

Evolution of random representable matroids: minors, circuits, connectivity and the critical number

Pu Gao

University of Waterloo
pu.gao@uwaterloo.ca

Jacob Mausberg

University of Waterloo
jmausberg@uwaterloo.ca

Peter Nelson

University of Waterloo
apnelson@uwaterloo.ca

Abstract

We study the evolution of random matroids represented by the sequence of random matrices over \mathbb{F}_q where columns are added one after the other, and each column vector is a uniformly random vector in \mathbb{F}_q^n , independent of each other. We study the appearance of matroid minors, the appearance of circuits, the evolution of the connectivities and the critical number. We settle several open problems in the literature.

1 Introduction

Random graphs were first introduced by Erdős and Rényi [6] in 1959, and in particular they studied the evolution of random graphs by studying a random process of graphs on a set of n vertices, where the edges are added one after the other in the process. The most astonishing phenomenon in the study of random graph evolution is the characterisation of phase transitions of various (often increasing) graph properties. For instance, initially, the graph is acyclic with each component being a tree of bounded order. Then, small cycles may start to appear whereas each component remains in small size and contains at most one cycle. At the time where the number of edges is around $n/2$, small components start to rapidly connect to each other and form a giant component in which more complicated graph structures start to appear. Long cycles and graph minors of fixed sizes simultaneously appear at the time the giant component emerges. Other well studied graph properties and graph parameters include connectivity, Hamiltonicity, the appearance of given subgraphs, chromatic number, etc.

The Erdős-Rényi graph process immediately induces a random process on graphical matroids, which motivates the generalisations to other classes of random matroid processes. One generalisation is to consider the matroids represented by the incidence matrices of random uniform hypergraphs, which was introduced by Cooper, Frieze and Pegden [4], and was more

formally described and studied as an evolutionary random process in [7]. The other generalisation is to consider uniformly random vectors over a finite field and add them one after the other independently, and consider the random matroids represented by these matrices. This model was first introduced by Kelly and Oxley [9, 10]. In their model, they consider random subsets of the elements in the complete projective geometry $PG(n-1, q)$ where q is a prime power. Two related models were introduced and studied. In the first one, $PG(n-1, q; p)$ where $p \in [0, 1]$, every element in $PG(n-1, q)$ is kept independently with probability p . In the second model $PG(n-1, q; m)$ where m is an integer between 0 and $(q^n-1)/(q-1)$, a uniformly random subset of m elements of $PG(n-1, q)$ is selected. Obviously, $PG(n-1, q; p)$ and $PG(n-1, q; m)$ are analogs of $\mathcal{G}(n, p)$ and $\mathcal{G}(n, m)$ for random graphs, and $PG(n-1, q; m)$ is precisely $PG(n-1, q; p)$ conditioned to $|PG(n-1, q; p)| = m$; i.e. exactly m elements are selected. After that, Kordecki [12, 13], Kordecki and Luczak [14, 15] further studied matroid properties of these models, including circuits, connectivity, and submatroids, etc. Soon after the introduction of $PG(n-1, q; m)$ and $PG(n-1, q; p)$, Kelly and Oxley [8] introduced a slightly different model $M([U_q]_{n \times m})$. In this model, $[U_q]_{n \times m}$ is a uniformly random $n \times m$ matrix over \mathbb{F}_q , and $M([U_q]_{n \times m})$ is the matroid represented by $[U_q]_{n \times m}$. Due to the independence of the column vectors, $[U_q]_{n \times m}$ is a little easier to analyse than $PG(n-1, q; p)$ and $PG(n-1, q; m)$. Indeed, as noted by all the authors that followed this study, these three models are asymptotically equivalent for all the problems (e.g. rank, circuits and connectivity) they were studying. We give a proof of their equivalence in Proposition 1 below.

In this paper, we use the last model introduced by Kelly and Oxley [8], however we stress that columns are added one by one and we are interested in the evolution of the matroids represented by this random process of matrices. Let \mathbb{F}_q denote the finite field of order q where q is a prime power. Let \mathbf{v} be a uniformly random vector in \mathbb{F}_q^n , and let $(v_i)_{i \geq 1}$ be a sequence of random vectors that are independent copies of \mathbf{v} . Finally, for every $m \geq 1$, let $A_m = [v_1, \dots, v_m]$ be the $n \times m$ matrix formed by including the first m vectors v_1, \dots, v_m in the sequence, and let $M[A_m]$ be the matroid represented by A_m . Notice that for every $m \geq 1$, A_m has the same distribution as $[U_q]_{n \times m}$. We study various matroid properties and parameters of $M[A_m]$ as m grows. In particular, we take a thorough study of the time when matroid minors of small ranks appear, the appearance of the projective geometry of growing rank of n as a minor, the appearance of the circuits of different lengths, the evolution of the connectivity, and the growth rate of the critical number. We discuss them in turn in the coming subsections. We start our discussions by unifying the notions of the three models $PG(n-1, q; p)$, $PG(n-1, q; m)$ and $M[A_m]$ and show their asymptotic equivalence. The following Gaussian binomial coefficients will be used throughout the paper, which counts the k -dimensional subspaces of \mathbb{F}_q^n :

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \prod_{i=0}^{k-1} \frac{q^{n-i} - 1}{q^{k-i} - 1}; \quad [n]_q := \begin{bmatrix} n \\ 1 \end{bmatrix}_q = \frac{q^n - 1}{q - 1}.$$

Notice that $PG(n-1, q)$ contains exactly $[n]_q$ elements. Given a sequence of probability spaces indexed by n , we say a sequence of events A_n occurs asymptotically almost surely (a.a.s.) if $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 1$. The standard Landau notation and basic matroid notions such as simple matroid, free matroid, and the rank of a matroid will be introduced in Section 2.

Proposition 1. *The three models (M1): $M[A_m]$, (M2): $PG(n-1, q; m)$ and (M3): $PG(n-1, q; p)$, are related as follows:*

- (a) (M1), conditioned on the event that $M[A_m]$ is simple, is equivalent to (M2).
- (b) (M3), conditioned on the event that precisely m elements of $PG(n-1, q)$ are included, is equivalent to (M2).
- (c) If $m = o(q^{n/2})$ then (M1) and (M2) are asymptotically equivalent.

Proof. Parts (a,b) are obvious. For part (c), suppose that $m = o(q^{n/2})$. By part (i), it suffices to show that a.a.s. $M[A_m]$ is simple. The probability that A_m has a zero column is at most $m q^{-n} = o(1)$. Each element in $PG(n-1, q)$ corresponds to exactly $q-1$ vectors in \mathbb{F}_q^n . Thus, the probability that A_m has two linearly dependent columns is at most $\binom{m}{2} [n]_q (q-1)^2 q^{-2n} = o(1)$. Hence, a.a.s. $M[A_m]$ is a simple matroid. ■

1.1 Minors

Let N be an \mathbb{F}_q -representable matroid. Let $\tau_{N\text{-minor}}$ be the smallest m such that $M[A_m]$ contains N as a minor (the definition of matroid minor is given in Section 2). Altschuler and Yang [1] determined the critical window in which $\tau_{N\text{-minor}}$ lies, provided that the size of N is fixed, i.e. independent of n . Given a matrix or a matroid M , let $\text{crk}(M)$ denote the co-rank of M . Given a random variable X_n and a real number x_n , we write $X_n = O_p(x_n)$ if

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} (\mathbb{P}(X_n < \varepsilon x_n) + \mathbb{P}(X_n > \varepsilon^{-1} x_n)) = 0.$$

Theorem 2. *Suppose N is a fixed non-free \mathbb{F}_q -representable matroid.*

- (a) (Theorems 5 and 6 of [1]) *Let $k \geq 0$ be a positive integer. There exist constants $C, D > 0$ depending on N, q and k such that*

$$\begin{aligned} \liminf_{n \rightarrow \infty} \mathbb{P}(\tau_{N\text{-minor}} \leq m) &> C \quad \text{if } m = n + k \text{ and } k \geq 1 \\ \limsup_{n \rightarrow \infty} \mathbb{P}(\tau_{N\text{-minor}} \leq m) &\leq D \quad \text{if } m = n - k. \end{aligned}$$

- (b) (Theorems 3 and 8 of [1]) $\tau_{N\text{-minor}} = n + O_p(1)$.

Note that if N is a free matroid with rank r such that $n - r \rightarrow \infty$ then it is easy to see (e.g. it follows easily from the proof of Lemma 22) that a.a.s. all the columns of A_r are linearly independent and thus a.a.s. $\tau_{N\text{-minor}} = r$. On the other hand, if N is not \mathbb{F}_q -representable then N can never be a minor of $M[A_m]$, no matter how large m is. Therefore, in the discussions of matroid minors we only focus on non-free \mathbb{F}_q -representable matroid.

Altschuler and Yang gave specific bounds C, D in Theorem 2(a). Interested readers can find them in [1]. We do not express them here, as these bounds only give qualitative information about $\mathbb{P}(|\tau_{N\text{-minor}} - n| > k)$ for $k \rightarrow \infty$, which they used to come to the conclusion of Theorem 2(b). However these bounds give little information about $\mathbb{P}(\tau_{N\text{-minor}} \leq n + k)$

when $|k|$ is small. Our first result is a strengthening of Theorem 2(a) by providing the precise limiting distribution of $\tau_{N\text{-minor}} - n$. For convenience, we define $\sum_{i=j}^h a_i$ to be 0 and $\prod_{i=j}^h a_i$ to be 1 if $h < j$ for any sequence of real numbers or real functions a_i . Given a power series $P(z)$ let $[z^n]P(z)$ denote the coefficient of z^n in $P(z)$.

Theorem 3. *Let N be a fixed \mathbb{F}_q -representable matroid and let r and c denote the rank and the co-rank of N respectively. Suppose that $c \geq 1$. Then, for any fixed $k \in \mathbb{Z}$,*

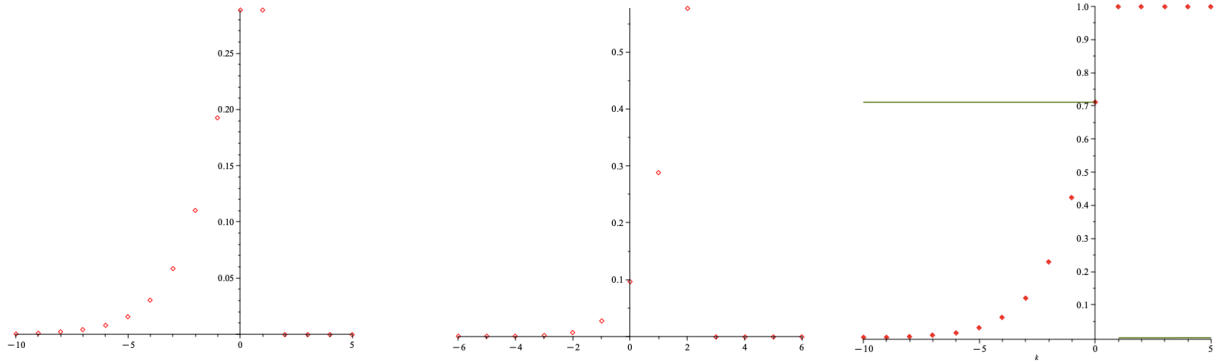
$$\lim_{n \rightarrow \infty} \mathbb{P}(\tau_{N\text{-minor}} = n + k) = C_{c,k},$$

where $C_{c,k} = 0$ if $k > c$; and if $k \leq c$ then

$$\begin{aligned} C_{c,k} &= \beta_{c,k} q^{k-c} \sum_{i=0}^{c-1} \left(\alpha_{c,k,0} + \sum_{i=1}^{c-1} \frac{1}{\prod_{j=1}^i (1 - q^{-j})} \alpha_{c,k,i} \right), \\ \beta_{c,k} &= \prod_{j=c+1-k}^{\infty} (1 - q^{-j}), \\ \alpha_{c,k,i} &= [z^{c-1-i}] \prod_{j=0}^{c-k-1} (1 - zq^{-j}) = (-1)^{c-1-i} \sum_{*} q^{-\sum_{r=1}^{c-1-i} j_r}, \end{aligned}$$

where the summation \sum_{*} in the expression for $\alpha_{c,k,i}$ is over all integers $0 \leq j_1 < j_2 < \dots < j_{c-1-i} \leq c - k - 1$.

