Hitting Time Theorems for Random Matrices

LOUIGI ADDARIO-BERRY and LAURA ESLAVA

Department of Mathematics and Statistics, McGill University
(e-mail: louigi@math.mcgill.ca, laura.eslavafernandez@mail.mcgill.ca)

Starting from an n-by-n matrix of zeros, choose uniformly random zero entries and change them to ones, one at a time, until the matrix becomes invertible. We show that with probability tending to one as $n \to \infty$, this occurs at the very moment the last zero row or zero column disappears. We prove a related result for random symmetric Bernoulli matrices, and give quantitative bounds for some related problems. These results extend earlier work by Costello and Vu [10].

2010 *Mathematics subject classification*: Primary 15A52
Secondary 15A15, 05C80

1. Introduction

In this paper, we initiate an investigation of hitting time theorems for random matrix processes. Hitting time theorems have their origins in the study of random graphs; we briefly review this history, then proceed to an overview of recent work on discrete and continuous random matrix models and a statement of our results.

To begin, consider the classical *Erdős–Rényi graph process* $\{G_{n,p}\}_{0 \leqslant p \leqslant 1}$, defined as follows. Independently for each pair $\{i, j\} \subset [n] = \{1, \ldots, n\}$, let U_{ij} be a Uniform$[0, 1]$ random variable. Then, for $p \in [0, 1]$, let $G_{n,p}$ have vertex set $[n]$ and edge set $\{\{i, j\} : U_{ij} \leqslant p\}$. In $G_{n,p}$, each edge is independently present with probability p, and for $p < p'$ we have that $G_{n,p}$ is a subgraph of $G_{n,p'}$.

Bollobás and Frieze [4] proved the following *hitting time theorem* for $G_{n,p}$, which is closely related to the main result of the present work. Let $\tau_{\delta \geqslant 1}$ be the first time p that $G_{n,p}$ has minimum degree one, and let τ_{PM} be the first time p that $G_{n,p}$ contains a perfect matching (or let $p = 1$ if n is odd). Then, as $n \to \infty$ along even numbers, we have

$$\mathbf{P}(\tau_{\delta \geqslant 1} = \tau_{\mathrm{PM}}) \to 1.$$

In other words, the first moment that the trivial obstacle to perfect matchings (isolated vertices) disappears, with high probability a perfect matching appears. Ajtai, Komlós and

Szemerédi [1] had slightly earlier shown a hitting time theorem for Hamiltonicity: the first time $G_{n,p}$ has minimum degree two, with high probability $G_{n,p}$ is Hamiltonian. In fact, [4] generalizes this, showing that if $\tau_{\delta \geqslant 2k}$ is the first time $G_{n,p}$ has minimum degree $2k$ and $\tau_{k-\mathrm{Ham}}$ is the first time $G_{n,p}$ contains k disjoint Hamilton cycles, then with high probability $\tau_{\delta \geqslant 2k} = \tau_{k-\mathrm{Ham}}$. Hitting time theorems have since been proved for a wide variety of other models and properties, including connectivity [5], k-edge-connectivity [3, 21] and k-vertex-connectivity [3] in random graphs and in maker–breaker games, connectivity in geometric graphs [19], and Hamiltonicity in geometric graphs and in the d-dimensional Gilbert model [2].

In this work we introduce the study of hitting time theorems for random discrete matrices. The study of random matrices is burgeoning, with major advances in our understanding over the last three to five years. Thanks to work by a host of researchers, the behaviour of the determinant [22], a wide range of spectral properties [11, 25], invertibility [6], condition numbers [23, 24], and singular values [20, 16] are now well (though not perfectly) understood. (This list of references is representative, rather than exhaustive.) The recent paper [14] provides a nice collection of open problems, with a focus on random discrete matrices.

In order to have a concept of hitting time theorems for matrices, we need to consider matrix *processes*, and we focus on two such processes. The first is the *Bernoulli* process $\mathcal{R}_n = \{R_{n,p}\}_{0 \leqslant p \leqslant 1}$, defined as follows. Independently for each ordered pair $\{ij : 1 \leqslant i \neq j \leqslant n\}$, let U_{ij} be a Uniform$[0, 1]$ random variable. Then let $R_{n,p}$ be an n-by-n matrix with i, j entry $R_{n,p}(i, j)$ equal to one if $U_{ij} \leqslant p$ and zero otherwise. For $n \geqslant 1$ and $0 \leqslant p \leqslant 1$ we let $H_{n,p}$ be the directed graph with adjacency matrix $R_{n,p}$, so $\{H_{n,p}\}_{0 \leqslant p \leqslant 1}$ is a *directed Erdős–Rényi graph process*. (We take $R_{n,p}$ to have zero diagonal entries as it is technically convenient for $H_{n,p}$ to have no loop edges; however, all our results for this model would still hold if the diagonal entries were generated by independent uniform random variables $\{U_{ii} : 1 \leqslant i \leqslant n\}$, and with essentially identical proofs.)

The second model we consider is the *symmetric* Bernoulli process $\mathcal{Q}_n = \{Q_{n,p}\}_{0 \leqslant p \leqslant 1}$: with U_{ij} as above, for $1 \leqslant i < j \leqslant n$ let $Q_{n,p}(i, j) = Q_{n,p}(j, i) = \mathbf{1}_{[U_{ij} \leqslant p]}$, and set all diagonal entries equal to zero. Throughout the paper, we work in a space in which $Q_{n,p}$ is the adjacency matrix of $G_{n,p}$ for each $0 \leqslant p \leqslant 1$. The principal result of this paper is to prove hitting time theorems for invertibility (or full rank) for both the Bernoulli matrix process and the symmetric Bernoulli process; we now proceed to state our new contributions in detail.

2. Statement of results

Given a real-valued matrix M, write

$$Z^{\mathrm{ROW}}(M) = \{i : \text{ all entries in row } i \text{ of } M \text{ equal zero}\},$$

define $Z^{\mathrm{COL}}(M)$ similarly, and let

$$z(M) = \max(|Z^{\mathrm{ROW}}(M)|, |Z^{\mathrm{COL}}(M)|).$$

Given a collection of matrices $\mathcal{M} = \{M_p\}_{0 \leqslant p \leqslant 1}$, let $\tau(\mathcal{M}) = \inf\{p : z(M_p) = 0\}$, with the convention that $\inf \emptyset = 1$. We write $\tau = \tau(\{M_p\}_{0 \leqslant p \leqslant 1})$ when the matrix process under

consideration is clear. We say that a square matrix M is singular if $\det M = 0$, and otherwise say that M is non-singular. Our main result is the following.

Theorem 2.1. *As $n \to \infty$ we have*

$$\mathbf{P}\big(R_{n,\tau(\mathcal{R}_n)} \text{ is non-singular}\big) \to 1,$$
$$\mathbf{P}\big(Q_{n,\tau(\mathcal{Q}_n)} \text{ is non-singular}\big) \to 1.$$

In proving Theorem 2.1, we also obtain the following new result, which states that for a wide range of probabilities p, with high probability there are no non-trivial linear dependencies in the random matrix $R_{n,p}$.

Theorem 2.2. *For any fixed $c > 1/2$, uniformly over $p \in (c \ln n/n, 1/2)$, we have*

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}) = n - z(R_{n,p})\big) = 1 - O((\ln \ln n)^{-1/2}).$$

The analogue of Theorem 2.2 for the symmetric process $\mathcal{Q}_n$ was established by Costello and Vu [10], and our analysis builds on theirs as well as that of [8]. The requirement that $c > 1/2$ in Theorem 2.2 is necessary, since for $p = c \ln n/n$ with $c < 1/2$, with probability $1 - o(1)$ the matrix $R_{n,p}$ will contain two identical rows, each with a single non-zero entry, as well as two identical columns, each with a single non-zero entry; in this case $\mathrm{rank}(R_{n,p}) < n - z(R_{n,p})$.

Our analyses of the processes $\mathcal{R}_n$ and $\mathcal{Q}_n$ are similar, but each presents its own difficulties. In the former, lack of symmetry yields a larger number of potential 'problematic configurations' to control; in the latter, symmetry reduces independence between the matrix entries. Where possible, we treat the two processes in a unified manner, but on occasion different proofs are required for the two models.

There are two main challenges in proving Theorem 2.1. First, there are existing bounds of the form of Theorem 2.2 for the symmetric Bernoulli process [10]. However, in both models, τ is of order $\ln n/n + \Theta(1/n)$. This is rather diffuse: it means that the moment when the last zero row/column disappears is spread over a region in which $\Theta(n)$ ones appear in the matrix ($\Theta(n)$ new edges appear in the associated graph). As such, a straightforward argument from Theorem 2.2 or its symmetric analogue, using a union bound, is impossible. This is not purely a weakness of our methods. Indeed, if the matrix contains two identical *non-zero* rows then it is singular, and the probability that there are two such rows (each containing a single non-zero entry, say) when $p \leqslant \ln n/n$ is $\Omega(\ln^2 n/n)$. This is already too large for a naive union bound to succeed. Moreover, with current techniques there seems no hope of replacing our bound by one that is even, say, $\Omega(n^\epsilon)$ for any positive ϵ, so another type of argument is needed.

The second challenge is that invertibility is not an increasing property (adding ones to a zero–one matrix can destroy invertibility). All existing proofs of hitting time theorems for graphs (of which we are aware) use monotonicity, usually in the following way. An *increasing graph property* is a collection $\mathcal{G}$ of graphs, closed under graph isomorphism, and such that if $G \in \mathcal{G}$ and G is a subgraph of H, then $H \in \mathcal{G}$. Suppose that $\mathcal{H}$ and $\mathcal{K}$ are

increasing graph properties with $\mathcal{H} \subset \mathcal{K}$. If there is a function $f(n, p)$ such that, uniformly in $0 < p < 1$,

$$\mathbf{P}\big(G_{n, f(n,p)} \in \mathcal{H}\big) = p + o(1) = \mathbf{P}\big(G_{n, f(n,p)} \in \mathcal{K}\big),$$

then with probability $1 - o(1)$, the first hitting times of $\mathcal{H}$ and of $\mathcal{K}$ coincide. This follows easily from the fact that $\mathcal{H}$ and $\mathcal{K}$ are increasing. However, it breaks down for non-increasing properties and there is no obvious replacement.

To get around these two issues, we introduce a method for *decoupling* the event of having full rank from the precise time the last zero row or column disappears. This method is most easily explained in graph terminology. We take a subset of the vertices of the graph under consideration, and replace their (random) out- and/or in-neighbourhoods with deterministic sets of neighbours. We prove results about the modified model, and then show that by a suitable averaging out, we can recover results about the original, fully random model. We believe the results about the partially deterministic models are independently interesting, and we now state them.

Definition 2.3. Given $n \geqslant 1$, a *template* (or *n-template*) is an ordered pair $\mathcal{L} = (\mathcal{L}^+, \mathcal{L}^-) = ((S_i^+)_{i \in I^+}, (S_j^-)_{j \in I^-})$, where

(1) I^+, I^- are subsets of $[n]$,
(2) $(S_i^+)_{i \in I^+}$ and $(S_j^-)_{j \in I^-}$ are sequences of non-empty, pairwise disjoint subsets of $[n]$,
(3) $\bigcup_{i \in I^+} S_i^+ \subset [n] \setminus I^-$ and $\bigcup_{j \in I^-} S_j^- \subset [n] \setminus I^+$.

The *size* of $\mathcal{L}$ is $\max(|I^+|, |I^-|, \max(|S_i^+|, i \in I^+), \max(|S_i^-|, i \in I^-))$. We write $\mathcal{I} = \mathcal{I}(\mathcal{L}) = (I^+, I^-)$. Also, we say $\mathcal{L}$ is *symmetric* if $\mathcal{L}^+ = \mathcal{L}^-$. Finally, for $l \in \mathbb{N}$, we let $\mathcal{M}^n(l)$ be the collection of *n*-templates of size at most l.

We remark that there is a unique template $\mathcal{L}$ of size zero, which satisfies $I^+ = \emptyset = I^-$; we call this template *degenerate*.

Given $n \geqslant 1$, an undirected or directed graph G on n vertices and a template $\mathcal{L}$ as defined above, let $G^{\mathcal{L}}$ be the graph obtained from G by letting each $i \in I^+$ have out-neighbours S_i^+ (and no others) and each $j \in I^-$ in-neighbours S_j^- (and no others). Note that if G is undirected and $\mathcal{L}$ is symmetric, then $G^{\mathcal{L}}$ is again undirected, provided we view a pair uv, vu of directed edges as a single undirected edge. We write $Q_{n,p}^{\mathcal{L}}$ and $R_{n,p}^{\mathcal{L}}$ for the adjacency matrix of $G_{n,p}^{\mathcal{L}}$ and $H_{n,p}^{\mathcal{L}}$. In $Q_{n,p}^{\mathcal{L}}$ and $R_{n,p}^{\mathcal{L}}$, for $i \in I^+$ (resp. $i \in I^-$), the non-zero entries of row i are precisely those with indices in S_i^+ (resp. in S_i^-).

Theorem 2.4. *Fix $K \in \mathbb{N}$ and $c > 1/2$. For any $p \in (c \ln n / n, 1/2)$ and any template $\mathcal{L} \in \mathcal{M}^n(K)$,*

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}^{\mathcal{L}}) = n - z(R_{n,p}^{\mathcal{L}})\big) = 1 - O((\ln \ln n)^{-1/2}).$$

If, moreover, $\mathcal{L}$ is symmetric then

$$\mathbf{P}\big(\mathrm{rank}(Q_{n,p}^{\mathcal{L}}) = n - z(Q_{n,p}^{\mathcal{L}})\big) = 1 - O((\ln \ln n)^{-1/4}).$$

We briefly remark on the assertions of the latter theorem. First, the first probability bound immediately implies Theorem 2.2, by taking I^+ and I^- to be empty. Next, the condition of pairwise disjointness is necessary. To see this, note that if vertices u, v have degree one and have a common neighbour w, then the rows of the adjacency matrix corresponding to u and w are identical, creating a non-degenerate linear relation. Finally, in proving the theorem we in fact only require that if $K = \max_{i \in I^+} |S_i^+|$ then $|I^+| \cdot K = o(p^{-2}/(n \ln n))$ (and similarly for the maximum size of S_i^-). As we do not believe this condition is optimal, we have opted for a more easily stated theorem. However, it would be interesting to know how far the boundedness condition could be weakened.

The proof of Theorem 2.4 is based on an analysis of an iterative exposure of minors; the first use of such a procedure to study the rank of random matrices was due to Komlós [13]. In brief, we first show that for suitably chosen $n' < n$, a fixed n'-by-n' minor of $R_{n,p}$ is likely to have nearly full rank. Adding the missing rows and columns one at a time, we then show that any remaining dependencies are likely to be 'resolved', *i.e.*, eliminated, by the added rows. Our argument is similar to that appearing in [10] for the study of $Q_{n,p}$, but there are added complications due to the fact that our matrices are partially deterministic on the one hand, and asymmetric on the other. The proof of Theorem 2.4 occupies a substantial part of the paper; a somewhat more detailed sketch appears in Section 4.

Vershynin [26] has very recently strengthened the bounds of Costello, Tao and Vu [8], showing that for a broad range of dense symmetric random matrices, the singularity probability decays at least as quickly as e^{-n^β}, for some (model-dependent) $\beta > 0$. It seems plausible (though not certain) that Vershynin's techniques could be transferred to the current, sparse setting, to yield bounds of the form $1 - O(e^{-(\ln \ln n)^\beta})$ in Theorems 2.2 and 2.4. However, we believe, and the results of [9] suggest, that aside from zero-rows, the most likely cause of singularity is two identical rows, each containing a single one. If the latter is correct then it should in fact be possible to obtain bounds of the form $1 - O(n^{1-2c})$ (for $c > 1/2$ as above); as alluded to earlier, for the moment such bounds seem out of reach.

Notation. Before proceeding with details, we briefly pause to introduce some terminology. Given an $m \times m$ matrix $M = (m_{ij})_{1 \leqslant i,j \leqslant m}$, the *deficiency* of M is the quantity

$$Y(M) := m - \operatorname{rank}(M) - z(M).$$

Also, for any $i, j \in [m]$ we let $M^{(i,j)}$ denote the matrix obtained by removing the ith row of M and the jth column; we refer to $M^{(i,j)}$ as the (i, j) *minor* of M. More generally, given $A, B \subset [m]$ we write $M^{(A,B)}$ for the matrix obtained from M by removing the rows indexed by A and the columns indexed by B. Also, for $1 \leqslant k \leqslant m$ we write $M[k] = (M_{ij})_{1 \leqslant i,j \leqslant k}$.

