, -1\}^n$, by perform a shuffling on $M$ at rows $(i_1, i_2)$ according to $\pi$, $\vec{\xi}$, we mean to replace the $2 \times 2$ minors $M_{(i_1, i_2) \times (j, \pi(j))}$ with\[ I_2 I_{\xi_j = +1} + J_2 I_{\xi_j = -1} \]
for each $j \in \text{Ex}_M(i_1, i_2)$, and to leave all other entries of $M$ unchanged.

The key to applying the shuffling operation in the proof of Theorem 1.2 will be to take $\pi$ and $\vec{\xi}$ to be random.

Lemma 3.1 (Shuffling coupling). Let $M$ be an rrd matrix, and fix $i_1, i_2 \in [n]$ distinct. Conditional on $M$, let $\pi : \text{Ex}_M(i_1, i_2) \to \text{Ex}_M(i_2, i_1)$ be a uniform random bijection. Draw a sequence $\vec{\xi} = (\xi_j)_{j=1}^n \in \{+1, -1\}$ of iid uniform signs, independent of all other variables.

Form $\tilde{M}$ by performing a shuffling on $M$ at rows $(i_1, i_2)$ according to $\pi$ and $\vec{\xi}$. Then $\tilde{M} \overset{d}{=} M$.

Proof. Condition on the rows $(R_i)_{i \notin \{i_1, i_2\}}$ (which are the same for $M$ and $\tilde{M}$). This fixes the set\[ E := \text{Ex}_M(i_1, i_2) \cup \text{Ex}_M(i_2, i_1). \]
As previously noted, the only remaining randomness of $M$ is in the uniform random equipartition of $E$ into the sets $\text{Ex}_M(i_1, i_2)$, $\text{Ex}_M(i_2, i_1)$. It hence suffices to show that $\text{Ex}_{\tilde{M}}(i_1, i_2)$ is also distributed uniformly over subsets of $E$ of size $|E|/2$.

Conditional on $(R_i)_{i \notin \{i_1, i_2\}}$, the randomness of $\pi$, $\vec{\xi}$ and the remaining randomness of $M$ amounts selecting a uniform random pairing of the elements of $E$, and independently and uniformly selecting one element of each pair to form the set $\text{Ex}_{\tilde{M}}(i_1, i_2)$. It follows that $\text{Ex}_{\tilde{M}}(i_1, i_2)$ is uniform over subsets of $E$ of size $|E|/2$ as desired. \hfill \Box

At one point in the proof we will need the following slightly more general coupling, where we are restricted to modifying entries of $M$ within a fixed set $B$ of columns.

Lemma 3.2 (Restricted shuffling). Let $M$ be an rrd matrix and fix $i_1, i_2 \in [n]$ distinct. Let $B \subset [n]$ be deterministic under conditioning on the rows $(R_i)_{i \notin \{i_1, i_2\}}$. Set\[ B^+ = B \cap \text{Ex}_M(i_1, i_2), \quad B^- = B \cap \text{Ex}_M(i_2, i_1) \]
and fix $r \leq |B^+| \land |B^-|$. Conditional on $M$ let\[ E^+ \subset B^+, \quad E^- \subset B^- \]
chosen independently and uniformly among subsets of size $r$. Conditional on $M, E^+, E^-$, let\[ \pi : E^+ \to E^- \]
be a uniform random bijection. Finally, let $\vec{\xi} = (\xi_j)_{j=1}^n \in \{+1, -1\}$ be a sequence of iid uniform signs.

Form $\tilde{M}$ from $M$ by replacing the $2 \times 2$ minors $M_{(i_1, i_2) \times (j, \pi(j))}$ with\[ I_2 I_{\xi_j = +1} + J_2 I_{\xi_j = -1} \]
for each $j \in E^+$, leaving all other entries of $M$ unchanged. Then $\tilde{M} \overset{d}{=} M$. 

Note that Lemma 3.1 follows from the above lemma by taking \( B = [n] \) and \( r = \text{ex}_M(i_1, i_2) \).

**Proof.** Condition on the rows \((R_t)_{i \notin \{i_1, i_2\}}\). This fixes \( B \) and the set

\[
E := \text{Ex}_M(i_1, i_2) \cup \text{Ex}_M(i_2, i_1).
\]

Condition also on the columns of \( M \) indexed by \([n] \setminus B\) – this fixes \(|B^+|\) and \(|B^-|\).

The only remaining randomness of \( M \) is in the uniform random partition of \( E \cap B \) into the sets \( B^+, B^- \) of prescribed sizes. It hence suffices to show that \( \tilde{B}^+ := B \cap \text{Ex}_{\tilde{M}}(i_1, i_2) \) is also distributed uniformly over subsets of \( E \) of size \(|B^+|\).

Conditional on \((R_t)_{i \notin \{i_1, i_2\}}\), the randomness of \( \pi, \tilde{\pi} \) and the remaining randomness of \( M \) amounts selecting a uniform random pairing of \( 2r \) elements of \( E \cap B \), and independently and uniformly selecting one element of each pair to form the set \( E^+ \). It follows that \( \text{Ex}_{\tilde{M}}(i_1, i_2) \)

is uniform over subsets of \( E \) of size \(|E|/2\) as desired. \(\square\)

### 3.2. Discrepancy properties.

In this section we collect various “good events” concerning the distribution of edges in the digraph \( \Gamma \) associated to \( M \). In all cases, the good event is shown to hold with overwhelming probability, for a suitable range of parameters and assuming \( d = \omega(\log^2 n) \). This will allow us to restrict to these events without further comment in subsequent stages of the proof. (Indeed, note that we are ultimately aiming for only a polynomially-small bound on the singularity probability, so the failure probabilities for the good events will be negligible.)

The results of this section are all corollaries of sharp tail estimates for codegrees and edge counts in random regular digraphs (Theorems 3.3 and 3.4 below). The proofs, which are too long for inclusion in the present work, are given in [13]. These results may also be of independent interest for graph theorists.

The shuffling coupling from Lemma 3.1 will only be useful if the sets \( \text{Ex}_M(i_1, i_2) \) are large (see Section 2.3). Hence, the following result from [13] will be essential for our arguments. Recall that \( p := d/n \) denotes the average edge density for the digraph.

**Theorem 3.3** (Concentration of codegrees [13]). Suppose \( 1 \leq d \leq n/2 \). For \( \delta \in (0, 1) \), let \( G^\text{ex}(\delta) \) denote the event that for every pair of distinct \( i_1, i_2 \in [n] \) we have

\[
\left| \frac{\text{ex}_M(i_1, i_2)}{p(1-p)n} - 1 \right| \leq \delta \quad \text{and} \quad \left| \frac{\text{ex}_M^r(i_1, i_2)}{p(1-p)n} - 1 \right| \leq \delta. \tag{3.3}
\]

Then

\[
\mathbb{P}(G^\text{ex}(\delta)) = 1 - O\left( n^2 \exp\left( -c \min\{ \delta d, \delta^2 n \} \right) \right). \tag{3.4}
\]

In particular, for any fixed \( \delta \in (0, 1) \) independent of \( n \) we have that \( G^\text{ex}(\delta) \) holds with overwhelming probability if \( d = \omega(\log n) \).

Our next result concerns the concentration of the number of edges \( e_M(A, B) \) passing from a set \( A \) to a set \( B \) (defined in (1.17)). We expect this random variable to be of size roughly \( \mu(A, B) := p|A||B| \). It is straightforward to check that from the \( d \)-regularity constraint, for any \( t \in \mathbb{R} \) we have the following equality of events:

\[
\{e_M(A, B) - p|A||B| \geq t\} = \{e_M(A^c, B^c) - p|A^c||B^c| \geq t\} \tag{3.5}
\]
where we denote $A^c = [n] \setminus A$. That is, a large deviation of $e_M(A, B)$ coincides with a large deviation of $e_M(A^c, B^c)$. It will hence be natural to express deviations of $e_M(A, B)$ at the scale
\[
\hat{\mu}(A, B) := p \min \{|A||B|, (n - |A|)(n - |B|)\}.
\] (3.6)

**Theorem 3.4** (Concentration of edge counts [13]). With $G^{\text{ex}}(\delta)$ as in Theorem 3.3, we have that for any $A, B \subset [n]$ and any $\tau \geq 0$,
\[
\mathbb{P}\left(\left\{ |e_M(A, B) - \mu(A, B)| \geq \tau \hat{\mu}(A, B) \right\} \cap G^{\text{ex}}(\delta)\right) \leq 2 \exp\left(-\frac{c\tau^2}{1 + \tau} \hat{\mu}(A, B)\right)
\] (3.7) provided $\delta \leq \min\left(\frac{1}{7}, \frac{\tau}{8}\right)$.

Combining the above theorems with a union bound over pairs of sets $(A, B)$ we can deduce that some uniform control on the density of all sufficiently large minors of $M$ holds with overwhelming probability.

**Corollary 3.5** (Discrepancy for large minors [13]). Let $C_0 > 0$ be a sufficiently large absolute constant. For $\varepsilon \in (0, 1)$, define the family of pairs of sets
\[
\mathcal{F}(\varepsilon) = \left\{ (A, B) : A, B \subset [n], |A| \wedge |B| \geq \frac{C_0}{\varepsilon^2} \frac{n \log n}{d} \right\}
\] (3.8)
and the event
\[
\mathcal{G}^\varepsilon(A, B) = \left\{ \forall (A, B) \in \mathcal{F}(\varepsilon), |e_M(A, B) - \mu(A, B)| \leq \varepsilon \hat{\mu}(A, B) \right\}.
\]
If $C_0 \varepsilon^{-1} \log n \leq d \leq n/2$, then
\[
\mathbb{P}(\mathcal{G}^\varepsilon(A, B)) \geq 1 - C \exp\left(-c \min\left\{ \varepsilon^2, \varepsilon^2 n, \frac{n}{\varepsilon^2 d} \log^2 n \right\}\right).
\] (3.9)
In particular, if $n/2 \geq d = \omega(\log n)$ and $\varepsilon \in (0, 1)$ is fixed independent of $n$, then $\mathcal{G}^\varepsilon(A, B)$ holds with overwhelming probability.

Note that for $S, B \subset [n]$ we have the deterministic bound
\[
e_M(S, B) \leq d|S|
\] (3.10) which is effective when $S$ is small (and we have equality when $B = [n]$). While this bound will be sufficient for many of our purposes, we will sometimes need a little more when $|B| = o(n)$.

Theorem 3.4 allows us to improve on (3.10) off a small event:

**Corollary 3.6** (Discrepancy for thin minors). For $\varepsilon_0 \in (0, 1], \gamma > 0$ and $s \in [n]$, denote
\[
s_0(\gamma) := \frac{n \log n}{2\gamma d}, \quad b_0(\varepsilon_0, \gamma, s) := \frac{\varepsilon_0 ds}{\log n}
\] (3.11) and define the family of “thin minors”
\[
\mathcal{F}_{\text{thin}}(\varepsilon_0, \gamma) = \left\{ (S, B) : s \leq s_0(\gamma), b \leq b_0(\varepsilon_0, \gamma, s) \right\}
\] (3.12) where as usual $s = |S|, b = |B|$. Let
\[
\mathcal{B}(\varepsilon_0, \gamma) = \left\{ \exists (S, B) \in \mathcal{F}_{\text{thin}}(\varepsilon_0, \gamma) : e_M(S, B) \lor e_M(B, S) \geq \varepsilon_0 ds \right\}.
\] (3.13) There are absolute constants $C_0, c_0 > 0$ such that if $d \geq C_0 \varepsilon_0^{-1} \log n$ and $\gamma \in (0, c_0)$, then
\[
\mathbb{P}(\mathcal{B}(\varepsilon_0, \gamma)) \ll \exp\left(-c\varepsilon_0 d\right).
\] (3.14)
In particular, if $d = \omega(\log n)$ and $\varepsilon_0$ is fixed independent of $n$, then $B(\varepsilon_0, \gamma)^c$ holds with overwhelming probability.

Proof. With $\varepsilon_0, \gamma$ as in statement of the corollary, fix $s \leq s_0(\gamma)$, $b \leq b_0(\varepsilon_0, \gamma, s)$. Note that we have

$$\mu(S, B) = \frac{d sb}{n} \leq \frac{b_0 ds}{n} \leq \frac{\varepsilon_0 ds}{2}. \quad (3.15)$$

Let $\delta > 0$ to be chosen. For fixed $S, B \subset [n]$ with $|S| = s, |B| = b$, from Theorem 3.4 we have

$$P\left( G^{\text{ex}}(\delta) \land \left\{ e_M(S, B) \geq \varepsilon_0 ds \right\} \right) \leq P\left( G^{\text{ex}}(\delta) \land \left\{ e_M(S, B) - \mu(S, B) \geq \frac{\varepsilon_0}{2} ds \right\} \right)$$

provided we take

$$\delta \leq \min\left( \frac{1}{4}, \frac{\varepsilon_0 ds}{16 \mu} \right).$$

Since $\varepsilon_0 ds/16 \mu \geq 1/8$ by (3.15), we can take $\delta = 1/8$. With this choice of $\delta$ we have

$$P(G^{\text{ex}}(\delta)^c) \ll n^2 \exp(-cd) \quad (3.17)$$

from Theorem 3.3. Now by a union bound, Equation (3.16) and the assumed lower bound on $d$,

$$P(G^{\text{ex}}(1/8) \land B(\varepsilon_0, \gamma)) \leq \sum_{s \leq s_0(\gamma)} \binom{n}{s} \sum_{b \leq b_0(\varepsilon_0, \gamma, s)} \binom{n}{b} \exp(-c\varepsilon_0 ds)$$

$$\ll \sum_{s \leq s_0(\gamma)} n^b \exp(s(\log n - c\varepsilon_0 d))$$

$$\leq \sum_{s \leq s_0(\gamma)} \exp(b_0 \log n - c\varepsilon_0 ds)$$

$$\leq \sum_{s \leq s_0(\gamma)} \exp(-c\varepsilon_0 ds)$$

$$\ll \exp(-c\varepsilon_0 d)$$

where in the fourth line we used the definition (3.11) of $b_0$ and took $\gamma$ sufficiently small. Combining with the bound (3.17) and our assumed lower bound on $d$ completes the proof. \qed

We have the following quick consequence that with high probability, the size of the neighborhood $N_M(S)$ of any small set $S$ is within a logarithmic factor of the upper bound $d|S|$.