We plot below the limiting point-wise probabilities $\mathbb{P}(\tau_{N\text{-minor}} = n + k)$ for $k \in [-10, 5]$ when $(q, c) = (2, 1)$ (the one on the left) and when $(q, c) = (2, 2)$ (the one in the middle). The last figure on the right compares the limiting cumulative distribution function $\mathbb{P}(\tau_{N\text{-minor}} \leq n + k)$, in red dots, with the bounds in Theorem 2(a) by Altschuler and Yang, in green curve. Recall that the bounds in Theorem 2(a) are upper bounds for non-positive k and lower bounds for positive k .



Our second result improving over Altschuler and Yang's is a hitting time result of $\tau_{N\text{-minor}}$ when the rank and the co-rank of N are fixed or slowly growing functions of n . Let $\tau_{\text{crk}=c}$ be the smallest m such that the co-rank of A_m is equal to c . This is well defined as the co-rank of A_m is non-decreasing as m grows.

Theorem 4. *Suppose that N is an \mathbb{F}_q -representable non-free matroid such that $rq^{rc} = o(n)$, where r and c denote the rank and the co-rank of N respectively. Then a.a.s. $\tau_{N\text{-minor}} = \tau_{\text{crk}=c}$.*

As a direct corollary, we prove that a.a.s. an N -minor is formed by step $n + \text{crk}(N)$, improving Theorem 2(b). On the other hand, there is a non-vanishing probability that the first N -minor is created precisely during the step $n + \text{crk}(N)$.

Corollary 5. *Suppose that N is a fixed non-free \mathbb{F}_q -representable matroid. Let c be the co-rank of N . Then, a.a.s. $\tau_{N\text{-minor}} = n + O_p(1)$. Moreover, letting $\gamma_{q,c} = \prod_{j=c}^{\infty} (1 - q^{-j})$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tau_{N\text{-minor}} \leq n + c) = 1; \quad \lim_{n \rightarrow \infty} \mathbb{P}(\tau_{N\text{-minor}} \leq n + c - 1) = 1 - \gamma_{q,c},$$

for every (q, c) . Moreover,

$$\lim_{q \rightarrow \infty} \gamma_{q,c} = 1 \quad \text{for every } c; \quad \lim_{c \rightarrow \infty} \gamma_{q,c} = 1 \quad \text{for every } q.$$

The next corollary shows that $\tau_{N\text{-minor}}$ has a 1-point concentration if $|N|$ is not too large, and $\text{crk}(N) = \omega(1)$. Here, $|N|$ denotes the size of N , which is the number of elements in N .

Corollary 6. *Suppose that N is an \mathbb{F}_q -representable non-free matroid. If $\text{crk}(N) = \omega(1)$ and $|N| \leq (2 - \varepsilon)\sqrt{\log_q n}$ for some fixed $\varepsilon > 0$, then a.a.s. $\tau_{N\text{-minor}} = n + \text{crk}(N)$.*

Note that the matroid N in Theorems 2, 3, 4 and Corollaries 5 and 6 is general, and does not need to be simple. In the next corollary of Theorem 4, we generalise Theorem 2 and study τ_{PGr} , the minimum integer m such that $M[A_m]$ contains the complete projective geometry $PG(r - 1, q)$ as a minor. All the logarithms in the paper are natural logarithms with base e , unless otherwise with a specified base.

Theorem 7. *Suppose that $r = \omega(1)$. Let $\zeta = [r]_q$.*

- (a) *A.a.s. $\tau_{PGr} \leq n + \zeta \log \zeta + \omega(\zeta)$.*
- (b) *If $r = \omega(\log n)$ then a.a.s. $\tau_{PGr} \sim \zeta \log \zeta$.*
- (c) *If N is an \mathbb{F}_q -representable simple non-free matroid with rank $r \leq \alpha \log n$ for any fixed $\alpha < 1/\log q$ then a.a.s. $\tau_{N\text{-minor}} \sim n$.*

Remark 8. *Theorem 7 determines the asymptotic value of τ_{PGr} provided that $r \leq (1 - \varepsilon) \log_q n$ for some fixed $\varepsilon > 0$, or $r = \omega(\log n)$. However, we did not manage to prove a lower bound for τ_{PGr} that matches its upper bound in Theorem 7(a) for $\log_q n \leq r = \Theta(\log n)$. There is a trivial lower bound $n + \zeta - r = n + (1 + o(1))\zeta$, since the co-rank of A_m must be at least the co-rank of $PG(r - 1, q)$ (see Lemma 23 below). It is possible that in the range $\log_q n \leq r = \Theta(\log n)$, both the upper bound (Theorem 7(a)) and the lower bound $(n + (1 + o(1))\zeta)$ are not tight.*

The proofs of Theorems 3, 4 and 7, and the proofs of Corollaries 5 and 6 will be presented in Section 3.

1.2 Circuits

The circuits in a matroid (see its formal definition in Section 2) are minimal dependent subsets and are analogs of cycles in a graph. The appearance of the circuits with constant length has been studied by Kelly and Oxley [8], whereas the longer circuits were investigated by Kordecki and Łuczak [15]. We collect and restate their results in the following theorem.

Theorem 9. (a) (Theorem 5.1 of [8]) Suppose that $k = o(m)$ and that $m^k q^{-n}$ is bounded. Then

$$\Pr(M[A_m] \text{ has no } k\text{-circuits}) \sim \exp\left(-\frac{(q-1)^{k-1} m^k}{k! q^n}\right).$$

(b) (Theorem 6 of [15]) For each $1 \leq k \leq n+1$, let $\mu_k = \binom{m}{k} (q-1)^k q^{-n}$.

(i) If $\mu_k \rightarrow 0$, then a.a.s. $M[A_m]$ does not contain a k -circuit.

(ii) If $\mu_k \rightarrow \infty$ and either $n-k \rightarrow \infty$ or $k(m-k)/m \rightarrow \infty$, then a.a.s. $M[A_m]$ contains a k -circuit.

Remark 10. Kordecki and Łuczak's result [15] was proved for $PG(n-1, q; m)$. The same holds for $M[A_m]$ by Proposition 1. The original statement of [15, Theorem 6] was slightly stronger, by considering the appearance of circuits whose lengths lie in a specified interval. For simplicity we stated a simplified version. The value μ_k is the asymptotic number of circuits with length k in $PG(n-1, q; m)$ and $M[A_m]$.

Theorem 9 provides the full information on the types of circuits that are likely or unlikely to appear in $M[A_m]$ for every m . We give a few interesting corollaries of Theorem 9 in terms of $\tau_{k\text{-circ}}$, the minimum integer m such that $M[A_m]$ has a circuit with length k . Given $0 < a \leq 1$, define

$$g_a(y) = y \log y + a \log(q-1) - a \log a - (y-a) \log(y-a) - \log q, \quad \text{for } y \geq a, \quad (1)$$

where $0 \log 0$ is defined to be 0 so that $g_a(y)$ is continuous at $y = a$.

Corollary 11. (a) Let k be a fixed positive integer. Then, $\tau_{k\text{-circ}} = \Theta_p(q^{n/k})$.

(b) If $k \rightarrow \infty$ and $k = o(n)$ then a.a.s. $\tau_{k\text{-circ}} \sim \frac{1}{e(q-1)} k q^{n/k}$.

(c) If $k \sim an$ for some fixed $0 < a \leq 1$ then $g_a(y)$ has a unique root b , and a.a.s. $\tau_{k\text{-circ}} \sim bn$.

Remark 12. For $0 < a \leq 1$, let $b = b(a)$ be the unique root of $g_a(y)$. Then b is strictly convex and has a minimum value of 1, achieved at $a^* := \frac{q-1}{q}$. Moreover, $b(1) < 2$ and $b(a) \rightarrow \infty$ as $a \rightarrow 0$.

Remark 12 follows from simple calculus and we include its proof in the Appendix for interested readers. A plot of $b(a)$ for $q = 2$ is given below in Figure 1. Following Remark 12 we immediately obtain the following interesting corollary about the length of the first appearing circuit and the appearing time of the first Hamilton circuit (i.e. the circuit with length n).

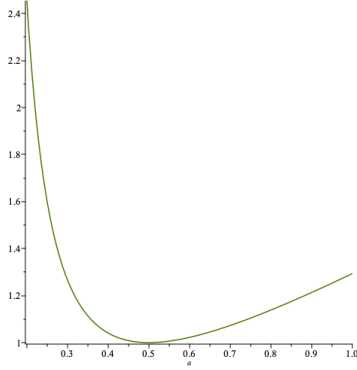


Figure 1: Plot of $b(a)$

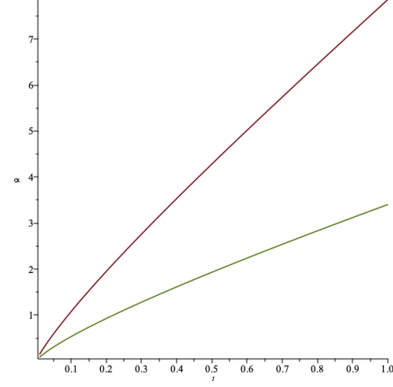


Figure 2: Upper bound of α from Kelly and Oxley and our lower bound

- Corollary 13.** (a) *The first circuit appearing in $M[A_m]$ has length asymptotic to $(1 - q^{-1})n$;*
- (b) *The first Hamilton circuit appears in less than $2n$ steps for every finite field \mathbb{F}_q .*

The proofs for Corollaries 11 and 13 are given in Section 4.

1.3 Connectivity

The concept of vertex-connectivity in graphs generalizes naturally to matroids. It is easily seen that a graph G is k -vertex-connected if and only if it is not the union of two edge-disjoint subgraphs G_1, G_2 such that $\min\{|V(G_1)|, |V(G_2)|\} > k$, and $|V(G_1) \cap V(G_2)| < k$; the latter is equivalent to $|V(G_1)| + |V(G_2)| - |V(G)| < k$.

Analogously, a *vertical k -separation* in a matroid M is a partition (A_1, A_2) of M so that

- (i) $\min\{\text{rk}(A_1), \text{rk}(A_2)\} \geq k$, and
- (ii) $\text{rk}(A_1) + \text{rk}(A_2) - \text{rk}(M) \leq k - 1$.

The *vertical connectivity* $\kappa(M)$ is defined to be the smallest k such that M has a vertical k -separation. It may be the case that M has no vertical k -separations for any k , in which case $\kappa(M) = \infty$ ¹; this holds, for instance, if $M = \mathbb{F}_q^n$.

From the perspective of graph theory, this is the most natural notion of matroid connectivity; indeed, if G is a graph with no isolated vertex, then the value of κ for the graphic matroid $M(G)$ is equal to the vertex connectivity of G : see [21], Theorem 8.6.1. (Incidentally, this fact is the reason for the unnatural-looking offset by one in condition (ii)).

