SINGULARITY OF THE k-CORE OF A RANDOM GRAPH

ASAF FERBER, MATTHEW KWAN, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. Very sparse random graphs are known to typically be singular (i.e., have singular adjacency matrix), due to the presence of "low-degree dependencies" such as isolated vertices and pairs of degree-1 vertices with the same neighbourhood. We prove that these kinds of dependencies are in some sense the only causes of singularity: for constants $k \geq 3$ and $\lambda > 0$, an Erdős–Rényi random graph $G \sim \mathbb{G}(n, \lambda/n)$ with n vertices and edge probability λ/n typically has the property that its k-core (its largest subgraph with minimum degree at least k) is nonsingular. This resolves a conjecture of Vu from the 2014 International Congress of Mathematicians, and adds to a short list of known nonsingularity theorems for "extremely sparse" random matrices with density $O(1/n)$. A key aspect of our proof is a technique to extract high-degree vertices and use them to "boost" the rank, starting from approximate rank bounds obtainable from (non-quantitative) spectral convergence machinery due to Bordenave, Lelarge and Salez.

1. INTRODUCTION

Let M be an $n \times n$ random matrix with i.i.d. Bernoulli(p) entries (meaning that each entry M_{ij} satisfies $\Pr(M_{ij} = 1) = p$ and $\Pr(M_{ij} = 0) = 1 - p$). It is a classical theorem of Komlós [31, 32] that for constant $p \in (0, 1)$ and $n \to \infty$, a random Bernoulli matrix is nonsingular *with high probability* ("whp" for short): that is, $\lim_{n \to \infty} \Pr(M \text{ is singular}) = 0$.

A huge number of strengthenings and variations of Komlós' theorem have been considered over the years. For example, Costello, Tao, and Vu [16] famously proved that a random *symmetric* Bernoulli matrix is also nonsingular whp. Of course, a symmetric binary matrix can be interpreted as the adjacency matrix of a graph, so the Costello–Tao–Vu theorem has an interpretation in terms of random graphs: for constant $p \in (0, 1)$, an Erdős–Rényi random graph $G \sim \mathbb{G}(n, p)$ has nonsingular adjacency matrix whp[1].

If p decays too rapidly with n (in particular, if $p \leq (1 - \varepsilon) \log n / n$ for some constant $\varepsilon > 0$), then for reasons related to the coupon collector problem, a typical $G \sim \mathbb{G}(n, p)$ has isolated vertices, meaning that its adjacency matrix has all-zero rows and is therefore singular. Costello and Vu [18] proved that $\log n / n$ is in fact a *sharp threshold* for singularity, in the sense that if $p \geq (1 + \varepsilon) \log n / n$ (and p is bounded away from 1) then a typical $G \sim \mathbb{G}(n, p)$ has nonsingular adjacency matrix. This fact was later refined and generalised by Basak and Rudelson [5] and Addario-Berry and Eslava [1]. In particular, the latter authors proved a sharp *hitting time* result: if we consider the random graph *process* where we start with the empty graph on n vertices and add random edges one-by-one, then whp at the very same moment where the last isolated vertex disappears our graph becomes nonsingular.

Even below the threshold $\log n / n$, it is natural to ask whether the only obstacles for singularity are "local dependencies" such as isolated vertices. In their aforementioned paper, Costello and Vu [18] actually proved that for $p \geq (1/2 + \varepsilon) \log n / n$, whp the subgraph obtained by deleting

[1]There is a slight difference between a random symmetric Bernoulli matrix and the adjacency matrix of a random graph: namely, the adjacency matrix of any graph has zeroes on the diagonal. However, the same techniques usually apply to both settings, and we will not further concern ourselves with this detail.

isolated vertices is nonsingular. In follow-up work [17], they extended analysis to the regime where $p \geq c \log n/n$ for any constant $c > 0$; this necessitated the consideration of more sophisticated types of local dependencies than isolated vertices. The most obvious example is *cherries*: pairs of degree-1 vertices with the same neighbour. Another obstruction is cycles of length divisible by 4 in which every second vertex has no neighbours outside the cycle.

A natural way to systematically eliminate these kinds of local dependencies is to iteratively delete low-degree vertices. The *k-core* $\mathrm{core}_k(G)$ of a graph G is the subgraph obtained by iteratively deleting vertices with degree less than k (in any order). Equivalently, it is the largest induced subgraph with minimum degree at least k. At the 2014 International Congress of Mathematicians, Vu conjectured [47, Conjecture 5.6] that for constants $k \geq 3$ and $\lambda > 0$, the k-core of a typical $G \sim \mathbb{G}(n, \lambda/n)$ is nonsingular[2]. He also repeated (a substantially weakened version of) this conjecture in his recent survey on combinatorial random matrix theory [46]. In this paper we prove Vu's conjecture.

Theorem 1.1. *Fix constants $k \geq 3$ and $\lambda > 0$. Then the adjacency matrix of the k-core of $G \sim \mathbb{G}(n, \lambda/n)$ is nonsingular with high probability.*

Remark 1.2. Vu actually conjectured that the k-core of the *giant component* of a typical $G \sim \mathbb{G}(n, \lambda/n)$ is nonsingular. This follows from Theorem 1.1, since a graph has a singular adjacency matrix if and only if at least one of its connected components has a singular adjacency matrix.

Remark 1.3. Many random graph properties of interest are *monotone*: if a graph satisfies the property, then adding edges can never destroy the property. In particular, this implies that if $G \sim \mathbb{G}(n, p)$ satisfies the property whp, and $p' \geq p$, then $G' \sim \mathbb{G}(n, p')$ satisfies the property whp as well. Nonsingularity is not a monotone property, so formally Theorem 1.1 leaves open the case when p grows asymptotically faster than $1/n$. However, the regime where $np = O(1)$ is the most challenging, and it is not difficult to adapt our proof methods to handle the case where $np \to \infty$ (we discuss this further in Remark 7.4). We remark that the analogue of Theorem 1.1 in the regime $np \to \infty$ was recently proved[3] in independent work by Glasgow [24] (see also [20]).

Before discussing the ideas in the proof of Theorem 1.1, we provide a bit more context about k-cores of random graphs. The notion of the k-core was introduced by Bollobás [8] in his study of the evolution of sparse random graphs.

It is not an easy task to understand the k-core of a random graph, but by now we have a very complete picture, at least concerning statistical aspects. In an influential paper, Pittel, Spencer and Wormald [43] proved that the property of having a nonempty k-core has a sharp threshold: for $k \geq 3$, there is an explicit constant $\lambda_k > 1$ such that if $np \leq \lambda_k - \varepsilon$ (respectively, if $np \geq \lambda_k + \varepsilon$), then a typical $G \sim \mathbb{G}(n, p)$ has an empty k-core (respectively, has a nonempty k-core). Moreover, they found formulas for the typical numbers of edges and vertices in the k-core above this critical threshold. Many alternative proofs of these results have since been discovered; see for example [10, 15, 19, 23, 26, 30, 42].

Another important milestone in this area is the counterpart of Theorem 1.1 for *Hamiltonicity*: Krivelevich, Lubetzky and Sudakov [33] proved that for $k \geq 3$, the k-core of a random graph has a Hamiltonian cycle whp. There are obvious parallels between Hamiltonicity and nonsingularity: both properties are of a "global" nature but are trivially obstructed by the existence of certain "local" low-degree configurations. It is worth remarking that other than considering the k-core, a different way to eliminate these kinds of local obstructions is to consider *random regular graphs*. That is, for some d, we consider a random graph drawn uniformly from the set of all n-vertex graphs in which

---

[2]The k-core can be empty (indeed, this is typically the case if λ is below a certain threshold, as we will soon discuss). We follow the standard convention that the empty 0×0 matrix, being an identity map, is nonsingular.

[3]The result in [24] is only stated for the 3-core, but the analysis also works for larger k; see [24, Remark 1].

every degree is exactly d. There are close connections between random regular graphs and the k-core of the Erdős–Rényi random graph (which we will discuss later in more detail), but random regular graphs have more symmetry and are therefore in many ways more tractable. It is a classical result of Robinson and Wormald [44] that for $d \geq 3$, random d-regular graphs are Hamiltonian whp, and in a recent breakthrough, Mészáros [41] and Huang [25] independently proved that for $d \geq 3$, random d-regular graphs are nonsingular whp. Before this, Landon, Sosoe and Yau [36] proved the same for $d = d_n = n^{\Omega(1)}$, while several other authors had previously considered the simpler case of random regular *directed* graphs: in particular, Cook [14] established nonsingularity for random d-regular directed graphs with $d \geq C(\log n)^2$ (for sufficiently large C), and Litvak, Lytova, Tikhomirov, Tomczak-Jaegermann and Youssef [37] proved the same only assuming that $d \to \infty$.

To our knowledge, the aforementioned recent results of Huang and Mészáros (and now Theorem 1.1) are the only known results establishing nonsingularity for any "extremely sparse" model of $n \times n$ random matrices, in which the number of nonzero entries is only $O(n)$. Many of the standard arguments (for example, the "row-by-row exposure" ideas of Komlós and Costello–Tao–Vu) do not work in this regime, and Huang and Mészáros developed quite different enumeration-based techniques in their work. Unfortunately, there are serious obstacles towards adapting Huang and Mészáros' ideas to the k-core setting.

Instead, the starting point for our proof of Theorem 1.1 is some spectral convergence machinery of Bordenave, Lelarge and Salez [9]. These authors showed that if a sequence of graphs "locally converges" to an appropriate Galton–Watson tree T, then the empirical spectral measures of the graphs converge weakly to the "local spectral measure" of T. They used this fact to obtain estimates on the rank of convergent sequences of graphs; in particular their main result was an asymptotic estimate for the rank of the adjacency matrix of a typical random graph[4] $G \sim \mathbb{G}(n, \lambda/n)$. It turns out that the k-core of a typical random graph $G \sim \mathbb{G}(n, \lambda/n)$ satisfies an appropriate local convergence property, and the machinery in Bordenave, Lelarge and Salez [9] can be used to show that the adjacency matrix A of $\mathrm{core}_k(G)$ is close to full rank: that is, $\mathrm{corank}\, A = o(n)$. So, our task is to "boost" the corank from $o(n)$ to zero. This is where our new ideas come in: we are able to "extract" high degree vertices from the k-core, in a way that does not bias too significantly the distribution of the rest of the graph, and we then add back these high-degree vertices one-by-one, typically decreasing the corank at each step. We discuss the ideas in our proof in more detail in Section 2.

Remark 1.4. With some very minor modifications, our proof also gives a *hitting time* result: if $k \geq 3$, then in the random graph process where we add random edges one at a time, whp the k-core is nonsingular at the very moment it becomes nonempty. We discuss this further in Remark 6.2.

Remark 1.5. As written, our proof does not give effective probability bounds. However, with certain modifications it is possible to prove for an explicit $C > 0$ that the conclusion of Theorem 1.1 holds with probability at least $1 - (\log \log n)^C / \sqrt{\log n}$ for sufficiently large n (though the cutoff implied by "sufficiently large" is an ineffective absolute constant). We discuss this further in Remark 9.1.

Notation. Our graph-theoretic notation is for the most part standard. In a graph G, we write $\deg_G(v)$ for the degree of a vertex v. Abusing notation slightly, we also write $\deg_S(v) = |N_S(v)|$ for the number of neighbours of a vertex v in a vertex subset S. We write $G[S]$ for the subgraph of G induced by the vertex subset S, and we write $\mathbb{G}(n, p)$ for the Erdős–Rényi random graph on n vertices with edge density p.

---

[4]It is worth remarking that related rank estimates for sparse random matrices have been obtained by many authors, in part due to connections to coding theory and constraint satisfaction problems. See for example [11, 12], and the references within.

Our use of asymptotic notation is standard as well. For functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leq C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f(n) \geq c|g(n)|$ for sufficiently large n, $f = \Theta(g)$ to mean that that $f = O(g)$ and $f = \Omega(g)$, and $f = o(g)$ to mean that $f/g \to 0$ as $n \to \infty$. We say that an event occurs with high probability ("whp") if it holds with probability $1 - o(1)$.

2. Outline of the proof, and comparison to previous work

In this section we outline the ideas in our proof and compare them to previous work.

2.1. **Degree-constrained random graphs.** An important observation about the k-core of a random graph $G \sim \mathbb{G}(n, p)$ is that for any set of vertices U and any degree sequence $\mathbf{d} = (d_v)_{v \in U}$, all the graphs on the vertex set U with degree sequence $(d_v)_{v \in U}$ have the same probability of occurring as $\text{core}_k(G)$. Writing $\mathbb{G}(\mathbf{d})$ for the uniform distribution on the set of all such graphs, it therefore would suffice to consider a set $\mathcal{D}$ of degree sequences such that $\text{core}_k(G)$ has degree sequence in $\mathcal{D}$ whp, and to show that for any $\mathbf{d} \in \mathcal{D}$, the degree-constrained random graph $\mathbb{G}(\mathbf{d})$ is nonsingular whp.

So, a reasonable plan of attack would be to try to generalise the recent work of Huang [25] and Mészáros [41] on random regular graphs, to the more general setting of random graphs with a given degree sequence. Roughly speaking, the approach in these papers is to view a random matrix of interest as a random matrix over a finite field $\mathbb{F}_q$, and to study the expected size of the kernel of this matrix using careful enumerative arguments. In the case of general degree sequences, while it is possible to use similar techniques to write a (very complicated) combinatorial expression for the expected kernel size, it appears intractable to analyse the associated variational problem without the symmetry present in the d-regular case (which is used crucially in the work of Huang [25] and Mészáros [41]). It is worth remarking that for certain reasonable-looking degree sequences $\mathbf{d}$ it is simply not true that a typical $G \sim \mathbb{G}(\mathbf{d})$ is nonsingular whp: a simple example (taken from [7]) is the case when we have n vertices, $9n/10$ of which have degree 3 and $n/10$ of which have degree 100. In this case, typically a majority of degree-3 vertices are only adjacent to degree-100 vertices, which implies that the adjacency matrix of G has corank $\Omega(n)$ whp. We further remark that there is not necessarily a direct two-way correspondence between expected kernel size and the singularity probability: it is unclear whether carefully analysing the expected kernel size is even a viable method to establish nonsingularity for general degree sequences.