**Corollary 3.7** (Expansion of small sets). For $\gamma > 0$, let $G^{\text{ex}}(\gamma)$ be the event that for every $S \subset [n]$ with $|S| \leq s_0(\gamma) = \frac{1}{2^7} \frac{n \log n}{d}$, we have

$$|N_M(S)| \geq \frac{\gamma}{\log n} |S|.$$  

There is a constant $c_0 > 0$ such that for all $\gamma \in (0, c_0]$,

$$P(G^{\text{ex}}(\gamma)) = 1 - O(e^{-cd}). \quad (3.18)$$
Proof. Let $\gamma > 0$. On $G^{\exp}(\gamma)^c$ there exist $S, B \subset [n]$ with $|S| \leq s_0(\gamma)$ and

$$|B| < \frac{\gamma}{\log n} d|S| = b_0(1, \gamma, |S|)$$

in the notation of (3.11), such that $e_M(S, B) = d|S|$ (simply from taking $B = N_M(S)$). Hence, $G^{\exp}(\gamma)^c$ is contained in the event $\mathcal{B}(1, \gamma)$ from Corollary 3.6, and the result follows by taking $\gamma$ sufficiently small. \hfill \square

3.3. Concentration of measure. The following concentration inequality for certain functions on the symmetric group will be useful when working with the bijections $\pi$ in the shuffling coupling of Lemma 3.1, and follows from the $d=1$ case of Theorem 1.18 in [13], or alternatively from Proposition 1.1 in [10].

Lemma 3.8 (Concentration for the symmetric group). For $m \geq 1$, $\pi \in \text{Sym}(m)$ a permutation on $m$ letters, and $A, B \subset [m]$, denote

$$e_\pi(A, B) = \left| \{ i \in A : \pi(i) \in B \} \right|.$$

If $\pi$ is a uniform random element of $\text{Sym}(m)$, we have that for any $\tau \geq 0$,

$$\mathbb{P}\left\{ |e_\pi(A, B) - |A||B|/m| \geq \tau|A||B|/m \right\} \leq 2 \exp\left\{ -\frac{c\tau^2 |A||B|}{1 + \tau |A||B|/m} \right\}. \quad (3.19)$$

Remark 3.9. Note that the above lemma is nearly the same as the $d=1$ case of Theorem 3.4, the only difference (apart from constants in the exponential) being that we do not need to restrict to any “good event” like $G^{\exp}(\delta)$.

4. Unstructured null vectors

In this section we prove Theorem 1.2 and Theorem 1.10, taking Propositions 2.5 and 2.4 as black boxes. These propositions are proved in Section 5.

4.1. Preliminary reductions. We first note the “good events” we have at our disposal:

- By Theorem 3.3, the event $G^{\exp}(\delta)$ holds with overwhelming probability for any fixed $\delta \in (0, 1)$ (independent of $n$), so we may freely restrict to this event.
- From Proposition 2.5 we have that $G^{\text{als}}(\eta)$, and hence $G^{\text{sp}}(\eta)$, holds with overwhelming probability for any $d^{-c} \ll \eta \leq 1$.
- Also, by Proposition 2.4 we have that $G_{\pm}^{\text{sp}}(\eta)$ holds w.o.p. for $M_{\pm}$ for any $\eta \in [C_0 d^{-1/4}, 1]$, where $C_0 > 0$ is a sufficiently large absolute constant.

We leave the parameters $\delta, \eta \in (0, 1)$ unspecified for now and will fix them sufficiently small later in the proofs.

For $k \in [n]$ define the event

$$\mathcal{R}_k := \{ \text{corank}(M) \geq k \}. \quad (4.1)$$

Our aim is to bound $\mathbb{P}(\mathcal{R}_1)$. By abuse of notation we will also denote $\mathcal{R}_k = \{ \text{corank}(M_{\pm}) \geq k \}$ – this will cause no confusion as we treat $M_{\pm}$ and $M$ in separate sections.

We obtain Theorem 1.10 by combining Proposition 2.4 and the following
Proposition 4.1. For all \( \eta > 0 \) sufficiently small we have
\[
P_{M_{\pm}} \left( R_{1} \land G_{\pm}^{sp} (\eta) \right) \ll \eta + d^{-1/2}. \tag{4.2} \]

It will turn out that for the rrd matrix \( M \) we need to separately handle \( R_{2} \) and \( R_{1} \setminus R_{2} \) by different arguments. The argument for \( R_{2} \) will follow the outline of the proof of Proposition 4.1 for \( M_{\pm} \), after invoking the shuffling coupling to inject iid signs. Controlling \( R_{1} \setminus R_{2} \) will require more care. Theorem 1.2 follows from the next proposition and Proposition 2.5.

Proposition 4.2. For all \( \eta > 0 \) sufficiently small we have
\[
P \left( R_{2} \land G^{\text{rls}} (\eta) \right) \ll \eta + d^{-1/2} \tag{4.3} \]
and
\[
P \left( R_{1} \land R_{2} \land G^{\text{rls}} (\eta) \right) \ll \eta + d^{-1/2}. \tag{4.4} \]

For both propositions we will use the following lemma for controlling the event that a random matrix with exchangeable rows has co-rank at least \( k \) by the event that \( k \) rows sampled randomly without replacement land in the span of the remaining rows.

Lemma 4.3 (Control by random sampling). Assume \( M \) is a random matrix with exchangeable rows \( R_{i}, 1 \leq i \leq n \). Let \( R_{k} \) be as in (4.1) and for \( \eta \in (0, 1) \) let \( G^{sp}(\eta) \) be the event that \( M \) has no \((1 - \eta)n\)-sparse left or right null vectors. For \( k \in [n/2] \), let \( I_{1}, \ldots, I_{k} \) be sampled uniformly without replacement from \([n]\). Define the event
\[
\hat{R}_{k} = \{ R_{I_{1}}, \ldots, R_{I_{k}} \in V_{\{I_{1}, \ldots, I_{k}\}} \}
\]
where for \( S \subset [n] \) we define
\[
V_{S} := \text{span} \left( R_{i} : i \notin S \right). \tag{4.5}
\]
Then if \( \eta \in (0, 1/2k) \), we have
\[
P \left( R_{k} \land G^{sp}(\eta) \right) \leq \frac{1}{1 - 2\eta k} P \left( \hat{R}_{k} \land G^{sp}(\eta) \right). \tag{4.6}
\]

Proof. Since
\[
P \left[ \hat{R}_{k} \land G^{sp}(\eta) \right] = P \left[ \hat{R}_{k} \mid R_{k} \land G^{sp}(\eta) \right] P \left[ R_{k} \land G^{sp}(\eta) \right]
\]
it suffices to show that for fixed \( M \) such that \( R_{k} \land G^{sp}(\eta) \) holds, we have
\[
P_{\mathcal{I}} \left( \hat{R}_{k} \right) \geq 1 - 2\eta k. \tag{4.6}
\]
where \( \mathcal{I} := (I_{1}, \ldots, I_{k}) \). We view the components of \( \mathcal{I} \) as being drawn sequentially from \([n]\).

Condition on \( M \) such that \( R_{k} \land G^{sp}(\eta) \) holds. For such \( M \), we may pick \( k \) linearly independent left null vectors \( y_{1}, \ldots, y_{k} \), so that for each \( j \in [k] \),
\[
\sum_{i \in [n]} y_{j}(i) R_{i} = 0.
\]
For \( \hat{R}_{k} \) to hold, it suffices that there exist \( Z = (z_{1}, \ldots, z_{k}) \in \mathbb{R}^{n \times k} \) with \( \text{span}(z_{1}, \ldots, z_{k}) = \text{span}(y_{1}, \ldots, y_{k}) \) such that the \( k \times k \) matrix
\[
Z_{\mathcal{I} \times [k]} = (z_{j}(I_{i}))_{1 \leq j, i \leq k} \tag{4.7}
\]
is upper triangular with nonzero diagonal entries. Indeed, this implies that \( R_{I_{1}}, \ldots, R_{I_{k}} \) can each be expressed as linear combinations of the rows \( \{ R_{i} : i \notin \{I_{1}, \ldots, I_{k}\} \} \).
Set \( z_1 = y_1 \). Let \( \mathcal{B}_1 = \{ z_1(I_1) = 0 \} \). For \( j \in [k-1] \), having defined linearly independent vectors \( z_1, \ldots, z_j \) and events \( \mathcal{B}_1, \ldots, \mathcal{B}_j \), on \( \bigwedge_{1 \leq l \leq j} \mathcal{B}_l \) we can find

\( z_{j+1} \in \text{span}(y_{j+1}, z_j, \ldots, z_1) \)

such that \( z_{j+1}(I_l) = 0 \) for all \( 1 \leq l \leq j \) (by linear independence). Let \( \mathcal{B}_{j+1} = \{ z_{j+1}(I_{j+1}) = 0 \} \).

Since \( z_{j+1} \in \ker(M^T) \), by our restriction to \( \mathcal{G}_{sp}(\eta) \) we have

\[
\mathbb{P}_{I_{j+1}} \left( \mathcal{B}_{j+1} \left\| \bigwedge_{1 \leq l \leq j} \mathcal{B}_l \right\| \right) \leq \frac{m}{n-j} \leq 2\eta
\]

(by the upper bound on \( k \)). By a union bound, \( \bigwedge_{1 \leq l \leq k} \mathcal{B}_l \) holds with probability at least \( 1 - 2\eta k \) in the randomness of \( \mathcal{I} \), and on this event the matrix (4.7) has the desired properties. \( \square \)

### 4.2. Warmup: Proof of Proposition 4.1.

As noted in Section 1.2, we may regard \( M_\pm \) as the Hadamard product \( M \odot \Xi \) of an rrd matrix \( M \) with an independent iid sign matrix \( \Xi \). We further note that all of the steps below apply for the more general case that \( M \) satisfies the properties listed in Theorem 1.11; here we will only use properties (1) (degrees are bounded below by \( d \)) and (4) (“no thin dense minors”).

We denote the rows of \( M \) by \( R_i \) and the rows of \( \Xi \) by \( Y_i \), so that the \( i \)th row of \( M \odot \Xi \) is \( R_i \odot Y_i \). For this section we will let

\[
\mathcal{R}_1 = \{ \text{corank}(M \odot \Xi) \geq 1 \}
\]

\[
V_I = \text{span}(R_i \odot Y_i : i \neq I),
\]

\[
\hat{\mathcal{R}}_1 = \{ R_I \odot Y_I \in V_I \}.
\]

From Lemma 4.3 (applied to \( M \odot \Xi \)) we have

\[
\mathbb{P}(\mathcal{R}_1) \leq \mathbb{P}(\mathcal{G}_{sp}(\eta)) + \frac{1}{1-2\eta} \mathbb{P}(\hat{\mathcal{R}}_1 \cap \mathcal{G}_{sp}(\eta)) \quad \text{(4.8)}
\]

where \( \mathcal{G}_{sp}(\eta) \) is the event that \( M \odot \Xi \) has no \((1-\eta)n\)-sparse left or right null vectors. By taking \( \eta \) sufficiently small, it now suffices to show

\[
\mathbb{P}(\hat{\mathcal{R}}_1 \cap \mathcal{G}_{sp}(\eta)) \ll \eta + d^{-1/2}. \quad \text{(4.9)}
\]

Draw \( u \) uniformly from the unit sphere in \( V_I^\perp \), in a way such that \( u, R_I \) and \( Y_I \) are jointly independent conditional on \( I \) and the remaining rows of \( M \) and \( \Xi \). Now it would be enough to show

\[
\mathbb{P}\left\{ (R_I \odot Y_I) \cdot u = 0 \right\} \ll \eta + d^{-1/2} \quad \text{(4.10)}
\]

From Theorem 2.1 we have

\[
\mathbb{P}_{Y_I}(\left( R_I \odot Y_I \right) \cdot u = 0 \mid M, I) \ll \left| \text{spt}(u) \cap \text{spt}(R_I) \right|^{-1/2} \quad \text{(4.11)}
\]

so we need to argue that \( \text{spt}(u) \) and \( \text{spt}(R_I) = \mathcal{N}_M(I) \) have large overlap.

Recall that \( |\mathcal{N}_M(i)| = d \) for all \( i \in [n] \). We identify the set of undesirable realizations of \( u \) as

\[
\mathcal{H}_{M,I} := \left\{ x \in \mathbb{R}^n : \left| \text{spt}(x) \cap \mathcal{N}_M(I) \right| \leq \frac{d}{2} \right\} \quad \text{(4.12)}
\]

and define the bad event

\[
\mathcal{B} = \{ \mathbb{P}_u(u \in \mathcal{H}_{M,I}) > 0 \}. \quad \text{(4.13)}
\]
We note that $B$ is decided by the randomness of $M$, $I$, and the rows $(Y_i)_{i \neq I}$ of $\Xi$. Now we divide our task as follows:

$$P(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B) = P\left(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B \right) + P\left(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B^c \right).$$

We begin with the first term on the right hand side of (4.14), which we will ultimately handle using the randomness of the index $I$. The crucial observation is that $H_{M,I}$ is a finite union of subspaces, each of co-dimension at least $d/2$:

$$H_{M,I} = \bigcup_{B \subset N_M(I): |B| \leq \frac{d}{2}} \mathbb{R}([n] \setminus N_M(I)) \cup B.$$

It follows that on $B$ we actually have $V^\perp_I \subset H_{M,I}$.

On $\mathcal{R}_1$ we may pick a nontrivial vector $x \in \ker(M \circ \Xi)$ (note that the kernel is nontrivial on this event) and we may do this independently of $I$. We have $x \in \ker(M \circ \Xi) \subset V^\perp_I$ and so on $B$ we have

$$x \in H_{M,I}$$

the key point being that the left hand side of the inclusion is independent of $I$. Summarizing our progress so far,

$$P\left(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B \right) \leq P\left(\mathcal{R}_1 \land G^\pm_{\pm}(\eta) \land \{x \in H_{M,I}\}\right)$$

where we used that $\hat{R}_1 \subset \mathcal{R}_1$.

We can bound the right hand side of (4.16) using the randomness of $I$ and our restriction to $G^\pm_{\pm}(\eta)$. Let

$$S_M(x) = \left\{ i \in [n]: |N_M(i) \cap \text{spt}(x)| \leq \frac{d}{2} \right\}.$$

Using the crude bound $e_M(A, B) \leq d|B|$ on edge counts we have

$$|S_M(x)| \frac{d}{2} \leq e_M(S_M(x), x^{-1}(0)) \leq d|x^{-1}(0)|$$

whence

$$|S_M(x)| \leq 2|x^{-1}(0)| \leq 2\eta n$$

by our restriction to $G^\pm_{\pm}(\eta)$. It follows that conditional on $M$ and $\Xi$ such that $\mathcal{R}_1$ holds,

$$P_I(x \in H_{M,I}) = P_I(I \in S_M(x)) \leq 2\eta$$

and so we conclude from (4.16) that

$$P\left(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B \right) \leq 2\eta.$$

It remains to bound the term $P\left(\hat{R}_1 \land G^\pm_{\pm}(\eta) \land B^c \right)$ from (4.14). Condition on $I$, $M$ and $(Y_i)_{i \neq I}$ such that $B$ does not hold. Off a null event we may assume that $u \notin H_{M,I}$. Now since

$$|\text{spt}(u) \cap N_M(I)| > \frac{d}{2}$$
applying (4.11) we have

\[
\mathbb{P}\left(\hat{R}_1 \land \mathcal{C}_{\pm}(\eta) \land \mathcal{B}^c\right) \leq \mathbb{P}\left(\left\{(R_I \odot Y_I) \cdot u = 0\right\} \land \left\{u \notin \mathcal{H}_{M,I}\right\}\right)
\]

\[
= \mathbb{E}\mathbb{P}_{Y_I}\left((R_I \odot Y_I) \cdot u = 0 \middle| M, I, (Y_i)_{i \neq I}\right)1_{u \notin \mathcal{H}_{M,I}}
\]

\[
\ll d^{-1/2}
\]

which combines with (4.19) to give the claim.