However, κ fails to have a certain natural matroid property: invariance under duality. Each matroid M has a dual matroid M^* , and the relationship between M and M^* is crucial in much of matroid theory – for instance, we have $M^{**} = M$, and matroid duality agrees with planar duality in the case of the graphic matroids of planar graphs. It is not necessary

¹some work defined $\kappa(M)$ to be the rank of M in this case

here to discuss matroid duality in detail, but we comment that the rank function of the dual matroid is given by $\text{rk}^*(X) = |X| + \text{rk}(E(M) \setminus X) - \text{rk}(M)$ (see [21], Proposition 2.1.9).

In general, we have $\kappa(M) \neq \kappa(M^*)$, so vertical connectivity has a dual notion. A *cyclic k -separation* of M is a vertical k -separation of M^* , and the *cyclic connectivity* $\kappa^*(M)$ is the smallest k such that M has a cyclic k -separation. Similarly as before, $\kappa^*(M)$ defined to be ∞ if no cyclic separation exists. Using the above formula for the dual rank function, one can easily show that a partition (A_1, A_2) of M is a cyclic k -separation of M if and only if

- (i) $\text{rk}(A_i) < |A_i|$ for each i , and
- (ii) $\text{rk}(A_1) + \text{rk}(A_2) - \text{rk}(M) \leq k - 1$.

Condition (i) can be replaced with the requirement that A_1 and A_2 are dependent, or by the condition $\min(\text{rk}^*(A_1), \text{rk}^*(A_2)) \geq k$. It is harder to relate this intuitively to graphs, except to comment that, if M is the graphic matroid of a planar graph G , then the cyclic connectivity of M is the vertex connectivity of a planar dual of G . One way to construct a small cyclic separation in a matroid M is simply to take a small circuit; if C is a circuit of M for which $E(M) \setminus C$ is a dependent set, then $(C, E(M) \setminus C)$ satisfies (i), and $\text{rk}(C) + (\text{rk}(E(M) \setminus C) - \text{rk}(M)) \leq |C| - 1 + 0$, so C gives a cyclic $|C|$ -separation, implying that $\kappa^*(M) \leq |C|$. In the setting of planar duality, this corresponds to a fact that a small cycle in the planar dual of G gives rise to a small cut in G .

Evidently cyclic connectivity is not invariant under matroid duality. However, there is a third notion of connectivity that is. A *Tutte k -separation* of M is a partition (A_1, A_2) of M so that

- (i) $\min\{|A_1|, |A_2|\} \geq k$, and
- (ii) $\text{rk}(A_1) + \text{rk}(A_2) - \text{rk}(M) < k - 1$.

The *Tutte connectivity* $t(M)$ is the smallest k so that G is Tutte k -connected. Each cyclic or vertical k -separation is also a Tutte k -separation, which implies that $t(M) \leq \min(\kappa(M), \kappa^*(M))$. In fact, one can show that if $|M| \geq 3$ and $t(M) < \infty$, then there is always either a vertical or a cyclic $t(M)$ -separation, which implies that equality holds. This fact is essentially stated in [21], Proposition 8.6.6, but the notation is slightly different from ours in edge cases, as $\kappa(M)$ and $\kappa^*(M)$ are defined in [21] to always be finite.

Proposition 14. *If M is a matroid with $|M| \geq 3$, then $t(M) = \min\{\kappa(M), \kappa^*(M)\}$.*

This fact shows that Tutte connectivity is invariant under matroid duality, so in a sense it is the most natural connectivity notion of the three; usually the unadorned ‘connectivity’ of a matroid refers to the Tutte connectivity.

The *girth* $\text{gir}(M)$ of a matroid M is the size of a smallest circuit of M , or ∞ if M is independent. We have seen that in nontrivial cases, small circuits give small cyclic separations. It follows that the girth usually provides an upper bound for the connectivity of a matroid. The following bound ([21], Theorem 8.6.4) will be useful.

Proposition 15. *If M is a matroid that is not a uniform matroid $U_{r,n}$ with $n \geq 2r - 1$, then $t(M) = \min(\kappa(M), \text{gir}(M))$.*

A striking difference between matroid and graph connectivities is the monotonicity. From the definitions it is easy to see that all of the Tutte, the vertical, and the cyclic connectivities are not monotone; i.e. adding elements to a matroid may decrease the connectivity. Our first observation is that $\kappa(M[A_m])$ is a.a.s. monotonely non-decreasing as m grows. On the other hand, the Tutte connectivity first follows $\kappa(M[A_m])$, and then after some linear number of steps, it becomes governed by $\text{gir}(M[A_m])$ and decreases as m grows. The evolutionary trajectory of $\kappa^*(M[A_m])$ is very different from $\kappa(M[A_m])$, which starts from ∞ and in the end is governed by $\text{gir}(M[A_m])$. We will study $\kappa^*(M[A_m])$ in a different paper.

Theorem 16. (a) *A.a.s. $\kappa(M[A_m])$ is monotonely non-decreasing as m increases.*

(b) *A.a.s. $t(M[A_m]) = \kappa(M[A_m])$ for all $m = n + o(n)$.*

(c) *A.a.s. there exists $\hat{m} = \Theta(n)$ such that $t(M[A_m]) = \kappa(M[A_m])$ for all $m < \hat{m}$, and $t(M[A_m]) < \kappa(M[A_m])$ for all $m \geq \hat{m}$.*

Thanks to the monotonicity of the vertical connectivity of $M[A_m]$ as shown in Theorem 16(a), it is natural to define $\tau_{k\text{-conn}}$ to be the smallest m such that $M[A_m]$ is vertically k -connected. Due to Proposition 15 and our good understanding of $\tau_{k\text{-circ}}$ from Corollary 11, $\tau_{k\text{-conn}}$ and $\tau_{k\text{-circ}}$ together would immediately determine the evolution of $t(M[A_m])$. The limiting distribution of $\tau_{k\text{-conn}}$ has been determined by Kordecki and Łuczak [15] when k is a constant; whereas an upper bound on $\tau_{k\text{-conn}}$ was provided by Kelly and Oxley [8] when k is linear in n .

Theorem 17. [15, Theorem 4] *Let $k \geq 2$ and $m - n - (k - 1) \log_q n \rightarrow c$ for some constant c . Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(M[A_m] \text{ is } k\text{-connected}) = \exp\left(- (q - 1)^{k-2} q^{-c} / (k - 1)!\right).$$

Kelly and Oxley give the following upper bounds for when $M[A_{n \times m}]$ becomes k -connected.

Theorem 18. (a) ([8, Theorem 4.4]) *A.a.s. $\kappa(M[A_m]) = \infty^2$ if $m \geq (1 + \alpha)n$ where α is any constant such that*

$$\alpha > \frac{\log(2q - 1)}{2 \log q - \log(2q - 1)}.$$

(b) ([8, Theorem 4.5]) *Suppose that $k \sim tn$ for some fixed $0 < t < 1$ then a.a.s. $\tau_{k\text{-conn}} \leq (1 + \alpha)n$ for any constant α such that*

$$t \log \left[\frac{(1 + t)\alpha}{t^2} \right] < (\alpha - t) \ln q - 2t.$$

Our contributions are the determination of the sharp phase transition of $\tau_{k\text{-conn}}$ for all $k = o(n)$, and a lower bound on $\tau_{k\text{-conn}}$ for $k = \Theta(n)$.

Theorem 19. (a) *Suppose $k \rightarrow \infty$ and $k = o(n)$. Then, a.a.s. $\tau_{k\text{-conn}} - n \sim k \log_q(n/k)$.*

²In the original work of [8], the theorem was phrased as “ $M[A_m]$ is vertically n -connected”; we rephrased it to be consistent with our definition of vertical connectivity when no separator exists.

(b) Suppose that $k \sim tn$ for some fixed $0 < t < 1$. Then, a.a.s. $\tau_{k\text{-conn}} \geq (1 + \alpha)n$ where α is any constant satisfying the following.

$$t \log \frac{1 + \alpha}{t} + (1 + \alpha - t) \log \frac{1 + \alpha}{1 + \alpha - t} + t \log(q - 1) - \alpha \log q > 0. \quad (2)$$

Our lower bound on $\tau_{k\text{-conn}}$ for linear k does not match the upper bound in Theorem 18. See Figure 2 for the plot in which we compare the upper bounds by Kelly and Oxley and the lower bounds in Theorem 19(b); the horizontal axis is for $t = k/n$, and the vertical axis is for $n^{-1}\tau_{k\text{-conn}} - 1$. Determining the asymptotic value of $\tau_{k\text{-conn}}$ for $k = \Theta(n)$ is an interesting open problem.

The proofs for Theorems 16 and 19 will be given in Section 5.

1.4 Critical number

The critical number is an extension of the notion of the chromatic number of graphs to matroids. It is easy to see that a graph G is 2^k -colourable if and only if $E(G)$ is the union of k edge cuts of G . For each set $U \subseteq V(G)$, the edge cut $\delta(U)$ corresponds to the set of support-2 vectors in \mathbb{F}_2^V that have a nonzero dot product with the characteristic vector 1_U . Hence, we can form an analogue of chromatic number by defining $\chi_q(M)$ for each matroid $M \subseteq \mathbb{F}_q^n$ to be the minimum k such that there are vectors $v_1, \dots, v_k \in \mathbb{F}_q^n$ such that, for all $w \in M$, the dot product $w^T v_i$ is nonzero for some $i \in [k]$. In other words, $\chi_q(M)$ is the minimum integer k such that there exists an $(n - k)$ -dimensional subspace S of \mathbb{F}_q^n such that $M \cap S = \emptyset$. We call $\chi_q(M)$ the *critical number* of M ; it has previously been called the critical exponent. It can be seen as the right analogue of chromatic number in a variety of contexts; see [21], p. 588 for a discussion.

Our goal is to determine when $M[A_m]$ has critical number k for all $1 \leq k \leq n$. (In the range of values for m that we consider, A_m will a.a.s. not contain a zero column, so the critical number is well-defined). Notice that the critical number starts at 1 (i.e. when $m = 1$) and increases monotonically with m (i.e. as more columns are added). Also, notice that the critical number cannot skip over any values of k , since adding one column can increase the critical number by at most one. Thus, we focus on determining the step when the critical number of $M[A_m]$ jumps from k to $k + 1$ for $1 \leq k \leq n - 1$.

Let $\tau_{k\text{-crt}}$ be the minimum integer m such that $\chi_q(M[A_m]) = k + 1$. In other words, $\tau_{k\text{-crt}}$ is the precise step where the critical number jumps from k to $k + 1$. We obtain the following theorem whose proof will be presented in Section 6.

Theorem 20. *Let k be a positive integer such that $k \leq n - \log_q n - \log_q \log n - \omega(1)$. Then, a.a.s. $\tau_{k\text{-crt}} \sim -k(n - k) \log q / \log(1 - q^{-k})$.*

Remark 21. *Note that our theorem covers all positive integers k up to distance $\log_q n + \log_q \log n + \omega(1)$ from n . For greater k up to $n - 1$ we have the trivial asymptotic upper bound $nq^n \log q$ on $\tau_{k\text{-crt}}$ from coupon collection (with q^n coupons).*

1.5 Other related work

Other than the rank, circuits, connectivity, critical number, minors that are discussed in this paper, the thresholds and limiting distributions of the number of small submatroids were

studied by Oxley [20], and further extended by Kordecki [12, 13].