However, there are certain things we can say about quite general degree-constrained random graphs: in particular, with very mild assumptions on $\mathbf{d}$, a typical random graph $G \sim \mathbb{G}(\mathbf{d})$ has the property that its local statistics can be very closely approximated by an appropriate Galton–Watson tree T (this is the well-known notion of local graph convergence introduced independently by Benjamini and Schramm [6] and by Aldous and Steele [2]). Using some machinery of Bordenave, Lelarge and Salez [9], one can estimate the rank of the adjacency matrix of such a graph in terms of an optimisation problem involving a probability generating function associated with $\mathbf{d}$. The typical degree sequence of the k-core of a random graph was computed by Cain and Wormald [10] (we state their result in Lemma 6.4), so one can solve an explicit optimisation problem to see that for a typical $G \sim \mathbb{G}(n, \lambda/n)$, with A as the adjacency matrix of $\text{core}_k(G)$, we have corank $A = o(n)$ whp.

2.2. **Boosting the corank.** At first glance, the above considerations may not seem particularly useful, as there is usually a big jump in difficulty between proving that a random matrix is close to full rank, and proving that it is actually of full rank. However, the k-core of a random graph has some properties that are very useful to us. In particular, even though the expected degrees are $(n-1)p = O(1)$, there are typically at least a few vertices in $\text{core}_k(G)$ which have quite high degree: it is easy to show that for any function $f(n) = o(n)$, there is some Δ growing to infinity such that whp at least $f(n)$ vertices in $\text{core}_k(G)$ have degree at least Δ.

The significance of high-degree vertices is that we can use them to "boost" the corank, adapting the ideas of Costello, Tao and Vu [16] first used to study singularity of random symmetric matrices. Indeed, we show (Lemma 5.1) that if we start with a graph satisfying a certain "unstructured kernel property", and we add a new vertex with a large random neighbourhood, then the corank typically decreases by 1 (unless it is already zero, in which case it typically stays at zero). This fact is proved with *Littlewood–Offord*-type inequalities. In particular, we prove and apply a quadratic Littlewood–Offord inequality for random vectors uniform on a "slice" of the Boolean hypercube (Proposition 3.4), which may be of independent interest.

The discussion so far suggests the following strategy. First, identify a small set T of high-degree vertices in $\mathrm{core}_k(G)$, and let $\mathrm{core}_k(G) - T$ be the graph obtained by deleting the vertices in T from $\mathrm{core}_k(G)$. Then, use the machinery of Bordenave, Lelarge and Salez to show that typically $\mathrm{core}_k(G) - T$ has corank at most $|T|/2$. Finally, add back the vertices in T one-by-one; as long as the relevant unstructuredness property continues to hold, the corank should follow a random walk heavily biased towards zero, implying that our final graph $\mathrm{core}_k(G)$ has nonsingular adjacency matrix whp.

On a very high level, this is indeed more or less our strategy to prove Theorem 1.1. For the fact that random walks biased towards zero typically end up at zero, we can adapt a simple argument of Costello, Tao and Vu (Lemma 5.2). However, the other parts of the argument are not so straightforward, and a number of additional ideas are required.

2.3. **Unstructuredness of the almost-kernel vectors.**
First, we need to know that k-cores of random graphs (and graphs obtained from k-cores by deleting a few vertices) typically satisfy an appropriate "unstructured kernel property". Specifically, for an adjacency matrix A, the property we need is that vectors $\mathbf{v}$ which are orthogonal to almost all rows of A ("almost-kernel vectors") are far from being constant vectors, in the sense that for all $x \in \mathbb{R}$ we have $|\{i : v_i \neq x\}| = \Omega(n)$. (Actually, due to certain issues that arise for small k, the precise property we need is a bit more technical than this; see Definition 4.1).

It is not a trivial matter to prove that properties of this type typically hold for degree-constrained random graphs (in the setting of random regular graphs, such a property follows from results of Backhausz and Szegedy [3], and one can also deduce such a property from the estimates in Huang's paper [25]). Fortunately in our case we do not actually need to consider degree-constrained random graphs at all, and it suffices to consider properties of $G \sim \mathbb{G}(n, p)$ directly. Indeed, as a consequence of some fairly simple expansion estimates, we are able to show (Lemma 4.2) that whp G has the property that *any* induced subgraph $G[U]$ vaguely resembling a k-core also satisfies an appropriate unstructuredness property. For example, this holds as long as $G[U]$ has minimum degree at least 3 and $\Omega(n)$ vertices of odd degree. Due to technical issues that arise when $k = 3$, we actually also need to permit a very small number of degree-2 vertices.

2.4. **Efficiently extracting high-degree vertices.**
The remaining issue with the strategy outlined earlier is that we need to reveal the identities of high-degree vertices in $\mathrm{core}_k(G)$ without revealing too much information about the rest of the graph. For example, suppose we were to naïvely reveal the degrees of all vertices in $\mathrm{core}_k(G)$, define T to contain the set of all vertices in $\mathrm{core}_k(G)$ whose degree is at least some appropriately chosen value Δ, and then reveal the subgraph $\mathrm{core}_k(G) - T$ (in preparation for adding back the vertices of T, and studying the effect on the corank). Unfortunately, conditioned on information revealed so far, the neighbourhoods of vertices in T are very far from being uniformly random: for example, we now know which vertices have different degrees in $\mathrm{core}_k(G) - T$ and in $\mathrm{core}_k(G)$, and these vertices must be the ones with neighbours in T.

We must therefore be extremely careful with the way we expose information. We fix a set S of αn vertices, for some small $\alpha > 0$, and let T be the set of vertices in S which have degree at

least Δ (with respect to $\mathrm{core}_k(G)$). Now, reveal the graph $\mathrm{core}_k(G) - T$ (which is uniform over graphs with its degree sequence, so we can estimate its corank using the machinery of Bordenave, Lelarge and Salez; we do this in Lemma 7.3). Due to the deletion of the vertices in T, this graph $\mathrm{core}_k(G) - T$ will typically have a few vertices with degree less than k (and we therefore know these vertices must have neighbours in T). However, other than this issue, by symmetry considerations the neighbourhoods of vertices in T still have quite a lot of usable randomness: for example, if we fix any two sets of vertices in $\mathrm{core}_k(G) - S$, neither of which contain a vertex which has degree less than k in $\mathrm{core}_k(G) - T$, then these two sets are equally likely to occur as the neighbourhood of a given vertex in T.

There are still certain complications remaining: for example, even though all the vertices in T were chosen to have high degree in $\mathrm{core}_k(G)$, we cannot assume they have high degree in $\mathrm{core}_k(G) - S$ (if we had chosen the vertices in T in that way, it would not have been true that $\mathrm{core}_k(G) - T$ locally converges to a Galton–Watson tree). Also, $\mathrm{core}_k(G) - T$ will typically have a few vertices of degree less than 2, which means its adjacency matrix typically does not actually satisfy our unstructured kernel property. So, before starting our corank-boosting random walk, we need to make certain small adjustments to the set T (being very careful to do this in a such a way that we retain useful randomness for the neighbourhoods of its vertices). The details of these adjustments are described in Lemma 8.1.

3. Anti-concentration on the Boolean slice

The *Erdős–Littlewood–Offord theorem* (see [45, Corollary 7.8]) is an important tool in random matrix theory. Given a sequence of nonzero real numbers $v_1, \ldots, v_n$, and independent zero-one random variables $x_1, \ldots, x_n \in \{0, 1\}$, it gives an *anti-concentration* bound for the random sum $X = x_1 v_1 + \cdots + x_n v_n$. We will need a variant for the case where $(x_1, \ldots, x_n) \in \{0, 1\}^n$ is constrained to have a certain sum (such vectors are said to lie on a *slice* of the Boolean hypercube).

Lemma 3.1. *Let $d \geq 1$, $\eta > 0$ and let $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ be a vector with no entry repeated more than $(1 - \eta)n$ times. Let $\mathbf{x} = (x_1, \ldots, x_n) \in \{0, 1\}^n$ be a random vector, uniformly selected from the zero-one vectors with exactly d ones, and consider any $x \in \mathbb{R}$. Then*

$$\Pr(\mathbf{v}^T \mathbf{x} = x) = O((\eta d)^{-1/2}).$$

Lemma 3.1 is a direct consequence of [22, Lemma 4.2]. A similar inequality for the case where $d = n/2$ was proved in [37].

We will also need an anti-concentration inequality for *quadratic* polynomials of random vectors $(x_1, \ldots, x_n) \in \{0, 1\}^n$ on the Boolean slice. Such an inequality does not seem to be available in the literature (though a special case appears in [35]) so we deduce it from a quadratic inequality for the case where each x_i is independent. Such an inequality was first proved by Costello, Tao and Vu [16], and the Costello–Tao–Vu inequality would already suffice for our proof of Theorem 1.1, but the following stronger inequality is now also available.

Theorem 3.2. *Let $d \geq 1$ and let $\mathbf{x} = (x_1, \ldots, x_n) \in \{0, 1\}^n$ be a uniformly random zero-one vector of length n. Let $M \in \mathbb{R}^{n \times n}$ be an $n \times n$ symmetric matrix with $\Omega(n^2)$ nonzero entries, let $\mathbf{v} \in \mathbb{R}^n$ be any vector, and let $x \in \mathbb{R}$ be any real number. Then*

$$\Pr(\mathbf{x}^T M \mathbf{x} + \mathbf{v}^T \mathbf{x} = x) \leq (\log n)^{O(1)} / \sqrt{n}.$$

Theorem 3.2 is a slight improvement over a result of Meka, Nguyen and Vu [40], and may be deduced from a result of Kane [29] on average sensitivity of polynomial threshold functions (see the journal version of [40] for this deduction).

To prove a Boolean slice version of Theorem 3.2 we will also need the following concentration inequality, which is a generalisation of [35, Lemma 2.1].

Lemma 3.3. *Let $m \in \mathbb{N}$, let S be any finite set of size $|S| \geq m$, let $\mathcal{F}$ be the set of functions $\{1,\ldots,m\} \to S$ and let $\mathcal{I} \subseteq \mathcal{F}$ be the set of injections $\{1,\ldots,m\} \to S$. Consider a function $f : \mathcal{F} \to \mathbb{R}$ satisfying the property that if π, π' agree except that $\pi(i) \neq \pi'(i)$, then $|f(\pi) - f(\pi')| \leq c_i$. Let $\pi \in \mathcal{I}$ be a uniformly random injection. Then for $t \geq 0$,*

$$\Pr[|f(\pi) - \mathbb{E}f(\pi)| \geq t] \leq 2\exp\left(-\frac{t^2}{8\sum_{i=1}^m c_i^2}\right).$$

Proof. We may assume without loss of generality that $c_1 \geq \cdots \geq c_m$. Consider the Doob martingale defined by $Z_i = \mathbb{E}[f(\pi)|\pi(1),\ldots,\pi(i)]$, so in particular $Z_0 = \mathbb{E}f(\pi)$ and $Z_m = f(\pi)$. Let $\mathcal{L}(x_1,\ldots,x_i)$ be the conditional distribution of π given $\pi(1) = x_1,\ldots,\pi(i) = x_i$. We want to show that

$$|\mathbb{E}[f(\mathcal{L}(x_1,\ldots,x_{i-1},x))] - \mathbb{E}[f(\mathcal{L}(x_1,\ldots,x_{i-1},y))]| \leq 2c_i$$

for any choice of distinct $x_1,\ldots,x_{i-1},x,y \in S$; this will imply that $|Z_i - Z_{i-1}|$ is uniformly bounded by $2c_i$, so the desired result will follow from the Azuma–Hoeffding inequality (see for example [28, Theorem 2.25]).

If π is distributed as $\mathcal{L}(x_1,\ldots,x_{i-1},x)$, we can change $\pi(i)$ to y, and if some $\pi(j)$ was already equal to y, change $\pi(j)$ to x; we thereby obtain the distribution $\mathcal{L}(x_1,\ldots,x_{i-1},y)$. This provides a coupling between $\mathcal{L}(x_1,\ldots,x_{i-1},x)$ and $\mathcal{L}(x_1,\ldots,x_{i-1},y)$ that differs in only two positions i and $j > i$, and since $c_j \leq c_i$ this implies the required bound. $\square$

Now, our Boolean slice version of Theorem 3.2 is as follows.

Proposition 3.4. *Let $M = (m_{ij})_{i,j}$ be an $n \times n$ matrix for which there are $\Omega(n^4)$ different 4-tuples (i,i',j,j') with $m_{ij} - m_{i'j} - m_{ij'} + m_{i'j'} \neq 0$. Let $\mathbf{x} = (x_1,\ldots,x_n)$ be a random zero-one vector, uniformly selected from the zero-one vectors with exactly $d \leq n/2$ ones. Then for any vector $\mathbf{v} \in \mathbb{R}^n$ and any $x \in \mathbb{R}$ we have*

$$\Pr(\mathbf{x}^T M \mathbf{x} + \mathbf{v}^T \mathbf{x} = x) \leq (\log d)^{O(1)}/\sqrt{d}.$$

Proof. We need a way to generate a random set of d ones in a way that involves independent random choices. Let π be a uniformly random injection $\{1,\ldots,2d\} \to \{1,\ldots,n\}$ and let $\xi = (\xi_1,\ldots,\xi_d)$ be a sequence of independent Rademacher random variables (satisfying $\Pr(\xi_i = 1) = \Pr(\xi_i = -1) = 1/2$, also independent from π). Then, we choose the positions of the d ones in $\mathbf{x}$ as follows. For each $i \in \{1,\ldots,d\}$, if $\xi_i = 1$ we set $x_{\pi(i)} = 1$, and if $\xi_i = -1$ we set $x_{\pi(i+d)} = 1$. That is to say, first we choose d random pairs of entries, and then for each we flip a coin to decide which of the two elements in that pair to set equal to 1.