4.3. **Injecting a random walk.** We turn to the proof of Proposition 4.2.

Without the randomness of the independent signs enjoyed by $M_{\pm}$, we must now use the shuffling coupling of Lemma 3.1 to express $\hat{R}_k$ as the event that a random walk lands at a particular point. We define a coupled pair of rrd matrices $(M, \tilde{M})$ as in that lemma, but with the pair of rows selected randomly. That is, we draw:

1. an rrd matrix $M$,
2. $I_1, I_2 \in [n]$ sampled uniformly without replacement from $[n]$, independently of $M$,
3. a uniform random bijection $\pi : \text{Ex}_M(I_1, I_2) \to \text{Ex}_M(I_2, I_1)$,
4. a sequence $\xi = (\xi_j)_{j=1}^n$ of iid signs independent of all other variables.

We form $\tilde{M}$ by performing a shuffling on $M$ at the rows $(I_1, I_2)$ with respect to $\pi$, $\tilde{\xi}$. By Lemma 3.1 and conditioning on $I_1, I_2$ we have that $M \overset{d}{=} \tilde{M}$.

Now we wish to control the events

\[\hat{R}_1 = \{R_{I_1} \in V_{\{I_1\}}\}\]

and

\[\hat{R}_2 = \{R_{I_1}, R_{I_2} \in V_{\{I_1, I_2\}}\}\]

with notation as in (4.5). Note that on $\hat{R}_2$, $R_{I_1}$ and $R_{I_2}$ are orthogonal to any vector in the orthocomplement of $V_{\{I_1, I_2\}}$. We are hence interested in the dot products $R_{I_1} \cdot u, R_{I_2} \cdot u$ for $u$ taken from the unit sphere (say) of $V^\perp_{\{I_1, I_2\}}$. Let us examine the joint distribution of these
dot products when we replace $M$ by $\tilde{M}$. Letting $\tilde{R}_i$ denote the $i$th row of $\tilde{M}$, we can express

\[
\begin{align*}
(\tilde{R}_{i_1} \cdot u) \\
(\tilde{R}_{i_2} \cdot u)
\end{align*}
\]

\[
= \left( \sum_{j \in N_{\tilde{M}}(i_1)} u(j) \right) \left( \sum_{j \in N_{\tilde{M}}(i_2)} u(j) \right)
\]

\[
= \left( \sum_{j \in \text{Co}_M(i_1, i_2)} u(j) + \sum_{j \in \text{Ex}_M(i_1, i_2)} u(j) \right) \left( \sum_{j \in \text{Co}_M(i_1, i_2)} u(j) + \sum_{j \in \text{Ex}_M(i_1, i_2)} u(j) \right)
\]

\[
= \left( \sum_{j \in \text{Co}_M(i_1, i_2)} u(j) \right) \left( \sum_{j \in \text{Ex}_M(i_1, i_2)} u(j) \right) + \left( \sum_{j \in \text{Co}_M(i_1, i_2)} u(j) \right) \left( \sum_{j \in \text{Ex}_M(i_1, i_2)} u(j) \right)
\]

\[
= \left( \frac{1}{2} (R_{i_1} + R_{i_2}) \cdot u \right) \left( \frac{1}{1} \right) + \left( \sum_{j \in \text{Ex}_M(i_1, i_2)} \xi_j \partial_j^\pi (u) \right) \left( \frac{1}{1} \right) (1)
\]

\[
= A(u) \left( \frac{1}{1} \right) + W(u) \left( \frac{1}{1} \right)
\]

where in the penultimate line we have defined

\[
\partial_j^\pi (u) = \frac{u(j) - u(\pi(j))}{2}
\]

and collected the coefficients of $\binom{1}{1}$ and $\binom{1}{1}$.

Note that the term $A(u)$ is fixed by conditioning on $M, I_1, I_2$. Furthermore, the $\partial_j^\pi (u)$ are fixed by additionally conditioning on $\pi$. Hence, conditional on $M, I_1, I_2, \pi$, in the randomness of the $\xi_j$ this pair of dot products is a random walk in the $(1, -1)$ direction with steps $\partial_j^\pi (u)$.

The following lemma isolates the role of the randomness of the signs $\xi_j$ and reduces the problem to the study of structural properties of the normal vector $u$.

**Lemma 4.4** (The role of the signs $\xi_j$). For $u \in \mathbb{R}^n$, $i_1, i_2 \in [n]$ distinct, and a bijection

\[
\pi : \text{Ex}_M(i_1, i_2) \rightarrow \text{Ex}_M(i_2, i_1)
\]

define

\[
\text{Steps}(u) = \text{Steps}^{(i_1, i_2)}_{M, \pi} (u) := \{ j \in \text{Ex}_M(i_1, i_2) : u(j) \neq u(\pi(j)) \}.
\]

Then with $(M, \tilde{M})$ coupled as in Lemma 3.1 and $u$ deterministic or random depending only on $(R_i : i \notin \{i_1, i_2\})$ we have

\[
\mathbb{P}^\xi (\tilde{R}_{i_1} \cdot u = 0) \ll |\text{Steps}(u)|^{-1/2}.
\]

**Proof.** From the representation (4.20) we have

\[
\tilde{R}_{i_1} \cdot u = A(u) + \sum_{j \in \text{Ex}_M(i_1, i_2)} \xi_j \partial_j^\pi (u)
\]

and the claim follows by conditioning on $M, \pi$ and applying Theorem 2.1. $\square$
4.4. **Corank** ≥ 2. In this section we establish the bound (4.3).

From Lemma 4.3, for η sufficiently small it suffices to bound \( Pr(\hat{R}_2) \). Conditional on \( M, I_1, I_2 \), let \( u \) be drawn from the uniform surface measure of the unit sphere in \( V_{\{I_1, I_2\}}^{\perp} \), independently of \( \pi, R_{I_1}, R_{I_2} \). We have

\[
Pr(\hat{R}_2) \leq Pr(\hat{R}_{I_1} \in V_{\{I_1, I_2\}}) \\
\leq Pr(\hat{R}_{I_1} \cdot u = 0).
\]

We want to bound this using Lemma 4.4, so we will need to argue that \( \left| \text{Steps}^M_{\{I_1, I_2\}}(u) \right| \) is large. For this task we use the randomness of \( u, I_1, I_2, \pi, \) and restrict \( M \) to the good events \( G_{\text{ex}}(\delta) \) and \( G_{\text{sls}}(\eta) \).

First we identify the set of undesirable realizations of \( u \). Let

\[
H'_{M,I} = \left\{ x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ with } \min_{i=1,2} \left| \mathcal{N}_M(I_i) \cap x^{-1}(\lambda) \right| > d/100 \right\}.
\]

That is, \( H'_{M,I} \) is the set of vectors with a level set intersecting at least 1% of the support of both \( R_{I_1} \) and \( R_{I_2} \). Note that \( H'_{M,I} \) is a finite union of proper subspaces of \( \mathbb{R}^n \). Indeed, we may express

\[
H'_{M,I} = \bigcup_{(T_1,T_2)} \mathcal{H}_{T_1 \cup T_2}
\]

where the union ranges over pairs of subsets \( T_1 \subset \mathcal{N}_M(I_1), T_2 \subset \mathcal{N}_M(I_2) \) of size at least \( d/100 \), and \( \mathcal{H}_T \) denotes the subspace of vectors that are constant on \( T \). Define the bad event

\[
B_{M,I} = \left\{ \Pr(u \in H'_{M,I}) > 0 \right\}.
\]

(4.25)

Since \( H'_{M,I} \) is a finite union of proper subspaces, and \( u \) is drawn from the uniform surface measure of the subspaces \( V_{\{I_1, I_2\}}^{\perp} \), it follows that if \( B_{M,I} \) holds then we actually have the inclusion

\[
V_{\{I_1, I_2\}}^{\perp} \subset H'_{M,I}.
\]

Further noting that \( \ker(M) \subset V_{\{I_1, I_2\}}^{\perp} \), we obtain an inclusion with the left side independent of \( \mathcal{I} \):

\[
\ker(M) \subset H'_{M,I}.
\]

On \( \mathcal{R}_2 \), we may fix an arbitrary nontrivial element \( x \in \ker(M) \) (note that the kernel is nonempty on this event), independent of \( \mathcal{I} \). Now we have

\[
Pr(\mathcal{R}_2 \land B_{M,I}) \leq Pr(\mathcal{R}_2 \land \{ x \in H'_{M,I} \}).
\]

We will bound the latter quantity using the randomness of \( \mathcal{I} \). For \( \lambda \in \mathbb{R} \), let

\[
S_M(x, \lambda) = \left\{ i \in [n] : \left| \mathcal{N}_M(i) \cap x^{-1}(\lambda) \right| \geq d/100 \right\}.
\]

We can control the size of these sets using only a crude bound on edge counts:

\[
|S_M(x, \lambda)| d/100 < e_M(S_M(x, \lambda), x^{-1}(\lambda)) \leq d|x^{-1}(\lambda)|
\]

whence

\[
|S_M(x, \lambda)| \leq 100|x^{-1}(\lambda)|.
\]
Now for $M$ such that $\mathcal{G}^\text{abs}(\eta)$ holds we have $|x^{-1}(\lambda)| \leq \eta n$ for all $\lambda \in \mathbb{R}$. Conditional on $M$ such that $R_2$ and $\mathcal{G}^\text{abs}(\eta)$ hold (which fixes $x \neq 0$), we can bound

$$\mathbb{P}_I(x \in \mathcal{H}_M') = \mathbb{P}_I(\exists \lambda \in \mathbb{R} : I_1, I_2 \in S_M(x, \lambda)) \leq \sum_{\lambda : x^{-1}(\lambda) \neq \phi} \mathbb{P}_I(I_1, I_2 \in S_M(x, \lambda)) \leq \mathbb{P}_I(I_1, I_2 \in S_M(x, \lambda)) \leq \eta.$$

Undoing the conditioning on $M$, we have shown that

$$\mathbb{P}(B_{M,I} \cap R_2 \cap \mathcal{G}^\text{abs}(\eta)) \ll \eta. \quad (4.26)$$

It remains to bound

$$\mathbb{P}(\{R_1 \cdot u = 0\} \setminus B_{M,I}).$$

Condition on $M, I$ such that $B_{M,I}$ does not hold. Off a null event we may assume that $u \notin \mathcal{H}_M'$. That is, for every $\lambda \in \mathbb{R}$, we may assume

$$|\mathcal{N}_M(I_1) \cap u^{-1}(\lambda)| \leq d/100 \quad \text{or} \quad |\mathcal{N}_M(I_2) \cap u^{-1}(\lambda)| \leq d/100. \quad (4.27)$$

We will now get a lower bound on $|\text{Steps}(u)|$. It will be more convenient to work with the complementary set

$$\text{Flats}(u) := \text{Ex}_M(I_1, I_2) \setminus \text{Steps}(u) = \{ j \in \text{Ex}_M(I_1, I_2) : u(j) = u(\pi(j)) \}. \quad \text{(4.27)}$$

We have

$$\mathbb{E}_\pi |\text{Flats}(u)| = \sum_{j \in \text{Ex}_M(I_1, I_2)} \mathbb{P}_\pi(\pi(j) \in u^{-1}(u(j))) = \sum_{j \in \text{Ex}_M(I_1, I_2)} \frac{|u^{-1}(u(j)) \cap \text{Ex}_M(I_2, I_1)|}{|\text{Ex}_M(I_2, I_1)|} \leq \frac{1}{|\text{Ex}_M(I_1, I_2)|} \sum_{\lambda : u^{-1}(\lambda) \neq \phi} |u^{-1}(\lambda) \cap \text{Ex}_M(I_1, I_2)| |u^{-1}(\lambda) \cap \text{Ex}_M(I_2, I_1)|$$

where the last line follows from double counting. Now we apply (4.27) to get

$$\mathbb{E}_\pi |\text{Flats}(u)| \leq \frac{d}{100 |\text{Ex}_M(I_1, I_2)|} \sum_{\lambda : u^{-1}(\lambda) \neq \phi} \max(|u^{-1}(\lambda) \cap \text{Ex}_M(I_1, I_2)|, |u^{-1}(\lambda) \cap \text{Ex}_M(I_2, I_1)|) \leq \frac{d}{100 |\text{Ex}_M(I_1, I_2)|} \sum_{\lambda : u^{-1}(\lambda) \neq \phi} |u^{-1}(\lambda) \cap [\text{Ex}_M(I_1, I_2) \cup \text{Ex}_M(I_2, I_1)]| \leq \frac{d}{50}. \quad (4.28)$$
We want to show that $|\text{Flats}(u)|$ is concentrated around its expectation (we only need control on the right tail). In the notation of Lemma 3.8 we have

$$|\text{Flats}(u)| = e_\pi(A, B)$$

with $A = \text{Ex}_M(I_1, I_2)$ and $B = u^{-1}(u(j_1))$ (which are fixed by conditioning on $M, I_1, I_2$). Applying Lemma 3.8 and (4.28) we conclude that for any $\varepsilon > 0$,

$$\mathbb{P}_\pi\left(|\text{Flats}(u)| \geq (1 + \varepsilon) \frac{d}{50}\right) \leq \exp\left(-\frac{c\varepsilon^2 d^2}{50^2}\right). \quad (4.29)$$

On the other hand, on $\mathcal{G}_{\text{ex}}(\delta)$ we have

$$|\text{Ex}_M(I_1, I_2)| \geq (1 - \delta)d\left(1 - \frac{d}{\eta}\right) \geq \frac{1 - \delta}{2}d$$

so that fixing $\delta \leq 1/2$ and $\varepsilon \leq 4$ (say), we conclude that on $\mathcal{B}^c_{M,I} \land \mathcal{G}_{\text{ex}}(\delta)$, except with probability at most $\exp(-cd)$ we have

$$|\text{Steps}(u)| = |\text{Ex}_M(I_1, I_2)| - |\text{Flats}(u)| \geq \frac{d}{10}. \quad (4.30)$$

Summarizing our work so far,

$$\mathbb{P}\left(\hat{\mathcal{R}}_2 \land \mathcal{G}_{\text{sls}}(\eta) \land \mathcal{G}_{\text{ex}}(\delta)\right) \leq \mathbb{P}\left(\hat{\mathcal{R}}_2 \land \mathcal{B}_{M,I} \land \mathcal{G}_{\text{sls}}(\eta)\right) + \mathbb{P}\left(\hat{\mathcal{R}}_2 \land \mathcal{B}^c_{M,I} \land \mathcal{G}_{\text{sls}}(\eta) \land \mathcal{G}_{\text{ex}}(\delta)\right)$$

$$\leq \mathbb{P}\left(\hat{\mathcal{R}}_2 \land \mathcal{B}_{M,I} \land \mathcal{G}_{\text{sls}}(\eta)\right) + \mathbb{P}\left(\mathcal{B}^c_{M,I} \land \left\{|\text{Steps}(u)| < \frac{d}{10}\right\} \land \mathcal{G}_{\text{ex}}(\delta)\right)$$

$$\ll \eta + e^{-cd} + d^{-1/2}$$

where in the last line we substituted our bounds (4.30) and (4.26) and applied Lemma 4.4. Combined with our estimates for $\mathcal{G}_{\text{sls}}(\eta)$ and $\mathcal{G}_{\text{ex}}(\delta)$, together with Lemma 4.3 we have

$$\mathbb{P}(\mathcal{R}_2) \ll \eta + d^{-1/2} \quad (4.31)$$

as desired.