The uniformly random matroid on n elements was introduced and studied by Mayhew, Newman, Welsh and Whittle [17]. Research in this direction focuses on enumeration of matroids and matroid extensions. We refer the readers to [16, 3, 19, 11, 22] for results in this field.

2 Preliminary

A matroid is defined on a pair $M = (E, \mathcal{I})$ where E is a set called the ground set of the matroid M , and $\mathcal{I} \subseteq 2^E$ denotes the set of independent sets of M . The size of M is $|E|$, the number of elements in the ground set. The rank of M , denoted by $\text{rk}(M)$, is the size of a largest independent set. The co-rank of M , denoted by $\text{crk}(M)$, is defined by $|E| - \text{rk}(M)$. Let \mathbb{F} be a field. A matroid $M = (E, \mathcal{I})$ is said \mathbb{F} -representable if there is a matrix $A = [a_p]_{p \in E}$ over \mathbb{F} , where the columns of A are indexed by elements in E , such that $S \subseteq E$ is an independent set if and only if $\{a_p : p \in S\}$ is a linearly independent set. A matroid M is free if $E(M)$ is an independent set, and M is simple if M does not have dependent subsets of cardinality one or two. Consequently, if M is represented by a matrix A over \mathbb{F} , then M is free if all columns of A are linearly independent, and M is simple if A does not contain the zero column, or two linearly dependent columns. The uniform matroid $U_{r,n} = ([n], \mathcal{I})$ is the matroid on ground set $[n]$ such that \mathcal{I} consists of all subsets of $[n]$ of cardinality at most r . A circuit of a matroid is a minimal dependent subset of elements in M . In other words, every proper subset of a circuit is an independent set. The length of a circuit is the number of elements in the circuit.

Suppose that $M = (E, \mathcal{I})$ and $X \subseteq E$. The deletion of X from M is defined by $M \setminus X = (E \setminus X, \{I \in \mathcal{I} : I \cap X = \emptyset\})$. The contraction of X from M is defined by $M/X = (E \setminus X, \mathcal{I}^*)$ where

$$\mathcal{I}^* = \{I \in \mathcal{I} : I \cap X = \emptyset, I \cup J \in \mathcal{I} \text{ for some maximal independent subset } J \text{ of } X\}.$$

A submatroid of M is any matroid obtained by deleting a subset of elements in M ; whereas a minor of M is a matroid obtained by deleting and contracting elements in M . If M is a matroid represented by a matrix A over a field \mathbb{F} , then there are matrix operations described as follows which yield representations for submatroids and minors of A .

Given a matrix A over \mathbb{F} where columns are indexed by E , and $X \subseteq E$, let A_X denote the matrix obtained from A by only including columns in X . Let $A_{E/X}$ be any matrix obtained by first obtaining matrix $B \sim A$ via row operations such that

$$B_X = \begin{bmatrix} I_{a \times a} & * \\ 0 & 0 \end{bmatrix}, \quad \text{where } a = \text{rk}(A_X),$$

and then deleting the a rows where $I_{a \times a}$ lies, together with all the columns in X . Suppose that $A = [a_p]_{p \in E}$ is a matrix representing $M = (E, \mathcal{I})$, and $X \subseteq E$. Then, $M \setminus X$ is represented by $A_{E \setminus X}$ and M/X is represented by $A_{E/X}$. We refer the readers to Oxley [21] for other basics in matroid theory.

Finally, we use standard Landau notation in this paper. Given sequences of real numbers a_n and b_n , we say $a_n = O(b_n)$ if there exists $C > 0$ such that $|a_n| \leq C|b_n|$ for every n . We say

$a_n = o(b_n)$ if $\lim_{n \rightarrow \infty} a_n/b_n = 0$. We say $a_n = \Theta(b_n)$ or $a_n \asymp b_n$ if $a_n, b_n > 0$ and $a_n = O(b_n)$ and $b_n = O(a_n)$. Finally, we write $a_n = \omega(b_n)$ if $a_n, b_n > 0$ and $b_n = o(a_n)$.

3 Minors

Lemma 22 (Lemma 2 of [1]). *For every $m \leq n$,*

$$\mathbb{P}(\text{rk}(A_m) = m) = \prod_{i=0}^{m-1} (1 - q^{i-n}).$$

Proof. It follows immediately by the fact that given the first $i \leq m - 1$ column vectors of A_m being linearly independent, the probability that the $(i + 1)$ -th column vector falls in the span of the first i column vectors is equal to q^{i-n} . ■

Lemma 23. *If M is a matroid containing N as a minor, then $\text{crk}(M) \geq \text{crk}(N)$.*

Proof. Let $e \in E(M)$. In both the case of deleting e or contracting e , the rank of M decreases by at most one, and the size of the ground set of M decreases by exactly one. Thus, $\text{crk}(M) \geq \text{crk}(M \setminus e)$ and $\text{crk}(M) \geq \text{crk}(M/e)$. The result follows. ■

Proof of Theorem 4. By Lemma 23, $\tau_{\text{N-minor}} \geq \tau_{\text{crk=c}}$. To prove that a.a.s. $\tau_{\text{N-minor}} \leq \tau_{\text{crk=c}}$, consider the following equivalent way of generating the process A_m for $m \leq \tau_{\text{crk=c}}$. As each column vector v_i is drawn, we only expose whether v_i lies in the subspace generated by v_1, \dots, v_{i-1} . If it is, we colour the column red; otherwise we colour it blue. Stop the process when there are exactly c red columns. The following claim follows by Lemma 22.

Claim 24. *A.a.s. the first $n/2$ columns are blue.*

Next, we expose all the column vectors corresponding to the blue columns. Let A_B denote the submatrix of A_m composed of all blue columns of A_m . Since the column vectors of A_B are linearly independent, A_B is row equivalent to the identity matrix, with possibly a few zero rows underneath. In other words, there is an invertible matrix P such that $PA_B = \begin{bmatrix} I \\ \mathbf{0} \end{bmatrix}$ where $\mathbf{0}$ are a set of all-0 vectors.

Finally, we expose the column vectors corresponding to the red columns. Note that each red column vector is a uniformly random vector in the span of the blue column vectors generated before it.

Claim 25. *Suppose that v is a red column vector and there are i blue columns before v . Then, $Pv \sim \begin{bmatrix} [U_q]^{i \times 1} \\ \mathbf{0} \end{bmatrix}$.*

In other words, Pv has the same distribution as the vector obtained by appending $n - i$ 0's after a uniformly random vector in \mathbb{F}_q^i .

Consider $M = M[A_m]$. Take the first $n/2$ blue columns of A_m and partition them into $t := \lceil n/2r \rceil$ groups I_1, \dots, I_t (by discarding the remaining columns if $n/2$ is not divisible by r), where I_1 denotes the first r blue columns, I_2 denotes the next r blue columns, etc. The a.a.s. existence of at least $n/2$ blue columns in A_m is guaranteed by Claim 24. Let X_j be the matrix obtained from A_m by contracting all blue columns except for the blue columns in I_j for $1 \leq j \leq t$. Then, each X_j has form $[I_{r \times r} \mid R_j]$.

Claim 26. $(R_j)_{j=1}^t$ are mutually independent, and each $R_j \sim [U_q]^{r \times c}$.

Since N is \mathbb{F}_q representable, we may represent N by a rank- r matrix of form $[I_{r \times r} \mid R]$ for some $r \times c$ matrix R over \mathbb{F}_q . By definition, M contains N as a minor if $R_j = R$ for some $1 \leq j \leq t$. By Claim 26, this occurs with probability q^{-rc} for each $1 \leq j \leq t$. Moreover, all R_j are independent. Thus, the probability that M has an N -minor is at least

$$1 - (1 - q^{-rc})^t \geq 1 - \exp\left(-\frac{n}{2r}q^{-rc}\right) = 1 - o(1),$$

since $rq^{rc} = o(n)$. ■

It remains to prove Claims 25 and 26.

Proof of Claim 25. Let v_1, \dots, v_i denote the i blue column vectors that appear before v . Let z_1, \dots, z_i be i.i.d. uniform random variables in \mathbb{F}_q . Since v is a uniform random vector in $\langle v_1, \dots, v_i \rangle$, the span of $v_1, \dots, v_i, v \sim \sum_{j=1}^i z_j v_j$. Hence,

$$Pv \sim \sum_{j=1}^i z_j P v_j.$$

The claim follows by the distribution of z_1, \dots, z_i and the fact that $P[v_1, \dots, v_i] = [I_{i \times i}]$. ■

Proof of Claim 26. Let A_R be the matrix formed by the red columns of A_m . By Claim 24 we may assume that the first $n/2$ columns of A_m are all blue, and thus by Claim 25, the submatrix of PA_R formed by the first $n/2$ rows has distribution $[U_q]^{n/2 \times c}$. The claim follows by noticing that R_1 is the first r rows of PA_R , R_2 is the next r rows of PA_R , etc. ■

Proof of Theorem 3. By Theorem 4, $\tau_{N\text{-minor}} = n+k$ if $\text{crk}(A_{n+k-1}) = c-1$ and $\text{crk}(A_{n+k}) = c$, which happens if $\text{rk}(A_{n+k-1}) = \text{rk}(A_{n+k}) = n+k-c$. Our derivation of $\mathbb{P}(\text{rk}(A_{n+k-1}) = \text{rk}(A_{n+k}) = n+k-c)$ is an easy adaptation of the proof of [14, Fact 3]. For each $1 \leq j \leq n+k-c$, let u_j be the number of column vectors v that are added in the process $(A_m)_{m=1}^{n+k}$ that lies in the subspace \mathcal{S} generated by the column vectors added before v when the dimension of \mathcal{S} is equal to j . Then, letting $u := \sum_{j=0}^{n+k-c} u_j$, $\text{rk}(A_{n+k-1}) = \text{rk}(A_{n+k}) = n+k-c$ if and only if $u = c$ and $u_{n+k-c} \geq 1$. Moreover, all u_j s are independent random variables, and for each $1 \leq j \leq n+k-c-1$, u_j has geometric distribution with probability q^{j-n} . It follows then that

$$\begin{aligned} & \mathbb{P}(\text{rk}(A_{n+k-1}) = \text{rk}(A_{n+k}) = n+k-c) \\ &= [z^c] \left(\prod_{j=0}^{n+k-c-1} \sum_{h=0}^{\infty} (zq^{j-n})^h (1 - q^{j-n}) \right) \left(\sum_{h=1}^{\infty} (zq^{k-c})^h \right) \\ &= [z^c] (1 + O(q^{-n})) \beta_{k,c} z q^{k-c} \prod_{t=c-k}^n \frac{1}{1 - zq^{-t}} \\ &= q^{k-c} \beta_{k,c} [z^{c-1}] (1 + O(q^{-n} + zq^{-n})) \prod_{t=c-k}^{\infty} \frac{1}{1 - zq^{-t}} \\ &= q^{k-c} \beta_{k,c} [z^{c-1}] (1 + O(q^{-n} + zq^{-n})) \prod_{t=0}^{c-k-1} (1 - zq^{-t}) \prod_{t=0}^{\infty} \frac{1}{1 - zq^{-t}}. \end{aligned}$$

By Euler's formula (see [2, Corollary 2.2]), if $|t| < 1$ and $|z| < 1$ then

$$\prod_{i=0}^{\infty} \frac{1}{1 - zt^i} = 1 + \sum_{i=1}^{\infty} \frac{z^i}{\prod_{j=1}^i (1 - t^j)}.$$