It is routine to check (see [35, Lemma 2.8] or the proof of [34, Lemma 2.3]) that $\mathbf{x}^T M \mathbf{x} + \mathbf{v}^T \mathbf{x}$ can be expressed by a quadratic polynomial of the form

$$f_\pi(\xi) = \sum_{1 \leq i < j \leq d} a_{ij}\xi_i\xi_j + \sum_{1 \leq i \leq d} a_i\xi_i + a_0,$$

where the coefficients a_{ij}, a_i, a_0 only depend on π (not on ξ), and in particular

$$a_{ij} = \frac{1}{4}(m_{\pi(i),\pi(j)} - m_{\pi(i),\pi(j+m)} - m_{\pi(i+m),\pi(j)} + m_{\pi(i+m),\pi(j+m)}).$$

Let X be the number of a_{ij} that are nonzero. For each $i < j$ we have $\Pr(a_{ij}) = \Omega(1)$, so $\mathbb{E}X = \Omega(d^2)$. For any i, modifying $\pi(i)$ affects X by only $O(d)$, so by Lemma 3.3 we have

$$\Pr(X \leq \mathbb{E}X/2) \leq \exp\left(-\Omega\left(\frac{(\mathbb{E}X/2)^2}{d^2 \cdot d}\right)\right) = e^{-\Omega(d)}.$$

For any outcome of π satisfying $X \geq \mathbb{E}X/2 = \Omega(d^2)$, by Theorem 3.2 we have $\Pr(f_\pi(\xi) = x) \leq (\log d)^{O(1)}/\sqrt{d}$. $\square$

4. Large level sets of kernel vectors

In this section we prove that for a typical outcome of $G \sim \mathbb{G}(n, p)$, vectors which are orthogonal to almost all the rows/columns of the adjacency matrix of the k-core cannot be "too structured". This will be accomplished using very minimal information about the k-core and some basic expansion properties in a typical $G \sim \mathbb{G}(n, p)$. Related expansion-type ideas have appeared previously in the random matrix literature (see for example [13, 14, 37, 38]), but the details here are much simpler. First, we define our structuredness property.

Definition 4.1. Define the support $\mathrm{supp}(\mathbf{w})$ of a vector $\mathbf{w}$ to be the set of indices i such that $w_i \neq 0$. Say that an $n \times n$ symmetric matrix A satisfies the *unstructured kernel property* $\mathrm{UKP}(\ell, \zeta, \eta)$ if there is a set of ζn indices $Q \subseteq \{1, \ldots, n\}$ such that every nonzero vector $\mathbf{v} \in \mathbb{Q}^n$ with $|\mathrm{supp}(A\mathbf{v})| \leq \ell$ and $\mathrm{supp}(A\mathbf{v}) \cap Q = \emptyset$ satisfies $\max_{x \in \mathbb{R}} |\{i : v_i = x\}| \leq (1 - \eta)n$.

For example, an $n \times n$ matrix satisfies $\mathrm{UKP}(0, 0, \eta)$ if and only if each of its nonzero kernel vectors have level sets of size at most $(1 - \eta)n$.

Now, the main result we prove in this section is as follows.

Lemma 4.2. *For constants $\theta, \lambda > 0$, there is $\eta > 0$ such that the following holds. Let $G \sim \mathbb{G}(n, \lambda/n)$, and say an induced subgraph $G[U]$ is* good *if*

- *it has at least θn odd-degree vertices,*
- *it has minimum degree at least 2,*
- *there are at most $(\eta/4)n$ vertices within distance 7 of a degree-2 vertex,*
- *there is no pair of degree-2 vertices within distance 4 of each other.*

Then whp G is such that the adjacency matrix of every good induced subgraph satisfies the unstructured kernel property $\mathrm{UKP}(2, \eta/3, \eta)$.

Remark 4.3. In order to prove Theorem 1.1 for $k \geq 5$, it would suffice to make the simpler definition that an induced subgraph is good if it has θn odd-degree vertices and minimum degree at least 4. With this weaker condition, it would be easy to show that whp every good induced subgraph satisfies $\mathrm{UKP}(2, 0, \eta)$ (meaning that we do not need to consider a set of exceptional vertices Q in the definition of the unstructured kernel property). Unfortunately, in the cases $k = 3$ and $k = 4$, with non-negligible probability there are a small number of "bad configurations" that force us to consider the more technical property $\mathrm{UKP}(2, \zeta, \eta)$.

Remark 4.4. Being a bit more careful with the estimates in this section, one can show that the property in Lemma 4.2 holds with probability $1 - n^{-\Omega(1)}$.

To prove Lemma 4.2 we will need the following basic expansion properties of $\mathbb{G}(n, p)$. For vertex sets S, T in a graph, let $e(S, T)$ be the number of ordered pairs $(x, y) \in S \times T$ such that xy is an edge (so, edges with endpoints in $S \cap T$ are counted twice).

Lemma 4.5. *For constants $\theta, \lambda > 0$, there is $\eta > 0$ such that the following holds. Let $G \sim \mathbb{G}(n, \lambda/n)$. Then whp there is no set S of fewer than ηn vertices such that at least $\theta n - 2$ vertices of G have a neighbour in S.*

Lemma 4.6. *For constant $\lambda > 0$ there is $\eta > 0$ such that the following holds. Let $G \sim \mathbb{G}(n, \lambda/n)$. Then whp:*

- *(1) there is no subgraph with fewer than 12 vertices and edge-density greater than 1,*
- *(2) there are at most $\log n$ vertices adjacent to a vertex in a 4-cycle, and*
- *(3) there is no pair of disjoint vertex sets S, W satisfying the inequalities*

$$5 \leq |S| < \eta n, \quad e(S, S \cup W) \geq \lceil 5|S|/2 \rceil, \quad |W| \leq \lceil 5|S|/2 \rceil /2 - e(S) + 1.$$

The proofs of the estimates in Lemmas 4.5 and 4.6 are quite standard, but involve some rather tedious calculations, so we first present the deduction of Lemma 4.2.

Proof of Lemma 4.2. Fix an outcome of G satisfying the conclusions of Lemmas 4.5 and 4.6 for the given constants $\theta, \lambda > 0$, choosing appropriate $\eta > 0$. Consider any good $G[U]$, and write $A[U]$ for its adjacency matrix. Let $Q \subseteq U$ be the set of vertices in $G[U]$ adjacent to a vertex in a 4-cycle or within distance 7 of a degree-2 vertex (so $|Q| \leq (\eta/3)n$).

We need to show that for each vector $\mathbf{v} \in \mathbb{Q}^U$ satisfying $|\mathrm{supp}(A[U]\mathbf{v})| \leq 2$ and $\mathrm{supp}(A[U]\mathbf{v}) \cap Q = \emptyset$, the level sets of $\mathbf{v}$ have size at most $(1-\eta)n$. Multiplying $\mathbf{v}$ by a suitable integer and dividing by the greatest common divisor of its entries, it actually suffices to show the same property for vectors $\mathbf{v} \in (\mathbb{F}_2)^U$, working mod 2.

Let $\mathbf{v} \in (\mathbb{F}_2)^U \setminus \{\mathbf{0}\}$ be such that $|\mathrm{supp}(A[U]\mathbf{v})| \leq 2$ and $\mathrm{supp}(A[U]\mathbf{v}) \cap Q = \emptyset$. For $x \in \mathbb{F}_2$ let $S_x = \{i \in U : v_i = x\}$. We will show that $|S_0|, |S_1| \geq \eta n$.

The 0-set. First, we consider S_0. Every vertex $i \in U \setminus \mathrm{supp}(A[U]\mathbf{v})$ has $(A[U]\mathbf{v})_i = 0 \pmod 2$, meaning that i has an even number of neighbours in $S_1 = U \setminus S_0$. So, all but at most two of the odd-degree vertices in $G[U]$ must have a neighbour in S_0. Since $G[U]$ is good, it has at least θn odd-degree vertices so by the property in Lemma 4.5 we have $|S_0| \geq \eta n$ as desired.

The 1-set. For the rest of the proof we consider S_1 (which is nonempty, since $\mathbf{v}$ is nonzero). Let F be an auxiliary graph on the vertex set S_1 with an edge uv when there is a path of length at most 2 between u, v in $G[U]$. Now, consider a connected component $F[S]$ of F. It suffices to prove that $|S| \geq \eta n$.

Let $s = |S|$, and let T be the set of vertices in U which have a neighbour in S. As before, every $i \in U \setminus \mathrm{supp}(A[U]\mathbf{v})$ has an even number of neighbours in S_1, so in fact all vertices of $T \setminus \mathrm{supp}(A[U]\mathbf{v})$ have at least two neighbours in S_1 (therefore at least two neighbours in S, noting that all its neighbors in S_1 are in S).

First, it is straightforward to rule out the case $s = 1$: in this case, all the neighbours of the single vertex in S would be in $\mathrm{supp}(A[U]\mathbf{v})$ (since they would have odd degree into S and hence S_1). Since $|\mathrm{supp}(A[U]\mathbf{v})| \leq 2$, this would imply the existence of a degree-2 vertex adjacent to a vertex in $\mathrm{supp}(A[U]\mathbf{v})$, which would contradict the choice of $\mathbf{v}$ with respect to Q.

So, we assume $s > 1$. Let $S^* \subseteq S$ be the set of vertices in S which have degree 2 with respect to $G[U]$. In $G[U]$, there are no degree-2 vertices within distance 4 of each other, so in $F[S]$, the neighbourhoods of vertices in S^* are disjoint (that is, S^* is a 2-independent set in $F[S]$). Since $F[S]$ is connected and has more than one vertex, each neighbourhood has at least one vertex, so $|S^*| \leq s/2$ and thus $e(S, T) \geq \lceil 5s/2 \rceil$. (This edge count, and every further edge count, is with respect to G.)

It follows that $e(S, T \setminus S) \geq \lceil 5s/2 \rceil - 2e(S)$. If $|T \setminus S| \leq \lceil (\lceil 5s/2 \rceil - 2e(S) + 1)/2 \rceil$ then let $W = T \setminus S$, and otherwise let W be a subset of $T \setminus S$ obtained by repeatedly deleting vertices (starting with vertices in $\mathrm{supp}(A[U]\mathbf{v})$, if any are present in $T \setminus S$) until exactly $\lceil (\lceil 5s/2 \rceil - 2e(S) + 1)/2 \rceil$ vertices remain. In the latter case, all but at most one of the vertices in W have at least two neighbours in S, implying that $e(S, W) \geq \lceil 5s/2 \rceil - 2e(S)$. Either way, property (3) in Lemma 4.6 is violated, unless $2 \leq s \leq 4$ or $s \geq \eta n$. So, it remains to rule out the cases $2 \leq s \leq 4$, with some more careful variants of the above argument.

- First suppose $|T \cap \mathrm{supp}(A[U]\mathbf{v})| = 0$. In this case, every vertex in T has at least two neighbours in S.
  - For $s = 2$, the sets S and T are disjoint, so $e(S, T) \geq \lceil 5s/2 \rceil = 5$. This is only possible if $|T| \geq 3$, but this would imply the existence of a complete bipartite subgraph $K_{2,3}$, violating property (1) in Lemma 4.6.
  - For $3 \leq s \leq 4$: starting from $T \setminus S$, delete vertices (if necessary) to obtain a set W containing at most $\lceil (\lceil 5s/2 \rceil - 2e(S))/2 \rceil \leq (\lceil 5s/2 \rceil + 1)/2 - e(S)$ vertices such that

$e(S, W) \geq \lceil 5s/2 \rceil - 2e(S)$. Then $S \cup W$ has at least $\lceil 5s/2 \rceil - e(S)$ edges and at most $s + (\lceil 5s/2 \rceil + 1)/2 - e(S)$ vertices, violating property (1) in Lemma 4.6.

- Otherwise, suppose $2 \leq s \leq 4$ and $1 \leq |T \cap \mathrm{supp}(A[U]\mathbf{v})| \leq 2$. Recall that $F[S]$ is connected, so all the vertices in S (being within distance 7 of a vertex in Q) have degree at least 3, so $e(S, T) \geq 3s$.
  - For $s = 2$, there are one or two vertices in T with exactly one neighbour in S (these vertices cannot be in Q), and to have $e(S, T) \geq 6$ there must then be at least two vertices with exactly two neighbours in S. But this is impossible, because it would imply that there is a vertex not in Q adjacent to a vertex in a 4-cycle.
  - For $3 \leq s \leq 4$: starting from $T \backslash S$, delete vertices (if necessary; starting from vertices in $\mathrm{supp}(A[U]\mathbf{v})$) to obtain a set W containing at most $\lceil (3s - 2e(S) + 1)/2 \rceil \leq (3s + 2)/2 - e(S)$ vertices such that $e(S, W) \geq 3s - 2e(S)$. Then $S \cup W$ has at least $3s - e(S)$ edges and at most $s + (3s + 2)/2 - e(S)$ vertices, violating property (1) in Lemma 4.6.

This completes the proof. $\qquad\square$

Now, we prove Lemmas 4.5 and 4.6.

Proof of Lemma 4.5. Let $s = \lfloor \eta n \rfloor$ and $f = \lfloor (\theta n - 2)/2 \rfloor$. It suffices to show that whp there is no "bad pair" (S, F), where S is a set of s vertices and F is a set of f edges each with at least one endpoint in S. Indeed, then for every set of size s there are fewer than $2f \leq \theta n - 2$ vertices incident to S.

The probability a bad pair exists is at most

$$\binom{n}{s}\left(\binom{\binom{s}{2} + s(n - s)}{f}\right)p^f \leq \left(\frac{en}{s}\right)^s \left(\frac{epns}{f}\right)^f = o(1)$$

for sufficiently small η. $\qquad\square$

Proof of Lemma 4.6. First, (1) and (2) trivially hold whp by Markov's inequality, observing that the expected number of subgraphs on 11 vertices with density greater than 1 is $O(n^{10}p^{11}) = o(1)$, and the expected number of vertices adjacent to a 4-cycle is $O(p^5 n^5) = O(1)$.

It remains to consider (3). This comes down to a union bound calculation, but we need to be rather careful with the estimates. It is convenient to handle the small s and large s cases separately; let $N_1 = \lfloor (\log n)^{1/10} \rfloor$ and $N_2 = \lfloor \eta n \rfloor$. First, we show that whp there is no set S of $N_1 \leq s \leq N_2$ vertices and set F of $\lceil 1.1s \rceil$ edges between vertices of S. Indeed, the probability that such a situation occurs is at most

$$\sum_{s=N_1}^{N_2} \binom{n}{s}\binom{\binom{s}{2}}{\lceil 1.1s \rceil}p^{\lceil 1.1s \rceil} \leq \sum_{s=N_1}^{N_2} \left(\frac{en}{s}\right)^s \left(\frac{esp}{\lceil 1.1s \rceil}\right)^{\lceil 1.1s \rceil} \leq \sum_{s=N_1}^{N_2} \left(O(np(sp)^{0.1})\right)^s = o(1).$$

To complete the proof that (3) holds whp, it suffices to show that whp there is no "bad quadruple" (S, R, F, D), where:

- S, R are disjoint vertex sets with sizes $3 \leq s \leq N_2$ and $r_{s,f} := \lfloor \lceil 5s/2 \rceil/2 - f + 1 \rfloor$, respectively;
- F is a set of f edges between vertices in S;
- D is a set of $\lceil 5s/2 \rceil - 2f$ edges between S and R;
- $f < \lceil 1.1s \rceil$ or $s < N_1$.