4.5. **Corank** = 1. In this section we establish the bound (4.4).

By Lemma 4.3 it suffices to bound

$$\mathbb{P}(\hat{\mathcal{R}}_1 \setminus \mathcal{R}_2) = \mathbb{P}\left(R_{I_2} \in V(I_1), \ \text{corank}(M) \leq 1\right)$$

(taking $\eta$ smaller if necessary). We cannot simply condition on all rows but $R_{I_1}$ and pick a normal vector $u \in V_{I_1}^\perp$, since this conditioning fixes $R_{I_2}$ as well by $d$-regularity. Instead, we will leave $R_{I_2}$ random and express the event $\hat{\mathcal{R}}_1$ in terms of a certain $2 \times 2$ determinant. We now have the advantage of being able to condition on a unique (up to dilation) null vector $x$ of $M$, which is independent of $I$. 

We turn to the details. Conditional on \( M, I_1, I_2 \), we pick a pair of orthonormal vectors \( u_1 \perp u_2 \in V_{\{I_1, I_2\}}^\perp \) uniformly at random, and independently of \( (R_{I_1}, R_{I_2}) \). (On \( \tilde{R}_1 \land R_2^\circ \) we have \( \dim(V_{\{I_1, I_2\}}) = n - 2 \) and so \( u_1, u_2 \) are an orthonormal basis for \( V_{\{I_1, I_2\}}^\perp \) on this event.) In terms of \( u_1, u_2 \) we may construct a vector which is also orthogonal to \( R_{I_2} \) as follows:
\[
z_I^{(1)} := (R_{I_2} \cdot u_2)u_1 - (R_{I_2} \cdot u_1)u_2 \in V_{\{I_1\}}^\perp.
\]
Since \( z_I^{(1)} \) lies in the orthocomplement of \( V_{\{I_1\}} \), on \( \tilde{R}_1 \) we have
\[
0 = R_{I_1} \cdot z_I^{(1)}
= (R_{I_1} \cdot u_1)(R_{I_2} \cdot u_2) - (R_{I_2} \cdot u_1)(R_{I_1} \cdot u_2)
= : D_M(I_1, I_2).
\]
Hence,
\[
\mathbb{P} (\tilde{R}_1 \land R_2^\circ) \leq \mathbb{P} (D_M(I_1, I_2) = 0, \ \text{corank}(M) = 1).
\] (4.33)

Substituting \( \tilde{M} \) for \( M \), may express the \( 2 \times 2 \) determinant using (4.21):
\[
D_{\tilde{M}}(I_1, I_2) = \begin{bmatrix} A(u_1) + W(u_1) \end{bmatrix} \begin{bmatrix} A(u_2) - W(u_2) \end{bmatrix} - \begin{bmatrix} A(u_2) + W(u_2) \end{bmatrix} \begin{bmatrix} A(u_1) - W(u_1) \end{bmatrix}
= 2A(u_2)W(u_1) - 2A(u_1)W(u_2)
= \sum_{j \in \text{Ex}_M(I_1, I_2)} \xi_j [2A(u_2)\partial_j^\circ(u_1) - 2A(u_2)\partial_j^\circ(u_2)]
= \sum_{j \in \text{Ex}_M(I_1, I_2)} \xi_j \partial_j^\circ(v_I)
= W(v_I)
\] (4.34) (4.35)

where we have defined
\[
v_I := 2A(u_2)u_1 - 2A(u_2)u_2
= \left[ (R_{I_1} + R_{I_2}) \cdot u_2 \right] u_1 - \left[ (R_{I_1} + R_{I_2}) \cdot u_1 \right] u_2
\] (4.36)
\[
\in V_{\{I_1, I_2\}}^\perp.
\]

We would like to replace \( M \) with \( \tilde{M} \) and bound (4.33) using the random walk representation (4.35) with Theorem 2.1. First we must reduce to an event on which many of the steps \( \partial_j^\circ(v_I) \) are nonzero. We will do this in two stages. First we must remove a bad event on which \( v_I = 0 \); in light of (4.36) this is the event
\[
\mathcal{B}_{M,I}^\prime := \{ R_{I_1} + R_{I_2} \in V_{\{I_1, I_2\}} \}.
\]

Once we have done this, we will be able to argue that \( v_I \) is unstructured in a manner similar to the way we controlled the event \( \mathcal{B}_{M,I} \) in Section 4.4.

We begin with \( \mathcal{B}_{M,I}^\prime \). Since we are free to restrict to \( R_2^\circ \land G^\text{dis}(\eta) \), let us condition on \( M \) such that these events hold. On \( \mathcal{B}_{M,I}^\prime \land R_2^\circ \), \( M \) has exactly one nontrivial left null vector (up to dilation) which we denote by \( y \); furthermore, on \( G^\text{dis}(\eta) \) the level sets of \( y \) are of size at most \( \eta n \). Now \( \mathcal{B}_{M,I}^\prime \) is the event that \( M \) has a left null vector \( y' \) with \( y'(I_1) = y'(I_2) \neq 0 \), so we have
\[
\mathcal{B}_{M,I}^\prime \land R_2^\circ \subset \{ y(I_1) = y(I_2) \}.
\]
It follows that
\[
\mathbb{P}_I \left( \mathcal{B}'_{M,I} \cap \mathcal{R}_x^d \cap \mathcal{G}_{\eta}^{\text{sln}} \right) \leq \mathbb{P}_I \left( y(I_1) = y(I_2) \right)
\leq \eta(1 + o(1))
\] (4.37)
which is small enough.

As in Section 4.4, let
\[
\mathcal{H}'_{M,I}(\varepsilon) = \left\{ x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ with } \min_{l=1,2} |x^{-1}(\lambda) \cap \mathcal{N}_M(I_l)| > \varepsilon d \right\}
\] (4.38)
and define also
\[
\mathcal{H}''_{M,I}(\varepsilon) = \left\{ x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ with } |x^{-1}(\lambda) \cap \left[ \text{Ex}_M(I_1, I_2) \cup \text{Ex}_M(I_2, I_1) \right]| > \varepsilon_0 d \left(1 - \frac{d}{n} \right) \right\}.
\] (4.39)
For \( \varepsilon \in (0,1) \) we have the inclusion
\[
\mathcal{H}''_{M,I}(1 + 2\varepsilon) \subset \mathcal{H}'_{M,I}(2\varepsilon(1 - p)) \subset \mathcal{H}''_{M,I}(\varepsilon).
\] (4.40)
(4.41)
The advantage of the set \( \mathcal{H}''_{M,I}(\varepsilon) \) is that it is the same for \( M \) and \( \bar{M} \):
\[
\mathcal{H}''_{M,I}(\varepsilon) = \mathcal{H}''_{\bar{M},I}(\varepsilon)
\] (4.42)
for any \( \varepsilon > 0 \), whereas this does not hold for \( \mathcal{H}'_{M,I}(\varepsilon) \).

Let \( \varepsilon > 0 \) to be chosen later. On \( \{ \text{corank}(M) = 1 \} \), let \( x \) denote a fixed nontrivial null vector of \( M \), so that \( \ker(M) = \langle x \rangle \). From (4.32), on \( \{ D_M(I_1, I_2) = 0 \} \) we have \( z^{(1)}_I, z^{(2)}_I \in \ker(M) \), where
\[
z^{(2)}_I = (R_{I_1} \cdot u_1)u_2 - (R_{I_1} \cdot u_2)u_1 \in V^\perp_{(I_2)}
\]
and (as before)
\[
z^{(1)}_I = (R_{I_2} \cdot u_2)u_1 - (R_{I_2} \cdot u_1)u_2 \in V^\perp_{(I_1)}.
\]
We have
\[
v_I = z^{(1)}_I - z^{(2)}_I.
\]
On \( \{ D_M(I_1, I_2) = 0 \} \land \{ \text{corank}(M) = 1 \} \), both \( z^{(1)}_I \) and \( z^{(2)}_I \) are nonzero and lie in \( \ker(M) = \langle x \rangle \). It follows that
\[
v_I \in \ker(M)
\]
on this event. Hence,
\[
\mathcal{B}'_{M,I} \land \{ D_M(I_1, I_2) = 0, \text{corank}(M) = 1, v_I \in \mathcal{H}'_{M,I}(\varepsilon) \} \subset \{ x \in \mathcal{H}'_{M,I}(\varepsilon) \}.
\] (4.43)
We may now argue exactly as in Section 4.4 to conclude
\[
\mathbb{P}_I \left( \mathcal{G}_{\eta}^{\text{sln}} \land \{ x \in \mathcal{H}'_{M,I}(\varepsilon) \} \right) \ll_{\varepsilon} \eta.
\] (4.44)
It only remains to bound
\[
\mathbb{P} \left( D_M(I_1, I_2) = 0, v_I \notin \mathcal{H}'_{M,I}(\varepsilon) \right).
\] (4.45)
From (4.40) this is bounded by
\[
\mathbb{P} \left( D_M(I_1, I_2) = 0, v_I \notin \mathcal{H}''_{M,I}(1 + 2\varepsilon) \right).
\]
Now we replace $M$ with $\tilde{M}$. We make the crucial observation that the second event is unchanged by this substitution. Indeed,

$$v_T = [(R_{I_1} + R_{I_2}) \cdot u_2] u_1 - [(R_{I_1} + R_{I_2}) \cdot u_1] u_2$$

and $R_{I_1} + R_{I_2} = \tilde{R}_{I_1} + \tilde{R}_{I_2}$ since the shuffling preserves the sets $N_M(I_1) \cap N_M(I_2)$ and $N_M(I_1) \cup N_M(I_2)$. Similarly note that $H''_{M,I}$ is determined by $E_M(I_1, I_2) \cup E_M(I_2, I_2) = E_{\tilde{M}}(I_1, I_2) \cup E_{\tilde{M}}(I_2, I_2)$. Hence, (4.45) is bounded by

$$\mathbb{P} \left( D_M(I_1, I_2) = 0, \ v_T \notin H''_{M,I}(1 + 2\varepsilon) \right) = \mathbb{P} \left( W(v_T) = 0, \ v_T \notin H''_{M,I}(1 + 2\varepsilon) \right). \tag{4.46}$$

In the final step of the argument, we must show that Steps$(v_T)$ is large on $\{ v_T \notin H''_{M,I}(1 + 2\varepsilon) \}$ with high probability in the randomness of $\pi$ (and taking $\varepsilon$ sufficiently small). We have

$$\mathbb{E}_\pi \left| \text{Flats}(v_T) \right| = \sum_{j \in E_M(I_1, I_2)} \mathbb{P}_\pi \left( \pi(j) \in v_T^{-1}(v_T(j)) \right)$$

$$= \sum_{j \in E_M(I_1, I_2)} \frac{\left| v_T^{-1}(v_T(j)) \cap E_M(I_2, I_1) \right|}{\left| E_M(I_2, I_1) \right|}$$

$$= \frac{1}{\text{ex}_M(I_1, I_2)} \sum_{\lambda: v_T^{-1}(\lambda) \neq \emptyset} \left| v_T^{-1}(\lambda) \cap E_M(I_1, I_2) \right| \left| v_T^{-1}(\lambda) \cap E_M(I_2, I_1) \right|$$

$$\leq \frac{1}{\text{ex}_M(I_1, I_2)} \sum_{\lambda: v_T^{-1}(\lambda) \neq \emptyset} \frac{1}{4} \left| v_T^{-1}(\lambda) \cap (E_M(I_1, I_2) \cup E_M(I_2, I_1)) \right|^2$$

$$\leq \frac{(1 + 2\varepsilon)p(1 - p)n}{4 \text{ex}_M(I_1, I_2)} \sum_{\lambda: v_T^{-1}(\lambda) \neq \emptyset} \left| v_T^{-1}(\lambda) \cap (E_M(I_1, I_2) \cup E_M(I_2, I_1)) \right|$$

$$\leq \frac{1 + 2\varepsilon}{2} d \left( 1 - \frac{d}{n} \right).$$

We can then argue exactly as in (4.29) that

$$\mathbb{P}_\pi \left( \left| \text{Flats}(v_T) \right| \geq \left( \frac{1}{2} + 2\varepsilon \right) d \left( 1 - \frac{d}{n} \right) \right) \leq \exp \left( -\frac{c \varepsilon^2}{1 + \varepsilon} d \right). \tag{4.47}$$

On the other hand, on $\mathcal{G}^{\text{ex}}(\delta)$ we have

$$\left| \text{Ex}_M(I_1, I_2) \right| \geq (1 - \delta) d \left( 1 - \frac{d}{n} \right)$$

so that if we take $\varepsilon$ and $\delta$ sufficiently small,

$$\left| \text{Steps}(v_T) \right| = \left| \text{Ex}_M(I_1, I_2) \right| - \left| \text{Flats}(v_T) \right| \gg d. \tag{4.48}$$

Applying Lemma 4.4,

$$\mathbb{P} \left( \{ W(v_T) = 0 \} \land \{ v_T \notin H''_{M,I}(1 + 2\varepsilon) \} \land \mathcal{G}^{\text{ex}}(\delta) \right) \ll e^{-cd} + d^{-1/2}$$

which combines with (4.37) and (4.44) to give

$$\mathbb{P}(\mathcal{R}_1 \land \mathcal{R}_2 \land \mathcal{G}^{\text{sls}}(\eta)) \ll \eta + d^{-1/2}$$

as desired. □
5. Structured null vectors

Our aim in the section is to prove Propositions 2.5 and 2.4. The latter appears first as it is easier and outlines the the approach to the former. The difference is that for Proposition 2.5 we will need Lemma 3.1 to inject random walks, and we will make heavier use of the discrepancy properties from Section 3.2.