Thus,

$$\begin{aligned} & \mathbb{P}(\tau_{\mathbf{N}\text{-minor}} = n + k) \\ &= q^{k-c} \beta_{k,c} [z^{c-1}] (1 + O(q^{-n} + zq^{-n})) \left(\prod_{t=0}^{c-k-1} (1 - zq^{-t}) \right) \left(1 + \sum_{i=1}^{\infty} \frac{z^i}{\prod_{j=1}^i (1 - q^{-j})} \right) \\ &\sim \beta_{k,c} q^{k-c} \left([z^{c-1}] \prod_{t=0}^{c-k-1} (1 - zq^{-t}) + \sum_{i=1}^{c-1} \frac{1}{\prod_{j=1}^i (1 - q^{-j})} [z^{c-1-i}] \prod_{t=0}^{c-k-1} (1 - zq^{-t}) \right). \end{aligned}$$

The theorem follows. \blacksquare

Proof of Corollary 5. The claim that $\mathbb{P}(\tau_{\mathbf{N}\text{-minor}} \leq n + c) = 1 - o(1)$ follows by Theorem 3. Moreover, $C_{c,c}$ in Theorem 3 is equal to $\gamma_{q,c} = \prod_{j=c}^{\infty} (1 - q^{-j}) > 0$. Hence, $\lim_{n \rightarrow \infty} \mathbb{P}(\tau_{\mathbf{N}\text{-minor}} \leq n + c - 1) = 1 - \gamma_{q,c}$. \blacksquare

Proof of Corollary 6. Since $|N| \leq (2 - \varepsilon) \sqrt{\log_q n}$, $rc \leq (1 - \varepsilon) \log_q^n$ where $r = \text{rk}(N)$ and $c = \text{crk}(N)$, and thus $rq^{rc} = o(n)$. By Theorem 4, a.a.s. $\tau_{\mathbf{N}\text{-minor}} = \tau_{\text{crk}=\mathbf{c}}$. Since $c = \omega(1)$, a.a.s. $\tau_{\text{crk}=\mathbf{c}} = n + \omega(1)$ by Lemma 22. It follows then that a.a.s. $\text{rk}(A_{\text{crk}=\mathbf{c}}) = n$ and consequently $\tau_{\text{crk}=\mathbf{c}} = n + \text{crk}(N)$. \blacksquare

Before proving Theorem 7, we present a probabilistic tool of Poisson approximation of the balls-into-bins model.

Lemma 27. *Suppose b balls are placed into k bins, independently and uniformly at random. Let \mathcal{E} be the event that every bin gets at least one ball. Set $\lambda = b/k$. Then*

$$\mathbb{P}(\mathcal{E}) \leq 2(1 - e^{-\lambda})^k \leq 2e^{-ke^{-\lambda}} \quad \text{and} \quad \mathbb{P}(\overline{\mathcal{E}}) \leq 2ke^{-\lambda}.$$

Proof. Let Y_1, \dots, Y_k be independent Poisson variables each with mean λ . Then, the distribution of the number of balls in bins is the same as (Y_1, \dots, Y_k) conditioned to $\sum_{i=1}^k Y_i = b$ (see e.g. [18, Theorem 5.6] for a proof). By Theorem 5.10 of [18] (with $f(x_1, \dots, x_k)$ be the indicator variable that $x_i \geq 1$ for every $1 \leq i \leq k$ or $f(x_1, \dots, x_k)$ be the indicator variable that $x_i = 0$ for some $1 \leq i \leq k$),

$$\mathbb{P}(\mathcal{E}) \leq 2\mathbb{P}(Y_i \geq 1 \forall i) = 2(1 - e^{-\lambda})^k; \quad \mathbb{P}(\overline{\mathcal{E}}) \leq 2\mathbb{P}(Y_i = 0 \text{ for some } i) \leq 2ke^{-\lambda}. \quad \blacksquare$$

We also need the following lemma from [1] concerning the distribution of a uniformly random vector in \mathbb{F}_q^n after a change of basis.

Lemma 28. *[1, Lemma 5] Suppose that $P \in \mathbb{F}_q^{n \times n}$ is invertible, then $PA_m \sim [U_q]^{n \times m}$.*

Next, we assume that $r = r(n) \rightarrow \infty$. Recall that

$$\zeta = \begin{bmatrix} r \\ 1 \end{bmatrix}_q = \frac{q^r - 1}{q - 1}.$$

Note that $PG(r-1, q)$ has ζ elements.

Proof of Theorem 7. For part (a), let $f = \omega(\zeta)$ and $f = o(\zeta \log \zeta)$. First we prove that a.a.s.

$$\tau_{\text{PGr}} \leq n + \zeta \log \zeta + 2f.$$

Set $m = n + \zeta \log \zeta + 2f$. By Lemma 22, a.a.s. the first $n + f$ columns of A_m have rank n . Following these, there are $b = \zeta \log \zeta + f$ columns. After a change of basis, we obtain the following matrix that is row equivalent to A_m :

$$[I_{n \times n} \quad * \quad B],$$

where $I_{n \times n}$ is the n by n identity matrix, $*$ is a set of f columns, and B is obtained from the above b columns after the change of basis. By Lemma 28, $B \sim [U_q]^{n \times b}$.

Deleting the f columns in $*$ and contracting all but the first r columns in $I_{n \times n}$ we obtain

$$[I_{r \times r} \quad B_r],$$

where B_r is the $r \times b$ matrix obtained from the first r rows of B . Hence, $B_r \sim [U_q]^{r \times b}$. By definition $M[I_{r \times r} \quad B_r]$ is a minor of A_m . It is thus sufficient to prove that B_r contains all elements in $PG(r-1, q)$. (It suffices to prove that B_r contains all elements other than those already contained in $I_{r \times r}$. However it does not change the bound in any significant way.)

Consider each element of $PG(r-1, q)$ as a bin and consider each column of B_r as a ball. We say a ball j is thrown into a bin z if the j -th column vector of B_r corresponds to one of the $q-1$ vectors associated to the z -th bin. Hence, B_r contains all elements in $PG(r-1, q)$ if and only if every bin receives at least one ball. A ball here corresponds to a nonzero column vector, and it is easy to show that a.a.s. at most $f/2$ of the b columns can be zero columns. Hence, the total number of balls is at least $b - f/2$, and the total number of bins is equal to ζ . Setting $\lambda = (b - f/2)/\zeta = \log \zeta + f/2\zeta$ and by Lemma 27,

$$\mathbb{P}(\tau_{\text{PGr}} > m) \leq 2\zeta e^{-\lambda} = O(q^r e^{-\log \zeta - f/2\zeta}) = O(\exp(-f/2\zeta + O(1))) = o(1),$$

as $\log \zeta = r \log q + O(1)$.

For part (b), the upper bound immediately follows from part (a). For the lower bound, fix $\varepsilon > 0$ and we prove that if $r = \omega(\log n)$ then a.a.s. $\tau_{\text{PGr}} \geq (1-\varepsilon)\zeta \log \zeta$. Set $m = (1-\varepsilon)\zeta \log \zeta$. By Lemma 22, we may assume that A_m has rank n . For each $J \subseteq [m]$ where $|J| = n - r$, let X_J be the indicator variable that A_J has rank $n - r$, and the contraction of columns in J produces a matroid that contains $PG(r-1, q)$ as a minor. Let $X = \sum_J X_J$ over all such subsets J . We claim that for every J ,

$$\mathbb{E}X_J \leq 2 \exp(-\zeta^\varepsilon). \tag{3}$$

Then,

$$\mathbb{E}X = 2 \binom{m}{n-r} \exp(-\zeta^\varepsilon) \leq \exp(n \log m - \zeta^\varepsilon) = \exp(-\zeta^\varepsilon + O(nr)) = o(1),$$

where the last equation above holds as $r = \omega(\log n)$. The lower bound for (b) follows by the Markov inequality. It only remains to prove (3).

Proof of (3). Similarly as before, the rank of A_m is a.a.s. n and the contraction of columns in J where $|J| = n - r$ produces a matrix $[I_{r \times r} \ B]$, where each column of B is a uniform random vector in \mathbb{F}_q^r , provided that A_J has rank $n - r$. Moreover, B has $b = m - (n - r) \leq (1 - \varepsilon)\zeta \log \zeta$ columns. By Lemma 27 (with $\lambda = b/\zeta \leq (1 - \varepsilon) \log \zeta$),

$$\mathbb{P}(X_J = 1) \leq 2 \exp(-\zeta e^{-(1-\varepsilon) \log \zeta}) = 2 \exp(-\zeta^\varepsilon).$$

Finally, for part (c), the upper bound is again implied by part (a), noticing that $\zeta \log \zeta = o(n)$ for the range of r in part (c), and the fact that $PG(r - 1, q)$ contains every \mathbb{F}_q -representable minors of rank r . The lower bound follows since a.a.s. A_m is a free matroid if $m - n \rightarrow -\infty$ and thus a.a.s. $\tau_{\text{N-minor}} \geq n - \omega(1)$. ■

4 Circuits

Proof of Corollary 11. For (a), set $m = cq^{n/k}$ where $c > 0$ is fixed. Then by Theorem 9(a),

$$\Pr(M[A_m] \text{ has no } k\text{-circuits}) \sim \exp\left(-\frac{(q-1)^{k-1}c^k}{k!}\right).$$

Moreover, the above probability tends to 1 if $c \rightarrow 0$, and tends to 0 if $c \rightarrow \infty$. Therefore, $\tau_{\text{k-circ}} = \Theta_p(q^{n/k})$.

For (b), assume that $k, m \rightarrow \infty$ and $k = o(m)$. Let μ_k be as defined in Theorem 9. Then,

$$\log \mu_k = \log \binom{m}{k} + k \log(q-1) - n \log q, \quad (4)$$

where, by Stirling's formula,

$$\begin{aligned} \log \binom{m}{k} &= m \log m - k \log k - (m-k) \log(m-k) + o(1) \\ &= m \log m - k(\log m + \log k/m) - (m-k)(\log m + \log(1 - k/m)) + o(1) \\ &= k \log \left(\frac{m}{k}\right) + (m-k) \frac{k}{m} + O(k^2/m) = k \log \left(\frac{em}{k}\right) + o(k). \end{aligned} \quad (5)$$

Fix $c > 0$. Setting $m = c \frac{1}{e(q-1)} k q^{n/k}$, we find that $k = o(m)$ and thus by (4) and (5) we obtain

$$\log \mu_k = k \log(cq^{n/k}) + o(k) - n \log q = k \log c + o(k).$$

It follows now that $\mu_k = o(1)$ if $c < 1$ and $\mu_k = \omega(1)$ if $c > 1$. Thus, part (b) follows by Theorem 9(b).

For part (c), we need the following claim about the function $g_a(y)$.

Claim 29. *For every $0 < a \leq 1$, there exists a unique b such that $g_a(b) = 0$. Moreover, $g'_a(b) > 0$.*

Let b be the unique root of $g_a(y)$. Set $m = cbn$ for some fixed $c > 0$. By Stirling's formula and a similar calculation as before, provided that $cb > a$,

$$\begin{aligned} \log \mu_k &= -k \log(k/m) - (m-k) \log(1 - k/m) + k \log(q-1) - n \log q + O(\log n) \\ &\sim n(-a \log a + cb \log cb - (cb-a) \log(cb-a) + a \log(q-1) - \log q) = ng_a(cb). \end{aligned}$$

By the definition of b and Claim 29, $\mu_k = o(1)$ if $c < 1$ and $\mu_k = \omega(1)$ if $c > 1$. Part (c) follows. It only remains to prove Claim 29 and verify that $b > a$ (hence $cb > a$ for $c = 1 \pm \varepsilon$ for every sufficiently small $\varepsilon > 0$).