For a graph $G = (V, E)$ and $v \in V$, we write $N_G^+(v)$ for the set of out-neighbours of v in G and $N_G^-(v)$ for the set of in-neighbours of v in G. (If G is undirected then we write $N_G^+(v) = N_G^-(v) = N_G(v)$ for the set of neighbours of G.) If M is the adjacency matrix of G, then for $1 \leqslant i \leqslant m$ we write $N_M^+(i) = N_G^+(i)$, and similarly for $N_M^-(i)$ (and $N_M(i)$ if M is symmetric). Note that in this case, $Z^{\mathrm{ROW}}(M)$ and $Z^{\mathrm{COL}}(M)$ correspond to the sets $\{v \in V : |N_G^+(v)| = 0\}$ and $\{v \in V : |N_G^-(v)| = 0\}$, respectively.

Given real random variables X and Y we write $X \preceq_{\mathrm{st}} Y$ if, for all $t \in \mathbb{R}$, we have $\mathbf{P}(X \geqslant t) \leqslant \mathbf{P}(Y \geqslant t)$, and in this case say that Y *stochastically dominates* X. Finally, we omit floors and ceilings for readability whenever possible.

Outline. The structure of the remainder of the paper is as follows. In Section 3 we explain how to 'decouple' the linear dependencies of the matrix process from the time at which the last zero row or zero column disappears. This decoupling allows us to prove Theorem 2.1 assuming that Theorem 2.4 holds.

In Section 4 we introduce the *iterative exposure of minors*, analogous to the vertex exposure martingale for random graphs, which we use to study how $Y(R_{n,p}[m])$ changes as m increases. We then state a key 'coupling' lemma (Lemma 4.3), which asserts that for some $n' < n$ with $n - n'$ sufficiently small, the process $(Y(R_{n,p}[m]), n' \leqslant m \leqslant n)$ is stochastically dominated by a reflected simple random walk with strongly negative drift. Postponing the proof of this lemma, we then show how Theorem 2.4 follows by standard simple hitting probability estimates for simple random walk.

In Section 5, we describe 'good' structural properties, somewhat analogous to graph expansion, that we wish the matrices $R_{n,p}[m]$ to possess. The properties we require are tailored to allow us to apply *linear and quadratic Littlewood–Offord bounds* on the concentration of random sums. Proposition 5.10, whose proof is postponed, states that these properties hold with high probability throughout the iterative exposure of minors. Assuming the properties hold, it is then a straightforward matter to complete the proof of Lemma 4.3.

In Section 6 we complete the proof of Proposition 5.10. This, the most technical part of the proof, is most easily described in the language of random graphs rather than of random matrices. It is largely based on establishing suitable expansion and intersection properties of the neighbourhoods that hold for all 'small' sets of vertices and in each of the graphs $R_{n,p}[m]$ considered in the iterative exposure of minors.

Finally, Appendix A states standard binomial tail bounds that we use in our proofs, and Appendix B contains proofs of results that are either basic but technical, or that essentially follow from previous work but do not directly apply in our setting.

3. Decoupling connectivity from rank estimates: the proof of Theorem 2.1 from Theorem 2.4

In this section we explain how the first assertion of Theorem 2.1 follows from the first assertion of Theorem 2.4. A similar argument applies to the second assertion of Theorem 2.1; we comment on the necessary adjustments to the argument in Section 3.1.

Recall that the process $\mathcal{R}_n$ is generated by a family $\{U_{ij} : 1 \leqslant i \neq j \leqslant n\}$ of independent Uniform$[0, 1]$ random variables. Given $I^+, I^- \subset [n]$, let $\mathcal{I} = (I^+, I^-)$ and write

$$\mathcal{F}_{\mathcal{I}} = \sigma(\{U_{ij} : i \in I^+ \text{ or } j \in I^-\}),$$
$$\mathcal{G}_{\mathcal{I}} = \sigma(\{U_{ij} : i \in [n] \setminus I^+ \text{ and } j \in [n] \setminus I^-\}).$$

Informally, $\mathcal{F}_{\mathcal{I}}$ contains all the information that can be determined from the process $\mathcal{R}_n$ by observing only rows with indices in I^+ and columns with indices in I^-. All the information about all the remaining entries is contained in $\mathcal{G}_{\mathcal{I}}$.

Next, given $p \in (0,1)$, let

$$A_{\mathcal{I}}(p) = \{z(R_{n,p}^{(I^+,I^-)}) = 0\},$$

and let

$$B_{\mathcal{I}}(p) = \{I^+ \subset Z^{\text{ROW}}(R_{n,p}), I^- \subset Z^{\text{COL}}(R_{n,p})\}.$$

In words, $A_{\mathcal{I}}(p)$ is the event that the matrix obtained from $R_{n,p}$ by deleting the rows indexed by I^+ and the columns indexed by I^- has neither zero rows nor zero columns. We remark that $A_{\mathcal{I}}(p)$ and $B_{\mathcal{I}}(p)$ are measurable with respect to $\mathcal{G}_{\mathcal{I}}$ and $\mathcal{F}_{\mathcal{I}}$, respectively, and that $A_{\mathcal{I}}(p) \cap B_{\mathcal{I}}(p)$ is precisely the event that $Z^{\text{ROW}}(R_{n,p}) = I^+$ and $Z^{\text{COL}}(R_{n,p}) = I^-$. We write $A_{\mathcal{I}}, B_{\mathcal{I}}$ instead of $A_{\mathcal{I}}(p), B_{\mathcal{I}}(p)$ when the dependence on p is clear from context.

Finally, let

$$\tau_{\mathcal{I}} = \tau_{\mathcal{I}}(\mathcal{R}_n) = \min\{p \in (0,1) : Z^{\text{ROW}}(R_{n,p}) \cap I^+ = \emptyset = Z^{\text{COL}}(R_{n,p}) \cap I^-\}.$$

Then, for any template $\mathcal{L} = ((S_i^+)_{i \in I^+}, (S_j^-)_{j \in I^-})$, let

$$C_{\mathcal{L}} = \{\forall\ i \in I^+, N_{R_{n,\tau_{\mathcal{I}}}}^+(i) = S_i^+\} \cap \{\forall\ j \in I^-, N_{R_{n,\tau_{\mathcal{I}}}}^-(j) = S_j^-\}.$$

Observe that $C_{\mathcal{L}}$ and $\tau_{\mathcal{I}}$ are measurable with respect to $\mathcal{F}_{\mathcal{I}}$. Furthermore, the entries that are random in $R_{n,p}^{\mathcal{L}}$ are precisely those corresponding to the random variables generated by $\mathcal{G}_{\mathcal{I}}$. Lemma 3.1, below, uses this fact in order to express the conditional distribution of τ given $A_{\mathcal{I}}, B_{\mathcal{I}}$, and $C_{\mathcal{L}}$, as an integral against a conditional density function.

Lemma 3.1. *Fix $1 \leqslant l \leqslant n$ and $p \in (0,1)$. Then, for any non-degenerate n-template*

$$\mathcal{L} = ((S_i^+)_{i \in I^+}, (S_i^-)_{i \in I^-}),$$

letting $\mathcal{I} = (I^+, I^-)$ and writing $\tau = \tau(\mathcal{R}_n)$, we have

$$\mathbf{P}\big(Y(R_{n,\tau}) = 0 \mid A_{\mathcal{I}}(p), B_{\mathcal{I}}(p), C_{\mathcal{L}}\big) = \int_0^1 \mathbf{P}\big(Y(R_{n,t}^{\mathcal{L}}) = 0 \mid A_{\mathcal{I}}(p)\big) f(t \mid B_{\mathcal{I}}(p), C_{\mathcal{L}})\, \mathrm{d}t,$$

where $f(\cdot \mid B_{\mathcal{I}}(p), C_{\mathcal{L}})$ is the conditional density of $\tau_{\mathcal{I}}$ given $B_{\mathcal{I}}(p)$ and $C_{\mathcal{L}}$.

In proving Lemma 3.1 we use the following basic observation. Given independent σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$, for every $E_1, F_1 \in \mathcal{F}_1$ and $E_2, F_2 \in \mathcal{F}_2$, we have

$$\mathbf{P}(E_1, E_2 \mid F_1, F_2) = \mathbf{P}(E_1 \mid F_1)\mathbf{P}(E_2 \mid F_2).$$

Proof. If $A_{\mathcal{I}}, B_{\mathcal{I}}$ and $C_{\mathcal{L}}$ all occur, we necessarily have $\text{rank}(R_{n,\tau}) = \text{rank}(R_{n,\tau}^{\mathcal{L}})$ and $\tau_{\mathcal{I}} = \tau$. For any $K \in \mathbb{N}$, we may thus rewrite $\mathbf{P}(Y(R_{n,\tau}) = 0 \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}})$ as

$$\sum_{i=0}^{K-1} \mathbf{P}\left(Y(R_{n,\tau_{\mathcal{I}}}^{\mathcal{L}}) = 0, \tau_{\mathcal{I}} \in \left[\frac{i}{K}, \frac{i+1}{K}\right) \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}\right).$$

For any $0 \leqslant i < K$, if $i/K \leqslant \tau_{\mathcal{I}} < (i+1)/K$ and no edges arrive in the interval $(\tau_{\mathcal{I}}, (i+1)/K]$, then $R_{n,\tau_{\mathcal{I}}}^{\mathcal{L}}$ and $R_{n,(i+1)/K}^{\mathcal{L}}$ are identical. Writing D for the event that a pair of distinct

edges arrive within one of the intervals $\{[i/K, (i+1)/K], 0 \leqslant i < K\}$, it follows that

$$\left| \mathbf{P}(Y(R_{n,\tau}) = 0 \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}) - \sum_{i=0}^{K-1} \mathbf{P}\left(Y(R_{n,(i+1)/K}^{\mathcal{L}}), \tau_{\mathcal{I}} \in \left[\frac{i}{K}, \frac{i+1}{K} \right) \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}} \right) \right|$$

$$\leqslant \mathbf{P}\left(R_{n,\tau_{\mathcal{I}}}^{\mathcal{L}} \neq R_{n,\tau_{\mathcal{I}}+1/K}^{\mathcal{L}} \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}} \right) \leqslant \frac{\mathbf{P}(D)}{\mathbf{P}(A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}})}. \tag{3.1}$$

For fixed edges e and e', we have $\mathbf{P}(|U_e - U_{e'}| \leqslant 1/K) \leqslant 2/K$. By a union bound it follows that $\mathbf{P}(D) \leqslant (2n(n-1))/K$, and so the final term in (3.1) tends to 0 as $K \to \infty$.

Finally, by the observation about conditional independence just before the start of the proof, for all K we have

$$\sum_{i=0}^{K-1} \mathbf{P}\left(Y(R_{n,(i+1)/K}^{\mathcal{L}}), \tau_{\mathcal{I}} \in \left[\frac{i}{K}, \frac{i+1}{K} \right) \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}} \right)$$

$$= \sum_{i=0}^{K-1} \mathbf{P}\left(Y(R_{n,(i+1)/K}^{\mathcal{L}}) \mid A_{\mathcal{I}} \right) \mathbf{P}\left(\tau_{\mathcal{I}} \in \left[\frac{i}{K}, \frac{i+1}{K} \right) \mid B_{\mathcal{I}}, C_{\mathcal{L}} \right),$$

and taking $K \to \infty$ completes the proof. $\qquad\square$

The next definition captures the event that, for a given $p < \tau$, the rows (resp. columns) of $R_{n,\tau}$ indexed by $Z^{\text{ROW}}(R_{n,p})$ (resp. $Z^{\text{COL}}(R_{n,p})$) are such that Theorem 2.4 can be applied.

Definition 3.2. Given $0 < p < p' < 1$ and $K \in \mathbb{N}$, let $\mathcal{D}_K(p, p')$ be the event that

$$\left((N_{R_{n,p'}}^+(i))_{i \in Z^{\text{ROW}}(R_{n,p})}, (N_{R_{n,p'}}^-(i))_{i \in Z^{\text{COL}}(R_{n,p})} \right)$$

is a non-degenerate n-template of size at most K.

Lemma 3.3. *For any $\varepsilon > 0$ there exists $a > 0$ and integer $K = K(a) > 0$ such that, setting $p_1 = (\ln n - a)/n$ and $p_2 = (\ln n + a)/n$, for all n sufficiently large,*

$$\mathbf{P}(\tau(\mathcal{R}_n) > p_2) \leqslant \varepsilon,$$
$$\mathbf{P}(\mathcal{D}_K(p_1, \tau(\mathcal{R}_n))) \geqslant 1 - \varepsilon.$$

The proof of Lemma 3.3 is straightforward but technical, and is presented in Appendix B.

Now recall the definition of $Y(M)$ from the notation section on page 639. Observe that the first claim of Theorem 2.1 is equivalent to the statement that with high probability $Y(R_{n,\tau(\mathcal{R}_n)}) = 0$. Therefore, to establish the first assertion of Theorem 2.1, it suffices to prove the following theorem.

Theorem 3.4. $\mathbf{P}(Y(R_{n,\tau(\mathcal{R}_n)}) = 0) \to 1$ *as* $n \to \infty$.

Proof. Fix $\varepsilon > 0$, let $a > 0$ and $K = K(a) > 0$ be as in Lemma 3.3. Throughout the proof write $p_1 = (\ln n - a)/n$, $p_2 = (\ln n + a)/n$, and $\tau = \tau(\mathcal{R}_n)$. Note that $\mathcal{D}_K(p_1, \tau)$ occurs

precisely if there exists a non-degenerate $\mathcal{L} \in \mathcal{M}^n(K)$ such that $A_{\mathcal{I}(\mathcal{L})}(p_1), B_{\mathcal{I}(\mathcal{L})}(p_1)$, and $C_{\mathcal{L}}$ all occur. Furthermore, if $\mathcal{L} \neq \mathcal{L}'$ then

$$A_{\mathcal{I}(\mathcal{L})}(p_1) \cap B_{\mathcal{I}(\mathcal{L})}(p_1) \cap C_{\mathcal{L}} \quad \text{and} \quad A_{\mathcal{I}(\mathcal{L}')}(p_1) \cap B_{\mathcal{I}(\mathcal{L}')}(p_1) \cap C_{\mathcal{L}'}$$

are disjoint events. Writing $\hat{\mathcal{M}}^n(K) = \{\mathcal{L} \in \mathcal{M}^n(K) : \mathcal{L} \text{ is non-degenerate}\}$, it follows that

$$
\begin{aligned}
\mathbf{P}\big(Y(R_{n,\tau}) = 0\big) &\geqslant \mathbf{P}\big(Y(R_{n,\tau}) = 0, \mathcal{D}_K(p_1, \tau)\big) \\
&= \sum_{\mathcal{L} \in \hat{\mathcal{M}}^n(K)} \mathbf{P}\big(Y(R_{n,\tau}) = 0, A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}\big).
\end{aligned}
$$

We will show that for any $\mathcal{L} \in \hat{\mathcal{M}}^n(K)$,

$$\frac{\mathbf{P}\big(Y(R_{n,\tau}) = 0, A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}\big)}{\mathbf{P}\big(A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}, \tau \leqslant p_2\big)} \geqslant 1 - o(1). \tag{3.2}$$

Assuming this, it follows that

$$
\begin{aligned}
\mathbf{P}\big(Y(R_{n,\tau}) = 0\big) &\geqslant (1 - o(1)) \sum_{\mathcal{L} \in \hat{\mathcal{M}}^n(K)} \mathbf{P}\big(A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}, \tau \leqslant p_2\big) \\
&= (1 - o(1))\mathbf{P}\big(\mathcal{D}_K(p_1, \tau), \tau \leqslant p_2\big) \\
&\geqslant 1 - 2\varepsilon
\end{aligned}
$$

for n large, the last inequality by Lemma 3.3. Since $\varepsilon > 0$ was arbitrary, it thus remains to prove (3.2), for which we use Lemma 3.1.