5.1. Warmup: no sparse null vectors for $M_{\pm}$. In this section we prove Proposition 2.4. As in Section 4.2, we regard $M_{\pm}$ as the Hadamard product $M \circ \Xi$ of an rrd matrix $M$ with an independent iid sign matrix $\Xi$.

Let $\varepsilon \in (0,1)$ and $\gamma > 0$ to be chosen later independent of $n$. By Corollary 3.5 and Corollary 3.7 we may restrict to the events $G^i(\varepsilon)$ and $G^{\text{exp}}(\gamma)$ for $M$.

For $k \in [n]$, let
\[ \mathcal{E}_k = \{ \exists x \in \mathbb{R}^n : 0 < |\text{spt}(x)| \leq k, \ (M \circ \Xi)x = 0 \} \]
We have
\[ \mathbb{P}(G_{\pm}^k(\eta)^c) = \mathbb{P}(\mathcal{E}_{(1-\eta)n}) = \sum_{k=2}^{(1-\eta)n} \mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \] (5.1)
(noting that $\mathcal{E}_1$ is empty). Fix $1 \leq k \leq (1-\eta)n$. Since the rows and columns of $M \circ \Xi$ are exchangeable, we can follow the same lines as the proof of Proposition 2.2 to bound
\[ \mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \leq \binom{n}{k} \binom{n}{k-1} \mathbb{P}(\mathcal{E}_k^c) \] (5.2)
where
\[ \mathcal{E}_k^c = \{ \exists x \in \mathbb{R}^n : (M \circ \Xi)x = 0, \text{spt}(x) = [k], \ {R_i \circ Y_i}_{i=1}^{k-1} \text{ are linearly independent} \} \]

Fix an arbitrary $v = \hat{v}1_{[k]} \in \mathbb{R}^n$ with support $[k]$. Let $\hat{R}_i = R_i1_{[k]}$ and $\hat{Y}_i = Y_i1_{[k]}$ denote the restrictions of the $i$th rows of $M$ and $\Xi$, respectively, to the first $k$ coordinates. Since conditioning on the first $k - 1$ rows of $M \circ \Xi$ fixes $x$ on $\mathcal{E}_k^c$, it suffices to get a bound on
\[ \mathbb{P}\left( (M \circ \Xi)[k,n] \times [k]\hat{v} = 0 \ \middle| \ M, Y_1, \ldots, Y_{k-1} \right) \]
uniform in $\hat{v}$.

Our approach is different depending on whether $k$ is small or large. In both cases, we use the fact that the rows $R_i \circ Y_i$ decouple after conditioning on $M$:
\[ \mathbb{P}\left( (M \circ \Xi)[k,n] \times [k]\hat{v} = 0 \ \middle| \ M \right) = \prod_{i=k+1}^{n} \mathbb{P}\left( (\hat{R}_i \circ \hat{Y}_i) \cdot \hat{v} = 0 \ \middle| \ M \right) \]
\[ = \prod_{i=k+1}^{n} \mathbb{P}\left( \hat{Y}_i \cdot (\hat{R}_i \circ \hat{v}) = 0 \ \middle| \ M \right). \]
For small $k$, it will be enough to show that there are many $i \in [k+1,n]$ such that
\[ |\text{spt}(R_i) \cap [k]| = |\text{spt}(\hat{R}_i)| \geq 1. \] (5.3)
For such $i$ we have $\mathbb{P}\left( \hat{Y}_i \cdot (\hat{R}_i \circ \hat{v}) = 0 \ \middle| \ M \right) \leq 1/2$; indeed, since $\hat{v}(j) \neq 0$ for all $j \in [k]$, (5.3) implies the random walk $\hat{Y}_i \cdot (\hat{R}_i \circ \hat{v})$ takes at least one non-zero step. To lower bound the
number of rows satisfying (5.3) we will use the “expansion of small sets” property enforced by our restriction to $G_{\exp}(\gamma)$.

For larger $k$ we will need to argue that the random walks $\hat{Y}_i \cdot (\hat{R}_i \odot \hat{v})|M$ take more steps. For this we prove a consequence of the discrepancy property enforced by $G_{\exp}(\varepsilon)$ (5.1 below), which essentially guarantees that the intersection of any sufficiently large set $B$ with the neighborhoods $N_M(i) = \text{spt}(R_i)$ has roughly the expected size $p|B|$ for most $i \in [k + 1, n]$. Applying this with $B = [k]$ gives $|\text{spt}(R_i) \cap [k]| \gg pk$ for most $i \in [k + 1, n]$. We will build on this idea in the proof of Proposition 2.5 for the unsigned rrd matrix $M$, where we will also need that a large set $B$ “sees” roughly the expected portion of the sets $E_{X_M}(i_1, i_2)$.

We turn to the details. First assume $k \leq \frac{1}{2\gamma} \frac{n \log n}{d}$. Let

$$A_0 = \{ i \in [k + 1, n] : \hat{R}_i \neq 0 \}.$$  

By our restriction to $G_{\exp}(\gamma)$ we have

$$|A_0| = N_M([k]) \setminus [k] \geq \frac{dk}{\log n} - k.$$ (5.4)

Since $\hat{v}(j) \neq 0$ for all $j \in [k]$, we have that for $i \in A_0$,

$$\mathbb{P} \left( \hat{Y}_i \cdot (\hat{R}_i \odot \hat{v}) = 0 \bigg| M, Y_1, \ldots, Y_{k-1} \right) \leq \frac{1}{2}$$

whence

$$\mathbb{P} \left( (M \odot \Xi)_{[k,n] \times [k]} \hat{v} = 0 \bigg| M, Y_1, \ldots, Y_{k-1} \right) \leq \prod_{i \in A_0} \mathbb{P} \left( \hat{Y}_i \cdot (\hat{R}_i \odot \hat{v}) = 0 \bigg| M, Y_1, \ldots, Y_{k-1} \right)$$

$$\leq \left( \frac{1}{2} \right)^{|A_0|} \left( \frac{1}{2} \right)^{\frac{dk}{\log n} - k}.$$  

Since this bound is uniform in $\hat{v}$, for $R_i \odot Y_i$ we conclude from (5.2) that

$$\mathbb{P}(E_k \setminus E_{k-1}) \leq \binom{n}{k} \binom{n}{k-1} \left( \frac{1}{2} \right)^{\frac{dk}{\log n} - k}$$

$$\leq \exp \left( Ck \log n - c\gamma \frac{kd}{\log n} \right)$$

$$\ll \exp \left( -\Omega(\gamma \frac{kd}{\log n}) \right)$$ (5.5)

where we used the assumption $d = \omega(\log^2 n)$. Now by the assumptions that $d = \omega(\log^2 n)$ and $\gamma > 0$ is independent of $n$, summing the bounds (5.5) gives

$$\mathbb{P}(E_k) \ll n^{-\omega(1)}$$ (5.6)

for any $k \leq \frac{1}{2\gamma} \frac{n \log n}{d}$.

Now assume $k > \frac{1}{2\gamma} \frac{n \log n}{d}$. For this case we apply the following consequence of our restriction to $G_{\exp}(\varepsilon)$:

Claim 5.1. For $A, B \subset [n]$ and $\varepsilon \in (0, 1)$ let

$$A_\varepsilon = \{ i \in A : |B(i)| \geq (1 - \varepsilon)p|B| \}$$
where we use the shorthand \( B(i) := B \cap N_M(i) \) and \( B^c = [n] \setminus B \). On \( G_\varepsilon^\epsilon(\varepsilon) \), we have
\[
|A \setminus A_\varepsilon| \ll \varepsilon p^{-1} \log n
\]
if \( |B| \geq \frac{\log n}{27p} \) with \( \gamma \) sufficiently small depending on \( \varepsilon \).

**Proof.** Denote \( S = A \setminus A_\varepsilon \). We claim \( (S, B) \notin F(\varepsilon) \). Indeed, if this were not the case we would have
\[
(1 - \varepsilon)p|S||B| \leq e_{M(S, B)} = \sum_{i \in S} |B(i)| < (1 - \varepsilon)p|S||B|
\]
a contradiction.

Suppose \( |B| \leq |S| \). Since \( (S, B) \notin F(\varepsilon) \) we have
\[
\frac{1}{2} \frac{n \log n}{d} \leq |B| \ll \varepsilon \frac{n \log n}{d}.
\]

Taking \( \gamma \) sufficiently small depending on \( \varepsilon \) we obtain a contradiction, and so \( |S| \leq |B| \); and by the definition of \( F(\varepsilon) \) we must have
\[
|S| \ll \varepsilon \frac{n \log n}{d}.
\]

□

Let us fix \( \varepsilon = 1/2 \). Applying the claim with \( A = [k + 1, n] \), \( B = [k] \), we have that for all \( i \in A_\varepsilon \), \( |\text{spt}(\hat{R}_i)| \geq pk/2 \), and so by Theorem 2.1,
\[
\mathbb{P}\left(\hat{Y}_i \cdot (\hat{R}_i \circ \hat{v}) = 0 \mid M, Y_1, \ldots, Y_{k-1}\right) \ll (pk)^{-1/2}.
\]

It follows that
\[
\mathbb{P}\left( (M \circ \Xi)[k,n]_{\times[k]} \hat{v} = 0 \mid M, Y_1, \ldots, Y_{k-1}\right) \leq \prod_{i \in A_\varepsilon} \mathbb{P}\left(\hat{Y}_i \cdot (\hat{R}_i \circ \hat{v}) = 0 \mid M, Y_1, \ldots, Y_{k-1}\right)
\]
\[
\leq \left[ \frac{C}{\sqrt{pk}} \right]^{n-k-O\left(\frac{n \log n}{d}\right)}.
\]

For \( \frac{n \log n}{d} \ll k \leq \frac{n}{2} \) this expression is bounded by
\[
O\left( \exp\left(-cn \log \log n\right)\right)
\]
which combines with Lemma 5.3 to give
\[
\mathbb{P}(E_k \setminus E_{k-1}) \ll 4^n \exp\left(-cn \log \log n\right)
\]
\[
= O\left( \exp\left(-cn \log \log n\right)\right).
\]

For \( \frac{n}{2} \leq k \leq (1 - \eta)n \) we instead bound (5.8) by
\[
O\left( \exp\left(-\frac{1}{2}(n - k) \log d\right)\right)
\]
\( \geq C \frac{\log n}{d} \) for \( C > 0 \) sufficiently large. With Lemma 5.3 we conclude

\[
\mathbb{P}(E_k \setminus E_{k-1}) \ll \left( \frac{n}{n-k} \right)^2 \exp \left( -\frac{1}{2} (n-k) \log d \right)
\]

\[
\leq \left( \frac{en}{n-k} \right)^{2(n-k)} d^{-(n-k)/2}
\]

\[
\leq \left( \frac{e}{\eta d^{1/4}} \right)^{2\eta n}
\]

\[
\ll \exp (-c\eta n)
\]

assuming \( \eta \) is at least a sufficiently large multiple of \( d^{-1/4} \).

Summing the bounds (5.5), (5.9), (5.10) over \( 1 \leq k \leq (1-\eta)n \) completes the proof.

5.2. Preliminary reductions. We now turn to the unsigned rrd matrix \( M \) and the proof of Proposition 2.5. Recall our notation for the level sets of a vector \( x \in \mathbb{R}^n \):

\[
x^{-1}(\lambda) := \{ i \in [n] : x(i) = \lambda \}
\]

for \( \lambda \in \mathbb{R} \). Our aim is to show that the good event

\[
G_{\text{sls}}^\eta := \left\{ \forall x \in \mathbb{R}^n : Mx = 0 \text{ or } M^T x = 0, \forall \lambda \in \mathbb{R}, \text{ we have } |x^{-1}(\lambda)| \leq \eta n \right\}
\]

holds with overwhelming probability for any \( d^{-c} \ll \eta \leq 1 \), for some sufficiently small absolute constant \( c > 0 \). Let

\[
G_1^\eta = \left\{ \forall 0 \neq x \in \mathbb{R}^n \text{ such that } Mx = 0 , \forall \lambda \in \mathbb{R}, \ |x^{-1}(\lambda)| \leq \eta n \right\}.
\]

(5.11)

Since \( M = M^T \), by a union bound it suffices to show that \( G_1^\eta(M) \) holds with overwhelming probability. The following claim recasts this event as the event that there is a sparse vector that is mapped by \( M \) to a constant vector.

**Claim 5.2.** For \( \eta \in (0, 1] \), let

\[
B(\eta) = \left\{ \exists 0 \neq y \in \mathbb{R}^n : |\text{spt}(y)| \leq (1-\eta)n, My \in \{0, 1\} \right\}
\]

where we recall that \( 1 \in \mathbb{R}^n \) is the vector with all components equal to 1. We have \( G_1^\eta(\eta)^c = B(\eta) \) for all \( \eta \in (0, 1] \).

(We actually only need the containment \( G_1^\eta(\eta)^c \subset B(\eta) \).)

**Proof.** Suppose that \( G_1^\eta(\eta) \) fails. Then there exists a nontrivial null vector \( x \in \mathbb{R}^n \) and \( \lambda \in \mathbb{R} \) such that \( |x^{-1}(\lambda)| > \eta n \). Let \( y = \lambda 1 - x \). Then \( y \) is nontrivial and \( |\text{spt}(y)| < (1-\eta)n \). Moreover,

\[
My = \lambda M 1 - M x = \lambda d 1 \in \{1\}
\]

so by dilating \( y \) we see that \( B(\eta) \) holds.

Conversely, suppose that \( B(\eta) \) holds. Then there exists a nontrivial vector \( y \) supported on at most \( (1-\eta)n \) coordinates such that \( My \) is either 0 or 1. If \( My = 0 \) then we are in \( G_1^\eta(\eta)^c \) (simply taking \( x = y \) and \( \lambda = 0 \)). So assume \( My = 1 \). Now letting \( x = y - \frac{1}{d} \), we have that \( x \) is a right null vector of \( M \) with \( |x^{-1}(1/d)| > \eta n \), so we are in \( G_1^\eta(\eta)^c \). \( \square \)
To prove Proposition 2.5 it now suffices to show that \( B(\eta)^c \) holds with overwhelming probability. Letting

\[
\mathcal{E}_k := \{ \exists y \in \mathbb{R}^n : |\text{spt}(y)| = k, MY \in \{0, 1\} \}
\]

we can decompose the bad event as

\[
\mathbb{P}(B(\eta)) = \sum_{k=2}^{(1-\eta)n} \mathbb{P}(\mathcal{E}_k \ \& \ \mathcal{E}_{k-1})
\]

(note that \( \mathcal{E}_1 \) is empty since no column can be parallel to 0 or 1). It now suffices to show that each of the events \( (\mathcal{E}_k \ \& \ \mathcal{E}_{k-1})^c \) holds with overwhelming probability for \( 1 \leq k \leq (1-\eta)n \), as long as \( \eta \gg n^{-c} \) for some \( c > 0 \).