Proof of $b > a$. This follows immediately from the facts that $g_a(y)$ is defined on $y \geq a$ and that $g_a(a) = a \log(q-1) - \log q < 0$.

Proof of Claim 29. We find that $g'_a(y) = \log y - \log(y-a) > 0$ for all $y > a$. Moreover, $\lim_{y \rightarrow \infty} g_a(y) = \infty$, and we have shown that $g_a(a) < 0$. It follows that $g_a(y)$ has a unique root $b > a$. ■

Proof of Corollary 13. Part (b) follows by Remark 12 that $b(1) < 2$ (the proof of Remark 12 is given in the Appendix). For part (a), let $a^* = 1 - q^{-1}$ and let $b = b(a^*) = 1$. By Remark 12 and Corollary 11(c), a.a.s. at some step $m = (1 + o(1))n$, $M[A_m]$ has a circuit whose length is asymptotic to a^*n . It remains to show that a.a.s. the first circuit cannot have length that is not asymptotic to a^*n . Fix $\varepsilon > 0$. We prove that there exists $c = c(\varepsilon) > 0$ such that a.a.s. there is no circuit of length greater than $(a^* + \varepsilon)n$ or shorter than $(a^* - \varepsilon)n$ by step $m = (1 + c)n$. Note that the expected number of circuits of length k in $M[A_m]$ is asymptotic to μ_k given in Theorem 9(b). For every $k \geq (a^* + \varepsilon)n$ or $k \leq (a^* - \varepsilon)n$, let $a_k = \lim_{n \rightarrow \infty} k/n$ (without loss of generality we may assume that a_k exists by the subsubsequence principle) and let b_k be the unique root of $g_{a_k}(y)$. By the condition on k and since b is strictly convex by Remark 12, it follows that $b_k > b(a^*) + \delta = 1 + \delta$ for some fixed $\delta = \delta(\varepsilon) > 0$. Let $m = (1 + \delta/2)n = c'b_k n$ for some $c' \leq 1 - \delta/4$. We have shown in the previous proof that $\log \mu_k \sim n g_{a_k}(c'b_k) < -c''n$ for some fixed $c'' = c''(\varepsilon) > 0$, as $c' < 1$. Hence, the probability that $M[A_m]$ has a circuit of length greater than $(a^* + \varepsilon)n$ or shorter than $(a^* - \varepsilon)n$ is at most $ne^{-c''n} = o(1)$ by the union bound (over all k such that $k \geq (a^* + \varepsilon)n$ and $k \leq (a^* - \varepsilon)n$) and the Markov inequality. ■

5 Connectivity

Lemma 30 (Proposition 3 of [5]). *If $M = (E, \mathcal{I})$ contains a vertically k -connected submatroid with the same rank as M , then M is vertically k -connected.*

Proof. Let $e \in E$ such that $M \setminus e$ has the same rank as M . That is, $\text{rk}(E \setminus \{e\}) = \text{rk}(E)$. It suffices to show that if M is not vertically k -connected, then neither is $M \setminus e$. Therefore, suppose M has a vertical ℓ -separation (X, Y) for some $\ell < k$. That is, $\text{rk}(X) + \text{rk}(Y) \leq \text{rk}(E) + \ell - 1$ and $\text{rk}(X), \text{rk}(Y) \geq \ell$. Without loss of generality we may assume that $e \in X$. Let $X' = X \setminus \{e\}$. If $\text{rk}(X') = \text{rk}(X)$, then (X', Y) is a vertical ℓ -separation of $M \setminus e$. If $\text{rk}(X') = \text{rk}(X) - 1$, then (X', Y) is a vertical $(\ell - 1)$ -separation of $M \setminus e$. In either case, $M \setminus e$ is not vertically k -connected. ■

Proof of Theorem 16(a). Let $f \rightarrow \infty$ be a slowly growing function of n . Obviously, a.a.s. $\kappa(M[A_m]) = 1$ for steps $m \leq n - f$, as $M[A_m]$ is a free matroid. We leave it as an easy exercise that a.a.s. $\kappa(M[A_m])$ remains one for all $n - f \leq m \leq n + f$. Finally, we know that a.a.s. the rank of A_m is n for $m = n + f$. Combining all, a.a.s. $\kappa(M[A_m]) = 1$ for all $m \leq n + f$, and $\kappa(M[A_m])$ is non-decreasing for all $m \geq n + f$ by Lemma 30. ■

Proof of Theorem 19. We first prove the upper bound in part (a) by extending the proof of Kelly and Oxley [8, Theorem 4.5]. Fix $\varepsilon > 0$. Set $m = n + (1 + 2\varepsilon)k \log_q(n/k)$. Let

$f = \omega(1)$ be a slowly growing function of n . Let D be the first $n - f$ columns of A_m . Let \mathcal{E}_ℓ be the event that $M[A_m]$ is vertically ℓ -separated, $\text{rk}(A_m) = n$, and all columns of D are independent. Since a.a.s. all columns of D are linearly independent and $\text{rk}(A_m) = n$, we have

$$\mathbb{P}(M[A_m] \text{ is not vertically } k\text{-connected}) \leq o(1) + \sum_{\ell \leq k-1} \mathbb{P}(\mathcal{E}_\ell). \quad (6)$$

The probability of \mathcal{E}_ℓ was upper bounded by Kelly and Oxley in the following lemma.

Lemma 31. ([8, Lemma 4.7]) $\mathbb{P}(\mathcal{E}_\ell) \leq \sum_{j=\ell}^{\lfloor \frac{1}{2}(n+\ell-1) \rfloor} b(\ell, j)$ where

$$\binom{m - |D|}{n + \ell - 1 - |D|} \binom{n + \ell - 1}{j} \left(\frac{q^j + q^{n+\ell-1-j} - q^{\ell-1}}{q^n} \right)^{m-(n+\ell-1)}.$$

We will prove the following claim.

Claim 32. $\mathbb{P}(M[A_m] \text{ is not vertically } k\text{-connected}) \leq O(nk \cdot b(k-1, k-1)) + o(1)$.

Proof of Claim 32. By (6) and Lemma 31, it suffices to show that $b(\ell, j)$ is maximized at $(\ell, j) = (k-1, k-1)$. Fix $1 \leq \ell \leq k-1$ and $\ell \leq j < \lfloor \frac{1}{2}(n+\ell-1) \rfloor$. Then

$$\frac{b(\ell, j)}{b(\ell, j+1)} = \frac{j+1}{n+\ell-1-j} \left(\frac{q^j + q^{n+\ell-1-j} - q^{\ell-1}}{q^{j+1} + q^{n+\ell-2-j} - q^{\ell-1}} \right)^{m-(n+\ell-1)}.$$

Let A be the fraction inside the parentheses above, and let $B = (q^2 + 1)/2q$. We first prove that $A \geq B$. Note that $A \geq B$ if and only if

$$2q^{j+1} + 2q^{n+\ell-j} - 2q^\ell \geq q^{j+3} + q^{n+\ell-j} - q^{\ell+1} + q^{j+1} + q^{n+\ell-2-j} - q^{\ell-1},$$

which holds if and only if

$$(q^2 - 1)q^{n+\ell-2-j} + q^{\ell-1}(q-1)^2 \geq (q^2 - 1)q^{j+1}.$$

Since $j < (n+\ell-1)/2$, we have $n+\ell-2-j \geq j+1$. This verifies that $A \geq B$. Therefore,

$$\frac{b(\ell, j)}{b(\ell, j+1)} \geq \frac{A^{m-(n+\ell-1)}}{n} \geq \frac{B^{m-n-k}}{n} \geq 1,$$

where the last inequality holds by the definition of B and by the setting of m . Thus, $b(\ell, j)$ is decreasing in j . Next, write $b(\ell, \ell) = XYZ$, where

$$X = \binom{m-n}{\ell-1}, \quad Y = \binom{n+\ell-1}{n-1}, \quad Z = \left(\frac{q^\ell + q^{n-1} - q^{\ell-1}}{q^n} \right)^{m-n-\ell+1}.$$

Then Y is increasing in ℓ . Since $2\ell < 2k \leq m-n$, so is X . Also, Z is increasing in ℓ since its base is less than one and increasing in ℓ , and its exponent is decreasing in ℓ . Therefore, $b(\ell, \ell)$ is increasing in ℓ , so $b(\ell, j) \leq b(\ell, \ell) \leq b(k-1, k-1)$, as required. ■

Observe that

$$\begin{aligned} b(k-1, k-1) &= \binom{m-n+f}{k-2+f} \binom{n+k-2}{k-1} \left(\frac{q^{k-1} + q^{n-1} - q^{k-2}}{q^n} \right)^{m-n-k+2} \\ &\leq \left(\frac{4k \log_q(n/k)}{k} \right)^{k+f} \left(\frac{4n}{k} \right)^k \left(\frac{1+q^{k-n}}{q} \right)^{(1+\varepsilon)k \log_q(n/k)}, \end{aligned}$$

where to derive the last inequality above we used the fact that $m-n-k+2 = (1+2\varepsilon)k \log_q(n/k) - k + 2 \geq (1+\varepsilon)k \log_q(n/k)$ since $k = o(n)$. Therefore, using that $n-k \rightarrow \infty$, we have

$$\begin{aligned} &\log [nk \cdot b(k-1, k-1)] \\ &\leq \log(nk) + (k+f) \log \log(n/k) + k \log(n/k) + O(k) + (1+\varepsilon)k \log_q(n/k)(-\log q + O(q^{k-n})) \\ &= -\varepsilon k \log(n/k) + O(\log n + k \log \log(n/k)) \end{aligned}$$

which goes to $-\infty$ since $k = o(n)$ and $k = \omega(1)$. Thus, $nk \cdot b(k-1, k-1)$ tends to zero. By Claim 32, a.a.s. $M[A_m]$ is vertically k -connected.

Next we prove the lower bounds in part (a) and (b) by the second moment method. A straight application of the second moment method to all possible separations would lead to failure due to heavy correlations. Instead we carefully craft the counting structures that imply the existence of a certain type of vertical $(k-1)$ -separations. For a pair (I, S) , where I is a subset of $k-1$ columns of A_m and S is an $(n-1)$ -dimensional subspace of \mathbb{F}_q^n , define $X_{I,S}$ to be the indicator variable that

- (i) $I \subseteq S^c$, and all columns of I are linearly independent, and
- (ii) All column vectors in $A_m \setminus I$ are in S , and $\text{rk}(A_m \setminus I) \geq k-1$.

Let I^c denote $A_m \setminus I$, the set of vectors not in I . Note that (ii) above implies that $k-1 \leq \text{rk}(I^c) \leq n-1$. Thus, $X_{I,S} = 1$ for some $(n-1)$ -dimensional subspace S immediately implies that (I, I^c) is an $(k-1)$ -separation. Therefore, it suffices to show that $X := \sum X_{I,S}$ is a.a.s. positive, where the summation is taken over all $(k-1)$ -subset of columns of A_m and all $(n-1)$ -dimensional subspaces of \mathbb{F}_q^n .