Fix $\mathcal{L} \in \hat{\mathcal{M}}^n(K)$ and let $N = n - |I^+| \geqslant n - K$. Then

$$p_1 = \frac{\ln n - a}{n} = \frac{\ln N - a}{N} + O\left(\frac{1}{N^2}\right).$$

For any fixed integer $i \geqslant 1$ and distinct $v_1, \ldots, v_i \in [n] \setminus (I^+ \cup \bigcup_{i \in I^-} S_i^-)$,

$$\mathbf{P}\left(\bigcap_{j=1}^{i} \{N_{R_{n,p_1}^{\mathcal{L}}}^+(v_j) = \emptyset\}\right) = (1 - p_1)^{i(N-i) + i(i-1)} = (1 + o(1))(1 - p_1)^{N \cdot i}.$$

Furthermore, for $j \in I^+ \cup \bigcup_{i \in I^-} S_i^-$ necessarily

$$N_{R_{n,p_1}^{\mathcal{L}}}^+(v_j) \neq \emptyset,$$

since $\mathcal{L}$ is non-degenerate. It follows by the method of moments (see [12], Chapter 6) that $|Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}})|$ is asymptotically Poisson(e^a). The same argument establishes that $|Z^{\text{COL}}(R_{n,p}^{\mathcal{L}})|$ has the same asymptotic distribution. It follows that $\mathbf{P}(A_{\mathcal{I}}) \leqslant 2e^{-e^a} + o(1)$, so by the first assertion of Theorem 2.4, for $p \in (p_1, p_2)$ we have

$$\mathbf{P}\big(Y(R_{n,p}^{\mathcal{L}}) > 0 \mid A_{\mathcal{I}}\big) \leqslant \frac{\mathbf{P}(Y(R_{n,p}^{\mathcal{L}}) > 0)}{\mathbf{P}(A_{\mathcal{I}})} = \frac{O(e^{e^a})}{(\ln \ln n)^{1/2}}. \tag{3.3}$$

Since $a = O(1)$, by Lemma 3.1 we thus have

$$\mathbf{P}\big(Y(R_{n,\tau}) = 0 \mid A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}\big) \geqslant \int_{p_1}^{p_2} \mathbf{P}\big(Y(R_{n,t}^{\mathcal{L}}) = 0 \mid A_{\mathcal{I}}\big) f(t \mid B_{\mathcal{I}}, C_{\mathcal{L}}) \, \mathrm{d}t$$

$$\geqslant (1 - o(1)) \int_{p_1}^{p_2} f(t \mid B_{\mathcal{I}}, C_{\mathcal{L}}) \, \mathrm{d}t$$

$$= (1 - o(1)) \mathbf{P}\big(\tau_{\mathcal{I}} \in [p_1, p_2] \mid B_{\mathcal{I}}, C_{\mathcal{L}}\big).$$

Multiply both sides of the preceding inequality by $\mathbf{P}(A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}})$. Since $A_{\mathcal{I}}$ is independent from $B_{\mathcal{I}}$ and $C_{\mathcal{L}}$, we obtain

$$\mathbf{P}\big(Y(R_{n,\tau}) = 0, A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}\big) \geqslant (1 - o(1)) \mathbf{P}\big(A_{\mathcal{I}}, B_{\mathcal{I}}, C_{\mathcal{L}}, \tau_{\mathcal{I}} \in [p_1, p_2]\big).$$

Finally, if $A_{\mathcal{I}}$, $B_{\mathcal{I}}$, and $C_{\mathcal{L}}$ all occur then necessarily $\tau_{\mathcal{I}} = \tau$ and $\tau > p_1$, so we may replace $\{\tau_{\mathcal{I}} \in [p_1, p_2]\}$ by $\{\tau \leqslant p_2\}$ in the final probability, and (3.2) follows. The proof is complete. $\square$

3.1. Notes on the proof of Theorem 2.1 for $G_{n,p}$

The decoupling of connectivity from the rank estimates of $R_{n,p}$ is not particularly sensitive to the structure of $R_{n,p}$ except through Theorem 2.4, and the broad strokes of the argument of this section are therefore unchanged. In particular, define $\mathcal{F}_{\mathcal{I}}$ and $\mathcal{G}_{\mathcal{I}}$ as before (but recall that only the variables $\{U_{ij}\}_{1 \leqslant i < j \leqslant n}$ are independent). Then Lemma 3.1 holds under the additional restriction that the template $\mathcal{L}$ is *symmetric*, as in this case $I^+ = I^-$ and the σ-algebras $\mathcal{F}_{\mathcal{I}}$ and $\mathcal{G}_{\mathcal{I}}$ are indeed symmetric. We replace the event $D_K(p, p')$ with the event that

$$\big((N_{Q_{n,p'}}(i))_{i \in Z(Q_{n,p})}, (N_{Q_{n,p'}}(i))_{i \in Z(Q_{n,p})}\big)$$

is a non-degenerate n-template of size at most K (in which case it is necessarily symmetric). Lemma 3.3 then holds with $\tau(\mathcal{R}_n)$ replaced by $\tau(\mathcal{Q}_n)$, with an essentially identical proof. Assuming the second bound in Theorem 2.4, the rest of the proof then follows without substantial changes.

4. Analysis of the iterative exposure process: the proof of Theorem 2.4 modulo a coupling lemma

To prove Theorem 2.4, we analyse an iterative exposure of minors of the matrix. (In other words, we will expose the edges incident to the vertices of $H_{n,p}^{\mathcal{L}}$ in a vertex-by-vertex fashion.) This strategy was first used in the context of random symmetric matrices in [8], to show that random symmetric Bernoulli$(1/2)$ matrices are almost surely non-singular.

For the remainder of the paper, $c \in (1/2, 1)$ is a fixed constant, and $c \ln n / n \leqslant p \leqslant 1/2$. Also, for the rest of the paper, let $\alpha < 1$ be such that $\alpha c \in (1/2, 3/4)$, write $\gamma = \alpha c - 1/2$, and let $n' = \lceil \alpha n \rceil$. Given any integer $K \geqslant 1$, for n sufficiently large and any n-template $\mathcal{L} = ((S_i^+)_{i \in I^+}, (S_i^-)_{i \in I^-}) \in \mathcal{M}^n(K)$, by permuting the rows and columns of $R_{n,p}^{\mathcal{L}}$ we may assume that

$$I^+ \cup I^- \cup \bigcup_{i \in I^+} S_i^+ \cup \bigcup_{i \in I^-} S_i^- \subset [n'].$$

We call such $\mathcal{L} \in \mathcal{M}^n(K)$ *permissible*. We work only with permissible templates to ensure that in the iterative exposure of minors in $R_{n,p}^{\mathcal{L}}$ starting from $R_{n,p}^{\mathcal{L}}[n']$, all new off-diagonal matrix entries whose row (resp. column) is not in $\bigcup_{i \in I^+} S_i^+$ (resp. $\bigcup_{i \in I^-} S_i^-$) are Bernoulli($p$) distributed. We begin by showing that $R_{n,p}^{\mathcal{L}}[n']$ is extremely likely to have quite large rank.

Lemma 4.1. *For any $\varepsilon > 0$ and $c \ln n/n \leqslant p \leqslant 1/2$, there exists a constant c_1 such that*

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}) \geqslant (1 - \varepsilon)n\big) \geqslant 1 - O(n^{-c_1 n}).$$

The proof of an analogous bound for random symmetric sparse Bernoulli matrices appears in [10].

Proof. Let $\mathbf{r}_1, \ldots, \mathbf{r}_n$ denote the rows of $R_{n,p}$. Let $S = \mathrm{span}(\mathbf{r}_1, \ldots, \mathbf{r}_{\lfloor (1-\varepsilon)n \rfloor})$ and $d = \dim S$. Let R be the event that $\mathbf{r}_i \in S$ for every $\lfloor (1 - \varepsilon)n \rfloor < i \leqslant n$. By symmetry we have

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}) \leqslant (1 - \varepsilon)n\big) \leqslant \binom{n}{\varepsilon n} \mathbf{P}(R), \tag{4.1}$$

so we now focus on bounding $\mathbf{P}(R)$. By relabelling (*i.e.*, replacing $R_{n,p}$ by $U R_{n,p} U^{-1}$ for some permutation matrix U), we may assume $R_{n,p}$ is a block matrix

$$R_{n,p} = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A is an $\lfloor (1 - \varepsilon)n \rfloor \times d$ matrix with $\mathrm{rank}(A) = d \leqslant (1 - \varepsilon)n$. Thus, the columns in B are in the span of the columns in A, so $AG = B$ for some (unique) matrix G.

On the other hand, D is an $\lceil \varepsilon n \rceil \times (n - d)$ matrix and $d \leqslant \varepsilon n$. It follows that there exists $c_2 > 0$ such that, for any fixed matrix M,

$$\mathbf{P}(D = M) \leqslant (1 - p)^{(\varepsilon n)^2 - \varepsilon n} \leqslant e^{-\varepsilon^2 cn \ln n + \varepsilon n/2} \leqslant n^{-c_2 n}.$$

The first inequality holds since D has at least $(\varepsilon n)^2 - \varepsilon n$ independent Bernoulli(p) entries.

Now, if R holds, then there exists a matrix F such that both $FA = C$ and $FB = D$ hold. Furthermore, note that if F' also satisfies $F'A = C$ then

$$FB = FAG = F'AG = F'B,$$

so if R occurs then $D = FB$ is uniquely determined by A, B and C. Consequently, for any such F we have

$$\mathbf{P}(R \mid A, B, C) \leqslant \mathbf{P}(D = FB \mid A, B, C) \leqslant n^{-c_2 n}.$$

Since $\mathbf{P}(R) = \mathbf{E}[\mathbf{P}(R \mid A, B, C)]$, by (4.1) it follows that

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}) \leqslant (1 - \varepsilon)n\big) \leqslant \binom{n}{\varepsilon n} \mathbf{P}(R) \leqslant \left(\frac{e}{\varepsilon}\right)^{\varepsilon n} n^{-c_2 n}. \qquad \square$$

In fact, Lemma 4.1 also holds for $R_{n,p}^{\mathcal{L}}[n']$ as well. To see this, observe that since $\mathcal{L}$ has size at most K,

$$|\text{rank}(R_{n,p}^{\mathcal{L}}[n']) - \text{rank}(R_{n,p}[n'])| \leqslant K^2 = O(1),$$

and $R_{n,p}[n']$ has the same distribution as $R_{n',p}$. We thus obtain the following corollary.

Corollary 4.2. *For any $K \in \mathbb{N}$ and $\varepsilon > 0$, there exists a constant c_1 such that uniformly over $\mathcal{L} \in \mathcal{M}^n(K)$ and $c \ln n / n \leqslant p \leqslant 1/2$,*

$$\mathbf{P}\big(\text{rank}(R_{n,p}^{\mathcal{L}}[n']) \geqslant (1-\varepsilon)n\big) \geqslant 1 - O(n^{-c_1 n}).$$

We next consider how the deficiency $Y(R_{n,p}^{\mathcal{L}}[m])$ drops as m increases from n' to n. We have

$$\begin{aligned}
Y(R_{n,p}^{\mathcal{L}}[m+1]) = {}& Y(R_{n,p}^{\mathcal{L}}[m]) + 1 \\
& + (z(R_{n,p}^{\mathcal{L}}[m]) - z(R_{n,p}^{\mathcal{L}}[m+1])) \\
& - (\text{rank}(R_{n,p}^{\mathcal{L}}[m+1]) - \text{rank}(R_{n,p}^{\mathcal{L}}[m])),
\end{aligned} \tag{4.2}$$

so Y decreases as the rank increases and, on the other hand, increases when zero rows or zero columns disappear causing that $z(R_{n,p}^{\mathcal{L}}[m]) > z(R_{n,p}^{\mathcal{L}}[m+1])$. To show that $Y(R_{n,p}^{\mathcal{L}})$ is likely zero, we will couple $(Y(R_{n,p}^{\mathcal{L}}[m]), n' \leqslant m \leqslant n)$ to a simple random walk with strongly negative drift, in such a way that with high probability the random walk provides an upper bound for $(Y(R_{n,p}^{\mathcal{L}}[m]), n' \leqslant m \leqslant n)$. Of course, showing that such a coupling exists involves control on the rank increase and on the decrease in $z(R_{n,p}^{\mathcal{L}}[m])$ as m increases from n' to n. Observe that we always have

$$\text{rank}(R_{n,p}^{\mathcal{L}}[m]) \leqslant \text{rank}(R_{n,p}^{\mathcal{L}}[m+1]) \leqslant \text{rank}(R_{n,p}^{\mathcal{L}}[m]) + 2$$

since $R_{n,p}^{\mathcal{L}}[m]$ may be obtained from $R_{n,p}^{\mathcal{L}}[m+1]$ by deleting a single row and column. It follows from (4.2) that if $z(R_{n,p}^{\mathcal{L}}[m]) = z(R_{n,p}^{\mathcal{L}}[m+1])$ then

$$Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m]) \in \{-1, 0, 1\}.$$

Also, if $z(R_{n,p}^{\mathcal{L}}[m]) = z(R_{n,p}^{\mathcal{L}}[m+1]) - 1$ then necessarily

$$\text{rank}(R_{n,p}^{\mathcal{L}}[m+1]) \geqslant \text{rank}(R_{n,p}^{\mathcal{L}}[m]) + 1,$$

and so

$$Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m]) \in \{0, 1\}.$$

Together, this shows that

$$|Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m])| \leqslant 1$$

whenever

$$z(R_{n,p}^{\mathcal{L}}[m]) - z(R_{n,p}^{\mathcal{L}}[m+1]) \leqslant 1.$$

Establishing further control on the rank increase is rather involved, and this is the primary work of Sections 5 and 6. It will turn out that typically

$$Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m]) = -1$$

when $Y(R_{n,p}^{\mathcal{L}}[m]) > 0$, and $Y(R_{n,p}^{\mathcal{L}}[m+1]) = 0$ when $Y(R_{n,p}^{\mathcal{L}}[m]) = 0$. More precisely, we have the following lemma.

Lemma 4.3. *For fixed $K \in \mathbb{N}$, there exists $C > 0$ such that the following holds. Given integer $n \geqslant 10$, let $\beta = \beta(n) = C(\ln \ln n)^{-1/2}$. Then uniformly over $\mathcal{L} \in \mathcal{M}^n(K)$ and $c \ln n/n \leqslant p \leqslant 1/2$, there exists a coupling of $(Y(R_{n,p}^{\mathcal{L}}[m]), n' \leqslant m \leqslant n)$ and a collection $(X_m, n' \leqslant m < n)$ of i.i.d. random variables with $\mathbf{P}(X_i = 1) = \beta$ and $\mathbf{P}(X_i = -1) = 1 - \beta$, such that with probability $1 - O(n^{-\gamma/2})$, for all $n' \leqslant m < n$,*

$$Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m]) \leqslant \begin{cases} X_m & \text{if } Y(R_{n,p}^{\mathcal{L}}[m]) > 0, \\ \max(X_m, 0) & \text{if } Y(R_{n,p}^{\mathcal{L}}[m]) = 0. \end{cases}$$

The proof of Lemma 4.3 occupies much of the remainder of the paper. We say the coupling in the preceding lemma *succeeds* if, for all $n' < m \leqslant n$, the final inequality holds.

Now fix $(X_i, i \geqslant 1)$ i.i.d. random variables with $\mathbf{P}(X_i = 1) = \beta$ and $\mathbf{P}(X_i = -1) = 1 - \beta$. Set $S_0 = 0$, and for $k \geqslant 1$ let $S_k = \sum_{i=1}^{k} X_i$. We call $(S_k, k \geqslant 0)$ a *β-biased simple random walk* (SRW). Also, for $k \geqslant 1$ let $M_k = \min_{0 \leqslant i \leqslant k} S_i$, and let $D_k = S_k - M_k$.

Observe that when S_k is not at a new global minimum, D_{k+1} is either $D_k + 1$ (with probability β) or $D_k - 1$. On the other hand, when S_k is at a new minimum then $D_k = 0$, and either $D_{k+1} = 1$ (again with probability β) or $D_{k+1} = 0$.

Now imagine for a moment that $Y(R_{n,p}^{\mathcal{L}}[n']) = 0$. In this case, in view of the preceding paragraph, if the coupling succeeds then we have $D_k \geqslant Y(R_{n,p}^{\mathcal{L}}[n' + k])$ for all $0 \leqslant k \leqslant n - n'$. It follows that if $Y(R_{n,p}^{\mathcal{L}}[n'])$ happens to equal zero then we can bound $\mathbf{P}(Y(R_{n,p}^{\mathcal{L}}[n]) > 0)$ by bounding $\mathbf{P}(D_{n-n'} > 0)$. This is accomplished by the following proposition and its corollary.