The probability a bad quadruple exists is at most

$$\sum_{s,f} \binom{n}{s}\binom{n}{r_{s,f}}\binom{\binom{s}{2}}{f}\binom{sr_{s,f}}{\lceil 5s/2 \rceil - 2f}p^{\lceil 5s/2 \rceil - f}, \tag{4.1}$$

10

where the sum is over all $3 \leq s \leq N_2$ and $0 \leq f \leq \binom{s}{2}$ satisfying $f < 1.1s$ or $s < N_1$. We now show that this sum is $o(1)$.

First, we handle the range where $3 \leq s < N_1$. The contribution to (4.1) from these terms is at most

$$\sum_{s=5}^{N_1} \sum_{f=0}^{\binom{s}{2}} n^s n^{r_{s,f}} 2^{s^2} 2^{sr_{s,f}} p^{\lceil 5s/2 \rceil - f} \leq \sum_{s=5}^{N_1} \sum_{f=0}^{\binom{s}{2}} n^{o(1)} n^{-\lceil 5s/2 \rceil / 2 + s + 1} \leq n^{-1/2 + o(1)} = o(1).$$

Second, we handle the range where $N_1 \leq s \leq N_2$. The contribution to (4.1) from these terms is at most

$$\sum_{s=N_1}^{N_2} \sum_{f=0}^{1.1s} \left(\frac{en}{s}\right)^s \left(\frac{en}{r_{s,f}}\right)^{r_{s,f}} \left(\frac{es^2}{f}\right)^f \left(\frac{sr_{s,f}}{\lceil 5s/2 \rceil - 2f}\right)^{\lceil 5s/2 \rceil - 2f} p^{\lceil 5s/2 \rceil - f}$$

$$\leq \sum_{s=N_1}^{N_2} \sum_{f=0}^{1.1s} e^{O(s)} \left(\frac{n}{s}\right)^{s + \lceil 5s/2 \rceil / 2 - f + 1} \left(\frac{s^2}{f}\right)^f s^{\lceil 5s/2 \rceil - 2f} p^{\lceil 5s/2 \rceil - f}$$

$$\leq \sum_{s=N_1}^{N_2} \sum_{f=0}^{1.1s} e^{O(s)} \left(n^{9/4} s^{1/4} p^{5/2} (n/s)^{2/s}\right)^s \left(\frac{s}{fnp}\right)^f. \tag{4.2}$$

Since $f \leq 1.1s$ we have $(s/(fnp))^f = e^{O(s)}$, and since $N_1 \leq s \leq N_2$ we have

$$n^{9/4} s^{1/4} p^{5/2} (n/s)^{2/s} \leq \lambda^{5/2} \eta^{1/4} e^{o(1)}.$$

So, if $\eta > 0$ is sufficiently small then the sum in (4.2) is at most $\sum_{s=N_1}^{N_2} (1.1s + 1) e^{-\Omega(s)} = o(1)$. $\square$

5. Boosting the corank

In this section we prove several lemmas closely related to the strategy of Costello, Tao and Vu in [16]. The first lemma shows that (under certain conditions) if one extends a symmetric matrix by adding a new random row and column (this corresponds to adding a new vertex to a random graph), then the corank typically decreases. Recall the definition of the unstructured kernel property $\mathrm{UKP}(\ell, \zeta, \eta)$ from Definition 4.1 (roughly speaking, this is the property that the vectors orthogonal to all but ℓ rows must have large level sets, allowing the possibility of a small number of exceptional rows).

Lemma 5.1. *Fix a constant $\eta > 0$, let $g \geq 1$ and let $d \geq 1$. Consider an $n \times n$ symmetric matrix A with the unstructured kernel property $\mathrm{UKP}(2, \eta/3, \eta)$ and rank at least $n - g$. Consider a subset $E \subseteq \{1, \ldots, n\}$ of size at least $n(1 - \eta/3)$. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a random zero-one vector, such that the restriction $\mathbf{x}_E$ to the entries indexed by E is a uniformly random zero-one vector with exactly d ones (and the restriction $\mathbf{x}_{\overline{E}}$ to entries not indexed by E is deterministic). Add $\mathbf{x}$ as a new row and column of A (and put a zero in the new diagonal entry) to obtain a new matrix A'. Then $\mathrm{rank}\, A' \geq n - g + 2$ with probability at least $1 - (\log d)^{O(1)} / \sqrt{d}$.*

Proof. Let $r = \mathrm{rank}\, A$. First suppose $r \leq n - 1$. Without loss of generality, we assume that the first r columns $\mathbf{v}_1, \ldots, \mathbf{v}_r$ of A are linearly independent. Thus, the last column $\mathbf{v}_n$ of A can be written as a linear combination of the first r rows in a unique way: $\mathbf{v}_n = c_1 \mathbf{v}_1 + \cdots + c_r \mathbf{v}_r$. Therefore, the vector $\mathbf{c} = (c_1, \ldots, c_r, 0, \ldots, 0, -1)$ is orthogonal to every row of A, hence all its level sets have size at most $(1 - \eta)n$ (for convenience, we say such a vector is *η-unstructured*). Note that any restriction of $\mathbf{c}$ to a set E of $n(1 - \eta/3)$ indices is still $(2\eta/3)$-unstructured, so by Lemma 3.1, $\mathbf{c}^T \mathbf{x} = \mathbf{c}_E^T \mathbf{x}_E + \mathbf{c}_{\overline{E}}^T \mathbf{x}_{\overline{E}} \neq 0$ with probability $1 - O(1/\sqrt{d})$. But if $\mathbf{c}^T \mathbf{x} \neq 0$ this means that adding $\mathbf{x}$ as a new row (thereby appending x_i as a new entry to each $\mathbf{v}_i$) increases the rank of A. That is to

11

say, $\mathbf{x}$ does not lie in the span of the rows of A, and by symmetry it also does not lie in the span of the columns of A. So, adding $\mathbf{x}$ as a new row and column increases the rank twice, meaning that $\operatorname{rank} A' \geq r + 2 \geq n - g + 2$.

On the other hand, suppose $r = n$. Observe that

$$\det A' = \mathbf{x}^T M \mathbf{x} = \mathbf{x}_E^T M_E \mathbf{x}_E + \mathbf{v}^T \mathbf{x}_E + x,$$

for some (non-random) $\mathbf{v} \in \mathbb{R}^E$ and $x \in \mathbb{R}$, where M is the adjugate matrix of A and M_E is its restriction to the rows and columns in E.

Since AM is a nonzero multiple of the identity matrix, the ith column of M is orthogonal to all rows of A except the ith, and the difference between the ith and jth columns of M is therefore orthogonal to all rows of A except the ith and the jth. Recalling the definition of $\mathrm{UKP}(2, \eta/3, \eta)$, it follows that each column of M (or row of M, by symmetry) not indexed by Q is η-unstructured, and every difference of two rows or two columns, neither of which are indexed by Q, is either zero or η-unstructured.

Fix any η-unstructured column $\mathbf{w} = (w_1, \ldots, w_n)$ of M (any column not indexed by Q will do). There are at least $(\eta/3)(1 - \eta)n^2$ pairs of indices $i, i' \in E \backslash Q$ such that $w_i \neq w_{i'}$. For each such pair, the difference $\mathbf{v}_i - \mathbf{v}_{i'}$ between the ith and i'th row vectors is nonzero, so is η-unstructured. It follows that there are at least $(2\eta/3)(1 - \eta)n^2$ pairs of indices $j, j' \in E$ such that the jth and j'th entries of $\mathbf{v}_i - \mathbf{v}_{i'}$ differ. For each of the resulting $\Omega(n^4)$ choices of i, i', j, j', we have $m_{ij} - m_{i'j} - m_{ij'} + m_{i'j'} \neq 0$. By Proposition 3.4 it follows that with probability at least $1 - (\log d)^{O(1)}/\sqrt{d}$ we have $\det A' \neq 0$ (meaning that $\operatorname{rank} A' = n + 1 \geq n + 2 - g$). □

The second lemma in this section shows that if a random process takes nonnegative values and at each time-step tends to move towards zero, then the process is likely to end at zero. If we construct a random graph by iteratively adding many vertices with random neighbourhoods, we can combine this lemma with Lemma 5.1 to show that the resulting random graph is likely to have an adjacency matrix with corank zero (i.e., be nonsingular).

Lemma 5.2. *Let $X_0, \ldots, X_n$ be a sequence of nonnegative integer random variables satisfying the following conditions.*

(1) $X_0 \leq n/2$.
(2) for all $1 \leq t \leq n$, we have $X_t \leq X_{t-1} + 1$.
(3) The sequence "drifts towards zero" in the sense that for any $x_0, \ldots, x_{t-1}$:
  (a) If $x_{t-1} \neq 0$ we have $\Pr(X_t < x_{t-1} \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}) \geq 1 - p$
  (b) If $x_{t-1} = 0$ we have $\Pr(X_t = 0 \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}) \geq 1 - p$.
Then $\Pr(X_n = 0) \geq 1 - 1000p$.

Proof. We may assume $p \leq 1/1000$. We may also assume $n \geq 10$, because otherwise $\Pr(X_n = 0) \geq (1 - p)^n \geq 1 - np \geq 1 - 9p$.

Define $Y_t = (1/\sqrt{p})^{X_t} - 1$. We claim that

$$\mathbb{E}[Y_{t+1} | Y_t, \ldots, Y_0] \leq (2\sqrt{p})Y_t + 2\sqrt{p}. \tag{5.1}$$

To see this, we distinguish cases: if $Y_t = 0$ then $\mathbb{E}[Y_{t+1} | Y_t, \ldots, Y_0] \leq (1/\sqrt{p} - 1)p \leq \sqrt{p}$, which implies (5.1) with room to spare. If $Y_t > 0$ then we have

$$\mathbb{E}[Y_{t+1} + 1 | Y_t, \ldots, Y_0] \leq \sqrt{p}(Y_t + 1) + p \cdot (1/\sqrt{p})(Y_t + 1),$$

which also implies (5.1). We deduce from (5.1) that

$$\mathbb{E}[Y_n] \leq \mathbb{E}[Y_0](2\sqrt{p})^n + (2\sqrt{p})^{n-1} + \cdots + (2\sqrt{p})^2 + 2\sqrt{p} \leq (1/\sqrt{p})^{n/2}(2\sqrt{p})^n + 4\sqrt{p} \leq 8\sqrt{p}$$

where the final inequality follows since $p \leq 1/1000$ and $n \geq 10$. We then apply Markov's inequality:

$$\Pr(X_n \neq 0) = \Pr(Y_n \neq 0) = \Pr(Y_n \geq 1/\sqrt{p} - 1) \leq \mathbb{E}[Y_n]/(1/\sqrt{p} - 1) \leq 9p. \qquad \square$$

6. Preliminaries on degree-constrained random graphs

For a set V and positive integers m, k, let $\mathcal{K}(V, m, k)$ be the uniform distribution on graphs with vertex set V, exactly m edges, and minimum degree at least k. The reason we are interested in this random graph model is the following symmetry property of the k-core. This property is well-known and appears implicitly for instance in [10].

Lemma 6.1. *For any p, let $G \sim \mathbb{G}(n, p)$, let V be the vertex set of $\mathrm{core}_k(G)$ and let m be the number of edges in $\mathrm{core}_k(G)$. If we condition on any outcome of V, m, then the conditional distribution of $\mathrm{core}_k(G)$ is $\mathcal{K}(V, m, k)$.*

Proof. Consider any two graphs H_1, H_2 on the vertex set V with m edges and minimum degree at least k. For any outcome of G yielding $\mathrm{core}_k(G) = G[V] = H_1$, we can simply replace $G[V]$ with H_2 to obtain an outcome of G yielding $\mathrm{core}_k(G) = H_2$ (iteratively deleting vertices with degree less than k yields $G[V]$ in both cases). This means that H_1 and H_2 are equally likely to occur as $\mathrm{core}_k(G)$. $\square$

Remark 6.2. The conclusion of Lemma 6.1 holds not just for $G \sim \mathbb{G}(n, p)$ but for any "sufficiently symmetric" random graph. For example, it holds when G is a random graph with a specified number of edges, and it holds when G is the graph obtained by running the random graph process until the moment the k-core becomes nonempty. Actually, it is straightforward to adapt the entire proof of Theorem 1.1 to either of these settings. Indeed, the only additional inputs we need are some very weak estimates on the number of vertices and edges in $\mathrm{core}_k(G)$, and analogues of the expansion estimates in Lemmas 4.5 and 4.6.

For convenience, we write $\mathcal{K}(n, m, k)$ instead of $\mathcal{K}(\{1, \ldots, n\}, m, k)$ The degree sequence of a typical $G \sim \mathcal{K}(n, m, k)$ was studied by Cain and Wormald [10], as follows[5].

Definition 6.3. Consider integers k, m, n satisfying $k \geq 3$ and $m \geq kn/2$. Let $Z_k(\lambda) = \sum_{t=k}^{\infty} \lambda^t / t!$, and let λ be the unique solution to $\lambda Z'_k(\lambda) / Z_k(\lambda) = 2m/n$. (This unique solution λ is easily seen to exist; see e.g. [10, Eq. (4)]). Define $\rho_t = (\lambda^t / t!) / Z_k(\lambda)$ for $t \geq k$; these values can be interpreted as probability masses associated with a truncated Poisson distribution.

Lemma 6.4 ([10, Theorem 2 and Lemma 1]). *Fix constants $k \geq 3$ and $\varepsilon > 0$. Suppose $\varepsilon n \leq m - kn/2 \leq n/\varepsilon$ and let $G \sim \mathcal{K}(n, m, k)$. Then, using the notation of Definition 6.3, whp the following properties hold.*

(1) for all $t \geq 3$, there are $\rho_t n + O(n^{3/4})$ vertices of degree t.
(2) all vertices have degree at most $\log n$.

Recall from the outline in Section 2 that we will need to extract some high-degree vertices from the k-core, which we will later use to "boost" the corank. So, we will also need to understand the degree sequence of the subgraph of a typical outcome of $G \sim \mathcal{K}(n, m, k)$ obtained by removing a subset of high-degree vertices. This degree sequence is a little more complicated to describe.