The following lemma is analogous to the bound (5.2) from the proof of Proposition 2.4. The proof is lengthier but follows similar reasoning, and is deferred to Appendix A.

**Lemma 5.3** (Passing to a large minor). For \( k \in [n] \), let

\[
W_k = \{ \hat{v} \in \mathbb{R}^k : v(j) \neq 0 \ \forall j \in [k] \}
\]

be the set of vectors in \( \mathbb{R}^k \) with full support. Suppose that we have a bound

\[
\mathbb{P}\left( M_{[k+1,n] \times [k]} \hat{v} = \alpha \mathbf{1} \ \bigg| \ R_1, \ldots, R_k \right) \leq Q_k
\]

that is uniform in the choice of \( \hat{v} \in W_k, \alpha \in \{0, 1\} \) and the realization \( R_1, \ldots, R_k \) of the first \( k \) rows of \( M \). Then we have

\[
\mathbb{P}(\mathcal{E}_k \ \& \ \mathcal{E}_{k-1}) \ll \left( \frac{n}{k} \right)^2 Q_k.
\]

As in the proof of Proposition 2.4, our approach to bounding \( Q_k \) will be different depending on whether \( k \) is small. We want to control the event

\[
\{ M_{[k+1,n] \times [k]} \hat{v} = \alpha \mathbf{1} \} = \bigwedge_{i=k+1}^n \{ R_i \cdot v = \alpha \}
\]

where \( v = (\hat{v} \ 0) \) extends \( \hat{v} \) to a vector in \( \mathbb{R}^n \). Recall that in Section 5.1 we did this by conditioning on \( M \) and using the randomness of the signs. We could then view (5.18) (with \( R_i \odot Y_i \) in place of \( R_i \)) as the event that several independent random walks all land at the same point. We then used the discrepancy and “expansion of small sets” properties enjoyed by the base matrix \( M \) to argue that a large number of the walks \( R_i \cdot v \) take a large number of steps. When \( k \) was small it was enough to know that most walks take at least one step, while for larger \( k \) we needed more steps.

Here we will “inject” random walks into the distribution of the dot products \( R_i \cdot v \) by applying the shuffling coupling of Lemma 3.1. For small \( k \), we will apply shufflings to pairs of columns \( X_{j_1}, X_{j_2} \), with \( j_1 \in [k] \) and \( j_2 \in [k+1, n] \), where the column pairs will be chosen so that the number of affected rows is large. Conditioning on \( M \), in the randomness of the switchings we will have that the events the on the right hand side of (5.18) are independent, and have probability at most 1/2 for the affected rows. For large \( k \) we will apply shufflings independently to several non-overlapping pairs of rows, and use Lemma 4.4 to bound the probabilities in (5.18).

In any case, we may restrict to the events \( G^\alpha(\delta) \) and \( G^c(\varepsilon) \) for some \( \delta, \varepsilon > 0 \) to be chosen sufficiently small and independent of \( n \). By Corollary 3.6 we may also restrict to the event
\( B(\varepsilon_0, \gamma)^c \) from Corollary 3.6 for some \( \varepsilon_0, \gamma > 0 \) to be chosen – this event will play a similar role to the one played by \( G^{\exp}(\gamma) \) in the proof of Proposition 2.4. For now let \( \eta \in (0,1] \) possibly depending on \( d \). We will put restrictions on the range of \( \eta \) as the proof develops, ultimately taking \( \eta \gg d^{-c} \) for some small constant \( c > 0 \).

5.3. High sparsity. Fix \( k \leq \frac{n \log n}{d} \). Towards an application of Lemma 5.3, we fix \( \hat{v} \in W_k \) and \( \alpha \in \{0,1\} \).

Pair off the first \( k \) columns of \( M \) with the last \( k \) columns according to some bijection
\[
\sigma : [k] \to [n - k + 1, n]
\]
chosen in some arbitrary fashion, say uniformly at random and independently of \( M \). We now give a greedy procedure which outputs for some \( m \geq 1 \):

1. an increasing sequence \( (j_l)_{l=1}^m \) of column indices in \([k]\), and
2. an associated increasing sequence of sets of row indices
\[
[k] = A_0 \subset A_1 \subset \cdots \subset A_m \subset [n]
\]
with certain properties.

**Step 0:** Set \( A_0 = [k] \).

**Induction step:**

For \( l \geq 1 \), let
\[
j_l = \min \left\{ j \in [k] : |\text{Ex}_{M^T}(j, \sigma(j)) \setminus A_{l-1}| \land |\text{Ex}_{M^T}(\sigma(j), j) \setminus A_{l-1}| \geq .01d \right\}.
\]
Let \( B_l^+ := \text{Ex}_{M^T}(j_l, \sigma(j_l)) \setminus A_{l-1} \) and \( B_l^- := \text{Ex}_{M^T}(\sigma(j_l), j_l) \setminus A_{l-1} \), and set
\[
A_l = A_{l-1} \cup B_l^+ \cup B_l^-.
\]

If no such \( j_l \) exists, set \( m := l - 1 \) and STOP.

The sequence of sets \( A_l \) produced by this procedure have the following two properties:

1. For each \( l \in [m] \), \( A_l \) is fixed by conditioning on the \( 2(l - 1) \) columns of \( M \) indexed by \( \bigcup_{\nu < l}\{j_{\nu}, \sigma(j_{\nu})\} \).
2. The \( m \) sets \( B_l := B_l^+ \cup B_l^- \) are pairwise disjoint, and
\[
|B_l^+| \land |B_l^-| \geq .01d \quad (5.19)
\]

for all \( l \in [m] \).

We want the halting time \( m \) of this procedure to be large. Note that \( m \) is a random variable fixed by conditioning on \( M \). We now show that \( m \) is large if \( M \) is such that all “thin minors” are sufficiently dense, in the sense of Corollary 3.6.

**Claim 5.4.** Assume \( B(\varepsilon_0, \gamma)^c \) holds for \( \varepsilon_0, \gamma \) sufficiently small, where \( B(\varepsilon_0, \gamma) \) was defined in Corollary 3.6. Then \( m \gg k / \log n \).
Proof. We abbreviate
\[ \text{Ex}^+(j) := \text{Ex}_{M^t}(j, \sigma(j)), \quad \text{Ex}^-(j) := \text{Ex}_{M^t}(\sigma(j), j). \]
We have that for all \( j \in [k] \), either \( |\text{Ex}^+(j) \setminus A_m| \) or \( |\text{Ex}^-(j) \setminus A_m| \) is < .01d. For each \( j \in [k] \), put \( j \in S \) if \( |\text{Ex}^+(j) \setminus A_m| < .01d \) and otherwise put \( \sigma(j) \in S \), so that \( |S| = k \).

Taking \( \delta \) sufficiently small, by our restriction to \( G^\text{ex}(\delta) \) we may assume
\[ |\text{Ex}^+(j)| \wedge |\text{Ex}^-(j)| \geq .1d \]
for all \( j \in [k] \). It follows that
\[ |\text{Ex}^+(j) \cap A_m| \vee |\text{Ex}^-(j) \cap A_m| \geq .09d \]
and since \( \text{Ex}^+(j) \subset N_{M^t}(j) \), \( \text{Ex}^-(j) \subset N_{M^t}(\sigma(j)) \), we have
\[ |N_{M^t}(j) \cap A_m| \geq .09d \]
for all \( j \in S \). Now
\[ e_M(A_m, S) = \sum_{j \in S} |N_{M^t}(j) \cap A_m| \geq .09d|S| \]
so taking \( \varepsilon_0 < .09 \) and \( \gamma \) sufficiently small, by our restriction to \( B(\varepsilon_0, \gamma)^c \) we must have
\[ |A_m| \geq b_0(\varepsilon_0, \gamma, k) \gg \frac{dk}{\log n}. \]
But by the inductive procedure to produce \( A_m \) we have
\[ |A_m| \leq k + 2md \]
from which it follows that
\[ m \gg \frac{k}{d} \left( \frac{d}{\log n} - 1 \right) \gg \frac{k}{\log n} \]
by our assumption \( d = \omega(\log^2 n) \).

With the sets \( B^+_l, B^-_l \), thus defined, we can form a coupling \((M, \tilde{M})\) of rrd matrices using Lemma 3.2 by performing independent shufflings on \( M \) at the columns \((j_l, \sigma(j_l))\). Specifically, letting \( r := \lfloor .01d \rfloor \), for each \( l \in [m] \) we draw \( E^+_l \subset B^+_l, E^-_l \subset B^-_l \) of size \( r \) independently and uniformly at random, and conditional on these \( 2m \) sets we draw \( m \) independent uniform random bijections
\[ \pi_l : E^+_l \to E^-_l. \]
We let \( \tilde{\xi} = (\xi_i)_{i=1}^n \) be a sequence of iid uniform signs independent of all other random variables. Then for each \( l \in [m] \) and each \( i \in E^+_l \), we replace the minor \( M(i, \pi_l(i)) \times (j_l, \sigma(j_l)) \) with the random \( 2 \times 2 \) matrix
\[ I_2\xi_{i_1=+1} + J_2\xi_{i_1=-1}. \]
By the independence of the \((\xi_i)_{i \in [n]}\) and the fact that the \(2m\) sets \(\{E_i^+\}_{i \in [m]} \cup \{E_i^-\}_{i \in [m]}\) are pairwise disjoint, we have

\[
P\left(\tilde{M}_{[k+1,n] \times [k]} \hat{v} = \alpha 1 \mid M\right) \leq \prod_{i=1}^{m} \prod_{i \in E_i^+} \prod_{j \in [k]} P\left(\sum_{j \in [k]} \tilde{M}(i,j)v(j) = \alpha \mid M\right)
\]

\[
\leq \left(\frac{1}{2}\right)^{\theta_{1}dm}
\]

\[
\leq \exp(-cdk/\log n).
\]

Since \(M \equiv \tilde{M}\), by Lemma 5.3, we conclude

\[
P(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \ll \left(\frac{n}{k}\right)^2 \exp\left(-\frac{cdk}{\log n}\right)
\]

\[
\leq \exp\left(2k\left(\log n - c\frac{d}{\log n}\right)\right)
\]

\[
\ll \exp\left(-\frac{dk}{\log n}\right)
\]

by our assumption \(d = \omega(\log^2 n)\). Summing the bounds (5.5) gives

\[
P(\mathcal{E}_k) \ll n^{-\omega(1)} \quad (5.20)
\]

for any \(k \leq \frac{1}{2\gamma} \frac{n \log n}{d}\).

5.4. Moderate sparsity. Now we fix \(k\) in the range \(\left[\frac{1}{2\gamma} \frac{n \log n}{d}, (1 - \eta)n\right]\).

The proof mirrors the proof for large \(k\) for the signed rrd matrix \(M_{\pm}\) in Section 5.1. The general idea is to express the event that \(M_{[k+1,n] \times [k]} \hat{v} = \alpha 1\) as the event that several independent random walks all land at \(\alpha\). Without the iid signs enjoyed by \(M_{\pm}\) we must use the shuffling coupling of Lemma 3.1 to create random walks. We use the discrepancy property enforced by our restriction to \(\mathcal{G}(\varepsilon)\) to argue that these walks take many steps (in particular we will need an extension of 5.1 used in Section 5.1), at which point we can apply the anti-concentration bound from Theorem 2.1 to each walk.

More precisely, we will fix disjoint sets of row indices \(A_1, A_2 \subset A := [k + 1,n]\) of equal size \(a_1 = |A_1| = |A_2| \gg n-k\), and pair off the elements of \(A_1\) with those of \(A_2\) according to a bijection \(\sigma : A_1 \to A_2\). For each \(i \in A_1\), we perform a shuffling on \(M\) at the row pair \((i, \sigma(i))\); we do this independently for each \(i \in A_1\) and denote the new matrix by \(\tilde{M}\). We have

\[
P\left(M_{[k+1,n] \times [k]} \hat{v} = 0 \mid R_1, \ldots, R_k\right) = \mathbb{E}_{R_{k+1}, \ldots, R_n} P\left(\tilde{M}_{[k+1,n] \times [k]} \hat{v} = \alpha 1 \mid M\right)
\]

so it suffices to bound

\[
P\left(\tilde{M}_{[k+1,n] \times [k]} \hat{v} = \alpha 1 \mid M\right) \leq \prod_{i \in A_1} P\left(\tilde{R}_i \cdot v = \alpha \mid M\right). \quad (5.21)
\]

As in Section 4, in order to bound the probabilities in (5.21) using Theorem 2.1, we will need to argue that these random walks take many steps, or at least that many of the walks do. It turns out this can be done if we take the pairing \(\sigma\) to be random – it is then possible
to show using our restriction to the edge discrepancy event \( G^a(\varepsilon) \) that with overwhelming probability most of the pairs \((i, \sigma(i))\) give walks that take a large number of steps.

We turn to the details. Fix disjoint sets \( A_1, A_2 \subset A := [k + 1, n] \) with
\[
a_1 := |A_1| = |A_2| \gg n - k.
\]

We create a new rrd matrix \( \widetilde{M} \) coupled to \( M \) from three additional sources of randomness:

1. a uniform random bijection \( \sigma : A_1 \to A_2 \) independent of all other variables;
2. a sequence \((\pi_i)_{i \in A_1}\) of uniform random bijections

\[
\pi_i : \text{Ex}_M(i, \sigma(i)) \to \text{Ex}_M(\sigma(i), i)
\]

which are jointly independent conditional on \( M \) and \( \sigma \);
3. an array of iid uniform random signs \((\xi_{ij})_{i,j \in [n]}\) independent of all other variables.

We form \( \widetilde{M} \) by performing a shuffling on \( M \) at \((i, \sigma(i))\) according to \( \pi_i \) and \((\xi_{ij})_{j \in [n]}\) for each \( i \in A_1 \). We have \( \widetilde{M} \overset{d}{=} M \) by Lemma 3.1 and the independence of the \( \pi_i \) and \( \xi_{ij} \).