Proposition 4.4. *Let $H = |\{k \geqslant 0 : S_k \geqslant 1\}|$. Then $\mathbf{E}(H) = \beta/(1-\beta)^2$.*

Proof. This is an elementary fact about hitting times for simple random walk, and in particular follows from Examples 1.3.3 and 1.4.3 of [17]. □

Corollary 4.5. *For $k \geqslant 0$, $\mathbf{P}(D_k > 0) < \beta/(1-\beta)^2$.*

Proof. For any $i \leqslant k$ we have

$$\mathbf{P}\left(D_k > 0, S_i = \min_{1 \leqslant j \leqslant k} S_j\right) \leqslant \mathbf{P}(S_{k-i} \geqslant 1),$$

and summing over i, plus a union bound, yields

$$\mathbf{P}(D_k > 0) < \sum_{i \geqslant 0} \mathbf{P}(S_i \geqslant 1) = \mathbf{E}(H). \qquad \square$$

In reality, $Y(R_{n,p}^{\mathcal{L}}[n'])$ may not equal zero, and so we should start the random walk S not from zero but from a positive height. The following corollary addresses this.

Corollary 4.6. *For any integers* $d, k \geqslant 1$,

$$\mathbf{P}\big(S_k + d > \min(M_k + d, 0)\big) \leqslant \mathbf{P}(S_k > -d) + \beta/(1-\beta)^2.$$

Proof. Let $\tau = \inf\{i : S_i = -d\}$. By the Markov property, for any $i \leqslant k$,

$$\mathbf{P}\big(S_k + d > \min(M_k + d, 0)|\tau = i\big) = \mathbf{P}(D_{k-i} > 0) < \beta/(1-\beta)^2$$

by the preceding corollary. On the other hand, $\mathbf{P}(\tau > k) \leqslant \mathbf{P}(S_k > -d)$, and the result follows. $\qquad\square$

On the other hand, if $t > Y(R_{n,p}^{\mathcal{L}}[n'])$ then when the coupling succeeds we have $S_k + t - \min(M_k + t, 0) \geqslant Y(R_{n,p}^{\mathcal{L}}[n' + k])$ for all $0 \leqslant k \leqslant n - n'$. It then follows from Lemma 4.3 and Corollary 4.6 that for any $\varepsilon > 0$,

$$\mathbf{P}\big(Y(R_{n,p}^{\mathcal{L}}) > 0\big) \leqslant \mathbf{P}\big(Y(R_{n,p}^{\mathcal{L}}[n']) > \varepsilon n\big) + \mathbf{P}\big(S_{n-n'} > -\varepsilon n\big) + \frac{\beta}{(1-\beta)^2} + O(n^{-\gamma/2})$$

$$\leqslant n^{-\Omega(n)} + e^{-\Omega(n)} + \frac{\beta}{(1-\beta)^2} + O(n^{-\gamma/2})$$

$$= O((\ln\ln n)^{-1/2}),$$

where the second inequality follows from a Chernoff bound for $\mathbf{P}(S_{n-n'} > -\varepsilon n)$ (assuming ε is chosen small enough), plus the bound from Corollary 4.2, and the last inequality follows from the definition of β in Lemma 4.3. This proves the first assertion of Theorem 2.4.

When treating the symmetric model $Q_{n,p}^{\mathcal{L}}$, the following modifications are required. First, Corollary 4.2 holds for all symmetric n-templates $\mathcal{L} \in \mathcal{M}^n(K)$ and with $Q_{n,p}^{\mathcal{L}}[n']$ in place of $R_{n,p}^{\mathcal{L}}[n']$. This was proved in [10] for $Q_{n,p}[n']$, but as $\mathcal{L}$ has size K,

$$|\text{rank}(Q_{n,p}[n']) - \text{rank}(Q_{n,p}^{\mathcal{L}}[n'])| \leqslant K^2 = O(1),$$

so the same bound holds for $Q_{n,p}^{\mathcal{L}}[n']$.

Second, we will likewise establish a coupling lemma for $Y(Q_{n,p}^{\mathcal{L}}[m])$.

Lemma 4.7. *For fixed* $K \in \mathbb{N}$, *there exists* $C > 0$ *such that the following holds. Given integer* $n \geqslant 10$, *let* $\beta = \beta(n) = C(\ln\ln n)^{-1/4}$. *Then uniformly over symmetric* $\mathcal{L} \in \mathcal{M}^n(K)$ *and* $c\ln n/n \leqslant p \leqslant 1/2$, *there exists a coupling of* $(Y(Q_{n,p}^{\mathcal{L}}[m]), n' \leqslant m \leqslant n)$ *and a collection* $(X_m, n' \leqslant m < n)$ *of i.i.d. random variables with* $\mathbf{P}(X_i = 1) = \beta$ *and* $\mathbf{P}(X_i = -1) = 1 - \beta$, *such that with probability* $1 - O(n^{-\gamma/2})$, *for all* $n' \leqslant m < n$,

$$Y((Q_{n,p}^{\mathcal{L}}[m+1]) - Y(Q_{n,p}^{\mathcal{L}}[m]) \leqslant \begin{cases} X_m & \text{if } Y(Q_{n,p}^{\mathcal{L}}[m]) > 0, \\ \max(X_m, 0) & \text{if } Y(Q_{n,p}^{\mathcal{L}}[m]) = 0. \end{cases}$$

Together, these two ingredients yield the second claim of Theorem 2.4 by a reprise of the arguments following Lemma 4.3. The remainder of the paper is therefore devoted to proving Lemmas 4.3 and 4.7.

5. Rank increase via iterative exposure

In this section we focus on understanding when and why the rank increases. In what follows, fix an $m \times m$ matrix $Q = (q_{i,j})_{1 \leqslant i,j \leqslant m}$. Given vectors $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_m)$, we write

$$\Gamma(Q, \mathbf{x}, \mathbf{y}) = \begin{pmatrix} q_{1,1} & \cdots & q_{1,m} & y_1 \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ q_{m,1} & \cdots & q_{m,m} & y_m \\ x_1 & \cdots & x_m & 0 \end{pmatrix}.$$

Now fix a matrix Q, and vectors $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_m)$. We remark that

$$\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) = \mathrm{rank}(Q) + 2$$

if and only if $\mathbf{x}$ is linearly independent of the non-zero rows of Q (*i.e.*, it does not lie in the row-span of Q) and $\mathbf{y}^T$ is linearly independent of the non-zero columns of Q. (In particular, if Q is symmetric and $\mathbf{x} = \mathbf{y}$ then $\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) = \mathrm{rank}(Q) + 2$ if and only if $\mathbf{x}$ lies outside the row-span of Q.) Note that for this to occur Q cannot have full rank.

We prove Lemmas 4.3 and 4.7 as follows. First, we describe structural properties of 0–1 matrices such that for any matrix $Q = (q_{ij})_{1 \leqslant i,j \leqslant m}$ satisfying such properties, for suitable random vectors $\mathbf{x}$ and $\mathbf{y}$, with high probability $\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) = \min(\mathrm{rank}(Q) + 2, m + 1)$. We then establish that with high probability, the matrices $(Q_{n,p}^{\mathcal{L}}[m], n' \leqslant m \leqslant n)$ and $(R_{n,p}^{\mathcal{L}}[m], n' \leqslant m \leqslant n)$ all have the requisite properties.

More precisely, for *fixed* Q and vectors $\mathbf{x}, \mathbf{y}$, we will see that

$$\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) = \min(\mathrm{rank}(Q) + 2, m + 1)$$

if and only if a suitable linear, bilinear or quadratic form in $\mathbf{x}$ and $\mathbf{y}$, with coefficients determined by the matrix Q, vanishes; we elaborate on this very shortly. When $\mathbf{x}$ and $\mathbf{y}$ are Bernoulli random vectors, this leads us to evaluate the probability that a particular random sum is equal to zero. To bound such probabilities, we use *Littlewood–Offord bounds* proved in [10, 7], which we now state.

Proposition 5.1. *Let* $x_1, \ldots, x_k, y_1, \ldots, y_k$ *be independent Bernoulli(p) random variables.*

(a) *Fix* $a_1, \ldots, a_k \in \mathbb{R} \setminus \{0\}$. *Then, uniformly over* $0 < p \leqslant 1/2$,

$$\sup_{r \in \mathbb{R}} \mathbf{P}\left(\sum_{i=1}^{k} a_i x_i = r \right) = O((kp)^{-1/2}).$$

(b) *Fix $l \geqslant 1$ and $(a_{ij}, 1 \leqslant i, j \leqslant k)$ such that there are at least l indices j for which $|\{i : a_{i,j} \neq 0\}| \geqslant l$. Then, uniformly over $0 < p \leqslant 1/2$,*

$$\sup_{r \in \mathbb{R}} \mathbf{P}\left(\sum_{1 \leqslant i,j \leqslant k} a_{ij} x_i y_j = r \right) = O((lp)^{-1/2}).$$

(c) *With $l \geqslant 1$ and $(a_{ij}, 1 \leqslant i, j \leqslant k)$ as in (b), if also $a_{ij} = a_{ji}$ for all $1 \leqslant i, j \leqslant k$, then uniformly over $0 < p \leqslant 1/2$,*

$$\sup_{r \in \mathbb{R}} \mathbf{P}\left(\sum_{1 \leqslant i,j \leqslant k} a_{ij} x_i x_j = r \right) = O((lp)^{-1/4}).$$

The matrix structural properties we require are precisely those that allow us to apply the bounds of Proposition 5.1. For this, the following definitions are germane.

Definition 5.2. Fix a matrix $Q = (q_{ij})_{1 \leqslant i,j \leqslant m}$.

- Given $S \subset [m]$, we say that $j \in [m]$ is an *S-selector* (for Q) if $|\{i \in S : q_{i,j} \neq 0\}| = 1$.
- Given $2 \leqslant b \leqslant m$, we say Q is *b-blocked* if any set $S \subset [m]$ with $S \cap Z^{\text{ROW}}(Q) = \emptyset$ and $2 \leqslant |S| \leqslant b$ has at least two S-selectors $j, l \in [m]$.

The final condition in the definition says that in the submatrix formed by only looking at the rows in S, there are at least two columns containing exactly one non-zero entry. We call j an S-selector as we think of j as 'selecting' the unique row i with $q_{ij} \neq 0$. We remark that if a matrix Q is b-blocked then any set S of non-zero rows of Q containing a linear dependency must have size at least $b + 1$. More strongly, this is true even after deleting any single column of Q.

Definition 5.3. We say that Q is *b-dense* if

$$|\{i \in [m] : \text{row } i \text{ of } Q \text{ has } > 1 \text{ non-zero entry}\}| \geqslant b.$$

We then have the following bounds, which are key to the proofs of Lemmas 4.3 and 4.7.

Proposition 5.4. *Fix integers $m \geqslant b \geqslant 1$ and a b-blocked $m \times m$ matrix Q with $\operatorname{rank}(Q) < m - z^{\text{ROW}}(Q)$. Then uniformly over $0 < p \leqslant 1/2$, if $\mathbf{y} = (y_1, \ldots, y_m)$ has i.i.d. Bernoulli(p) entries, then $\mathbf{y}^T$ is independent of the non-zero columns of Q with probability at least $1 - O((bp)^{-1/2})$.*

Proof. Let $k = \operatorname{rank}(Q)$, and note that if $\operatorname{rank}(Q) < m - z^{\text{ROW}}(Q)$ then $1 \leqslant k < m$. Write $\mathbf{r}_1, \ldots, \mathbf{r}_m$ for the rows of Q. By relabelling, we may assume that $\mathbf{r}_1, \ldots, \mathbf{r}_k$ are linearly independent and that $\mathbf{r}_{k+1}$ is non-zero. It follows that there exist unique coefficients $a_1, \ldots, a_k$ for which $\mathbf{r}_{k+1} = \sum_{i=1}^{k} a_i \mathbf{r}_i$. Then $\{\mathbf{r}_i : a_i \neq 0\} \cup \{\mathbf{r}_{k+1}\}$ forms a set of linearly dependent non-zero rows, and so has size at least $b + 1$ by the observation just after Definition 5.2.

Let $\hat{Q}$ be the matrix obtained from Q by adding $\mathbf{y}^T$ as column $m+1$. If $\mathbf{y}^T$ lies in the column-span of Q then $\operatorname{rank}(Q) = \operatorname{rank}(\hat{Q})$, so necessarily

$$y_{k+1} = \sum_{i=1}^{k} a_i y_i.$$

Since $|\{\mathbf{r}_i : a_i \neq 0\}| \geqslant b$, by Proposition 5.1(a) we have

$$\mathbf{P}\left(\sum_{i=1}^{k} a_i y_i = 0\right) = O((kp)^{-1/2}) = O((bp)^{-1/2}).$$

Therefore, the vector $\mathbf{y}^T$ is independent of the non-zero columns of Q with probability at least $1 - O((bp)^{-1/2})$. □

Proposition 5.5. *Fix integers $m \geqslant b \geqslant 1$ and a b-blocked, b-dense, $m \times m$ matrix Q with $\operatorname{rank}(Q) = m$. Then, uniformly over $0 < p \leqslant 1/2$, if $\mathbf{x} = (x_1, \dots, x_m)$ and $\mathbf{y} = (y_1, \dots, y_m)$ have i.i.d. Bernoulli(p) entries, then*

$$\mathbf{P}\big(\operatorname{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) = m\big) = O((bp)^{-1/2}).$$

Proof. Let $A = (a_{ij})_{1 \leqslant i,j \leqslant m}$ be the cofactor matrix of Q; that is,

$$a_{ij} = (-1)^{i+j+1} \det(Q^{(i,j)}),$$

where $Q^{(i,j)}$ is the (i,j) minor of Q. A double-cofactor expansion of the determinant of $Q' = \Gamma(Q, \mathbf{x}, \mathbf{y})$ yields

$$\det(Q') = \sum_{i,j=1}^{m} a_{ij} x_i y_j.$$

Note that $a_{ij} = 0$ when $Q^{(i,j)}$ is singular, so we want to lower-bound the number of non-singular minors $Q^{(i,j)}$. To do so, fix $j \in [m]$ and write $Q^{(\emptyset,j)}$ for the $m \times m-1$ matrix obtained by deleting the jth column of Q. Since Q has full rank it has no zero rows. We claim that if $Q^{(\emptyset,j)}$ also has no zero rows then $|\{i \in [m] : a_{ij} \neq 0\}| > b$. To see this, note that $Q^{(\emptyset,j)}$ has rank $m-1$, and so since there are no zero rows, there exists (up to scaling factors) a unique vanishing linear combination

$$\sum_{i=1}^{m} c_i \mathbf{r}_i^j = 0,$$

where $\mathbf{r}_i^j$ is the ith row of $Q^{(\emptyset,j)}$. Now, $Q^{(i,j)}$ is invertible (and thus $a_{ij} \neq 0$) if and only if $c_i \neq 0$. But the rows $\{\mathbf{r}_i^j : c_i \neq 0\}$ are linearly dependent, and by the remark just after Definition 5.2, since Q is b-blocked, we therefore have $|\{i \in [m] : c_i \neq 0\}| > b$.

Finally, since Q is b-dense, there are at most $m-b$ rows of Q with exactly one non-zero entry. Thus, $|\{j \in [m] : Q^{(\emptyset,j)} \text{ has no zero rows}\}| \geqslant b$, and for any such j we have $|\{i \in [m] : a_{ij} \neq 0\}| > b$ by the preceding paragraph. By Proposition 5.1(b) it follows that, uniformly in $0 < p \leqslant 1/2$, we have $\mathbf{P}(\det(Q') = 0) \leqslant O((bp)^{-1/2})$ as claimed. □

The following proposition is an analogue of Proposition 5.5 which we use in analysing the symmetric Bernoulli process.

Proposition 5.6. *Fix integers $m \geqslant b \geqslant 1$ and a b-blocked, b-dense, $m \times m$ symmetric matrix Q with $\mathrm{rank}(Q) = m$. Then uniformly over $0 < p \leqslant 1/2$, if $\mathbf{x} = (x_1, \ldots, x_m)$ has i.i.d. Bernoulli(p) entries, then*

$$\mathbf{P}(\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{x})) = m) = O((bp)^{-1/4}).$$

Proof sketch. The proof is nearly identical to that of Proposition 5.5. However, in this case the double cofactor expansion of $\det(\Gamma(Q, \mathbf{x}, \mathbf{x}))$ has the form $\sum_{i,j=1}^{m} a_{ij} x_i x_j$. Consequently, we conclude by applying part (c), rather than part (b), of Proposition 5.1. We omit the details. $\qquad\square$

We will apply Propositions 5.4 and 5.5 via the following lemma.