Definition 6.5. Consider integers Δ, k, m, n satisfying $\Delta, k \geq 3$ and $m \geq kn/2$. Let $\lambda, Z_k(\lambda), \rho_t$ be as in Definition 6.3, and further define:

$$\beta = \alpha \sum_{t=\Delta}^{\infty} \rho_t, \qquad \gamma = \frac{\alpha n}{2m} \sum_{t=\Delta}^{\infty} t \rho_t = \frac{\alpha}{Z'_k(\lambda)} \sum_{t=\Delta-1}^{\infty} \frac{\lambda^t}{t!},$$

$$\zeta_{j,t} = \binom{j}{t} \gamma^{j-t} (1-\gamma)^t, \qquad \delta_t = (1-\alpha) \sum_{j=k}^{\infty} \rho_j \zeta_{j,t}, \qquad \delta'_t = \alpha \sum_{j=k}^{\Delta-1} \rho_j \zeta_{j,t} \qquad \text{for } j, t \geq 0.$$

---

[5]Strictly speaking part (2) of Lemma 6.4 does not follow from the statements of [10, Theorem 2 and Lemma 1], but it is a much simpler fact than (1) and can be easily derived using Cain and Wormald's methods.

Lemma 6.6. *Fix $\Delta \geq k \geq 3$ and $\varepsilon, \alpha > 0$. Suppose $\varepsilon n \leq m - kn/2 \leq n/\varepsilon$ and let $G \sim \mathcal{K}(n, m, k)$. Let $V = \{1, \ldots, n\}$, $S = \{1, \ldots, \lfloor \alpha n \rfloor\}$ and $T = \{v \in S : \deg_G(v) \geq \Delta\}$. Then whp the following hold.*

    (1) $|T| = \beta n + O(n^{3/4})$.
    (2) For all $t \geq 0$, there are $\delta_t n + O(n^{4/5})$ vertices $v \in V \setminus S$ with $\deg_{V \setminus T} = t$.
    (3) For all $t \geq 0$, there are $\delta_t' n + O(n^{4/5})$ vertices $v \in T$ with $\deg_{V \setminus T} = t$.

For the proof of Lemma 6.6 (and for other computations in this and later sections), we work with a random graph model called the *configuration model*.

Definition 6.7. For a degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$, consider a set of $r = d_1 + \cdots + d_n$ points, grouped into n labelled "buckets" of sizes $d_1, \ldots, d_n$. A *configuration* is a perfect matching on the r points, consisting of $r/2$ disjoint edges. Now, given a configuration, contracting each of the buckets to a single point gives rise to a multigraph with degree sequence $d_1, \ldots, d_n$ (where we use the convention that loops contribute 2 to the degree of a vertex).

If we consider the multigraph G^* arising from a uniformly random configuration, and condition on the event that this multigraph is a simple graph, then we obtain the uniform distribution on graphs with degree sequence $\mathbf{d}$. Moreover, if $d_1^2 + \cdots + d_n^2 = O(n)$ (as is typically the case for the degree sequence of any subgraph of $G \sim \mathcal{K}(n, m, k)$, as follows from Lemma 6.4), it is well known that this simplicity probability is $\Omega(1)$. So, any property that holds whp for G^* also holds whp for a uniformly random graph with degree sequence $\mathbf{d}$.

Proof of Lemma 6.6. Let $\mathbf{d} = (d_1, \ldots, d_n)$ be the degree sequence of $G \sim \mathcal{K}(n, m, k)$. By symmetry, all reorderings of $\mathbf{d}$ are equally likely to occur, so (1) holds whp by Lemma 6.4 and a Chernoff bound for hypergeometric distributions (see for example [28, Theorem 2.10]). Similarly, $\mathbf{d}$ has the following properties whp.

    • For all $t \geq k$, there are $(1 - \alpha)\rho_t n + O(n^{3/4})$ vertices $v \in V \setminus S$ with $d_v = t$.
    • For all $k \leq t < \Delta$, there are $\alpha \rho_t n + O(n^{3/4})$ vertices $v \in S \setminus T$ with $d_v = t$.
    • For all $t \geq \Delta$, there are $\alpha \rho_t n + O(n^{3/4})$ vertices $v \in T$ with $d_v = t$.

Condition on an outcome of $\mathbf{d}$ satisfying the above properties and such that each $d_i \leq \log n$ (this degree sequence determines T). Now, G is now a uniformly random graph with degree sequence $\mathbf{d}$, and $d_1 + \cdots + d_n = 2m$. Consider a set of $2m$ points grouped into buckets of sizes $d_1, \ldots, d_n$. Let σ be a uniformly random bijection from $\{1, \ldots, 2m\}$ into these $2m$ points, so the perfect matching with edges $\sigma(1)\sigma(2)$, $\sigma(3)\sigma(4)$, $\ldots$, $\sigma(2m-1)\sigma(2m)$ is a uniformly random configuration. Let G^* be the multigraph obtained by contracting this configuration. It suffices to show that $G^*[V \setminus T]$ satisfies properties (2) and (3) whp.

For each $t \in \mathbb{N}$, let X_t (respectively X_t') be the number of vertices $v \in V \setminus S$ (respectively, $v \in S \setminus T$) with $\deg_{V \setminus T}(v) = t$ in G^*. Modifying σ by a transposition affects X_t (respectively, X_t') by at most 4, so by a concentration inequality for random permutations (see for example [39, Eq. (29)]), whp $X_t = \mathbb{E}X_t + O(n^{3/4})$ and $X_t' = \mathbb{E}X_t' + O(n^{3/4})$ for each t. Therefore, it suffices to show that $\mathbb{E}X_t = \delta_t n + O(n^{3/4}(\log n)^3)$ and $\mathbb{E}X_t' = \delta_t' n + O(n^{3/4}(\log n)^3)$ for each t.

Consider the set P_T of all the points in buckets corresponding to vertices in T. The number of such points is
$$\sum_{t \geq \Delta} \frac{t}{Z_k(\lambda)} \frac{\lambda^t}{t!} \alpha n + O(n^{3/4}(\log n)^2).$$

That is to say, the fraction of the $2m$ points occupied by P_T is $\gamma + O(n^{-1/4}(\log n)^2)$. Now, for a vertex $v \in V \setminus T$, there are $d_v \leq \log n$ points in the bucket corresponding to v (let b_v be the set of points in this bucket), and the number of points in this bucket that are matched with points in

14

P_T is very nearly binomially distributed with parameters d_v and $1 - \gamma$. Indeed, the only deviation from this distribution comes from the fact that if one knows whether some subset of points in b_v are matched with points in P_T, this very slightly biases (by a factor of $1 + O(1/n)$) the probability that a further point in b_v is matched with a point in P_T. The probability that v has exactly $d_v - t$ neighbours in T is therefore $\zeta_{d_v,t} + O((\log n)^2/n^{-1/4} + \log n/n)$. Linearity of expectation then yields the desired estimates. $\qquad\square$

Next, we observe that in the setting of Lemma 6.6, the subgraph $G[V \setminus T]$ is uniformly random given its degree sequence.

Lemma 6.8. *For any integers m, k, Δ and sets V, S, let $G \sim \mathcal{K}(V, m, k)$ and let T be the set of vertices in S which have at least Δ neighbours in V. If we condition on any outcome of $V \setminus T$ and any outcome of the degree sequence $\mathbf{d} = (\deg_{V \setminus T}(v))_{v \in V \setminus T}$, then the conditional distribution of $G[V \setminus T]$ is uniform over graphs on the vertex set $V \setminus T$ with degree sequence $\mathbf{d}$.*

Proof. Consider any two graphs H_1, H_2 on the same vertex set with the same degree sequence. For any outcome of G yielding $G[V \setminus T] = H_1$, we can simply replace $G[V \setminus T]$ with H_2 to obtain an outcome of G yielding $G[V \setminus T] = H_2$. So, H_1 and H_2 are equally likely to occur as $G[V \setminus T]$. $\qquad\square$

The final fact we record in this section is that sparse random graphs with given degree sequences are typically well-approximated by Galton–Watson trees.

Definition 6.9. For a probability distribution μ with nonnegative integer support, let $\tilde{\mu}$ be the probability distribution defined by

$$\tilde{\mu}(t) = (t + 1)\mu(t + 1)\Big/ \sum_{\ell=0}^{\infty} \ell\mu(\ell)$$

for positive integer t. A Galton–Watson tree T_μ with *degree distribution* μ is a random rooted tree obtained by a Galton–Watson branching process, where the root has offspring size distribution μ, and all other generations have offspring size distributions $\tilde{\mu}$. Let $T_\mu^{(r)}$ consist of the first r generations of T_μ (with all vertices unlabelled except the root).

Also, for a graph G, let $G^{(r)}$ be the random unlabelled rooted graph obtained by fixing a uniformly random vertex v as the root, and including the subgraph of all vertices within distance $r - 1$ of v. Let $\mathrm{d}_\mu^{(r)}(G)$ be the total variation distance $\mathrm{d}_{\mathrm{TV}}(G^{(r)}, T_\mu^{(r)})$ between $G^{(r)}$ and $T_\mu^{(r)}$ (if G is a random graph, then $\mathrm{d}_\mu^{(r)}(G)$ is a random variable).

Lemma 6.10. *Let μ be a probability distribution with nonnegative integer support and fix any positive integer r. Let $\mathbf{d} = (d_1, \ldots, d_n)$ be a degree sequence such that each $d_v \leq \log n$, and such that $|\{v : d_v = t\}| = \mu(t)n + O(n^{4/5})$ for each $t \geq 0$. Let G be a uniformly random graph with degree sequence $\mathbf{d}$. Then, whp $\mathrm{d}_\mu^{(r)}(G) = o(1)$.*

Proof. Consider a uniformly random configuration with degree sequence $\mathbf{d}$ (obtained via a random permutation σ, as in Lemma 6.6). Note that for $t \geq 1$, the number of points in buckets of size t is $\tilde{\mu}(t - 1)n + O(n^{4/5}\log n)$.

Let G^* be the multigraph obtained by contracting our random configuration. For a vertex v, let $G^*(v)$ be the rooted graph consisting of all vertices within distance $r - 1$ of v. For each possible outcome H of $G^*(v)$ which is a tree (there are at most $(\log n)^r$ such outcomes), let X_H be the number of vertices v for which $G^*(v) = T$. Also, let X_0 be the number of vertices v for which $G^*(v)$ is not a tree. Modifying σ by a transposition changes X_0 and each X_H by at most $4(\log n)^k = n^{o(1)}$, so by a concentration inequality (see for example [39, Eq. (29)]), whp $X_0 = \mathbb{E}X_H + O(n^{5/6})$ and $X_H = \mathbb{E}X_H + O(n^{5/6})$ for each H. Now, to finish the proof we claim that $\mathbb{E}X_0 = O(n^{4/5}\log n +$

$(\log n)^r/n)$ and $\mathbb{E}X_H = \Pr(T_\mu^{(r)} = H) + O(n^{4/5}\log n + (\log n)^r/n)$ for all H. This follows from a routine computation in the configuration model (choose a random vertex, and iteratively explore the neighbourhood of that vertex to a depth of r, at each step exposing a new matched pair in our random configuration). $\qquad\square$

7. Initial rank estimate

In this section we apply the machinery of Bordenave, Lelarge and Salez [9].

Definition 7.1. Say that a sequence of random graphs $G_1, G_2, \ldots$ is μ-*convergent* if, in the notation of Definition 6.9, the sequence of random variables $d_\mu^{(r)}(G_1), d_\mu^{(r)}(G_2), \ldots$ converges to zero in probability, for every fixed r. For a function $f : \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ with $\sum_{i=0}^\infty f(i) < \infty$, let $\varphi_f = [0,1] \to \mathbb{R}$ be the function $\sum_{i=0}^\infty f(i)x^i$, and let $M_f : [0,1] \to \mathbb{R}_\infty$ be the function defined by

$$M_f(x) = x\varphi_f'(1-x) + \varphi_f(1-x) + \varphi_f(1 - \varphi_f'(1-x)/\varphi_f'(1)) - 1.$$

Theorem 7.2 (See [9, Theorem 13 and Eq. (19)]). *Let μ be a probability distribution with nonnegative integer support and finite second moment, and let $G_1, G_2, \ldots$ be a μ-convergent sequence of graphs with adjacency matrices $A_1, A_2, \ldots$. Then*

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}[\operatorname{corank} A_n] \leq \max_{x\in[0,1]} M_\mu(x).$$

Using Theorem 7.2 together with Lemmas 6.4 and 6.10, one can prove that if $k \geq 3$ and $|V|, m = \Theta(n)$ then a random graph $G \sim \mathcal{K}(V, m, k)$ has corank $o(n)$ whp. We will need a variation on this fact, in which some high-degree vertices are first removed.

Lemma 7.3. *Fix $k \geq 3$ and $\varepsilon, \alpha, \Delta > 0$, such that α is sufficiently small with respect to ε and Δ is sufficiently large with respect to α. Consider sets $S \subseteq V$ and an integer m such that $|V| = \Theta(n)$ and $\varepsilon n \leq m - k|V|/2 \leq n/\varepsilon$ and $|S| = \lfloor \alpha|V| \rfloor$, and consider a random graph $G \sim \mathcal{K}(V, m, k)$. Let $T = \{v \in S : \deg_G(v) \geq \Delta\}$ be the set of vertices in S with degree at least Δ. Then, whp the corank of the adjacency matrix of $G[V \setminus T]$ is at most $|T|/8$.*

Remark 7.4. Essentially all parts of our proof of Theorem 1.1 straighforwardly generalise to the setting where np slowly tends to infinity (we only need to consider the regime where say $np \leq 2\log n$; if np grows faster than this, then it is well-known that whp $\operatorname{core}_k(G) = G$ is nonsingular). The exception is Lemma 7.3, since it is proved using Theorem 7.2 (which can only ever apply to sequences of graphs for which the number of edges is proportional to the number of vertices). It may be possible to quantify the methods of Bordenave, Lelarge and Salez, to obtain some version of Theorem 7.2 that applies when $n = o(m)$, but this is actually not necessary, due to various simplifications that can be made in the regime $np \to \infty$.