For each given (I, S) , the events in (i) and (ii) are independent. Let v_1, \dots, v_{k-1} be the column vectors in I . For each $1 \leq i \leq k-1$, the probability that $v_i \in \langle v_1, \dots, v_{i-1} \rangle \cup S$ conditional on that v_1, \dots, v_{i-1} are linearly independent, and that none of them are in S is

$$\begin{cases} q^{-n}(q^{i-1} + q^{n-1} - q^{i-2}) & \text{if } i \geq 2 \\ q^{-1} & \text{if } i = 1. \end{cases}$$

Thus, the probability of event (i) is (provided that $k/n < 1 - \varepsilon$ for some $\varepsilon > 0$)

$$(1 - q^{-1}) \prod_{i=2}^{k-1} (1 - q^{-n}(q^{i-1} + q^{n-1} - q^{i-2})) \sim (1 - q^{-1})^{k-1}.$$

Similarly, the probability of events (ii) is asymptotic to $q^{-(m-k+1)}$. Consequently,

$$\mathbb{E}X_{I,S} \sim \mu := (q-1)^{k-1}q^{-m} \quad \text{and} \quad \mathbb{E}X \sim \binom{m}{k-1} \begin{bmatrix} n \\ 1 \end{bmatrix}_q \mu.$$

For part (a), set $m = n + (1 - \varepsilon)k \log_q(n/k)$. For part (b), set $m = (1 + \alpha)n$ where α satisfies (2). We first verify that $\log \mathbb{E}X = \omega(1)$ in both parts. Suppose that $k = o(n)$ and $k = \omega(1)$. Then, $m = n + (1 - \varepsilon)k \log_q(n/k)$ and hence,

$$\begin{aligned} \log \mathbb{E}X &= k \log(n/k) + n \log q + (k - 1) \log(q - 1) - m \log q + O(k + \log n) \\ &= \varepsilon k \log(n/k) + O(k + \log n) \rightarrow \infty, \end{aligned}$$

implying that $\mathbb{E}X = \omega(1)$. Suppose that $k \sim tn$ for some $0 < t < 1$. Then, by (2),

$$\log \mathbb{E}X = \left(t \log \frac{1 + \alpha}{t} + (1 + \alpha - t) \log \frac{1 + \alpha}{1 + \alpha - t} + t \log(q - 1) - \alpha \log q + o(1) \right) n \rightarrow \infty.$$

Next, we prove that $\mathbb{E}X(X - 1) \leq (1 + o(1))(\mathbb{E}X)^2$. Consider a pair $(X_{I_i, S_i}, X_{I_j, S_j})$. Let $h = |I_i \cap I_j|$. Note that if $X_{I_i, S_i} X_{I_j, S_j} = 1$ then $I_i \subseteq S_i^c$, $I_j \subseteq S_j^c$, $I_i^c \subseteq S_i$, $I_j^c \subseteq S_j$. It follows that $I_i \cap I_j \subseteq (S_i \cup S_j)^c$, $I_i \setminus I_j \subseteq S_j \setminus S_i$, $I_j \setminus I_i \subseteq S_i \setminus S_j$ and $(I_i \cup I_j)^c \subseteq S_i \cap S_j$.

We further consider two cases. If $S_i \neq S_j$, then $\dim(S_i \cap S_j) = n - 2$, and so $|S_i \cap S_j| = q^{n-2}$. Thus,

$$\begin{aligned} \mathbb{E}X_{I_i, S_i} X_{I_j, S_j} &\leq \mathbb{P}(I_i \cap I_j \subseteq (S_i \cup S_j)^c) \mathbb{P}(I_i \setminus I_j \subseteq S_j \setminus S_i) \mathbb{P}(I_j \setminus I_i \subseteq S_i \setminus S_j) \mathbb{P}((I_i \cup I_j)^c \subseteq S_i \cap S_j) \\ &= \left(\frac{|(S_i \cup S_j)^c|}{q^n} \right)^{|I_i \cap I_j|} \left(\frac{|S_j \setminus S_i|}{q^n} \right)^{|I_i \setminus I_j|} \left(\frac{|S_i \setminus S_j|}{q^n} \right)^{|I_j \setminus I_i|} \left(\frac{|S_i \cap S_j|}{q^n} \right)^{|(I_i \cup I_j)^c|} \\ &= (1 - 2q^{-1} + q^{-2})^h (q^{-1} - q^{-2})^{k-1-h} (q^{-1} - q^{-2})^{k-1-h} (q^{-2})^{m-2(k-1-h)-h} \\ &= \left(\frac{q-1}{q} \right)^{2h} \left(\frac{q-1}{q^2} \right)^{k-1-h} \left(\frac{q-1}{q^2} \right)^{k-1-h} \left(\frac{1}{q^2} \right)^{m-2(k-1-h)-h} = \mu^2. \end{aligned}$$

On the other hand, if $S_i = S_j$, then $X_{I_i, S_i} X_{I_j, S_j} = 1$ implies that $I_i \cap I_j \subseteq S_i^c$, $I_i \setminus I_j \subseteq \emptyset$, $I_j \setminus I_i \subseteq \emptyset$, $(I_i \cup I_j)^c \subseteq S_i$, which implies that $I_i = I_j$ and $X_{I_i, S_i} = 1$. Therefore, $\mathbb{E}X_i X_j$ is nonzero only if $i = j$, in which case, $\mathbb{E}X_{I_i, S_i} X_{I_j, S_j} = \mu$. Thus,

$$\begin{aligned} \mathbb{E}X^2 &= \sum_{(I_i, S_i), (I_j, S_j)} X_{I_i, S_i} X_{I_j, S_j} = \sum_{I_i, I_j, S_i \neq S_j} \mathbb{E}X_{I_i, S_i} X_{I_j, S_j} + \sum_{I_i = I_j, S_i = S_j} \mathbb{E}X_{I_i, S_i} X_{I_j, S_j} \\ &= \binom{m}{k-1} \begin{bmatrix} n \\ 1 \end{bmatrix} \binom{m}{k-1} \left(\begin{bmatrix} n \\ 1 \end{bmatrix} - 1 \right) \mu^2 + \mathbb{E}X \sim (\mathbb{E}X)^2. \end{aligned}$$

By Chebyshev's inequality, a.a.s. $X > 0$, and therefore there exists a vertical $(k - 1)$ -separation. Now Theorem 19 follows by the definition of vertical connectivity and Theorem 16(a). ■

Proof of Theorem 16(b,c). By Theorem 19, $\kappa(M[A_m]) = o(n)$ if $m = n + o(n)$. On the other hand, $\text{gir}(M[A_m]) = \Theta(n)$ for all $m = n + o(n)$, by Corollary 11. It is easy to see that a.a.s. for every step m after the creation of the first circuit, A_m is not isomorphic to any uniform matroid. Thus, part (b) follows now by Proposition 15.

Part (c) follows by Proposition 15, and the fact that a.a.s. $\kappa(M[A_m])$ is monotonely non-decreasing by part (a), and that $\text{gir}(M[A_m])$ is monotonely non-increasing. ■

6 Critical number

The following two lemmas follow from well known results in counting subspaces of a vector space (see e.g. page 162 of [21]). We include proofs for self-containment. Recall that we write $a_n \asymp b_n$ if $a_n = \Theta(b_n)$.

Lemma 33. For any $0 \leq k \leq n$,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q \asymp q^{k(n-k)},$$

which is the number of k -dimensional (and $(n-k)$ -dimensional) subspaces of \mathbb{F}_q^n .

Proof. By definition,

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q = \prod_{i=0}^{k-1} \frac{q^{n-i} - 1}{q^{k-i} - 1} \asymp \prod_{i=0}^{k-1} \frac{q^{n-i}}{q^{k-i}} = q^{k(n-k)}. \blacksquare$$

Lemma 34. Let S be any k -dimensional subspace of \mathbb{F}_q^n , and let $N(n, k, j, \ell)$ be the number of j -dimensional subspaces of \mathbb{F}_q^n that intersect with S in dimension ℓ . Then

$$N(n, k, j, \ell) = q^{(k-\ell)(j-\ell)} \begin{bmatrix} k \\ \ell \end{bmatrix}_q \begin{bmatrix} n-k \\ j-\ell \end{bmatrix}_q.$$

Proof. There are $\begin{bmatrix} k \\ \ell \end{bmatrix}_q$ ways to choose an ℓ -dimensional subspace U of S .

Given U , there are $q^n - q^k$ vectors that are not in S . Adding any such vector into U results in an extension of U into an $(\ell+1)$ -dimensional subspace U_1 such that $S \cap U_1 = U$. Then, there are $q^n - q^{k+1}$ vectors that are not in $S + U_1$ whose addition extends U_1 to an $(\ell+2)$ -dimensional subspace U_2 . Repeat this, and we find that there are

$$\prod_{i=0}^{j-\ell-1} (q^n - q^{k+i})$$

ways to extend U to a j -dimensional subspace $W = U_{j-\ell}$ whose intersection with S is U . However, by the same counting scheme (by considering vectors in $W \setminus U_i$, $0 \leq i \leq j-\ell-1$ with $U_0 = U$ this time), each such j -dimensional subspace W can be constructed in $(q^j - q^\ell)(q^j - q^{\ell+1}) \cdots (q^j - q^{j-1})$ ways. It follows now that

$$N(n, k, j, \ell) = \begin{bmatrix} k \\ \ell \end{bmatrix}_q \prod_{i=0}^{j-\ell-1} \frac{q^n - q^{k+i}}{q^j - q^{\ell+i}} = q^{(k-\ell)(j-\ell)} \begin{bmatrix} k \\ \ell \end{bmatrix}_q \begin{bmatrix} n-k \\ j-\ell \end{bmatrix}_q. \blacksquare$$

By Lemma 34,

$$N(n, n-k, n-k, n-2k+h) = q^{(k-h)^2} \begin{bmatrix} n-k \\ n-2k+h \end{bmatrix}_q \begin{bmatrix} k \\ k-h \end{bmatrix}_q = q^{(k-h)^2} \begin{bmatrix} n-k \\ k-h \end{bmatrix}_q \begin{bmatrix} k \\ h \end{bmatrix}_q. \quad (7)$$

By Lemma 33,

$$N(n, n-k, n-k, n-2k+h) \asymp q^{(k-h)^2} q^{(k-h)(n-2k+h)} q^{h(k-h)} = q^{(n-k+h)(k-h)}.$$

Lemma 35. Let N, M, k be positive integers such that $N \geq M \geq k$. Then

$$\begin{bmatrix} N \\ k \end{bmatrix}_q = q^{(N-M)k} \begin{bmatrix} M \\ k \end{bmatrix}_q \left(1 + O(q^{k-M})\right).$$

Proof.

$$\begin{aligned}
\begin{bmatrix} N \\ k \end{bmatrix}_q &= \frac{(q^N - 1)(q^{N-1} - 1) \cdots (q^{N-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)} = \frac{q^N q^{N-1} \cdots q^{N-k+1}}{(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)} \left(1 + O(q^{k-N})\right) \\
&= q^{(N-M)k} \frac{q^M q^{M-1} \cdots q^{M-k+1}}{(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)} \left(1 + O(q^{k-N})\right) \\
&= q^{(N-M)k} \frac{(q^M - 1)(q^{M-1} - 1) \cdots (q^{M-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \cdots (q - 1)} \left(1 + O(q^{k-N} + q^{k-M})\right) \\
&= q^{(N-M)k} \begin{bmatrix} M \\ k \end{bmatrix}_q \left(1 + O(q^{k-M})\right). \blacksquare
\end{aligned}$$

Proof of Theorem 20. Fix $0 < \varepsilon < 1$. Let S_1, \dots, S_r , where $r = \begin{bmatrix} n \\ k \end{bmatrix}_q$, be the set of $(n - k)$ -dimensional subspaces of \mathbb{F}_q^n . Let E be the set of column vectors in A_m . For every $1 \leq i \leq r$, let X_i be the indicator variable for $S_i \cap E = \emptyset$, and let $X = \sum_{i=1}^r X_i$. Hence $M[A_m]$ is k -colourable if and only if $X > 0$. It is easy to see that

$$\mathbb{E}X_i = \mu := (1 - q^{-k})^m \quad \text{and} \quad \mathbb{E}X = \begin{bmatrix} n \\ k \end{bmatrix}_q \mu.$$

By Lemma 33,

$$\log \mathbb{E}X = k(n - k) \log q + m \log(1 - q^{-k}),$$

which goes to ∞ if $m = -(1 - \varepsilon)k(n - k) \log q / \log(1 - q^{-k})$, and goes to $-\infty$ if $m = -(1 + \varepsilon)k(n - k) \log q / \log(1 - q^{-k})$.