Lemma 5.7. *Fix integers $m \geqslant b \geqslant 1$ and an $m \times m$ matrix Q for which both Q and Q^T are b-blocked and b-dense. Then uniformly over $0 < p \leqslant 1/2$, if $\mathbf{x} = (x_1, \ldots, x_m)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ have i.i.d. Bernoulli(p) entries, then*

$$\mathbf{P}\big(\mathrm{rank}(\Gamma(Q, \mathbf{x}, \mathbf{y})) \neq \mathrm{rank}(Q) + 1 + \mathbf{1}_{[Y(Q)>0]}\big) = O((bp)^{-1/2}).$$

Proof. In what follows we write $Q' = \Gamma(Q, \mathbf{x}, \mathbf{y})$. Recall that if $\mathbf{x}$ and $\mathbf{y}$ lie outside the row-span and column-span of Q, respectively, then $\mathrm{rank}(Q') = \mathrm{rank}(Q) + 2$. Note also that $Y(Q) = Y(Q^T)$ always holds.

If $Y(Q) > 0$, then by the definition of $Y(Q)$ we have $\mathrm{rank}(Q) < m - z^{\mathrm{ROW}}(Q)$ and

$$\mathrm{rank}(Q^T) = \mathrm{rank}(Q) < m - z^{\mathrm{COL}}(Q) = m - z^{\mathrm{ROW}}(Q^T).$$

In this case the lemma follows by applying Proposition 5.4 twice, once to Q and $\mathbf{y}$ and once to Q^T and $\mathbf{x}$.

We now treat the case $Y(Q) = 0 = Y(Q^T)$. By replacing Q by Q^T if necessary, we may assume that Q has s non-zero rows and t non-zero columns, for some $0 \leqslant t \leqslant s \leqslant m$; in particular note that $\mathrm{rank}(Q) = t$. By relabelling the rows and columns, we may assume that Q' has the form

$$Q' = \begin{pmatrix} A & \mathbf{0} & (\mathbf{y}')^T \\ \mathbf{0} & \mathbf{0} & (\mathbf{y}^+)^T \\ \mathbf{x}' & \mathbf{x}^- & 0 \end{pmatrix},$$

where A is an $s \times t$ matrix with no zero rows or columns, $\mathbf{0}$ represents a block of zeros, and where $\mathbf{x} = (\mathbf{x}', \mathbf{x}^-)$ and $\mathbf{y} = (\mathbf{y}', \mathbf{y}^+)$.

If $t = s$ then A is b-blocked and b-dense and $\mathrm{rank}(A) = t = \mathrm{rank}(Q)$. Since

$$\mathrm{rank}(Q') \geqslant \mathrm{rank}(\Gamma(A, \mathbf{x}', \mathbf{y}'))$$

and $\mathbf{x}', \mathbf{y}'$ have i.i.d. Bernoulli(p) entries, in this case the lemma follows by applying Proposition 5.5 to A, $\mathbf{x}'$ and $\mathbf{y}'$.

Finally, if $t < s$ then $\operatorname{rank}(Q) = t < s = m - z^{\text{ROW}}(Q)$. Proposition 5.4 applied to Q and $\mathbf{y}$ then yields that $\mathbf{y}$ lies outside the column-span of Q with probability $1 - O((bp)^{-1/2})$. If the latter occurs then $\operatorname{rank}(Q') \geqslant \operatorname{rank}(Q) + 1$; this completes the proof. $\qquad\square$

The analogous result for symmetric matrices is as follows.

Lemma 5.8. *Fix integers $m \geqslant b \geqslant 1$ and a b-blocked, b-dense symmetric $m \times m$ matrix Q. Then uniformly over $0 < p \leqslant 1/2$, if $\mathbf{x} = (x_1, \ldots, x_m)$ has i.i.d. Bernoulli(p) entries, then*

$$\mathbf{P}\big(\operatorname{rank}(\Gamma(Q, \mathbf{x}, \mathbf{x})) < \operatorname{rank}(Q) + 1 + \mathbf{1}_{[Y(Q)>0]}\big) = O((bp)^{-1/4}).$$

The proof is practically identical to that of Lemma 5.7 but slightly easier, as for symmetric matrices we always have $z(Q) = z^{\text{ROW}}(Q) = z^{\text{COL}}(Q)$. The resulting bound is weaker because we must use Proposition 5.6 rather than Proposition 5.5. We omit the details.

To make future formulas more concise, we introduce the following shorthand.

Definition 5.9. *For $n \geqslant 1$, let $k = k(n, p) = \ln \ln n/(2p)$. We say that a square matrix Q is n-robust if both Q and Q^T are k-blocked and k-dense.*

The following proposition, whose proof is the most technical part of the paper, says that robustness is very likely to hold throughout the final $n - n'$ steps of the iterative exposure of minors in $R_{n,p}^{\mathcal{L}}$. In the following proposition, recall from the start of Section 4 the definition of permissible templates, and also the fact that $\gamma \in (0, 1/4)$ is a fixed constant depending only on c.

Proposition 5.10. *Fix $K \in \mathbb{N}$. For any $p \in (c \ln n/n, 1/2)$ and any permissible template $\mathcal{L} \in \mathcal{M}^n(K)$, we have*

$$\mathbf{P}\big(\forall\, m \in [n', n] : R_{n,p}^{\mathcal{L}}[m] \text{ is } n\text{-robust}\big) = 1 - O(n^{-\gamma}), \quad \text{and}$$
$$\mathbf{P}\big(\forall\, m \in [n', n] : Q_{n,p}^{\mathcal{L}}[m] \text{ is } n\text{-robust}\big) = 1 - O(n^{-\gamma}).$$

We provide the proof of Proposition 5.10 in Section 6.1 but now use it to complete the proofs of Lemma 4.3 and Lemma 4.7 (and so of Theorem 2.4). We begin by controlling the probability that $z(R_{n,p}^{\mathcal{L}}[m])$ ever decreases by more than one in a single step of the minor exposure process.

Lemma 5.11. *For fixed $K \in \mathbb{N}$, uniformly over permissible $\mathcal{L} \in \mathcal{M}^n(K)$ and $c \ln n/n \leqslant p \leqslant 1/2$ we have*

$$\mathbf{P}\big(\exists n' \leqslant m < n : z(R_{n,p}^{\mathcal{L}}[m+1]) < z(R_{n,p}^{\mathcal{L}}[m]) - 1\big) = O(n^{-\gamma/2}).$$

Proof. First, by symmetry this probability is at most twice

$$\mathbf{P}\big(\exists n' \leqslant m \leqslant n : |Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[m])| - |Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[m+1])| > 1\big).$$

For $p \geqslant (20 \ln n)/n$, with high probability $|Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[m])| = 0$ for each $n' \leqslant m \leqslant n$, so we assume that $(c \ln n)/n \leqslant p \leqslant (20 \ln n)/n$. The matrix $R_{n,p}[n']$ is distributed as $R_{n',p}$, and we have $p \geqslant \alpha c \ln n'/n' = (1/2 + \gamma) \ln n'/n'$. Since $R_{n,p}^{\mathcal{L}}$ contains at most K^2 rows with deterministic or partially deterministic coordinates, it follows that for n large,

$$\mathbf{P}\big(|Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[n'])| \geqslant n^{1/2-\gamma/2}\big) \leqslant \binom{n'}{n^{1/2-\gamma/2}}\big((1-p)^{n'-K^2}\big)^{n^{1/2-\gamma/2}} \leqslant e^{-n^{\gamma/2}}. \qquad (5.1)$$

We next bound the probability that $z = |Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[n'])| < n^{1/2-\gamma/2}$ and at least two zero rows disappear in a single step. For fixed m with $n' \leqslant m < n$, we have

$$\mathbf{P}\big(|Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[m+1])| < |Z^{\text{ROW}}(R_{n,p}^{\mathcal{L}}[m])| - 1, z < n^{1/2-\gamma/2}\big) \leqslant \binom{\lfloor n^{1/2-\gamma/2} \rfloor}{2}p^2, \qquad (5.2)$$

which is at most $n^{-1-\gamma/2}$ for n large and $p \leqslant 20 \ln n/n$. By (5.1), (5.2), and a union bound, the result follows. $\qquad \square$

We state the symmetric analogue of Lemma 5.11 for later use.

Lemma 5.12. *Under the conditions of Lemma 5.11, if $\mathcal{L}$ is symmetric then*

$$\mathbf{P}\big(\exists n' \leqslant m \leqslant n : z(Q_{n,p}^{\mathcal{L}}[m]) - z(Q_{n,p}^{\mathcal{L}}[m+1]) > 1\big) \leqslant O(n^{-\gamma/2}).$$

Proof. In this case, the desired probability is equal to the probability that

$$|Z(Q_{n,p}^{\mathcal{L}}[m])| - |Z(Q_{n,p}^{\mathcal{L}}[m+1])| > 1$$

for some $n' \leqslant m \leqslant n$. The proof then follows that of Lemma 5.11. $\qquad \square$

Proof of Lemma 4.3. For $n' \leqslant m \leqslant n$ let

$$\mathcal{F}_{n,m} = \sigma(\{R_{n,p}^{\mathcal{L}}[i] : n' \leqslant i \leqslant m\})$$

and let

$$E_{n,m} = \{\forall n' \leqslant i \leqslant m : R_{n,p}^{\mathcal{L}}[i] \text{ is } n\text{-robust}\}.$$

Note that $E_{n,m} \in \mathcal{F}_{n,m}$ for all $n' \leqslant m \leqslant n$. Also, $E_{n,j} \subset E_{n,i}$ for all $n' \leqslant i < j \leqslant n$.

Now, for $n' \leqslant m \leqslant n - 1$ let

$$C_m = \{\text{rank}(R_{n,p}^{\mathcal{L}}[m+1]) = \text{rank}(R_{n,p}^{\mathcal{L}}[m]) + 1 + \mathbf{1}_{[Y(R_{n,p}^{\mathcal{L}}[m])>0]}\}.$$

Then, since $R_{n,p}^{\mathcal{L}}[m]$ is $\mathcal{F}_{n,m}$-measurable and $E_{n,m} \in \mathcal{F}_{n,m}$, we have

$$\begin{aligned}
\mathbf{P}(C_m \mid \mathcal{F}_{n,m}) &\geqslant \mathbf{P}(C_m \mid \mathcal{F}_{n,m})\mathbf{1}_{[E_{n,m}]} \\
&\geqslant \inf\{\mathbf{P}(C_m | R_{n,p}^{\mathcal{L}}[m] = Q) : Q \text{ is } n\text{-robust}\} \cdot \mathbf{1}_{[E_{n,m}]} \\
&\geqslant (1 - O((kp)^{-1/2}))\mathbf{1}_{[E_{n,n}]},
\end{aligned}$$

the last inequality by Lemma 5.7 and since $\mathbf{1}_{[E_{n,m}]} \geqslant \mathbf{1}_{[E_{n,n}]}$. Therefore, there exists $C > 0$ such that, for all $n' \leqslant m < n$, writing $\beta = C(\ln \ln n)^{1/2}$, we have

$$\mathbf{P}\big(\mathrm{rank}(R_{n,p}^{\mathcal{L}}[m+1]) < \mathrm{rank}(R_{n,p}^{\mathcal{L}}[m]) + 1 + \mathbf{1}_{[Y_m > 0]} \mid \mathcal{F}_{n,m}\big)$$
$$\leqslant C(2kp)^{-1/2}\mathbf{1}_{[E_{n,n}]} + \mathbf{1}_{[E_{n,n}^c]}$$
$$\leqslant \beta + \mathbf{1}_{[E_{n,n}^c]}.$$

For $n' \leqslant m < n$ let

$$I_m = \mathbf{1}_{[\mathrm{rank}(R_{n,p}^{\mathcal{L}}[m+1]) < \mathrm{rank}(R_{n,p}^{\mathcal{L}}[m]) + 1 + \mathbf{1}_{[Y_m > 0]}]}.$$

It follows from the preceding bound that, for $m \in [n', n-1]$,

$$\mathbf{P}\big(I_m = 1 \mid I_{n'}, \ldots, I_{m-1}, E_{n,n}\big) \leqslant \beta.$$

We may therefore couple $(I_m, n' \leqslant m < n)$ with a family $(B_m, n' \leqslant m < n)$ of independent Bernoulli(β) random variables such that, for all $n' < m \leqslant n$,

$$I_m \leqslant B_m + (1 - B_m)\mathbf{1}_{[E_{n,n}^c]}.$$

Finally, for $n' \leqslant m < n$ let $X_m = 2B_m - 1$, so that $\mathbf{P}(X_m = 1) = \beta = 1 - \mathbf{P}(X_m = -1)$. By the identity (4.2) for $Y(R_{n,p}[m+1])$, if

$$Y(R_{n,p}[m+1]) - Y(R_{n,p}[m]) \leqslant \max(X_{m+1}, -Y(R_{n,p}[m]))$$

then either $I_m > B_m$ (in which case $E_{n,n}^c$ occurs) or $\{z(R_{n,p}^{\mathcal{L}}[m+1]) \leqslant z(R_{n,p}^{\mathcal{L}}[m]) - 2\}$. It follows that

$$\mathbf{P}\big(\forall n' \leqslant m < n : Y(R_{n,p}^{\mathcal{L}}[m+1]) - Y(R_{n,p}^{\mathcal{L}}[m]) \leqslant \max(X_{m+1}, -Y(R_{n,p}[m]))\big)$$
$$\geqslant 1 - \mathbf{P}(E_{n,n}^c) - \mathbf{P}\big(\exists n' \leqslant m < n : z(R_{n,p}^{\mathcal{L}}[m+1]) \leqslant z(R_{n,p}^{\mathcal{L}}[m]) - 2\big)$$
$$= 1 - O(n^{-\gamma/2}),$$

the final bound by Lemma 5.11 and Proposition 5.10. This completes the proof. $\qquad\square$

The proof of Lemma 4.7 is practically identical, using the second rather than the first bound of Proposition 5.10 and using Lemmas 5.12 and 5.8 rather than Lemmas 5.11 and 5.7, respectively. We omit the details.

6. Structural properties that guarantee rank increase

In this section we prove Proposition 5.10. For the remainder of the paper, fix $K \in \mathbb{N}$, $n \in \mathbb{N}$ large, let $k = k(n, p) = \ln \ln n/(2p)$, and fix a permissible template

$$\mathcal{L} = (\mathcal{L}^+, \mathcal{L}^-) = ((S_i^+)_{i \in I^+}, (S_j^-)_{j \in I^-}) \in \mathcal{M}^n(K).$$

For $i \in [n]$ write $R_i = R_{n,p}^{\mathcal{L}}[i]$, $H_i = H_{n,p}^{\mathcal{L}}[i]$. Also, for the remainder of the paper, $T = [n] \setminus (I^- \cup \bigcup_{i \in I^+} S_i^+)$ and let $U = \bigcup_{i \in I^+} S_i^+$. These definitions are illustrated in Figure 1. Finally, recall that $c \in (1/2, 1)$ and α are fixed so that $\alpha c \in (1/2, 3/4)$, that $\gamma = \alpha c - 1/2$ and that $n' = \lceil \alpha n \rceil$.
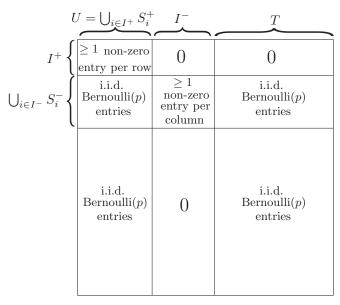
	$U = \bigcup_{i\in I^+} S_i^+$	I^-	T
I^+	≥ 1 non-zero entry per row	0	0
$\bigcup_{i\in I^-} S_i^-$	i.i.d. Bernoulli(p) entries	≥ 1 non-zero entry per column	i.i.d. Bernoulli(p) entries
	i.i.d. Bernoulli(p) entries	0	i.i.d. Bernoulli(p) entries

Figure 1. The deterministic and random structure of the matrix R_n.

Before proceeding to details, we pause to describe the broad strokes of our proof. Our arguments are more straightforwardly described in the language of graphs rather than matrices, so we shall begin to switch to the language of graphs.

We separately bound the probability that for some $n' \leqslant m \leqslant n$, R_m either is not $k(n, p)$-blocked or is not $k(n, p)$-dense. Bounding the latter probability is straightforward: this is essentially the event there are too many vertices with low out-degree in $H_{n'}$, and $p \geqslant c \ln n/n$ is sufficiently large that such vertices are rare.