For fixed \( i \in A_1 \), denote
\[
\text{Steps}_i(\hat{v}) := \text{Steps}^{(i, \sigma(i))}(v)
\]
\[
:= \{ j \in \text{Ex}_M(i, \sigma(i)) : v(j) \neq v(\sigma(j)) \}
\]
where we recall the notation \( v = (\hat{v} \ 0) \in \mathbb{R}^n \) with \( \hat{v} = v_{[k]} \). Now since \( \text{spt}(v) = [k] \), we have that for each \( i \in A_1 \),
\[
|\text{Steps}_i(\hat{v})| \geq |\text{Cross}_i(k)| \tag{5.22}
\]
where we define
\[
\text{Cross}_i(k) = \text{Cross}^{(i, \sigma(i))}(k)
\]
\[
:= \{ j \in \text{Ex}_M(i, \sigma(i)) \cap [k] : \pi_i(j) \in [k + 1, n] \}
\]
the number of pairs \((j, \pi_i(j))\) which are in \([k] \times [k + 1, n]\), i.e. pairs which cross the partition \([n] = [k] \cup [k + 1, n]\) going from left to right. (We could also include pairs crossing right to left, but this will tend to only improve the lower bound (5.22) by a constant factor.)

Hence, for \( m \geq 1 \), defining to the good events
\[
\mathcal{G}_i(m) := \{ |\text{Cross}_i(k)| \geq m \} \tag{5.23}
\]
for each \( i \in A_1 \), by Lemma 4.4 we have
\[
P_{\tilde{\xi}} \left\{ \tilde{R}_i \cdot v = \alpha \right\} 1_{\mathcal{G}_i(m) \cap \{ \alpha \}} = O(m^{-1/2}) \tag{5.24}
\]
where we denote \( \tilde{\xi}_i = (\xi_{ij})_{j=1}^n \), the \( i \)th rows of the array of signs.

In the remainder of the proof, we show that with overwhelming probability in the randomness of the bijections \( \sigma \) and \((\pi_i)_{i \in A_1}\), for most \( i \in A_1 \) and for a reasonably large value of \( m \), \( \mathcal{G}_i(m) \) holds except on an exponentially small event. (Hence we are done with the iid signs \( \xi_{ij} \).) The randomness of \( M \) will only enter through our restriction to the events \( G^a(\varepsilon) \) and \( G^{\text{ex}}(\delta) \).

Lemma 5.5 below summarizes what we need from the discrepancy property enforced on \( G^a(\varepsilon) \) – it is an extension of 5.1 from the proof for \( M_{\pm} \). While for \( M_{\pm} \) it was enough to know that the intersections \( B(i) \) of a large set \( B \) with the neighborhoods \( \mathcal{N}_M(i) \) were of size
roughly \( p|B| \), here we will need intersections of \( B \) with the sets \( \text{Ex}_M(i_1,i_2), \text{Ex}_M(i_2,i_1) \) to be at least a constant factor of their expected size.

For \( \varepsilon \in (0,1) \) and a set of column indices \( B \subset [n] \), say that an ordered pair \((i_1, i_2)\) of distinct row indices in \( A \) is \( \varepsilon \)-bad for \( B \) if either
\[
|\text{Ex}_M(i_1, i_2) \cap B| \leq \varepsilon p|B| \quad \text{or} \quad |\text{Ex}_M(i_2, i_1) \cap B^c| \leq \varepsilon p(n - |B|).
\] (5.25)
The following lemma shows that on \( \mathcal{G}^n(\varepsilon) \) with \( \varepsilon \) sufficiently small, only a small number of pairs of elements of \([k + 1, n]\) are \( \varepsilon \)-bad for \([k]\).

**Lemma 5.5.** Let \( A, B \subset [n] \). For \( i \in [n] \), denote \( B(i) := N_M(i) \cap B \). For \( \varepsilon \in (0,1) \), define
\[
A_\varepsilon = \left\{ i \in A : \frac{|B(i)|}{p|B|} - 1 \leq \varepsilon, \frac{|B^c(i)|}{p(n - |B|)} - 1 \leq \varepsilon \right\},
\] (5.26)
and for \( i \in A \), let
\[
S_\varepsilon(i) = \left\{ i' \in A_\varepsilon : (i, i') \text{ is } \varepsilon \text{-bad for } B \right\}
\] (5.27)
on \( \mathcal{G}^n(\varepsilon) \) we have
\[
|A \setminus A_\varepsilon| = \ll_\varepsilon p^{-1} \log n
\] (5.28)
and for every \( i \in A_\varepsilon \),
\[
|S_\varepsilon(i)| \ll_\varepsilon p^{-1} \log n
\] (5.29)
assuming \( |B| \wedge |B^c| \geq \frac{\gamma}{27} p^{-1} \log n \) for \( \gamma \) sufficiently small depending on \( \varepsilon \).

**Proof.** We begin with (5.28).

Define the sets
\[
S_1 = \{ i \in A : |B(i)| < (1 - \varepsilon) pb \}
\]
\[
S_2 = \{ i \in A : |B(i)| > (1 + \varepsilon) pb \}
\]
\[
S_3 = \{ i \in A : |B^c(i)| < (1 - \varepsilon) p(n - b) \}
\]
\[
S_4 = \{ i \in A : |B^c(i)| > (1 + \varepsilon) p(n - b) \}
\]
so that
\[
A \setminus A_\varepsilon = \bigcup_{k=1}^{4} S_k.
\]

From 5.1 we have
\[
|S_1| \ll_\varepsilon \frac{n \log n}{d}.
\]
By replacing \( B \) with \( B^c \) we obtain the same bound on \( |S_3|, |S_2| \) and \( |S_4| \) are bounded similarly.

We turn to the estimate (5.29). Fix \( i \in A_\varepsilon \). We can write
\[
S_\varepsilon(i) = S^1_\varepsilon(i) \cup S^2_\varepsilon(i)
\]
where
\[
S^1_\varepsilon(i) = \left\{ i' \in A_\varepsilon : |\text{Ex}(i, i') \cap B| \leq \varepsilon p|B| \right\}
\]
\[
S^2_\varepsilon(i) = \left\{ i' \in A_\varepsilon : |\text{Ex}(i', i) \cap B^c| \leq \varepsilon p(n - |B|) \right\}.
\]
We first bound \( |S^1_\varepsilon(i)| \). For \( i' \in S^1_\varepsilon(i) \), we have
\[
|\text{Ex}(i, i') \cap B| \leq \varepsilon p|B| \leq \frac{\varepsilon}{1 - \varepsilon} |B(i)|
\] (5.30)
since $i \in A_\varepsilon$. It follows that
\[
e_M(S^1_\varepsilon(i), B(i)) = \sum_{i' \in S^1_\varepsilon(i)} |\text{Co}(i, i') \cap B| \\
= \sum_{i' \in S^1_\varepsilon(i)} |B(i)| - |\text{Ex}(i, i') \cap B| \\
\geq |S^1_\varepsilon(i)| \left(1 - \frac{\varepsilon}{1 - \varepsilon}\right)|B(i)| \\
= \frac{1 - 2\varepsilon}{1 - \varepsilon}|S^1_\varepsilon(i)||B(i)|.
\] (5.31)

On the other hand, if $(S^1_\varepsilon(i), B(i)) \in \mathcal{F}(\varepsilon)$ we have
\[
e_M(S^1_\varepsilon(i), B(i)) \leq (1 + \varepsilon)p|S^1_\varepsilon(i)||B(i)|.
\] (5.32)

From (5.32) it follows that
\[
e_M(S^1_\varepsilon(i), B(i)) \leq \frac{(1 + \varepsilon)}{2}|S^1_\varepsilon(i)||B(i)|
\]
which contradicts (5.31) if $\varepsilon$ is a sufficiently small absolute constant. We may hence assume $(S^1_\varepsilon(i), B(i)) \notin \mathcal{F}(\varepsilon)$. Similarly to how we argued in the bound for $|S_1|$, we can deduce from the lower bound
\[|B(i)| \gg p|B| \gg \gamma^{-1} \log n
\]
(since $i \in A_\varepsilon$) that taking $\gamma$ sufficiently small, we must have $|S^1_\varepsilon(i)| = |S^1_\varepsilon(i)| \land |B(i)|$ (for $n$ sufficiently large), and hence
\[|S^1_\varepsilon(i)| = \ll \varepsilon p^{-1} \log n.
\]
The proof that $|S^2_\varepsilon(i)| = O(\varepsilon p^{-1} \log n)$ follows similar lines and is omitted. \hfill \Box

Now we define the subset of $A_1$ of “good” row indices to be
\[A'_1 = \left\{ i \in A_1 \cap A_\varepsilon : \sigma(i) \in A_\varepsilon \setminus S_\varepsilon(i) \right\}
\] (5.33)
i.e. the set of $i \in A_1$ such that $i$ and $\sigma(i)$ are both in $A_\varepsilon$, and such that the pair $(i, \sigma(i))$ is not bad for $|k|$. Note that this is a random set depending on $M$ and $\sigma$. We can now use Lemma 5.5 and the randomness of $\sigma$ to show that with overwhelming probability, $A'_1$ constitutes most of $A_1$.

For $\kappa > 0$, let
\[B'(\kappa) = \left\{ |A_1 \setminus A'_1| \geq \kappa |A_1| \right\}.
\] (5.34)

Now for arbitrary $m \geq 1$ we have
\[
P\left(\tilde{M}_{[k+1,n] \times [k]} \overset{\circ}{=} \alpha \mathbf{1} \mid M\right) \leq P_\sigma(B'(\kappa)) + E_\sigma 1_{B'(\kappa)} \cdot \mathbb{P}\left(\tilde{M}_{[k+1,n] \times [k]} \overset{\circ}{=} \alpha \mathbf{1} \mid M, \sigma\right) \\
\leq P_\sigma(B'(\kappa)) + E_\sigma 1_{B'(\kappa)} \mathbb{P} \left(\tilde{R}_i \cdot v = \alpha \mid M, \sigma\right) \\
\leq P_\sigma(B'(\kappa)) + E_\sigma 1_{B'(\kappa)} \mathbb{P} \left(\tilde{R}_i \cdot v = \alpha \right) 1_{G_i(m)}.
\] (5.35)
The term \( \mathbb{P}_{\xi} \{ \tilde{R}_i \cdot v = \alpha \} 1_{G_i(m)} \) is \( O(m^{-1/2}) \) by (5.24). It remains to bound \( \mathbb{P}_{\sigma}(B'(\kappa)) \) for some \( \kappa \in (0, 1) \) (we will eventually fix \( \kappa \) independent of \( n \)), and \( \mathbb{P}_{\pi_r}(G_i(m)^c) \) (for some large value of \( m \)).

From Lemma 5.5 we have
\[
|A \setminus A_\varepsilon| \vee \left( \max_{i \in A_1} |S_\varepsilon(i)| \right) \leq s_0
\]
for some \( s_0 = O(p^{-1} \log n) \) (assuming \( \eta \geq \frac{1}{27} \log \frac{n}{d} \)). By crudely estimating the number of bad realizations of \( \sigma \), we can bound
\[
\mathbb{P}_{\sigma}(B'(\kappa)) \leq \frac{s_0^{\kappa a_1} \lfloor (1 - \kappa) a_1 \rfloor!}{a_1!} \leq \left( \frac{s_0}{(1 - \kappa) a_1} \right)^{\kappa a_1} \leq \left( \frac{C n \log n}{d(n - k)} \right)^{\kappa a_1}
\]
where in the last line we substituted the bound on \( s_0 \) and assumed \( \kappa \leq 1/2 \), say.

Now we estimate the terms \( \mathbb{P}_{\pi_i}(G_i(m)^c) \). For fixed \( i \in A'_1 \) we have
\[
E_{\pi_i} |\text{Cross}_i(k)| = \left[ \frac{\text{Ex}_M(i, \sigma(i)) \cap [k]}{|\text{Ex}_M(i, \sigma(i))|} \right] \frac{\text{Ex}_M(\sigma(i), i) \cap [k + 1, n]}{\text{Ex}_M(\sigma(i), i)}. \tag{5.38}
\]
From our restriction to \( G^\text{ex}(\delta) \) we know the denominator is of size \( \Theta_\delta(d) \), and since \( i \in A'_1 \) the numerator is of size \( \Omega \left( p^2 k(n - k) \right) \) whence,
\[
E_{\pi_i} |\text{Cross}_i(k)| \gg \frac{d k(n - k)}{n} \gg p[k \land (n - k)]. \tag{5.39}
\]
From Lemma 3.8 it follows that
\[
|\text{Cross}_i(k)| \gg p[k \land (n - k)] \tag{5.40}
\]
except with probability at most \( \exp \left( -c p[k \land (n - k)] \right) \) in the randomness of \( \pi_i \). We have hence shown that for \( i \in A'_1 \),
\[
\mathbb{P}_{\pi_i}(G_i(m)^c) \leq \exp \left( -c p[k \land (n - k)] \right) \tag{5.41}
\]
where set
\[
m := c p[k \land (n - k)]
\]
and \( c > 0 \) is a sufficiently small absolute constant. In particular, this bound is of lower order than the bound \( \mathbb{P}_{\xi} \{ \tilde{R}_i \cdot v = \alpha \} 1_{G_i(m)} = O(m^{-1/2}) \).

Substituting our bounds (5.24), (5.37) and (5.41) into (5.35), we have
\[
\mathbb{P} \left( M_{[k+1,n] \times [k]}^\varepsilon = \alpha \left| M \right. \right) \leq \mathbb{P}_{\sigma}(B'(\kappa)) + E_{\sigma} 1_{G'(\kappa)^c} O(m^{-1/2}) \delta A'_1 \leq \left( \frac{C n \log n}{d(n - k)} \right)^{\kappa a_1} + m^{-(1 - \kappa + o(1)) a_1/2}.
\]

Applying Lemma 5.3 we have
\[
\mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \ll (I)_k + (II)_k \tag{5.42}
\]
where

\[(I)_k = \left( \binom{n}{k} \right)^2 \left( \frac{Cn \log n}{d(n-k)} \right)^{\kappa a_1}\]

\[(II)_k = \left( \binom{n}{k} \right)^2 m^{-(1-\kappa+o(1))a_1/2}.\]

First assume \( \frac{n \log n}{d} \ll k \leq \frac{n}{2} \). In this case we have

\[m = c_0 pk \gg \log n.\] (5.43)

and

\[a_1 \gg n - k \geq n/2.\] (5.44)

\[(I)_k \leq 4^n \left( \frac{Cn \log n}{d(n-k)} \right)^{\kappa a_1}\]

\[\leq 4^n \left( \frac{C \log n}{d} \right)^{\kappa n} = o(1)^n\]

by our assumption \( d = \omega(\log^2 n). \)

\[(II)_k \leq 4^n m^{-(1-\kappa+o(1))a_1/2}\]

\[= (\log n)^{-\Omega(n)}\]

by the lower bounds (5.43), (5.43) and the assumption \( \kappa \asymp 1. \)

Now assume \( \frac{n}{2} \leq k \leq (1-\eta)n \). In this case we have

\[m = c_0 p(n-k) = c_0 d \frac{n-k}{n} \gg \eta d.\] (5.45)

Assuming \( a_1 \geq c_1(n-k), \)

\[(I)_k \leq \left( \frac{en}{n-k} \right)^{2(n-k)} \left( \frac{Cn \log n}{d(n-k)} \right)^{\kappa a_1}\]

\[\leq \left[ \frac{C \log n}{d} \left( \frac{n}{n-k} \right)^{1+\frac{2}{c_1 \kappa}} \right]^{c_1 \kappa (n-k)}\]

\[\leq \left( \frac{C \log n}{d \eta^{1+\frac{2}{c_1 \kappa}}} \right)^{c_1 \kappa (n-k)}.