We (very) briefly sketch how to adapt and simplify our proof of Theorem 1.1: if $np \to \infty$ then typically there are only $o(n)$ vertices outside of $\operatorname{core}_k(G)$, and (by estimating the expected size of the kernel of the adjacency matrix of G, interpreted as a matrix over $\mathbb{F}_2$; see [21]), it is not difficult to show that $\operatorname{corank} G = o(n)$ whp. So, we can simply fix a set T of βn vertices, for some small β (without worrying about the set S). Since $G[V \setminus T]$ is an Erdős–Rényi random graph on its vertex set, whp $\operatorname{core}_k(G) - T$ has corank $o(n)$, and it is also easy to see that whp all but $o(n)$ vertices in T have $(1 - \beta + o(1))np \to \infty$ neighbours in $\operatorname{core}_k(G) - T$.

There are various other minor changes that would have to be made in various parts of this paper to handle the case $np \to \infty$ (for example, we would need a slightly more involved analysis for the unstructured kernel property than is currently in Lemma 4.2). However, since the analogue of Theorem 1.1 in the regime $np \to \infty$ was recently proved by Glasgow [24] (for $k = 3$, but the extension to all $k \geq 3$ is immediate), we do not elaborate further on these changes.

The proof of Lemma 7.3 basically amounts to analysis of the function M_μ, for an appropriate distribution μ. Indeed, we will be able to prove Lemma 7.3 via a very short deduction from the following lemma.

Lemma 7.5. *Fix $\Delta \geq k \geq 3$ and $0 < \varepsilon, \alpha < 1/2$, such that Δ is sufficiently large with respect to k, ε, α. Suppose $\varepsilon n \leq m - kn/2 \leq n/\varepsilon$. For $t \in \mathbb{Z}_{\geq 0}$, let $\mu(t) = (\delta_t + \delta'_t)/(1 - \beta)$, in the notation of Definition 6.5. Then $\sup_{x \in [0,1]} M_\mu(x) \leq \beta/16$.*

Proof of Lemma 7.3 given Lemma 7.5. First note that it suffices to consider the case where $V = \{1, \ldots, n\}$. Also, note that μ defines a probability distribution (one can directly compute that $\sum_{t=0}^{\infty} \mu(t) = 1$, or deduce this indirectly from Lemma 6.6). Let A be the adjacency matrix of $G[V \setminus T]$. By part (1) of Lemma 6.6, we have $|T| = \beta n + O(n^{3/4})$ whp. By part (2) of Lemma 6.4 and parts (2) and (3) of Lemma 6.6, together with Lemmas 6.8 and 6.10, we see that $G[V, T]$ is μ-convergent, so $\mathbb{E} \operatorname{corank} A \leq (\beta/16)n + o(n)$.

Now, we finish the proof by noting that $\operatorname{corank} A$ is tightly concentrated: indeed, if we condition on the degree sequence $\mathbf{d}$ of $G[V \setminus T]$, then Lemma 6.8 tells us that $G[V \setminus T]$ is uniform over graphs with degree sequence $\mathbf{d}$. Consider a random configuration with degree sequence $\mathbf{d}$, defined in terms of a random permutation σ (as in the proof of Lemma 6.6), and let A^* be the adjacency matrix of the corresponding contracted multigraph. Then, modifying σ by a transposition changes at most 4 rows of A^*, so changes $\operatorname{corank} A^*$ by at most 4. Therefore, by a concentration inequality (see for example [39, Eq. (29)]), we have $\operatorname{corank} A^* = \mathbb{E} \operatorname{corank} A^* + o(n)$ whp and therefore $\operatorname{corank} A = \mathbb{E} \operatorname{corank} A^* + o(n)$ whp, since G^* is simple with probability $\Omega(1)$. Since $\operatorname{corank} A^*$ always lies between 0 and n, it follows that $\mathbb{E} \operatorname{corank} A = \mathbb{E} \operatorname{corank} A^* + o(n)$, and therefore $\operatorname{corank} A \leq |T|/8$ whp. $\qquad\square$

We will spend the rest of this section proving Lemma 7.5.

Proof of Lemma 7.5. First, recall the definitions of $\lambda, Z_k(\lambda), \gamma$ from Definitions 6.3 and 6.5, and define $\nu \colon \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by

$$\nu(t) = \frac{(1-\gamma)^t \lambda^t}{(1-\beta)Z_k(\lambda)t!}((\mathbb{1}_{t \geq k} - \alpha \mathbb{1}_{t \geq \Delta}) + \gamma\lambda(\mathbb{1}_{t+1 \geq k} - \alpha \mathbb{1}_{t+1 \geq \Delta})).$$

This does not define a probability distribution, but we do have $\sum_{i=0}^{\infty} i^t \nu(i) < \infty$ for all $t \in \mathbb{N}$. The idea is that ν is an approximation of μ that is more tractable to analyse.

For any $f : \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$ satisfying $\sum_{i=0}^{\infty} i^2 f(i) < \infty$ and $f(0) = f(1) = 0$, we can differentiate term-by-term to compute

$$M'_f(0) = \frac{\varphi''_f(1)}{\varphi'_f(1)}\varphi'_f(0) = 0$$

By [9, Theorem 13] and the remarks following, if $\varphi''_f(x)$ is log-concave on $[0, 1]$ then the global maximum of M_f over $[0, 1]$ is the first local extremum, namely $x = 0$.

The remainder of the proof proceeds in two steps. In one step, we show that $\varphi''_\nu(x)$ is log-concave, hence $M_\nu(x) \leq M_\nu(0)$. In the other step, we show that ν and μ are "close" in the sense that

$$\sup_{x \in [0,1]} |M_\mu(x) - M_\nu(x)| \leq \beta/32. \tag{7.1}$$

Thus, for all $x \in [0, 1]$, it will follow that

$$M_\mu(x) \leq M_\nu(x) + \beta/32 \leq M_\nu(0) + \beta/32 \leq M_\mu(0) + \beta/16 = \beta/16$$

as desired. Here we used the fact that $M_\mu(0) = \varphi_\mu(1) - 1 = 0$.

Step 1: Small difference. We first show (7.1). Recall from Definition 6.5 that $\zeta_{j,t}$ is the probability that a Binomial$(j, 1-\gamma)$ random variable is equal to t (so $\zeta_{j,t} = 0$ if $j < t$), write $x \vee y = \max(x, y)$, and observe that

$$(1 - \beta)Z_k(\lambda)(\mu(t) - \nu(t)) = \left(\sum_{j=k\vee t}^{\infty} \frac{\lambda^j}{j!}\zeta_{j,t} - \alpha \sum_{j=\Delta\vee t}^{\infty} \frac{\lambda^j}{j!}\zeta_{j,t} \right)$$
$$- \left((\mathbb{1}_{t\geq k} - \alpha\mathbb{1}_{t\geq\Delta})\zeta_{t,t}\frac{\lambda^t}{t!} + (\mathbb{1}_{t+1\geq k} - \alpha\mathbb{1}_{t+1\geq\Delta})\zeta_{t+1,t}\frac{\lambda^{t+1}}{(t+1)!} \right)$$
$$= \sum_{j=k\vee(t+2)}^{\infty} \frac{\lambda^j}{j!}\zeta_{j,t} - \alpha \sum_{j=\Delta\vee(t+2)}^{\infty} \frac{\lambda^j}{j!}\zeta_{j,t}.$$

Hence, using the assumption $\Delta \geq k$, we have

$$|(1 - \beta)Z_k(\lambda)(\mu(t) - \nu(t))| \leq \sum_{j=k\vee(t+2)}^{\infty} \frac{\lambda^j}{j!}\zeta_{j,t} = \sum_{j=k\vee(t+2)}^{\infty} \frac{\lambda^j}{t!(j-t)!}(1-\gamma)^t\gamma^{j-t} \leq \frac{2e^\lambda\lambda^t\gamma^2}{t!}.$$

For the rest of the proof we use asymptotic notation, letting $\Delta \to \infty$ while ε, k, α are fixed (so $\beta, \gamma = o(1)$ and $\lambda = O(1)$). In asymptotic notation, we have just shown that $|\mu(t) - \nu(t)| = O(\gamma^2\lambda^t/t!)$, and therefore $|\varphi_\mu(x) - \varphi_\nu(x)|, |\varphi'_\mu(x) - \varphi'_\nu(x)| \leq O(\gamma^2)$ for all $x \in [0, 1]$. Also, note that $\mu(t) \leq e^\lambda\lambda^t/t!$, so $\varphi'_\mu(x) = O(1)$ for all $x \in [0, 1]$, and $\varphi'_\mu(1) \geq \mu(k) = \Omega(1)$. So, we have

$$\varphi_\mu\left(1 - \frac{\varphi'_\mu(1 - x)}{\varphi'_\mu(1)} \right) - \varphi_\nu\left(1 - \frac{\varphi'_\nu(1 - x)}{\varphi'_\nu(1)} \right) = O(\gamma^2) + O\left(\frac{\varphi'_\mu(1 - x)}{\varphi'_\mu(1)} - \frac{\varphi'_\nu(1 - x)}{\varphi'_\nu(1)} \right) = O(\gamma^2)$$

for each $x \in [0, 1]$. Recalling that $M_f(x) = x\varphi'_f(1-x) + \varphi_f(1-x) + \varphi_f(1 - \varphi'_f(1-x)/\varphi'_f(1)) - 1$, it follows that $\sup_{x\in[0,1]} |M_\mu(x) - M_\nu(x)| = O(\gamma^2)$. Finally, we claim that $\gamma^2 = o(\beta)$, which would imply that if Δ is sufficiently large in terms of the other parameters, then (7.1) holds. To see this, we observe that $\gamma = O(\mathbb{E}[X\mathbb{1}_{X\geq\Delta}])$ and $\beta = \Omega(\mathbb{E}[\mathbb{1}_{X\geq\Delta}])$, for $X \sim \text{Poisson}(\lambda)$, and $\mathbb{E}[X\mathbb{1}_{X\geq\Delta}] \leq (\mathbb{E}X^4)^{1/4}(\mathbb{E}[\mathbb{1}_{X\geq\Delta}])^{3/4} = O((\mathbb{E}[\mathbb{1}_{X\geq\Delta}])^{3/4})$ by Hölder's inequality.

Step 2: Log-concavity. Next we show the desired log-concavity of $\varphi''_\nu(x)$, for $x \in [0, 1]$. By rescaling ν and x, we see that it suffices to show ϕ'' is log-concave in the interval $x \in [0, (1 - \gamma)\lambda]$, where $\phi : [0, 1] \to \mathbb{R}$ is defined by

$$\phi(x) = \sum_{t=k-1}^{\infty} \frac{x^t}{t!}((\mathbb{1}_{t\geq k} - \alpha\mathbb{1}_{t\geq\Delta}) + \gamma\lambda(\mathbb{1}_{t+1\geq k} - \alpha\mathbb{1}_{t+1\geq\Delta})).$$

Now, it is well-known (see for example [4, Lemma 3]) that if a nonnegative function h is log-concave on an interval $[a, b]$, then its antiderivative $x \mapsto \int_a^x h(y)\, dy$ is also log-concave on that same interval. Since $k \geq 3$, it therefore suffices to show that the k-fold derivative $f = \phi^{(k-1)}$ is log-concave on the interval $[0, (1 - \gamma)\lambda]$. We claim that in fact it suffices to prove that $g = f + 1$ is log-concave on this interval. Indeed, first observe that f is convex (e.g., by differentiating term-by-term). If we could show that $g = f + 1$ were log-concave on $[0, (1 - \gamma)\lambda]$, it would follow that for any $x, y \in [0, (1 - \gamma)\lambda]$,

$$f(x)f(y) = (1 + f(x))(1 + f(y)) - f(x) - f(y) - 1$$
$$\leq \left(1 + f\left(\frac{x+y}{2} \right) \right)^2 - 2f\left(\frac{x+y}{2} \right) - 1 = f\left(\frac{x+y}{2} \right)^2$$

18

implying that f itself is log-concave on $[0, (1-\gamma)\lambda]$, as desired. So, we study the function g, which we may write as

$$g(x) = (1 + \gamma\lambda) \sum_{t=0}^{\Delta-k-1} \frac{x^t}{t!} + (1 + \gamma\lambda(1-\alpha))\frac{x^{\Delta-k}}{(\Delta-k)!} + (1+\gamma\lambda)(1-\alpha) \sum_{t=\Delta-k+1}^{\infty} \frac{x^t}{t!}.$$

Let $h(x) = e^x - g(x)/(1 + \gamma\lambda)$, so

$$h(x) = \frac{\alpha\gamma\lambda}{1+\gamma\lambda} \cdot \frac{x^{\Delta-k}}{(\Delta-k)!} + \alpha \sum_{t=\Delta-k+1}^{\infty} \frac{x^t}{t!}.$$

We have

$$\frac{\mathrm{d}^2}{\mathrm{d}\,x^2} \log\left(\frac{g(x)}{1+\gamma\lambda}\right) = \frac{(e^x - h(x))(e^x - h''(x)) - (e^x - h'(x))^2}{(e^x - h(x))^2}$$

$$= -\frac{e^x(h(x) + h''(x) - 2h'(x)) + (h'(x)^2 - h(x)h''(x))}{(e^x - h(x))^2}.$$

Furthermore, for sufficiently large Δ we have $h(x) - 2h'(x) + h''(x) \geq 0$ for $x \in [0, (1-\gamma)\lambda]$. To see this, observe that $h(x)$ has a nonzero coefficient of x^t only when $t \geq \Delta - k + 1$, and note that $x^t - 2tx^{t-1} + t(t-1)x^{t-2} = x^{t-2}(x^2 - tx + t(t-1)) \geq 0$ when $x \leq t - \sqrt{t}$. Thus, it suffices to show that $h(x)h''(x) \leq h'(x)^2$; that is, it suffices to show that h is log-concave on $[0, (1-\gamma)\lambda]$. We will show that in fact h is log-concave everywhere.

To this end, it suffices to show that the $(\Delta - k)$-fold derivative $r = h^{(\Delta-k)}$ is log-concave. We write

$$r(x) = \frac{\alpha\gamma\lambda}{1+\gamma\lambda} + \alpha \sum_{t=1}^{\infty} \frac{x^t}{t!}$$

and now we can use the same trick as before: r is convex, so we finish the proof by observing that the function $r + \alpha/(1+\gamma\lambda)$, which is precisely the exponential function $x \mapsto e^x$, is log-concave. $\qquad\square$

8. EXTRACTING HIGH-DEGREE VERTICES WITH RANDOM NEIGHBOURHOODS

In the last section we showed that if we take a random graph $G \sim \mathcal{K}(V, m, k)$ and remove a set of high-degree vertices T, then the result $G[V \setminus T]$ typically has small corank. As sketched in Section 2, the next step is to add back the high-degree vertices, and show that they "boost" the corank to zero. Unfortunately, after deleting the high-degree vertices we are not yet in a position to apply the rank-boosting lemmas in Section 5, mainly because $G[V \setminus T]$ does not satisfy the unstructured kernel property defined in Section 4 (in the language of Lemma 4.2, $G[V \setminus T]$ is not "good"). So, we need to strategically add back a few of the vertices in T (but we need to be careful to do so in such a way that we retain useful randomness). Say that a pair of vertices (u, v) in a (multi)graph is *r-joined* if $u \neq v$ and there is a path of length at most r between u and v, or if $u = v$ and there is a cycle of length at most r containing $u = v$.