Bounding the probability that R_m is not $k(n, p)$-blocked for some m is more involved, and we pause to develop some intuition. Recall that for R_m to be $k(n, p)$-blocked, we need that for any $S \subset [m]$ with $2 \leqslant |S| \leqslant k(n, p)$, there are at least two S-selectors in R_m. In the language of graphs, an S-selector is a vertex v such that v has exactly one in-neighbour in S. For $n' \leqslant m \leqslant n$ and $i \in S$, conditional on $\{N_{H_m}^+(j) : j \in S, j \neq i\}$, the larger the degree of i the more likely it is that $N_{H_m}^+(i)$ contains a vertex v lying outside $\bigcup_{j\in S\setminus\{i\}} N_{H_m}^+(j)$, and such a vertex v is an S-selector. For this reason, low-degree vertices pose a potential threat to the existence of S-selectors. (Indeed, low-degree vertices in a sense pose the greatest difficulty for the proof; it is precisely out-degree one vertices that cause Theorems 2.2 and Theorem 2.4 to be fail for $c < 1/2$.) We neutralize this threat by showing that with high probability all (sufficiently) low-degree vertices have pairwise disjoint out-neighbourhoods, and so sets S of *exclusively* low-degree vertices have many S-selectors.

We now turn to details. We begin by bounding the probability that some R_m is not $k(n, p)$ dense.

Lemma 6.1. *Uniformly over $c \ln n/n \leqslant p \leqslant 1/2$ we have*

$$\mathbf{P}\big(\forall\ n' \leqslant m \leqslant n : R_m\ \text{is}\ k(n, p)\text{-dense}\big) = 1 - O(n^{-1}),$$

and if $\mathcal{L}$ is symmetric then

$$\mathbf{P}\big(\forall\ n' \leqslant m \leqslant n : Q_{n,p}^{\mathcal{L}}[m] \text{ is } k(n,p)\text{-dense}\big) = 1 - O(n^{-1}).$$

Proof. For $n' \leqslant m \leqslant n$, let

$$A_m = \{i \in [n'] \setminus I^+ : |([m] \cap N_{R_m}^+(i)) \setminus I^-| > 1\}.$$

Observe that $A_{n'} \subset A_m$ for all $n' \leqslant m \leqslant n$. On the other hand, if R_m is $k(n,p)$-dense, then $|A_m| \geqslant k(n,p) - K(K+1) \geqslant k(n,p)/2$, the latter for n large. It follows that

$$\mathbf{P}\big(\exists\ n' \leqslant m \leqslant n : R_{n,p}^{\mathcal{L}}[m] \text{ is not } k(n,p)\text{-dense}\big) \leqslant \mathbf{P}\big(|A_{n'}| < k(n,p)/2\big).$$

For $i \in [n'] \setminus (I^+ \cup \bigcup_{i \in I^-} S_i^-)$, let B_i be the event that $i \notin A_{n'}$. The event B_i is monotone decreasing, so in bounding its probability from above we may assume that p is equal to $p_{\min} = c \ln n/n$. Write $s = |I^-| \leqslant K$. For $i \in [n'] \setminus (I^+ \cup \bigcup_{i \in I^-} S_i^-)$ we then have

$$\begin{aligned}
\mathbf{P}(B_i) &= \mathbf{P}(\text{Binomial}(n' - s, p_{\min}) \leqslant 1) \\
&= \left(1 + \frac{(n' - s)p_{\min}}{1 - p_{\min}}\right)(1 - p_{\min})^{n'-s} \\
&\leqslant 2(1 + \alpha c \ln n)n^{-\alpha c},
\end{aligned}$$

so

$$\mathbf{E}[|[n'] \setminus A_{n'}|] \leqslant n' \cdot 2(1 + \alpha c \ln n)n^{-\alpha c} \leqslant 2n^{1-\alpha c}(1 + \alpha c \ln n).$$

A similar calculation shows that

$$\mathbf{E}[|[n'] \setminus A_{n'}|^4] \leqslant (2n^{1-\alpha c}(1 + \alpha \ln n))^4 16(1 + \alpha c \ln n),$$

and so by Markov's inequality

$$\mathbf{P}(|A_{n'}| < k(n,p)/2) \leqslant \frac{\mathbf{E}[|[n'] \setminus A_{n'}|^4]}{(n' - k(n,p)/2)^4} \leqslant \frac{(4n^{1-\alpha c})^4(1 + \alpha \ln n)^5}{(n' - k(n,p)/2)^4} < \frac{1}{n^2},$$

the last inequality holding for n large since $1 - \alpha c < 1/2$ and $n' - k(n,p)/2 = \Omega(n)$. The lemma follows by a union bound. An identical proof establishes the stated bound for $Q_{n,p}^{\mathcal{L}}$ in the case that $\mathcal{L}$ is symmetric. $\qquad\square$

Next, we address the probability that the minors R_m, $n' \leqslant m \leqslant n$, are not $k(n,p)$-blocked. The next definition allows us to avoid the (partially) deterministic neighbourhoods of H_m.

Definition 6.2. Fix $b \geqslant 2$. A matrix $Q = (q_{ij})_{1 \leqslant i,j \leqslant m}$ is $(b, \mathcal{L})$-*blocked* if any set $S \subset [m]$ with $2 \leqslant |S| \leqslant b$ satisfies the following conditions.

- If $S \cap (Z^{\text{ROW}}(Q) \cup I^+) = \emptyset$ then there exist distinct $j, l \in [m] \setminus (I^- \bigcup_{i \in I^+} S_i^+)$ that are S-selectors for Q.
- If $S \cap (Z^{\text{COL}}(Q) \cup I^-) = \emptyset$ then there exist distinct $j, l \in [m] \setminus (I^+ \bigcup_{i \in I^-} S_i^-)$ that are S-selectors for Q^T.

(In the last bullet point of the preceding definition, it may be useful to note that if M is the adjacency matrix of a directed graph then M^T is the adjacency matrix of the graph with all edge orientations reversed.) We then have the following lemma.

Proposition 6.3. *Uniformly in* $c \ln n/n \leqslant p \leqslant 1/2$,

$$\mathbf{P}\big(\forall n' \leqslant m \leqslant n : R_m \text{ is } (k, \mathcal{L})\text{-blocked}\big) = 1 - O(n^{-\gamma}),$$

and if $\mathcal{L}$ *is symmetric then also*

$$\mathbf{P}\big(\forall n' \leqslant m \leqslant n : Q_{n,p}^{\mathcal{L}}[m] \text{ is } (k, \mathcal{L})\text{-blocked}\big) = 1 - O(n^{-\gamma}).$$

The proof of Proposition 5.10 assuming Proposition 6.3 is straightforward, and largely consists of showing that *if* R_m *is* $(k, \mathcal{L})$-*blocked then most sets* $S \in [m] \setminus Z^{\text{ROW}}(R_m)$ *deterministically have at least two* S-*selectors (even if* $S \cap I^+ \neq \emptyset$). An easy probability bound then polishes off the proof.

Proof of Proposition 5.10. Let

$$C_1^m = \{E \subset [m] \setminus Z^{\text{ROW}}(R_m) : 2 \leqslant |E| \leqslant k(n,p), E \subset I^+\},$$
$$C_2^m = \{E \subset [m] \setminus Z^{\text{ROW}}(R_m) : 2 \leqslant |E| \leqslant k(n,p), |E \setminus I^+| \geqslant 2\},$$
$$C_3^m = \{E \subset [m] \setminus Z^{\text{ROW}}(R_m) : 2 \leqslant |E| \leqslant k(n,p), |E \setminus I^+| = 1\},$$

and for $k = 1, 2, 3$ let

$$A_k^m = \{\forall E \in C_k^m : \text{there are two } E\text{-selectors in } T \cap [m]\}.$$

Note that if R_m is $(k, \mathcal{L})$-blocked for all $n' \leqslant m \leqslant n$ but is not k-blocked for some $n' \leqslant m \leqslant n$, then one of the events A_k^m, $k \in \{1, 2, 3\}$, $n' \leqslant m \leqslant n$ must fail to occur. We consider the events A_k^m, $k = 1, 2, 3$ in turn.

First, note that since the sets $(S_i^+, i \in I^+)$ are disjoint and non-empty, for every $E \in C_1^m$, for all $i \in E$, every $j \in S_i^+$ is an E-selector. Thus, A_1^m holds deterministically.

Second, if R_m is $(k, \mathcal{L})$-blocked then for any $E \in C_2^m$ there are $(E \setminus I^+)$-selectors $\ell_1, \ell_2 \in T \cap [m]$. Since $\bigcup_{i \in I^+} S_i^+$ is disjoint from T, it follows that ℓ_1, ℓ_2 are not in the out-neighbourhoods of any vertex in I^+. Therefore ℓ_1 and ℓ_2 are also E-selectors, so if R_m is $(k, \mathcal{L})$-blocked then A_2^m holds. It follows by Proposition 6.3 that $\mathbf{P}\big(\bigcap_{m=n'}^n A_2^m\big) = 1 - O(n^{-\gamma})$.

Third, fix $E \in C_3^m$ and write $E \setminus I^+ = \{v\}$. Note that v must have at least one neighbour in H_m as $E \cap Z^{\text{ROW}}(R_m) = \emptyset$. If $|N_{H_m}^+(v) \cap T| \geqslant 2$ then any two $\ell_1, \ell_2 \in N_{H_m}^+(v) \cap T$ are E-selectors since $N_{H_m}^+(I^+) \cap T = \emptyset$. Also, if $|N_{H_m}^+(v) \cap T| \geqslant 1$ and $N_{H_m}^+(v) \cap N_{H_m}^+(I^+) = \emptyset$ then choose $\ell_1 \in S_i^+ \subset N_{H_m}^+(I^+)$ for some $i \in E \cap I^+$, and $\ell_2 \in N_{H_m}^+(v) \cap T$; both ℓ_1 and ℓ_2 are again E-selectors.

It follows that

$$\mathbf{P}\left(\bigcup_{m=n'}^n (A_3^m)^c\right) \leqslant \mathbf{P}\big(\exists v \in [n] \setminus I^+ : |N_{H_n}^+(v) \cap U| \geqslant 1, |N_{H_n}^+(v) \cap T \cap [n']| \leqslant 1\big). \quad (6.1)$$

For fixed $v \in [n] \setminus I^+$, since $|[n'] \cap T| \geqslant n' - K(K+1)$, we have

$$\mathbf{P}\big(|N_{H_{n'}}^+(v) \cap T \cap [n']| \leqslant 1\big) \leqslant \mathbf{P}\big(\mathrm{Binomial}(n' - K(K+1) - 1, p) \leqslant 1\big)$$
$$\leqslant (1 + n'p)(1 - p)^{n' - (K+1)^2}.$$

Furthermore,

$$\mathbf{P}(|N_{H_n}^+(v) \cap U| \geqslant 1) \leqslant K^2 p.$$

The events in the two preceding probabilities are independent since U and T are disjoint. By a union bound over $v \in [n] \setminus I^+$, it follows that the probability in (6.1) is bounded by

$$n \cdot (1 + n'p)(1 - p)^{n' - (K+1)^2} \cdot K^2 p.$$

If $p \geqslant 4 \ln n / (\alpha n)$ then

$$(1 - p)^{n' - (K+1)^2} \leqslant e^{-p(n' - (K+1)^2)} \leqslant n^{-3}$$

for n large, proving the result in this case. If $p \leqslant 4 \ln n / (\alpha n)$ then $np \leqslant (4/\alpha) \ln n$, and since $p \geqslant c \ln n / n$ and $n' \geqslant \alpha n$, this expression is bounded by

$$K^2 (1 + (4/\alpha) \ln n)^2 (1 - p)^{-(K+1)^2} e^{-pn'} = O(\ln^2 n / n^{\alpha c}).$$

Since $\alpha c = \gamma + 1/2$ we conclude that

$$\mathbf{P}\big(\exists v \in [n] \setminus I^+ : |N_{H_n}^+(v) \cap U| \geqslant 1, |N_{H_n}^+(v) \cap T \cap [n']| \leqslant 1\big) = O(n^{-\gamma}). \qquad (6.2)$$

Combining this bound with our bound on $\mathbf{P}\big(\bigcap_{m=n'}^n A_2^m\big)$ and our deterministic observation about the events A_1^m, it follows that

$$\mathbf{P}\big(\forall n' \leqslant m \leqslant n : R_m \text{ is } k\text{-blocked}\big) = 1 - O(n^{-\gamma}).$$

A symmetric argument for R_m^T (using the second bullet from Definition 6.2 instead of the first) shows that

$$\mathbf{P}\big(\forall n' \leqslant m \leqslant n : R_m^T \text{ is } k\text{-blocked}\big) = 1 - O(n^{-\gamma}),$$

which completes the proof of the first assertion of the Proposition 5.10. The second assertion of the proposition follows by a practically identical argument using the second bound of Proposition 6.3 (the only difference is that in this case there is no need to conclude by a 'symmetric argument' as $Q_{n,p}^{\mathcal{L}}[m] = (Q_{n,p}^{\mathcal{L}}[m])^T$). $\qquad \square$

The remainder of the paper is devoted to the proof of Proposition 6.3.

6.1. Proof of Proposition 6.3
First, suppose $\mathcal{L}$ is symmetric and write

$$D = I^- \cup \bigcup_{i \in I^+} S_i^+ = I^+ \cup \bigcup_{i \in I^-} S_i^-.$$

Then, for $n' \leqslant m \leqslant n$, $Q_{n,p}^{\mathcal{L}}[m]$ is $(k, \mathcal{L})$-blocked if and only if $Q_{n,p}[[m] \setminus D]$ is k-blocked. It follows that

$$\mathbf{P}(\forall n' \leqslant m \leqslant n : Q_{n,p}^{\mathcal{L}}[m] \text{ is } (k, \mathcal{L})\text{-blocked})$$
$$= \mathbf{P}(\forall n' \leqslant m \leqslant n : Q_{n,p}[[m] \setminus D] \text{ is } k\text{-blocked})$$
$$= 1 - O(n^{-\gamma}),$$

the last bound by Lemma 2.10 of [10] (note that in that paper, the first property in the definition of 'good' is equivalent to our property 'k-blocked'). This establishes the second assertion of the lemma, so we may now focus exclusively on the first.

We say a set $E \subset [m] \setminus I^+$ is *blocked* if $T \cap [m]$ contains two distinct E-selectors. Given $n' \leqslant m \leqslant n$ and $2 \leqslant s \leqslant k(n, p)$, let

$$D_{m,s} = \{\exists E \subset [m] \setminus (Z^{\text{ROW}}(R_m) \cup I^+) : |E| = s, E \text{ is not blocked}\}.$$

To prove Proposition 6.3 it suffices to show that

$$\mathbf{P}\left(\bigcup_{m=n'}^{n} \bigcup_{s=2}^{k(n,p)} D_{m,s}\right) = O(n^{-\gamma}). \tag{6.3}$$

Since $n' = \Theta(n)$, our arguments are mostly insensitive to the value of $m \in [n', n]$. The value of s plays a more significant role, and we tailor our arguments for different values.

The region where $(p \ln^{1/2} n)^{-1} \leqslant s \leqslant k(n, p)$ is rather straightforward; fix such s and $n' \leqslant m \leqslant n$, and fix $E \subset [m] \setminus (Z^{\text{ROW}}(R_m) \cup I^+)$ with $|E| = s$. Then, for fixed

$$j \in [m] \setminus \left(I^- \cup \bigcup_{i \in I^+} S_i^+\right),$$

we have

$$\mathbf{P}(j \text{ is an } E\text{-selector}) = \mathbf{P}(|N_{R_m}^-(j) \cap E| = 1) = sp(1-p)^{s-1} \geqslant spe^{-sp}.$$

These events are independent for distinct $j \in [m] \setminus (I^- \cup \bigcup_{i \in I^+} S_i^+)$, and it follows that

$$\mathbf{P}(E \text{ is not blocked}) \leqslant \mathbf{P}(\text{Bin}(m - K(K+1), spe^{-sp}) \leqslant 1) \leqslant n(1 - spe^{-sp})^{n/2},$$

the last inequality since $m - K(K+1) \geqslant n/2$ for n large. Since $\binom{n}{s} \leqslant \exp(s \ln(ne/s))$, it follows by a union bound over $E \subset [m] \setminus I^+$ that

$$\mathbf{P}(D_{m,s}) \leqslant \exp(s \ln(ne/s) + \ln n)(1 - spe^{-sp})^{n/2}$$
$$\leqslant \exp(s \ln(ne/s) + \ln n - nspe^{-sp}/2)$$
$$= \exp(\ln n + s(\ln(ne/s) - npe^{-sp}/2))$$
$$\leqslant \exp(\ln n + s(\ln(npe \ln^{1/2} n) - np/\ln^{1/2} n)), \tag{6.4}$$

the last bound following since $(p \ln^{1/2} n)^{-1} \leqslant s \leqslant \ln \ln n/(2p) = k(n, p)$. Using that $x/y \geqslant 2\ln(xy)$ when $x \geqslant y^2/2 \geqslant 2e^6$, since $np \geqslant c \ln n \geqslant (\ln^{1/2} n)^2/2$ it follows that $np/\ln^{1/2} n \geqslant 2\ln(npe \ln^{1/2} n)$ for n sufficiently large, and so (6.4) yields

$$\mathbf{P}(D_{m,s}) \leqslant \exp(\ln n - snp/(2 \ln^{1/2} n)) \leqslant \exp(\ln n - n/(2 \ln n)),$$

where in the final inequality we use that $s \geqslant (p \ln^{1/2} n)^{-1}$. A union bound and the fact that $(n - n' + 1)(\ln \ln n/(2p) - (p \ln^{1/2} n)^{-1}) \leqslant n^2$ then yields

$$\mathbf{P}\left(\bigcup_{n' \leqslant m \leqslant n} \bigcup_{(p \ln^{1/2} n)^{-1} \leqslant s \leqslant \ln \ln n/(2p)} D_{m,s} \right) \leqslant \exp(3 \ln n - n/(2 \ln n)). \quad (6.5)$$

This takes care of the range $(p \ln^{1/2} n)^{-1} \leqslant s \leqslant k(n, p)$, which for small p is the lion's share of the values of s under consideration (though the smaller values of s require slightly more work).