By our assumption \( d = \omega(\log^2 n) \) we conclude that

\[(I)_k \leq o(1)^{\Omega(\eta n)}\] (5.46)

provided \( \eta \geq d^{-c} \) for some \( c > 0 \) depending on \( \kappa, c_1 \).
RANDOM REGULAR DIGRAPHS: SINGULARITY OF THE ADJACENCY MATRIX

\[(II)_k \leq \left( \frac{en}{n-k} \right)^{2(n-k)} m^{-(1-\kappa+o(1))a_1/2} \]
\[\leq \left[ \frac{C}{d\eta^{1+4/(c_1(1-\kappa+o(1)))}} \right]^{\Omega_{\kappa,c_1}(n-k)} \leq o(1)^{\Omega(\eta n)} \quad (5.47)\]

for fixed \(\kappa, c_1\), assuming \(\eta \geq d-1/20\) for some \(c' > 0\) depending on \(\kappa, c_1\).

**Remark 5.6.** Optimizing the choice of \(c_1, \kappa\) above, one can check that for (5.46) and (5.47) to hold it is enough to assume \(\eta \gg d-1/20\). Specifically, we can clearly take \(c_1 = \frac{1}{2} + o(1)\), by letting \(A_1 \cup A_2\) saturate \(A\). Taking \(\kappa \rightarrow 1\) leads to a better value of \(c\), while taking \(\kappa \rightarrow 0\) improves \(c'\). Tuning \(\kappa\) to balance these two constants leads to an exponent a little larger than \(1/20\).

**Appendix A. Passing to a large minor**

In this section we proof Lemma 5.3.

By column exchangeability and a union bound, we have
\[P(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \leq \binom{n}{k} P(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \quad (A.1)\]

where \(\mathcal{E}_k = \mathcal{E}_k^0 \vee \mathcal{E}_k^1\), with
\[\mathcal{E}_k^0 := \{ \exists x \in \mathbb{R}^n: \text{spt}(x) = [k], Mx = 0 \}\]
and
\[\mathcal{E}_k^1 := \{ \exists x \in \mathbb{R}^n: \text{spt}(x) = [k], Mx = 1 \}\].

From
\[\mathcal{E}_k = \mathcal{E}_k^0 \vee (\mathcal{E}_k^1 \setminus \mathcal{E}_k^0)\]
we may bound
\[P(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \leq P(\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}) + P(\mathcal{E}_k^1 \setminus (\mathcal{E}_k^0 \vee \mathcal{E}_{k-1})). \quad (A.2)\]

For the first term on the right hand side, note that on \(\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}\) the minor \(M_{[n]} \times [k]\) has \(k-1\) linearly independent rows. Indeed, if this were not the case we would have \(\text{rank}(M_{[n]} \times [k]) \leq k-2\), so that \(M_{[n]} \times [k]\) has 2 linearly independent right null vectors \(x_1, x_2 \in \mathbb{R}^k\). But there is a \(k-1\)-sparse linear combination of \(x_1, x_2\), putting us in \(\mathcal{E}_{k-1}\).

For the second term in (A.2), note that on the complement of \(\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}\) the minor \(M_{[n]} \times [k]\) has full rank, and hence has \(k\) linearly independent rows.

Now we spend some symmetry to fix the linearly independent rows. Let \(L_i\) denote the event that \(R_1, \ldots, R_i\) are linearly independent. By row exchangeability we have
\[P(\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}) \leq \binom{n}{k-1} P(\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1} \wedge L_{k-1}) \quad (A.3)\]
and
\[ \mathbb{P}\left( \mathcal{E}_k^1 \setminus \left( \mathcal{E}_k^0 \cup \mathcal{E}_{k-1} \right) \right) \leq \binom{n}{k} \mathbb{P}\left( \left( \mathcal{E}_k^1 \setminus \left( \mathcal{E}_k^0 \cup \mathcal{E}_{k-1} \right) \right) \cap \mathcal{L}_k \right). \] (A.4)

In (A.3), \((\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}) \cap \mathcal{L}_{k-1}\) is the event that the first \(k-1\) rows of \(M\) are linearly independent, that there is a null vector \(x\) of \(M\) supported on \([k]\), and that there are no \(k-1\)-sparse null vectors of \(M\). Now on this event there is actually only one possibility for \(x\) up to dilation. Indeed, on \(\mathcal{L}_{k-1}\) the system
\[ M_{[k-1] \times [k]} z = 0 \] (A.5)
has a unique solution up to dilation, by the linear independence of the first \(k-1\) rows. Let us pick a nontrivial solution \(\hat{x} \in \mathbb{R}^k\) of (A.5) arbitrarily, and set \(x^* = (\hat{x} \cdot 0)^T \in \mathbb{R}^n\). On the complement of \(\mathcal{E}_{k-1}\), each component of \(\hat{x}\) is nonzero. Hence, \((\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}) \cap \mathcal{L}_{k-1}\) is contained in the event
\[ \mathcal{E}'_k := \mathcal{L}_{k-1} \setminus \{ \hat{x}(j) \neq 0 \text{ for all } j \in [k] \} \cap \{ Mx^* = 0 \} \]
where we have let \(\mathcal{L}'_{k-1} = \mathcal{L}_{k-1} \setminus \{ \hat{x}(j) \neq 0 \text{ for all } j \in [k] \}\). We emphasize that \(\hat{x}\) is a random vector in \(\mathbb{R}^k\), defined only on the event \(\mathcal{L}_{k-1}\), and fixed by conditioning on the first \(k-1\) rows of \(M\) through (A.5).

We may similarly fix the vector in the preimage of \(1\) on the event \((\mathcal{E}_k^1 \setminus (\mathcal{E}_k^0 \cup \mathcal{E}_{k-1})) \cap \mathcal{L}_k\) from (A.4). This event is disjoint from the event \((\mathcal{E}_k^0 \setminus \mathcal{E}_{k-1}) \cap \mathcal{L}_{k-1}\) from (A.3), and on it we may define \(\hat{y} \in \mathbb{R}^k\) as the unique solution of
\[ M_{[k] \times [k]} y = 1. \] (A.6)
Setting
\[ \mathcal{E}''_k := \mathcal{L}'_k \cap \{ M_{[k+1,n] \times [k]} \hat{y} = 1 \} \] (A.7)
where
\[ \mathcal{L}'_k := \mathcal{L}_k \cap \{ \hat{y}(j) \neq 0 \text{ for all } j \in [k] \} \]
we similarly conclude that
\[ (\mathcal{E}_k^1 \setminus (\mathcal{E}_k^0 \cup \mathcal{E}_{k-1})) \cap \mathcal{L}_k \subset \mathcal{E}''_k. \]
Here also, \(\hat{y} \in \mathbb{R}^k\) is a random vector defined only on the event \(\mathcal{L}_k\) via (A.6), fixed by conditioning on the first \(k\) rows of \(M\).

Combined with (A.3), (A.4), (A.2) and (A.1), we have
\[ \mathbb{P}(\mathcal{E}_k \setminus \mathcal{E}_{k-1}) \leq \binom{n}{k} \binom{n}{k-1} \mathbb{P}(\mathcal{E}'_k) + \binom{n}{k} \mathbb{P}(\mathcal{E}''_k). \] (A.8)

By conditioning on a realization of \(R_1, \ldots, R_k\) such that \(\mathcal{L}'_k\) holds, which fixes \(\hat{x}\), we see that \(\mathbb{P}(\mathcal{E}'_k) \leq Q_k\), with \(Q_k\) as in (5.16). We similarly have that \(\mathbb{P}(\mathcal{E}''_k) \leq Q_k\), and the result follows.

References

[1] N. Alon and J. Spencer, The probabilistic method. Third edition. With an appendix on the life and work of Paul Erdős. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[2] R. Adamczak and D. Chafaï, Circular law for random matrices with unconditional log-concave distribution. Preprint available at arXiv:1303.5838.
[3] R. Adamczak, D. Chafaï and P. Wolff, Circular law for random matrices with exchangeable entries. Preprint available at arXiv:1402.3660.

[4] A. Basak and A. Dembo, Limiting spectral distribution of sum of unitary and orthogonal matrices. Electron. Commun. Probab. 18 (2013), no. 69, 19 pp.

[5] C. Bordenave, P. Caputo and D. Chafaï, Circular law theorem for random Markov matrices. Probab. Theory Related Fields 152 (2012), no. 3-4, 751–779.

[6] C. Bordenave and D. Chafaï, Around the circular law. Probability Surveys 9 (2012) 1–89.

[7] J. Bourgain, P. Wood and V. Vu, On the singularity probability of random discrete matrices. J. Funct. Anal. 258 (2010), no. 2, p. 559–603.

[8] E. R. Canfield and B. D. McKay, Asymptotic enumeration of dense 0-1 matrices with equal row sums and equal column sums. Electron. J. Combin. 12 (2005), Research Paper 29, 31 pp. (electronic).

[9] E.R. Canfield and B.D. McKay, Asymptotic enumeration of integer matrices with large equal row and column sums. Combinatorica 30 (2010), no. 6, 655–680.

[10] S. Chatterjee, Stein’s method for concentration inequalities. Probab. Theory Related Fields 138 (2011), no. 2, 1189–1217.

[11] 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.

[12] J. Komlós, On the determinant of (0, 1) matrices. Studia Sci. Math. Hungar. 2 (1967) 7–22.

[13] J. Komlós, On the determinant of random matrices. Studia Sci. Math. Hungar. 3 (1968) 387–399.

[14] J. Komlós, Circulated manuscript, 1977. edited version available online at: http://www.math.rutgers.edu/~komlos/01short.pdf

[15] M. Krivelevich, B. Sudakov, Pseudo-random graphs. More sets, graphs and numbers, 199–262, Bolyai Soc. Math. Stud., 15, Springer, Berlin, 2006.

[16] M. Krivelevich, B. Sudakov, V. Vu and N. Wormald, Random regular graphs of high degree. Random Structures Algorithms 18 (2001), no. 4, 346–363.

[17] V. A. Marchenko and L. A. Pastur, The distribution of eigenvalues in certain sets of random matrices, Mat. Sb. 72 (1967), 507–536.

[18] B. D. McKay, The expected eigenvalue distribution of a large regular graph. Linear Algebra Appl. 40 (1981), p. 203–216.

[19] B.D. McKay, Subgraphs of random graphs with specified degrees. Congressus Numerantium, 33 (1981), 213–223.

[20] B.D. McKay and X. Wang, Asymptotic enumeration of 0-1 matrices with equal row sums and equal column sums. Special issue on the Combinatorial Matrix Theory Conference (Pohang, 2002). Linear Algebra Appl. 373 (2003), 273–287.

[21] H. Nguyen, On the singularity of random combinatorial matrices. SIAM J. Discrete Math. 27 (2013), no. 1, 447–458.

[22] H. Nguyen, Random doubly stochastic matrices: the circular law. To appear in Ann. Probab.

[23] H. Nguyen and V. Vu, Circular law for random discrete matrices of given row sum. J. Comb. 4 (2013), no. 1, 1–30.

[24] M. Rudelson, Lower estimates for the singular values of random matrices, C. R. Math. Acad. Sci. Paris 342 (2006), 247–252.

[25] M. Rudelson and R. Vershynin, The Littlewood-Offord Problem and invertibility of random matrices. Advances in Mathematics 218 (2008), 600–633.

[26] M. Rudelson and R. Vershynin, Invertibility of random matrices: unitary and orthogonal perturbations. To appear in J. Amer. Math. Soc.
[34] E. Shamir and E. Upfal, Large regular factors in random graphs. Convexity and graph theory (Jerusalem, 1981), 271–282. North-Holland Math. Stud., 87, North-Holland, Amsterdam, 1984.

[35] R. P. Stanley, Enumerative combinatorics. Vol. 2. With a foreword by Gian-Carlo Rota and appendix by Sergey Fomin. Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999.

[36] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, pp. 583–602. Univ. California Press, Berkeley, Calif., 1972.

[37] T. Tao, Topics in random matrix theory. Graduate Studies in Mathematics, 132. American Mathematical Society, Providence, RI, 2012.

[38] T. Tao, Outliers in the spectrum of iid matrices with bounded rank perturbations. Probab. Theory Related Fields 155 (2013), no. 1-2, 231–263.

[39] T. Tao and V. Vu, On random ±1 matrices: singularity and determinant. Random Structures Algorithms 28 (2006), no. 1, 1–23.

[40] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices. J. Amer. Math. Soc. 20 (2007), no. 3, 603–628.

[41] T. Tao and V. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. Annals of Math. 169 (2009), 595–632.

[42] T. Tao and V. Vu, Random Matrices: The circular Law. Commun. Contemp. Math. 10 (2008), no. 2, 261–307.

[43] T. Tao and V. Vu, Random matrices: universality of ESDs and the circular law. Ann. Probab. 38 (2010), no. 5, p. 2023–2065, With an appendix by Manjunath Krishnapur.

[44] T. Tao and V. Vu, Additive Combinatorics. Cambridge University Press, 2006.

[45] A. Thomason, Pseudo-random graphs. In Proceedings of Random Graphs, Poznań 1985, M. Karoński, ed., Annals of Discrete Math. 33 (North Holland 1987), 307–331.

[46] A. Thomason, Random graphs, strongly regular graphs and pseudo-random graphs. Surveys in Combinatorics, 1987, C. Whitehead, ed., LMS Lecture Note Series 123 (1987), 173–195.

[47] L. Tran, Local law for eigenvalues of random regular bipartite graphs. Available at arXiv:math/1310.4606.

[48] L. Tran, V. Vu, and K. Wang, Sparse random graphs: eigenvalues and eigenvectors. Random Structures Algorithms 42 (2013), no. 1, 110–134.

[49] V. Vu, Discrete random matrices. Available at arXiv:math/0611321.

[50] N.C. Wormald, Models of random regular graphs. Surveys in combinatorics, 1999 (Canterbury), 239–298, London Math. Soc. Lecture Note Ser., 267, Cambridge Univ. Press, Cambridge, 1999.

[51] http://mathoverflow.net/questions/86118/non-enumerative-proof-that-there-are-many-derangements

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES, CA 90095-1555

E-mail address: nickcook@math.ucla.edu