Lemma 8.1. *Fix $\Delta \geq k \geq 3$ and $\varepsilon, \alpha > 0$, such that Δ is sufficiently large with respect to ε, α. Consider sets $S \subseteq V$ and an integer m such that $|V| = \Theta(n)$ and $\varepsilon n \leq m - k|V|/2 \leq n/\varepsilon$ and $|S| = \lfloor \alpha|V| \rfloor$, and consider a random graph $G \sim \mathcal{K}(V, m, k)$.*

- *Let $T = \{v \in S : \deg_G(v) \geq \Delta\}$.*
- *Let $B_{\mathrm{bias}} \subseteq V \setminus T$ be the set of vertices which either have $\deg_{V \setminus T}(v) < k$ or are within distance 2 of such a vertex (in $G[V \setminus T]$).*
- *Let T_{bad} be the set of vertices appearing in a pair $(u, v) \in T^2$ which is 6-joined in $G[T \cup B_{\mathrm{bias}}]$.*
- *Let $E = V \setminus (B_{\mathrm{bias}} \cup S)$.*
- *Let $T_{\mathrm{low}} = \{v \in T : \deg_E(v) < \sqrt{\Delta}\}$.*

19

- Let $T' = T \setminus (T_{\text{bad}} \cup T_{\text{low}})$.

Then, the following hold.

(1) *For every superset $U \supseteq V \setminus T'$, the graph $G[U]$ has minimum degree at least 2 and has no pair of degree-2 vertices within distance 4 of each other.*
(2) *Every vertex $v \in T'$ has $\deg_E(v) \geq \sqrt{\Delta}$.*
(3) *Whp $|T'| \geq 7|T|/8$.*
(4) *Suppose we reveal the sets T', E, and reveal the status of every edge in G not between T' and E, and reveal the degree sequence $(\deg_E(v))_{v \in T'}$. Then, conditionally, for each $v \in T'$, the neighbourhood $N_E(v)$ is a uniformly random subset of $\deg_E(v)$ vertices of E, and all such neighbourhoods are jointly independent.*

Some of the parts of Lemma 8.1 are straightforward. In particular, (2) is immediate by the definitions of T_{low} and T'. Part (1) is essentially true by definition as well.

Proof of Lemma 8.1(1). Consider $U \supseteq V \setminus T'$. First, note that $T_{\text{bad}} \subseteq U$, so the vertices in $U \cap T$ have all of their (at least $\Delta \geq k \geq 3$) neighbours inside U. So, the only vertices that could possibly have degree less than 3 are the vertices in $V \setminus T$.

Now, since $k \geq 3$, if a vertex $v \in V \setminus T$ has degree at most 2 in $G[U]$, then $v \in B_{\text{bias}}$. Such a v has exactly one neighbour x_v in $T \setminus U$ (if it had more than one neighbour in T, all of those neighbours would be in $T_{\text{bad}} \subseteq U$), so in fact this is only possible if $k = 3$ and $\deg_{V \setminus T}(v) = 2$.

Finally, for two degree-2 vertices u, v, each having unique neighbours x_u, x_v in $T \setminus U$, there cannot be a path of length at most 4 between u, v in $G[U]$. Indeed, all the vertices of this path would be in B_{bias}, which would imply a path of length at most 6 between $x_u, x_v \in T \setminus T_{\text{bad}}$ (or a cycle of length at most 6 involving $x_u = x_v$) in $G[T \cup B_{\text{bias}}]$. □

For (4), we proceed in essentially the same way as Lemmas 6.1 and 6.8.

Proof of Lemma 8.1(4). Let $G[T', E]$ be the bipartite graph of edges between T' and E. Let H_1, H_2 be bipartite graphs with the same bipartition $X \cup Y$, such that every vertex in X has the same degree in H_1 as it does in H_2. Then, for any outcome of G such that $G[T', E] = H_1$, we can swap $G[T', E]$ with H_2 to obtain an outcome of G such that $G[T', E] = H_2$. So, H_1 and H_2 are equally likely to occur as $G[T', E]$. (It is important that this swap can never change the sets E or T', and can never cause the degree of any vertex to drop below k.) □

The most involved of the four parts of Lemma 8.1 is (3), for which we perform some calculations in the configuration model. The calculations are similar to the ones in Lemmas 6.6 and 6.10, so we will be brief with the details.

Proof of Lemma 8.1(3). Let $\mathbf{d} = (d_1, \ldots, d_n)$ be the degree sequence of G, so as in the proof of Lemma 6.6, by Lemma 6.4 and a concentration inequality, the following properties hold whp. (Recall Definitions 6.3 and 6.5).

- $T = \beta n + o(n)$.
- Each $d_v \leq \log n$.
- For all t, there are $\rho_t n + O(n^{3/4})$ vertices $v \in V$ with $d_v = t$.
- For all $t \geq k$, there are $(1 - \alpha)\rho_t n + O(n^{3/4})$ vertices $v \in V \setminus S$ with $d_v = t$.
- For all $t \geq \Delta$, there are $\alpha \rho_t n + O(n^{3/4})$ vertices $v \in T$ with $d_v = t$.

Condition on an outcome of $\mathbf{d}$ satisfying the above properties. Now, G is a uniformly random graph with degree sequence $\mathbf{d}$, which we may study with the configuration model. So, consider a uniformly random configuration with degree sequence $\mathbf{d}$, defined in terms of a random permutation σ. Let G^* be the resulting contracted multigraph. We may interpret the random sets $B_{\text{bias}}, T_{\text{bad}}, E, T_{\text{low}}, T'$ as being defined with respect to G^*.

Define B_{safe} to be the set of vertices v in $G^*[V \setminus S]$ for which $d_v = k + 1$ and $\deg_{V \setminus S}(v) = k$, and for which all vertices w within distance 2 of v in $G^*[V \setminus S]$ have $d_w = \deg_{V \setminus S} = k$. Note that $E \supseteq B_{\text{safe}}$. Let X be the number of vertices in T with fewer than $\sqrt{\Delta}$ neighbours in B_{safe}, and let Y be the number of paths of length at most 6 (in G^*) between an ordered pair of vertices $(u, v) \in T^2$ (where we count a cycle of length greater than zero as a path, in the case $u = v$). It suffices to show that $X, Y \leq (\beta/16)n$ whp. By a concentration inequality for random permutations, we have $X = \mathbb{E}X + o(n)$ and $Y = \mathbb{E}Y + o(n)$ whp, so it actually suffices to show that $\mathbb{E}X, \mathbb{E}Y \leq (\beta/20)n$.

First, we consider Y. For any sequence $t_1, \ldots, t_{r-1}$ (with $r \leq 6$) there are

$$(2m\gamma + O(n^{3/4} \log n))^2 \prod_{i=1}^{r-1} t_i(t_i - 1)(\rho_{t_i} n + O(n^{3/4}))$$

ways to choose a sequence of distinct points $u, x_1, y_1, \ldots, x_{r-1}, y_{r-1}, v$, such that u, v are each in a bucket corresponding to a vertex in T, and for each i, the points x_i, y_i are in a common bucket corresponding to a vertex v with $d_v = t_i$. The probability any given sequence of r disjoint pairs are chosen in our random configuration is $(2m-1)^{-1}(2m-3)^{-1} \ldots (2m-2r+1)^{-1} = (1+o(1))(2m)^{-r}$. It follows that

$$\mathbb{E}Y \leq \sum_{r=1}^{6} (2m\gamma + o(n))^2 \left(\sum_{t=k}^{\log n} t(t-1)\rho_t n + O(n^{3/4}(\log n)^2) \right)^{r-1} \frac{1 + o(1)}{(2m)^r}$$

$$\leq \gamma^2 n \sum_{r=1}^{6} 2^{2-r} \varepsilon^{-|2-r|} \left(\sum_{t=k}^{\infty} t(t-1)\rho_t \right)^{r-1} + o(n).$$

Recalling the interpretation of ρ_t as a probability mass of a truncated Poisson distribution, we have $\sum_{t=k}^{\infty} t(t-1)\rho_t \leq \lambda^2/Z(\lambda)$ (in particular, the sum is convergent and does not depend on Δ). Also, recall from the proof of Lemma 7.5 that γ^2/β can be made arbitrarily small by taking sufficiently large Δ (holding ε, α, k constant). So for large Δ we have $\mathbb{E}Y \leq (\beta/20)n$, as desired.

It remains to consider X. Consider any vertex $v \in T$, and let b_v be a set of Δ points in the bucket corresponding to v. Let Q_v be the set of points in b_v which are matched to a point corresponding to a vertex in B_{safe} in our random configuration. Note that $x \in Q_v$ if and only if x is matched with a point corresponding to a degree-$(k+1)$ vertex in $V \setminus S$, and the entire 2-neighbourhood of that vertex consists of vertices in $V \setminus S$ which have degree k. There are at least $(1-\alpha)\delta_k n + o(n)$ vertices $v \in V \setminus S$ with $d_v = k$, and at least $(1-\alpha)\delta_{k+1} n + o(n)$ vertices $v \in V \setminus S$ with $d_v = k+1$, and in total our configuration model has $2m \leq n/\varepsilon$ points. So, for each $x \in b_v$, we have

$$\Pr(x \in Q_v) \geq (\varepsilon(1-\alpha)\delta_{k+1})(\varepsilon(1-\alpha)\delta_k)^{k^3} + o(1).$$

Let $p = (\varepsilon(1-\alpha)\rho_{k+1})(\varepsilon(1-\alpha)\rho_k)^{k^3}/2$; crucially, this quantity does not depend on Δ. Note that the events $x \in Q_v$, for $x \in b_v$, are very nearly independent (we only need to inspect $O(1)$ points to determine whether $x \in Q_v$, and there are only Δ points in b_v), so $|Q_v|$ stochastically dominates the binomial distribution with parameters Δ and p. By a Chernoff bound, the probability that $|Q_v| \leq \sqrt{\Delta}$ (i.e., that v has fewer than $\sqrt{\Delta}$ neighbours in B_{safe}) is at most $e^{-p\Delta/8} \leq 1/21$ for large Δ. It follows that $\mathbb{E}X \leq (\beta/20)n$, as desired. $\qquad \square$

9. Finishing the proof

In this section we combine all the ingredients collected so far and prove Theorem 1.1. As part of this proof, we will need the following simple observation.

Proof of Theorem 1.1. For a small constant $\varepsilon > 0$, let $\mathcal{E}_{\text{large}}$ be the event that $\text{core}_k(G)$ has at least εn vertices, at least $k|\text{core}_k(G)|/2 + \varepsilon^2 n$ edges, and at most n/ε edges (i.e., it has linear size,

bounded average degree, and linearly many vertices have degree strictly greater than k). It follows from the work of Pittel, Spencer and Wormald [43] that if $\varepsilon = \varepsilon_k > 0$ is a small enough constant, then whp either $\mathrm{core}_k(G) = \emptyset$ or $\mathcal{E}_{\mathrm{large}}$ holds (see for example [27] for explicit theorem statements which imply this claim).

For a small constant $\theta' > 0$, let $\mathcal{E}_{\mathrm{odd}}$ be the event that $\mathrm{core}_k(G)$ has at least $\theta' n$ odd-degree vertices. If we condition on an outcome of the numbers of vertices and edges in $\mathrm{core}_k(G)$ such that $\mathcal{E}_{\mathrm{large}}$ holds, then Lemma 6.4 tells us that for sufficiently small $\theta' > 0$, the event $\mathcal{E}_{\mathrm{odd}}$ holds whp (considering vertices of degree k or $k+1$, depending on the parity of k).

Let $\theta = \theta'/2$, and for a small constant $\eta > 0$, let $\mathcal{E}_{\mathrm{UKP}}$ be the event that the conclusion of Lemma 4.2 holds (so if η is sufficiently small, $\mathcal{E}_{\mathrm{UKP}}$ holds whp). For the convenience of the reader we recall that in the statement of Lemma 4.2 we defined an induced subgraph of G to be *good* if it has minimum degree at least 2, every pair of degree-2 vertices are at distance at least 4, there are fewer than $(\eta/4)n$ vertices within distance 7 of a degree-2 vertex, and there are at least θn odd-degree vertices. The conclusion of Lemma 4.2 was that every good induced subgraph of G satisfies the unstructured kernel property $\mathrm{UKP}(2, \eta/3, \eta)$.

For a small constant $\alpha > 0$, let $\mathcal{E}_{\mathrm{expand}}$ be the event that for every set S of αn vertices, there are at most $\min(\theta'/2, \eta/4)$ vertices within distance 8 of a vertex in S. If α is sufficiently small, then $\mathcal{E}_{\mathrm{expand}}$ holds whp (for example, one can iterate Lemma 4.5).

Let $\mathcal{S}$ be the event that $\mathrm{core}_k(G)$ is singular, and fix $\Delta \in \mathbb{N}$ (which we view as a constant for the purpose of "whp" statements and asymptotic notation). It now suffices to show that

$$\Pr(\mathcal{S} \cap \mathcal{E}_{\mathrm{UKP}} \cap \mathcal{E}_{\mathrm{odd}} \cap \mathcal{E}_{\mathrm{expand}} \cap \mathcal{E}_{\mathrm{large}}) \le f(\Delta) + o(1), \tag{9.1}$$

where f is some function satisfying $\lim_{\Delta \to \infty} f(\Delta) = 0$. Indeed, recalling that the empty 0×0 matrix is nonsingular, it will follow from the above considerations that $\Pr(\mathcal{S}) \le f(\Delta) + o(1)$ and therefore $\limsup_{n \to \infty} \Pr(\mathcal{S}) \le f(\Delta)$. Since this is true for all Δ, and $f(\Delta) \to 0$ as $\Delta \to \infty$, the desired result will follow.