For $2 \leqslant s \leqslant 1/(p \ln^{1/2} n)$, our approach to bounding $\mathbf{P}(D_{m,s})$ is based on the pigeonhole principle and a simple stochastic relation, and we now explain both. For convenience set $\hat{n} = n' - K(K+1)$, and note that $|T \cap [m]| \geqslant \hat{n}$. Note that for a set $E \subset [m] \setminus (Z^{\text{row}}(H_m) \cup I^+)$, if E is not blocked, then at most one vertex in $N_{H_m}^+(E) \cap T$ has less than two in-neighbours in E, so

$$\sum_{i \in E} |N_{H_m}^+(i) \cap T| \geqslant 2|N_{H_m}^+(E) \cap T| - 1.$$

It follows that

$$\mathbf{P}(E \text{ is not blocked}) \leqslant \mathbf{P}\left(|N_{H_m}^+(E) \cap T| \leqslant \frac{\sum_{i \in E} |N_{H_m}^+(i) \cap T| + 1}{2} \right). \quad (6.6)$$

Second, the number of distinct objects obtained by sampling with replacement is always smaller than when taking the same number of samples without replacement. We use this to couple each set $N_{H_m}^+(i) \cap T$ with a set (of smaller or equal size) obtained by sampling with replacement $|N_{H_m}^+(i) \cap T|$ times from T. It follows that, conditional on $\sum_{i \in E} |N_{H_m}^+(i) \cap T|$, the size $|N_{H_m}^+(E) \cap T|$ stochastically dominates $|S|$, where S is a set of $\sum_{i \in E} |N_{H_m}^+(i) \cap T|$ independent, uniformly random elements of T. On the other hand, for such S and for fixed $b < |T \cap [m]|$, if $\sum_{i \in E} |N_{H_m}^+(i) \cap T| = b$ then $|S|$ stochastically dominates a $\text{Binomial}(b, 1 - b/|T \cap [m]|)$ random variable. It thus follows from standard Binomial tail estimates (see Proposition A.1) and the fact that $|T \cap [m]| \geqslant \hat{n}$, that if $\hat{n} \geqslant 4b$ then

$$\mathbf{P}\left(|N_{H_m}^+(E) \cap T| \leqslant (b+1)/2 \; \Big| \; \sum_{i \in E} |N_{H_m}^+(i) \cap T| = b \right)$$

$$\leqslant \mathbf{P}(\text{Binomial}(b, b/\hat{n}) \geqslant (b-1)/2)$$

$$\leqslant \exp\left(-\frac{b-1}{2} \log \frac{\hat{n}}{4eb} \right).$$

We note that this upper bound is decreasing in b for $b < \hat{n}/(4e^2)$, as can be straightforwardly checked. With (6.6), this yields

$$\mathbf{P}\left(\sum_{i \in E} |N_{H_m}^+(i) \cap T| = b, E \text{ is not blocked} \right)$$

$$\leqslant \mathbf{P}\left(\sum_{i \in E} |N_{H_m}^+(i) \cap T| = b \right) \cdot \exp\left(-\frac{b-1}{2} \log \frac{\hat{n}}{4eb} \right), \quad (6.7)$$

from which Proposition 6.3 will follow essentially by union bounds and Binomial tail estimates. Some such estimates are encoded in the following straightforward bound, whose proof we defer to Appendix B.

Lemma 6.4. *Let G be the event that for all $m \in [n', n]$ and all $E \subset [m] \setminus (Z^{\mathrm{ROW}}(H_m) \cup I^+)$ with $|E| \leqslant 1/(p \ln^{1/2} n)$, it is the case that $|E| \leqslant \sum_{i \in E} |N^+_{H_m}(i) \cap T| < \hat{n}/(4e^2)$. Then*

$$\mathbf{P}(G) = 1 - O(n^{-\gamma}).$$

Now fix $E \subset [m] \setminus I^+$, and write $s = |E|$. Then $\sum_{i \in E} |N^+_{H_m}(i) \cap T|$ stochastically dominates a Binomial$(\hat{n}s, p_{\min})$ random variable (writing $p_{\min} = c \log n / n$). It follows from the binomial tail bounds stated in Proposition A.1 that

$$\mathbf{P}\left(\sum_{i \in E} |N^+_{H_m}(i) \cap T| \leqslant 20s \right) \leqslant \left(\frac{e\hat{n}p_{\min}}{20e^{\hat{n}p_{\min}/20}} \right)^{20s} \leqslant \left(\frac{\alpha c \log n}{n^{\alpha c/20}} \right)^{20s}, \tag{6.8}$$

the last inequality holding for n large since $\hat{n} = n' - K(K+1) = \alpha n - K(K+1)$ and so $e^{\hat{n}p_{\min}} \geqslant (e/20)e^{n'p_{\min}} = (e/20)n^{\alpha c}$. Next, observe that

$$\mathbf{P}(E \text{ is not blocked}, G) \leqslant \mathbf{P}\left(s \leqslant \sum_{i \in E} |N^+_{H_m}(i) \cap T| \leqslant 20s, E \text{ is not blocked} \right) \tag{6.9}$$

$$+ \mathbf{P}\left(20s \leqslant \sum_{i \in E} |N^+_{H_m}(i) \cap T| < \hat{n}/(4e^2), E \text{ is not blocked} \right).$$

Taking a union bound over sets $E \subset [m] \setminus I^+$ with $|E| = s$ and using the bounds (6.7) and (6.8) in (6.11), it follows that

$$\mathbf{P}(D_{m,s}, G) \leqslant \underbrace{\binom{m}{s} \exp\left(-\frac{s-1}{2} \log \frac{\hat{n}}{4es} \right) \cdot \left(\frac{e\alpha c \log n}{n^{\alpha c/20}} \right)^{20s}}_{(A)} + \underbrace{\binom{m}{s} \exp\left(-4s \log \frac{\hat{n}}{80es} \right)}_{(B)}.$$

Using that $\hat{n} \geqslant n/2$ for n large and that $\binom{m}{s} \leqslant (en/s)^s$, we have

$$(A) \leqslant \left(\frac{n}{4es} \right)^{1/2} \left(\frac{(2e)^{3/2}(e\alpha c \log n)^{20}}{n^{\alpha c - 1/2}s^{1/2}} \right)^s \leqslant \left(\frac{(2e)^{3/2}(e\alpha c \log n)^{20}}{16n^{7\gamma/8}s^{1/2}} \right)^s.$$

For $s \geqslant 4/\gamma = 4/(\alpha c - 1/2)$, we then have $n^{(\alpha c - 1/2)s - 1/2} = n^{\gamma s - 1/2} \geqslant n^{7\gamma/8}$, so for such s and for n large,

$$(A) \leqslant \left(\frac{(2e)^{3/2}(e\alpha c \log n)^{20}}{n^{7\gamma/8}} \right)^s.$$

Again using that $\hat{n} \geqslant n/2$ for n large and that $\binom{m}{s} \leqslant (en/s)^s$, we have

$$(B) \leqslant \left(\frac{80^4 e^5 s^3}{n^3} \right)^s.$$

The preceding bounds on (A) and on (B) are decreasing in s for $4/\gamma \leqslant s \leqslant (p \ln^{1/2} n)^{-1}$, as can be verified by differentiation; it follows that

$$
\mathbf{P}\left(G \cap \bigcup_{n' \leqslant m \leqslant n} \bigcup_{4/\gamma < s < (p \ln^{1/2} n)^{-1}} D_{m,s} \right)
$$

$$
\leqslant \sum_{4/\gamma < s < (p \ln^{1/2} n)^{-1}} \sum_{n' \leqslant m \leqslant n} [(A) + (B)]
$$

$$
\leqslant (n - n' + 1)(p \ln^{1/2} n)^{-1} \left[\left(\frac{4(e\alpha c \log n)^{20}}{n^{7\gamma/8}} \right)^{4/\gamma} + \left(\frac{80^4 e^5 (4/\gamma)^3}{n^3} \right)^{4/\gamma} \right]
$$

$$
\leqslant n^2 \left[\frac{4^{4/\gamma}(e\alpha c \log n)^{80/\gamma}}{n^{7/2}} + \frac{O(1)}{n^{12/\gamma}} \right]
$$

$$
= O(n^{-1}), \tag{6.10}
$$

since $\gamma = \alpha c - 1/2 \in (0, 1/4)$.

We now treat the range $2 \leqslant s \leqslant 4/\gamma$; for this we require a final lemma. We say H_n is *well separated* if for all $n' \leqslant m \leqslant n$, for any distinct $u, v \in [m] \setminus I^+$, if $|N_{H_m}^+(u)| \leqslant \ln \ln n$ and $|N_{H_m}^+(v)| \leqslant \ln \ln n$, then there is no 2-edge path (with edges of any orientation) joining u and v in H_m.

Lemma 6.5. $\mathbf{P}(H_n \text{ is well separated}) = 1 - O(n^{-\gamma})$.

We defer the proof of Lemma 6.5 to Appendix B, as it is essentially a reprise of an argument found in [8] (though the results of [8] do not themselves directly apply).

Fix $n' \leqslant m \leqslant n$, $2 \leqslant s \leqslant 4/\gamma$ and $E \subset [m] \setminus I^+$ with $|E| = s$. Write

$$
\hat{G} = G \cap \{H_m \text{ is well separated for all } n' \leqslant m \leqslant n\}.
$$

Arguing as at (6.11), we have

$$
\mathbf{P}(E \text{ is not blocked}, \hat{G}) \leqslant \mathbf{P}\left(s \leqslant \sum_{i \in E} |N_{H_m}^+(i) \cap T| \leqslant 20s, E \text{ is not blocked}, \hat{G} \right)
$$

$$
+ \mathbf{P}\left(20s \leqslant \sum_{i \in E} |N_{H_m}^+(i) \cap T| < \hat{n}/(4e^2), E \text{ is not blocked} \right).
$$

Now note that if

$$
\sum_{i \in E} |N_{H_m}^+(i) \cap T| \leqslant 20s \leqslant 80/\gamma,
$$

then all vertices in E have degree at most

$$
80/\gamma + |[m] \setminus T| \leqslant 80/\gamma + K(K+1).
$$

For n sufficiently large that $80/\gamma + K(K+1) \leqslant \ln \ln n$, if H_m is well separated then the sets $\{|N_{H_m}^+(i) \cap T| : i \in E\}$ are disjoint. Since $s \geqslant 2$, it follows that in this case E *is* blocked, so the first probability on the right-hand side of the preceding bound is zero. By a union

bound and the same argument used to bound (B), above, it follows that

$$\mathbf{P}\left(\hat{G} \cap \bigcup_{n' \leqslant m \leqslant n} \bigcup_{2 \leqslant s \leqslant 4/\gamma} D_{m,s}\right) \leqslant (n+1) \cdot \frac{4}{\gamma} \cdot \left(\frac{80^4 e^5 2^3}{n^3}\right)^2 = O(n^{-5}).$$

Combining this bound with (6.5), (6.10) and Lemmas 6.4 and 6.5 then yields

$$\mathbf{P}\left(\bigcup_{n' \leqslant m \leqslant n} \bigcup_{2 \leqslant s \leqslant \ln\ln n/(2p)} D_{m,s}\right)$$
$$\leqslant \exp\left(3\ln n - n/(2\ln n)\right) + O(n^{-1}) + O(n^{-5}) + O(n^{-\gamma}) + O(n^{-\gamma}) = O(n^{-\gamma}),$$

which, recalling (6.3), completes the proof of Proposition 6.3. □

Acknowledgements

Appendix A: Binomial tail bounds

In this section we recall standard binomial tail bounds. The bounds in the following proposition are contained in [18], Lemma 1.1 and [15], Theorem 1.1.

Proposition A.1 *For $m \in \mathbb{N}$ and $0 \leqslant q \leqslant 1$, if $X \stackrel{d}{=} \text{Binomial}(m, q)$ then writing $\mu = mq$, for $k \geqslant \mu$ we have*

$$\mathbf{P}(X \geqslant k) \leqslant \exp(-\mu - k\ln(k/(e\mu))),$$

for $k \leqslant \mu$ we have

$$\mathbf{P}(X \leqslant k) \leqslant \exp(-\mu - k\ln(k/(e\mu))),$$

and for $\varepsilon > 0$ we have

$$\mathbf{P}(X - \mu > \varepsilon m) \leqslant \exp(-2\varepsilon^2 m), \quad \mathbf{P}(X - \mu < -\varepsilon m) \exp(-2\varepsilon^2 m).$$

Appendix B: Remaining proofs

Proof of Lemma 3.3. Fix $\varepsilon > 0$ and let $a > 1$ sufficiently large that $1 - e^{-e^{-a}} < \varepsilon/4$ and that $6(e/4)^{e^a} < \varepsilon/2$. Recall that $p_1 = (\ln n - a)/n$ and $p_2 = (\ln n + a)/n$.

Note that $\tau(\mathcal{H}_n) > p_2$ if either $Z^{\text{ROW}}(H_{n,p_2})$ or $Z^{\text{COL}}(H_{n,p_2})$ is non-empty. As $|Z^{\text{ROW}}(H_{n,p_2})|$ and $|Z^{\text{COL}}(H_{n,p_2})|$ are both asymptotically Poisson(e^{-a}), our choice of a yields that

$$\mathbf{P}\left(\tau(\mathcal{H}_n) > p_2\right) \leqslant 2(1 - e^{-e^{-a}}) + o(1) \leqslant \frac{\varepsilon}{2} + o(1).$$

This gives the first bound of the lemma. For the second bound, we claim that it suffices to prove $\mathbf{P}(\mathcal{D}_K(p_1, p_2)) \geqslant 1 - \varepsilon/2$.