Let V be the vertex set of $\mathrm{core}_k(G)$, and let m be the number of edges in $\mathrm{core}_k(G)$. Condition on outcomes of V, m satisfying $\mathcal{E}_{\mathrm{large}}$. By Lemma 6.1, $\mathrm{core}_k(G)$ is now a uniformly random graph on the vertex set V with m edges and minimum degree at least k. That is to say, $\mathrm{core}_k(G) \sim \mathcal{K}(V, m, k)$.

Fix a subset $S \subseteq V$ of $\lfloor \alpha |V| \rfloor$ vertices. As in the statement of Lemma 7.3, let T be the set of vertices in S which have at least Δ neighbours in V, and let $T' \subseteq T$ be the set defined in Lemma 8.1. Let A be the adjacency matrix of $G[V \setminus T]$, and let A' be the adjacency matrix of $G[V \setminus T']$. By Lemma 7.3, whp $\mathrm{corank}\, A \le |T|/8$, and by Lemma 8.1(3), whp $|T'| \ge 7|T|/8$. It follows that in fact whp

$$\mathrm{corank}\, A' \le |T|/8 + |T|/8 \le |T'|/2. \tag{9.2}$$

With E as defined in Lemma 8.1, condition on all information except the edges between E and T'. Moreover condition on the degree of every vertex in T' into E. Let $\mathcal{I} = (V, m, \tilde{G}, (\deg_E v)_{v \in T'})$ encode the information revealed so far, where $\tilde{G}$ is the graph consisting of all edges of G not between T' and E. Assume that (9.2) holds.

By Lemma 8.1(2), every vertex in T' has at least $\sqrt{\Delta}$ neighbours in E, and by Lemma 8.1(4), these neighbours are uniformly random. Now, let $t = |T'|$, and enumerate the vertices in T' as $v_1, \ldots, v_t$. Let $T_i = \{v_{i+1}, \ldots, v_t\}$, and write A_i for the adjacency matrix of $G[V \setminus T_i]$. Also, write $\mathcal{E}_{\mathrm{UKP}}^i$ for the event that A_i satisfies $\mathrm{UKP}(2, \eta/3, \eta)$. We also define a sequence of random variables $X_0, \ldots, X_t$ as follows.

$$X_i = \begin{cases} \mathrm{corank}\, A_i & \text{if } \mathcal{E}_{\mathrm{UKP}}^0, \ldots, \mathcal{E}_{\mathrm{UKP}}^i \text{ all hold} \\ 0 & \text{otherwise.} \end{cases}$$

22

Note that each X_i is determined by $G[V \setminus T_i]$. By Lemma 5.1,

$$\Pr(X_t \geq \max(X_{t-1}, 1) \,|\, X_1, \ldots, X_{t-1}, \mathcal{I}) \leq \frac{(\log \Delta^{1/2})^C}{\Delta^{1/4}}$$

for an absolute constant C. Let $f(\Delta) = 11(\log \Delta)^C \Delta^{-1/4}$, so by (9.2) and Lemma 5.2, we have $\Pr(X_t = 0 \,|\, \mathcal{I}) \geq 1 - f(\Delta)$ for sufficiently large n.

Now we claim that if $\mathcal{E}_{\mathrm{UKP}} \cap \mathcal{E}_{\mathrm{odd}} \cap \mathcal{E}_{\mathrm{expand}}$ holds then corank $A_i = X_i$ for all i (so if additionally $X_t = 0$ then $\mathrm{core}_k(G)$ is nonsingular). Indeed, for each $G[V \setminus T_i]$, first $\mathcal{E}_{\mathrm{odd}} \cap \mathcal{E}_{\mathrm{expand}}$ implies that there are at least $\theta'n - (\theta'n/2) = \theta n$ odd-degree vertices, then $\mathcal{E}_{\mathrm{expand}}$ implies that there are fewer than $(\eta/4)n$ vertices within distance 7 of a degree-2 vertex (since each degree-2 vertex has a neighbour in S with respect to G), and then Lemma 8.1(1) implies that the minimum degree is at least 2 and there is no pair of degree-2 vertices within distance 4 of each other. That is to say, each $G[V \setminus T_i]$ is good, so $\mathcal{E}_{\mathrm{UKP}}$ implies that each $\mathcal{E}^i_{\mathrm{UKP}}$ holds, so corank $A_i = X_i$ for all i, as claimed.

Now we are ready to conclude the proof. For all $\mathcal{I}$ such that V, m satisfy $\mathcal{E}_{\mathrm{large}}$ and such that (9.2) holds, we have

$$\Pr(\mathcal{S} \cap \mathcal{E}_{\mathrm{UKP}} \cap \mathcal{E}_{\mathrm{odd}} \cap \mathcal{E}_{\mathrm{expand}} \,|\, \mathcal{I}) \leq \Pr(\mathcal{S} \cap \mathrm{corank}\, A_t = X_t \,|\, \mathcal{I}) \leq \Pr(X_t = 0 \,|\, \mathcal{I}) \leq f(\Delta).$$

Using the law of total probability and the fact that (9.2) holds (uniformly) whp conditional on any V, m satisfying $\mathcal{E}_{\mathrm{large}}$, we find

$$\Pr(\mathcal{S} \cap \mathcal{E}_{\mathrm{UKP}} \cap \mathcal{E}_{\mathrm{odd}} \cap \mathcal{E}_{\mathrm{expand}} \,|\, V, m) \leq f(\Delta) + o(1).$$

Finally, (9.1) follows from the law of total probability applied over pairs V, m satisfying $\mathcal{E}_{\mathrm{large}}$. As discussed, this concludes the proof. $\square$

Remark 9.1. The reason that our proof of Theorem 1.1 does not provide an effective probability bound is that we must take $\Delta \to \infty$ sufficiently slowly in terms of the (ineffective) convergence in Theorem 7.2. Apart from the ineffective $f(\Delta)$ probability bound, all relevant events can be shown to hold with probability $1 - n^{-\Omega(1)}$.

We now briefly sketch how to modify the proof in order to obtain a probability bound of the form $1 - (\log \log n)^{O(1)}/\sqrt{\log n}$. Instead of extracting vertices of degree at least Δ, for some Δ that is slowly taken to infinity, we separately extract vertices with degree between Δ and $\log n/(\log \log n)^2$ (call these "medium-degree vertices"), and vertices with degree at least $\log n/(\log \log n)^2$ (call these "high-degree vertices"), for some large constant Δ. One can show that there are very likely to be at least (say) $n^{4/5}$ high-degree vertices.

Now, let T' be the set of medium and high degree vertices (after being "cleaned up" in a similar way to Lemma 8.1). Add these vertices back to $G[V \setminus T']$ one-by-one, starting with the medium-degree vertices. If we do not take $\Delta \to \infty$ then it is no longer true that after adding back the medium-degree vertices we end up with corank zero whp, but if Δ is a sufficiently large constant it is still possible to show that the corank is very likely at most (say) $n^{1/3}$. Then, adding back the high-degree vertices boosts the corank all the way to zero, with probability $1 - (\log \log n)^{O(1)}/\sqrt{\log n}$.

Of course, this modification does not avoid ineffectivity entirely. Indeed, one can show that after adding the medium-degree vertices, the corank is at most $n^{1/3}$ with probability at least say $1 - 1/n$, for sufficiently large n, but the cutoff for "sufficiently large" is an absolute constant depending on the (ineffective) convergence in Theorem 7.2. It may be possible to quantify the methods used by Bordenave, Lelarge and Salez [9] to prove Theorem 7.2, and thereby obtain completely effective probability bounds, but we have not attempted to do so.

References

[1] L. Addario-Berry and L. Eslava, *Hitting time theorems for random matrices*, Combin. Probab. Comput. **23** (2014), no. 5, 635–669. 1

[2] D. Aldous and J. M. Steele, *The objective method: probabilistic combinatorial optimization and local weak convergence*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 1–72. 4

[3] A. Backhausz and B. Szegedy, *On the almost eigenvectors of random regular graphs*, Ann. Probab. **47** (2019), no. 3, 1677–1725. 5

[4] M. Bagnoli and T. Bergstrom, *Log-concave probability and its applications*, Economic theory **26** (2005), no. 2, 445–469. 18

[5] A. Basak and M. Rudelson, *Sharp transition of the invertibility of the adjacency matrices of sparse random graphs*, arXiv:1809.08454. 1

[6] I. Benjamini and O. Schramm, *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), no. 23, 13. 4

[7] T. Bohman and A. Frieze, *Karp-Sipser on random graphs with a fixed degree sequence*, Combin. Probab. Comput. **20** (2011), no. 5, 721–741. 4

[8] B. Bollobás, *The evolution of sparse graphs*, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 35–57. 2

[9] C. Bordenave, M. Lelarge, and J. Salez, *The rank of diluted random graphs*, Ann. Probab. **39** (2011), no. 3, 1097–1121. 3, 4, 16, 17, 23

[10] J. Cain and N. Wormald, *Encores on cores*, Electron. J. Combin. **13** (2006), no. 1, Research Paper 81, 13. 2, 4, 13

[11] A. Coja-Oghlan, A. A. Ergür, P. Gao, S. Hetterich, and M. Rolvien, *The rank of sparse random matrices*, Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2020, pp. 579–591. 3

[12] A. Coja-Oghlan and P. Gao, *The rank of random matrices over finite fields*, arXiv:1810.07390. 3

[13] N. Cook, *The circular law for random regular digraphs*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 4, 2111–2167. 8

[14] N. A. Cook, *On the singularity of adjacency matrices for random regular digraphs*, Probab. Theory Related Fields **167** (2017), no. 1-2, 143–200. 3, 8

[15] C. Cooper, *The cores of random hypergraphs with a given degree sequence*, Random Structures Algorithms **25** (2004), no. 4, 353–375. 2

[16] K. P. Costello, T. Tao, and V. Vu, *Random symmetric matrices are almost surely nonsingular*, Duke Math. J. **135** (2006), no. 2, 395–413. 1, 5, 6, 11

[17] K. P. Costello and V. Vu, *On the rank of random sparse matrices*, Combin. Probab. Comput. **19** (2010), no. 3, 321–342. 2

[18] K. P. Costello and V. H. Vu, *The rank of random graphs*, Random Structures Algorithms **33** (2008), no. 3, 269–285. 1

[19] R. W. R. Darling and J. R. Norris, *Differential equation approximations for Markov chains*, Probab. Surv. **5** (2008), 37–79. 2

[20] P. DeMichele, M. Glasgow, and A. Moreira, *Distances to the span of sparse random matrices, with applications to gradient coding*, arXiv:2105.11718. 2

[21] A. Ferber, *Singularity of random symmetric matrices – simple proof*, arXiv:2006.07439. 16

[22] A. Ferber, M. Kwan, and L. Sauermann, *Singularity of sparse random matrices: simple proofs*, arXiv:1807.06465. 6

[23] D. Fernholz and V. Ramachandran, *The giant k-core of a random graph with a specified degree sequence*, manuscript,, 2003. 2

[24] M. Glasgow, *Invertibility of the 3-core of Erdős Rényi graphs with growing degree*, arXiv:2106.00963. 2, 16

[25] J. Huang, *Invertibility of adjacency matrices for random d-regular graphs*, arXiv:1807.06465. 3, 4, 5

[26] S. Janson and M. J. Luczak, *A simple solution to the k-core problem*, Random Structures Algorithms **30** (2007), no. 1-2, 50–62. 2

[27] S. Janson and M. J. Luczak, *Asymptotic normality of the k-core in random graphs*, Ann. Appl. Probab. **18** (2008), no. 3, 1085–1137. 22

[28] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. 7, 14

[29] D. M. Kane, *The correct exponent for the Gotsman-Linial conjecture*, Comput. Complexity **23** (2014), no. 2, 151–175. 6

[30] J. H. Kim, *Poisson cloning model for random graphs*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 873–897. 2

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

[32] J. Komlós, *On the determinant of random matrices*, Studia Sci. Math. Hungar. **3** (1968), 387–399. 1

[33] M. Krivelevich, E. Lubetzky, and B. Sudakov, *Cores of random graphs are born Hamiltonian*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 1, 161–188. 2

[34] M. Kwan and L. Sauermann, *An algebraic inverse theorem for the quadratic Littlewood-Offord problem, and an application to Ramsey graphs*, Discrete Anal. (2020), Paper No. 12, 34. 7

[35] M. Kwan, B. Sudakov, and T. Tran, *Anticoncentration for subgraph statistics*, J. Lond. Math. Soc. (2) **99** (2019), no. 3, 757–777. 6, 7

[36] B. Landon, P. Sosoe, and H.-T. Yau, *Fixed energy universality of Dyson Brownian motion*, Adv. Math. **346** (2019), 1137–1332. 3

[37] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, *Adjacency matrices of random digraphs: singularity and anti-concentration*, J. Math. Anal. Appl. **445** (2017), no. 2, 1447–1491. 3, 6, 8

[38] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, *The smallest singular value of a shifted d-regular random square matrix*, Probab. Theory Related Fields **173** (2019), no. 3-4, 1301–1347. 8

[39] C. McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248. 14, 15, 17

[40] R. Meka, O. Nguyen, and V. Vu, *Anti-concentration for polynomials of independent random variables*, Theory Comput. **12** (2016), Paper No. 11, 16. 6

[41] A. Mészáros, *The distribution of sandpile groups of random regular graphs*, Trans. Amer. Math. Soc. **373** (2020), no. 9, 6529–6594. 3, 4

[42] M. Molloy, *Cores in random hypergraphs and Boolean formulas*, Random Structures Algorithms **27** (2005), no. 1, 124–135. 2

[43] B. Pittel, J. Spencer, and N. Wormald, *Sudden emergence of a giant k-core in a random graph*, J. Combin. Theory Ser. B **67** (1996), no. 1, 111–151. 2, 22

[44] R. W. Robinson and N. C. Wormald, *Almost all cubic graphs are Hamiltonian*, Random Structures Algorithms **3** (1992), no. 2, 117–125. 3

[45] T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2010. 6

[46] V. Vu, *Recent progress in combinatorial random matrix theory*, arXiv:2005.02797. 2

[47] V. H. Vu, *Combinatorial problems in random matrix theory*, Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV, Kyung Moon Sa, Seoul, 2014, pp. 489–508. 2

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, IRVINE.
Email address: `asaff@uci.edu`

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA.
Email address: `mattkwan@stanford.edu`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
Email address: `{asah,msawhney}@mit.edu`