Indeed, assuming this bound, since

$$\mathcal{D}_K(p_1, p_2) \cap \{\tau(\mathcal{H}_n) = p_2\} \subset \mathcal{D}_K(p_1, \tau(\mathcal{H}_n)),$$

we have

$$\mathbf{P}(\mathcal{D}_K(p_1, \tau(\mathcal{H}_n))) \geqslant \mathbf{P}(\mathcal{D}_K(p_1, p_2)) - \mathbf{P}(\tau(\mathcal{H}_n) > p_2) \geqslant 1 - \varepsilon + o(1).$$

Given $i, j \in [n]$, if $U_{ij} > p_1$, then $e = ij \notin H_{n,p_1}$. We have

$$\mathbf{P}(e \in H_{n,p_2} \mid e \notin H_{n,p_1}) = \mathbf{P}(U_{ij} \leqslant p_2 \mid U_{ij} > p_1) \leqslant \frac{p_2 - p_1}{1 - p_1} \leqslant \frac{4a}{n}.$$

We use this estimate to study $(N_{H_{n,p_2}}^+(i))_{i \in Z^{\text{ROW}}(H_{n,p_1})}$. For $i, j \in [n]$, by the preceding bound and a union bound,

$$\mathbf{P}\left(N_{H_{n,p_2}}^+(i) \cap N_{H_{n,p_2}}^+(j) \neq \emptyset \mid i, j \in Z^{\text{ROW}}(H_{n,p_1})\right) \leqslant \frac{n(4a)^2}{n^2} = \frac{(4a)^2}{n}.$$

By another union bound, for any fixed set $Z^+ \in [n]$ it follows that

$$\mathbf{P}\left(\{N_{H_{n,p_2}}^+(i)\}_{i \in Z^+} \text{ are not pairwise disjoint} \mid Z^{\text{ROW}}(H_{n,p_1}) = Z^+\right) \leqslant \frac{8a^2|Z^+|^2}{n}. \tag{B.1}$$

Similarly, for any fixed $Z^+, Z^- \in [n]$,

$$\mathbf{P}\left(\bigcup_{i \in Z^+} N_{H_{n,p_2}}^+(i) \cap Z^- \neq \emptyset \mid Z^{\text{ROW}}(H_{n,p_1}) = Z^+, Z^{\text{COL}}(H_{n,p_1}) = Z^-\right) \leqslant \frac{4a|Z^+||Z^-|}{n}. \tag{B.2}$$

Finally, given that $i \in Z^{\text{ROW}}(H_{n,p_1})$ we have $|N_{H_{n,p_2}}^+(i)| \leqslant_{\text{st}} \text{Bin}(n, 4a/n)$. It follows by a Chernoff bound that

$$\mathbf{P}\left(|N_{H_{n,p_2}}^+(i)| > 8a \mid i \in Z^{\text{ROW}}(H_{n,p_1})\right) \leqslant e^{-16a^2},$$

By a union bound, for any $Z+ \subset [n]$,

$$\mathbf{P}\left(\bigcup_{i \in Z^+} |N_{H_{n,p_2}}^+(v)| > 8a \mid Z^{\text{ROW}}(H_{n,p_1}) = Z^+\right) \leqslant |Z^+| e^{-16a^2}. \tag{B.3}$$

Observe that a similar argument conditioning on $Z^{\text{COL}}(H_{n,p_1}) = Z^-$ gives the same bounds in (B.1) and (B.3) for the sequence $(N_{H_{n,p_2}}^-(i))_{i \in Z^{\text{COL}}(H_{n,p_1})}$. Furthermore,

$$\bigcup_{i \in Z^{\text{ROW}}(H_{n,p_1})} N_{H_{n,p_2}}^+(i) \cap Z^- \neq \emptyset$$

implies

$$\bigcup_{i \in Z^{\text{COL}}(H_{n,p_1})} N_{H_{n,p_2}}^-(i) \cap Z^{\text{ROW}} \neq \emptyset.$$

So far the bounds obtained depend on the size of fixed sets $Z^+, Z^- \in [n]$. Now let $\mathcal{K}$ be the event that both $Z^{\text{ROW}}(H_{n,p_1})$ and $Z^{\text{COL}}(H_{n,p_1})$ have size at most $K = K(a) = \lfloor 2e^a \rfloor$.

We claim that

$$\mathbf{P}(\mathcal{K}) = \mathbf{P}\big(|Z^{\text{ROW}}(H_{n,p_1})|, |Z^{\text{COL}}(H_{n,p_1})| \leqslant K\big) \geqslant 1 - 2(e/4)^{e^a} + o(1). \tag{B.4}$$

For this, we use that if $X \stackrel{d}{=} \text{Poisson}(\lambda)$, then $\mathbf{P}(X \geqslant 2\lambda) \leqslant (e/4)^{\lambda}$: see, *e.g.*, Lemma 1.2 of [18]. Since $|Z^{\text{ROW}}(H_{n,p_1})|, |Z^{\text{COL}}(H_{n,p_1})|$ are asymptotically Poisson(λ), (B.4) follows by a union bound.

We now bound $\mathcal{D}_K(p_1, p_2)$ using the above inequalities. If $\mathcal{K}$ occurs, then there exist (possibly empty) sets $Z^+, Z^- \in [n]$ of size at most K. By (B.1), (B.2), (B.3) we then obtain

$$\mathbf{P}\big(\mathcal{D}_K(p_1, p_2) \cup \{\tau(\mathcal{H}_n) \leqslant p_1\} \mid \mathcal{K}\big) \geqslant 1 - \frac{16a^2K^2 + 4aK^2}{n} - 2Ke^{-16a^2}. \tag{B.5}$$

Note that if $\tau(\mathcal{H}_n) \leqslant p_1$ then $|Z^{\text{ROW}}(H_{n,p_1})| + |Z^{\text{COL}}(H_{n,p_1})| = 0$. Thus,

$$\begin{aligned}
\mathbf{P}(\mathcal{D}_K(p_1, p_2)) &\geqslant \mathbf{P}\big(\mathcal{D}_K(p_1, p_2) \cup \{\tau(\mathcal{H}_n) \leqslant p_1\} \mid \mathcal{K}\big)\mathbf{P}(\mathcal{K}) - \mathbf{P}(\tau(\mathcal{H}_n) \leqslant p_1) \\
&\geqslant (1 - 2Ke^{-16a^2} - o(1))(1 - 2(e/4)^{e^a}) - 2e^{-e^a} + o(1) \\
&\geqslant 1 - 6(e/4)^{e^a}.
\end{aligned}$$

In the second inequality we use (B.4), (B.5) and the fact that $|Z^{\text{ROW}}(H_{n,p})|$ and $|Z^{\text{ROW}}(H_{n,p})|$ are asymptotically Poisson(e^a); the final inequality then follows from the fact that $a > 1$ and $K = \lfloor 2e^a \rfloor$ by straightforward calculation. By our choice of a, the final bound is at least $1 - \varepsilon/2 + o(1)$, completing the proof. $\qquad\square$

Proof of Lemma 6.4. We first bound the maximum degree of vertices in $[n] \setminus I^+$. A union bound together with a Chernoff bound yields

$$\mathbf{P}\big(\exists v \in [n] \setminus I^+ : |N_{H_n}^+(v)| > 2np\big) \leqslant ne^{-(np)^2} \leqslant n^{1 - c^2 \ln n},$$

where the last inequality uses that $p \geqslant p_{\min} = c \ln n / n$. It follows that with high probability, for any $n' \leqslant m \leqslant n$,

$$\sum_{i \in E} |N_{H_m}^+(i) \cap T| \leqslant 2|E|np \leqslant 2n/\ln^{1/2} n \leqslant \hat{n}/4e^2.$$

To obtain the lower bound note that

$$\sum_{i \in E} |N_{H_m}^+(i) \cap T| \geqslant \sum_{i \in E} |N_{H_m}^+(i) \cap T \cap [n']|.$$

Thus, it suffices to show that

$$\mathbf{P}\big(\forall v \in [n] \setminus I^+ : |N_{H_n}^+(v) \cap T \cap [n']| \neq \emptyset\big) \geqslant 1 - O(n^{-\gamma}).$$

For fixed $v \in [n] \setminus I^+$, since $|[n'] \cap T| \geqslant n' - K(K+1)$, we have

$$\mathbf{P}\big(|N_{H_n}^+(v) \cap T \cap [n']| = 0\big) \leqslant (1 - p)^{n' - (K+1)^2} \leqslant Ce^{-\alpha pn} = O(n^{-\gamma - 1/2}).$$

The bound above applies in particular for vertices in $\cup_{i \in I^-} S_i^-$, and there are at most $K^2 = O(1)$ such vertices. On the other hand, if $v \in [n] \setminus (\cup_{i \in I^-} S_i^- \cup I^+)$ and $N_{H_n}^+(v) \cap [n'] \neq \emptyset$, then either $|N_{H_n}^+(v) \cap T \cap [n']| \neq \emptyset$ or $|N_{H_n}^+(v) \cap U \cap [n']| \neq \emptyset$. It thus remains to

bound

$$\mathbf{P}\big(\exists v \in [n] \setminus I^+ : |N_{H_n}^+(v) \cap U| \geqslant 1, |N_{H_n}^+(v) \cap T \cap [n']| = 0\big),$$

which is $O(n^{-\gamma})$ by (6.2). $\qquad\square$

Proof of Lemma 6.5. We say that a vertex v has *low degree* in H_m if $N_{H_m}^+(v) \leqslant d := \ln \ln n$.

First consider the graph $H_{n'}$. The event that fixed vertices v_1 and v_2 are connected by a 2-path is monotone increasing, while the event that both vertices have low out-degree is monotone decreasing. By the FKG inequality, these events are negatively correlated and so the probability that both events hold is bounded from above by the product of their probabilities.

The random variable $|N_{H_{n'}}^+(v_1)|$ is Binomial$(n'-1, p)$ distributed; we will bound

$$\mathbf{P}(|N_{H_{n'}}^+(v_1)| \leqslant d).$$

We have

$$\mathbf{P}(|N_{H_{n'}}^+(v_1)| \leqslant d) = \sum_{i=0}^{d} \binom{n'-1}{i} (p)^i (1-p)^{n'-i-1}$$

$$\leqslant (1-p)^{n'} \sum_{i=0}^{d} 2(2n'p)^i$$

$$\leqslant e^{-n'p}(2n'p)^{d+1},$$

where in the first inequality we use that $1 - p \geqslant 1/2$. Also, the random variables $|N_{H_{n'}}^+(v_1)|$ and $|N_{H_{n'}}^+(v_2)|$ are i.i.d. and hence

$$\mathbf{P}\big(|N_{H_{n'}}^+(v_1)| \leqslant d, |N_{H_{n'}}^+(v_2)| \leqslant d\big) \leqslant e^{-2n'p}(2n'p)^{2(d+1)}.$$

Since $\mathbf{P}(|N_{H_{n'}}^+(v_1)| \leqslant d)$ is decreasing in p, the last bound implies that if $p \geqslant (2/\alpha) \ln n/n$ then

$$\mathbf{P}\big(|N_{H_{n'}}^+(v_1)| \leqslant d, |N_{H_{n'}}^+(v_2)| \leqslant d\big) \leqslant n^{-4}(4 \ln n)^{2(d+1)}.$$

Also, the same bound holds for

$$\mathbf{P}\big(|N_{H_m}^+(v_1)| \leqslant d, |N_{H_m}^+(v_2)| \leqslant d\big)$$

for any $m \in [n', n]$, since $|N_{H_m}^+(v_1)|$ and $|N_{H_m}^+(v_2)|$ are increasing in m. In this case the bound claimed in the statement of the lemma holds by a union bound over $m \in [n', n]$ and over pairs of vertices v_1, v_2 of H_m.

We thus assume for the remainder of the proof that $p \leqslant (4/\alpha) \ln n/n$. We still also have $p \geqslant p_{\min} = c \ln n/n$, so the preceding argument yields that for any $m \in [n', n]$,

$$\mathbf{P}\big(|N_{H_m}^+(v_1)| \leqslant d, |N_{H_m}^+(v_2)| \leqslant d\big) \leqslant \frac{(2 \ln n)^{2d+2}}{n^{2c\alpha}} \leqslant \frac{\ln^{3d} n}{n^{1+2\gamma}}.$$

On the other hand, the probability that v_1 and v_2 are adjacent or have a common neighbour is at most $2p + 4np^2$. By a union bound and the upper bound on p, the

probability that $H_{n'}$ contains two low-degree vertices with a common neighbour is therefore at most

$$\binom{n'}{2}\frac{\ln^{3d} n(2p + 4np^2)}{n^{1+2\gamma}} \leqslant \frac{72\ln^{3d+2} n}{n^{2\gamma}}.$$

We next consider H_m with $m > n'$. Let $z = m$ be the unique vertex of H_m not in H_{m-1}. If H_m is the first graph which is not well separated, then it either (a) z is adjacent to two low-degree vertices v_1, v_2 which are neither connected by a two-edge path nor adjacent in H_{m-1} or (b) z is itself a low-degree vertex at (undirected) distance 1 or 2 of a second vertex v_0 of low out-degree. By a union bound over the pairs of vertices in H_{m-1} we obtain that (a) occurs with probability at most

$$\binom{m-1}{2}\frac{4p^2 \ln^{3d} n}{n^{1+2\gamma}} \leqslant \frac{(64/\alpha^2)\ln^{3d+2} n}{n^{1+2\gamma}}.$$

Similarly, since z and any fixed vertex $v_0 \in [m-1]$ are at distance 1 or 2 with probability less than $2p + 4np^2$, a union bound over the vertices of H_{m-1} implies that (b) occurs with probability at most

$$\frac{(m-1)\ln^{3d} n(2p + 4np^2)}{n^{1+2\gamma}} \leqslant \frac{(72/\alpha^2)\ln^{3d+2} n}{n^{1+2\gamma}}.$$

Combining these bounds and summing over $m \in [n', n]$, we obtain that

$$\mathbf{P}(H_n \text{ is well separated}) = O(\ln^{3d+2} n/n^{2\gamma}),$$

and the latter term is $O(n^{-\gamma})$, as required. □

References

[1] Ajtai, M., Komlós, J. and Szemerédi, E. (1985) First occurrence of Hamilton cycles in random graphs. In *Cycles in Graphs: Burnaby, BC, 1982*, Vol. 115 of *North-Holland Mathematics Studies*, North-Holland, pp. 173–178.

[2] Balogh, J., Bollobás, B., Krivelevich, M., Müller, T. and Walters, M. (2011) Hamilton cycles in random geometric graphs. *Ann. Appl. Probab.* **21** 1053–1072.

[3] Ben-Shimon, S., Ferber, A., Hefetz, D. and Krivelevich, M. (2012) Hitting time results for maker–breaker games. *Random Struct. Alg.* **41** 23–46.

[4] Bollobás, B. and Frieze, A. M. (1985) On matchings and Hamiltonian cycles in random graphs. In *Random Graphs '83*, Vol. 118 of *North-Holland Mathematics Studies*, North-Holland, pp. 23–46.

[5] Bollobás, B. and Thomason, A. (1985) Random graphs of small order. In *Random graphs '83*, Vol. 118 of *North-Holland Mathematics Studies*, North-Holland, pp. 47–97.

[6] Bourgain, J., Vu, V. H. and Wood, P. M. (2010) On the singularity probability of discrete random matrices. *J. Funct. Anal.* **258** 559–603.

[7] Costello, K. P. (2013) Bilinear and quadratic variants on the Littlewood–Offord problem. *Israel J. Math.* **194** 359–394.

[8] Costello, K. P., Tao, T. and Vu, V. H. (2006) Random symmetric matrices are almost surely nonsingular. *Duke Math. J.* **135** 395–413.

[9] Costello, K. P. and Vu, V. H. (2010) On the rank of random sparse matrices. *Combin. Probab. Comput.* **19** 321–342.

[10] Costello, K. P and Vu, V. H. (2008) The rank of random graphs. *Random Struct. Alg.* **33** 269–285.

[11] Erdős, L. (2011) Universality of Wigner random matrices: A survey of recent results. *Uspekhi Mat. Nauk* **66** 67–198.

[12] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization.

[13] Komlós, J. (1967) On the determinant of $(0, 1)$ matrices. *Studia Sci. Math. Hungar* **2** 7–21.

[14] Linh, T. V. and Vu, V. H. (2010) Random matrices I: Combinatorial problems. *Acta Math. Vietnam.* **35** 335–354.

[15] McDiarmid, C. (1998) Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, pp. 1–46.

[16] Mendelson, S. and Pajor, A. (2006) On singular values of matrices with independent rows. *Bernoulli* **12** 761–773.

[17] Norris, J. R. (1998) *Markov Chains*, Vol. 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press. Reprint of 1997 original.

[18] Penrose, M. (2003) *Random Geometric Graphs*, Vol. 5 of *Oxford Studies in Probability*, Oxford University Press.

[19] Penrose, M. D. (1997) The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.* **7** 340–361.

[20] Rudelson, M. and Vershynin, R. (2010) Non-asymptotic theory of random matrices: Extreme singular values. In *Proc. International Congress of Mathematicians*, Vol. III, Hindustan Book Agency, pp. 1576–1602.

[21] Stojaković, M. and Szabó, T. (2005) Positional games on random graphs. *Random Struct. Alg.* **26** 204–223.

[22] Tao, T. and Vu, V. H. (2005) On random ± 1 matrices: Singularity and determinant. In *STOC'05: Proc. 37th Annual ACM Symposium on Theory of Computing*, ACM, pp. 431–440.

[23] Tao, T. and Vu, V. H. (2007) The condition number of a randomly perturbed matrix. In *Proc. 39th Annual ACM Symposium on Theory of Computing: STOC '07* (D. S. Johnson and U. Feige, eds), ACM, pp. 248–255.

[24] Tao, T. and Vu, V. H. (2009) Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math.* (2) **169** 595–632.

[25] Tao, T. and Vu, V. H. (2012) Random matrices: The universality phenomenon for Wigner ensembles. `arXiv:1202.0068`

[26] Vershynin, R. (2014) Invertibility of symmetric random matrices. *Random Struct. Alg.* **44**(2), 135–182.