It suffices now to show that $\mathbb{E}X(X - 1) \sim (\mathbb{E}X)^2$ if $m = -(1 - \varepsilon)k(n - k) \log q / \log(1 - q^{-k})$. Consider a pair of distinct $(n - k)$ -dimensional subspaces (S_i, S_j) . Let $h = \dim(S_i \cap S_j) - (n - 2k)$. Thus, $\max\{0, 2k - n\} \leq h \leq k - 1$. Note that $X_i X_j = 1$ if and only if $(S_i \cup S_j) \cap E = \emptyset$, and

$$|S_i \cup S_j| = |S_i| + |S_j| - |S_i \cap S_j| = 2q^{n-k} - q^{n-2k+h}.$$

Therefore,

$$\mathbb{E}X_i X_j = \left(1 - \frac{|S_i \cup S_j|}{q^n}\right)^m = \pi_h := (1 - 2q^{-k} + q^{-2k+h})^m.$$

Notice that $\pi_0 = \mu^2$. For $0 \leq h \leq k$, let N_h denote the number of $(n - k)$ -dimensional subspaces whose intersection with S_i is $n - 2k + h$ (notice that this number is independent of S_i). Then

$$\mathbb{E}X(X - 1) = \sum_{i,j} \mathbb{E}X_i X_j = \sum_{i=1}^r \sum_{h=\max\{0, 2k-n\}}^{k-1} \sum_{\substack{\dim(S_i \cap S_j) \\ = n-2k+h}} \mathbb{E}X_i X_j = \begin{bmatrix} n \\ k \end{bmatrix}_q \sum_{h=\max\{0, 2k-n\}}^{k-1} N_h \pi_h.$$

Claim 36. Suppose that $(q, k) \neq (2, 1)$. Let $\Lambda = 2 \log n$. If $k \leq n/2 - \Lambda$ then $\sum_{h=0}^{k-1} N_h \pi_h \sim N_0 \pi_0$. If $k \geq n/2 + \Lambda$ then $\sum_{h=2k-n}^{k-1} N_h \pi_h \sim N_{2k-n} \pi_{2k-n}$. If $|k - n/2| \leq \Lambda$ then $\sum_{h=\max\{0, 2k-n\}}^{k-1} N_h \pi_h \sim \sum_{h=\max\{0, 2k-n\}}^{2\Lambda} N_h \mu^2$.

We first consider the case that $(q, k) \neq (2, 1)$. By (7) and Lemma 35,

$$\begin{aligned}
N_0 &= N(n, n - k, n - k, n - 2k) = q^{k^2} \begin{bmatrix} n - k \\ k \end{bmatrix}_q \sim \begin{bmatrix} n \\ k \end{bmatrix}_q \quad \text{if } n/2 - k = \omega(1) \\
N_{2k-n} &= N(n, n - k, n - k, 0) = q^{(n-k)^2} \begin{bmatrix} k \\ n - k \end{bmatrix}_q \sim \begin{bmatrix} n \\ n - k \end{bmatrix}_q \quad \text{if } k - n/2 = \omega(1).
\end{aligned}$$

Thus, if $n/2 - k \geq \log n$ then

$$\mathbb{E}X^2 \sim \left[\begin{matrix} n \\ k \end{matrix} \right]_q N_0 \pi_0 \sim \left[\begin{matrix} n \\ k \end{matrix} \right]_q^2 \mu^2 = (\mathbb{E}X)^2.$$

On the other hand, if $k - n/2 \leq \log n$ and $n - k \geq \log_q n + \log_q \log n + \omega(1)$ then

$$\begin{aligned} \mathbb{E}X(X-1) &\sim \left[\begin{matrix} n \\ k \end{matrix} \right]_q N_{2k-n} \pi_{2k-n} \sim \left[\begin{matrix} n \\ k \end{matrix} \right]_q^2 \left(1 - 2q^{-k} + q^{-n}\right)^m \\ &= \left[\begin{matrix} n \\ k \end{matrix} \right]_q^2 \left((1 - q^{-k})^2 + O(q^{-n})\right)^m = (\mathbb{E}X)^2 (1 + O(mq^{-n})) \sim (\mathbb{E}X)^2. \end{aligned}$$

We know that

$$\left[\begin{matrix} n \\ k \end{matrix} \right]_q = \sum_{h=\max\{0, 2k-n\}}^k N_h.$$

However,

$$\frac{N_h}{N_{h-1}} \asymp q^{-2h+2k-n+1} \leq 1/q \quad \text{for all } \max\{0, 2k-n\} + 1 \leq h \leq k.$$

It follows that

$$\left[\begin{matrix} n \\ k \end{matrix} \right]_q \sim \sum_{h=\max\{0, 2k-n\}}^{\log n} N_h.$$

Thus, $\mathbb{E}X(X-1) \sim \left[\begin{matrix} n \\ k \end{matrix} \right]_q^2 \mu^2 = (\mathbb{E}X)^2$. The theorem for the case that $(q, k) \neq (2, 1)$ follows by combining all the three ranges of k .

In the case $(q, k) = (2, 1)$ the critical number jumps from one to two when the first even circuit appears, which occurs in some step $n + O_p(1) \sim n$. Hence the theorem holds for the case $(q, k) = (2, 1)$ as well. ■

The proof of Claim 36 uses the following inequality.

Lemma 37. $(1 - q^{-k})^2 > kq^{-k}$ for all positive integers $q \geq 2$ and $k \geq 1$ except that $(q, k) = (2, 1)$.

Proof. Let $f_q(x) = (1 - q^{-x})^2 - xq^{-x}$ for $x \geq 1$. We find that $f'_q(x) = q^{-x}((2 - 2q^{-x} + x) \log q - 1) > 0$. Moreover, $f_q(1) = 1 - 3q^{-1} + q^{-2} > 0$ except that $q = 2$, and $f_q(2) = 1 - 4q^{-2} + q^{-4} > 0$ for all $q \geq 2$. The first assertion follows. Similarly letting $g(x) = (1 - 2^{-k})^2 \log 2 - k2^{-2k+1}$ the second assertion follows by $g'(x) = 2^{-2k+1}(2k \log 2 - 1 + 2^k - (\log 2)^2) > 0$ and $g(2) > 0$. ■

Proof of Claim 36. Note that

$$\pi_h \leq ((1 - q^{-k})^2 + q^{-2k+h})^m \leq \mu^2 \exp \left(m \frac{q^{-2k+h}}{(1 - q^{-k})^2} \right), \quad \text{for all } \max\{0, 2k-n\} \leq h \leq k.$$

Note also that $\pi_h \sim \mu^2$ if $k - h \geq 4 \log n$.

$$\frac{N_h}{N_0} \asymp q^{-h(n-2k+h)} \quad \frac{N_h}{N_{2k-n}} \asymp q^{-h(n-2k+h)}.$$

It suffices to show that for every h ,

$$N_h \pi_h / N_0 \pi_0 = N_h \pi_h / N_0 \mu^2 = o(1/k). \tag{8}$$

We consider the following cases of k :

In the first case we consider $k = O(1)$. For every $1 \leq h \leq k$

$$\log \frac{N_h \pi_h}{N_0 \pi_0} = -hn \log q + m \frac{q^{-2k+h}}{(1-q^{-k})^2} + O(1) \leq -n \log q \left(h - k \frac{q^{-k}}{(1-q^{-k})^2} + o(1) \right),$$

where the above inequality used $-\log(1-q^{-k}) \geq q^{-k}$. Since $h \geq 1$, the above is $-\Theta(n)$ by Lemma 37. Thus, (8) holds when $k = O(1)$.

In the second case we consider $k = \omega(1)$ and $n/2 - k \geq \log n$:

$$\begin{aligned} \log \frac{N_h \pi_h}{N_0 \pi_0} &= -h(n-2k+h) \log q + m \frac{q^{-2k+h}}{(1-q^{-k})^2} + O(1) \\ &\leq -h(n-2k+h) \log q + (1+O(q^{-k}))k(n-k)q^{-k+h} \log q. \end{aligned} \quad (9)$$

It suffices to prove that

$$h(n-2k+h) > (1+\alpha)k(n-k)q^{-k+h} \quad (10)$$

for some constant $\alpha > 0$. Let $0 < \varepsilon < 1/8$. We further discuss two cases:

(a). $h/k \leq 1 - \varepsilon$. If $n-2k = \Theta(n)$. Then the left hand side of (10) is $\Theta(n)$, whereas the right hand side above is $o(n)$ since $k = \omega(1)$. Thus (10) holds. If $n-2k = o(n)$. Then, $k = \Theta(n)$. Hence the right hand side above is $o(1)$ whereas the left hand side is $\Theta(\log n)$. Thus (10) holds.

(b). $h/k \geq 1 - \varepsilon$. Let $c = k - h$. Then, $1 \leq c \leq \varepsilon k$. The inequality (10) is equivalent to

$$q^c(k(n-k) - c(n-c)) > (1+\alpha)k(n-k). \quad (11)$$

Since $c \leq \varepsilon k$ and $k \leq n/2$, $c(n-c) \leq 2\varepsilon k(n-k)$. Hence, the left hand side of (11) is at least $(1-2\varepsilon)q^c k(n-c) \geq (3/4)q^c k(n-k)$, which is greater than the right hand side as $q \geq 2$ and $c \geq 1$.

In the third case we consider $k - n/2 \geq \log n$. As $k = \omega(1)$ we have that (9) still holds after replacing $N_0 \pi_0$ by $N_{2k-n} \pi_{2k-n}$, and thus it suffices to prove (10), i.e.

$$q^{k-h}h(n-2k+h) > (1+\alpha)k(n-k) \quad \text{for all } 2k-n+1 \leq h \leq k-1. \quad (12)$$

Similarly as before, if $h/k \leq 1 - \varepsilon$ then we can easily verify (12). Suppose that $h \geq (1-\varepsilon)k$. Consider the derivative of $f(h) = q^{-h}h(n-2k+h)$ we find that $f'(h) = q^{-h}((\log q)h(2k-n-h) + 2h-2k+n)$. Taking the derivative of $(\log q)h(2k-n-h) + 2h-2k+n$ we obtain $(2k-n-2h) \log q + 2 < 0$ for all $(1-\varepsilon)k \leq h \leq k-1$. Hence, f is concave in the interval $(1-\varepsilon)k \leq h \leq k-1$. Thus, to verify (12) it suffices to show that $f(h) > k(n-k)$ for $h = (1-\varepsilon)k$ and for $h = k-1$. We have already argued that (12) holds for $h = (1-\varepsilon)k$. For $h = k-1$, the left hand side of (12) is $q(k-1)(n-k-1)$ which is clearly greater than the right hand side with sufficiently small α as $q \geq 2$, $k \geq n/2$ and $n-k = \omega(1)$.

In the final case let's consider the case that $k = n/2 + O(\Lambda)$. It is easy to see that for k in this range and for h between $\max\{0, 2k-n\} + 1$ and 2Λ , $\pi_h \sim \mu^2$. It is thus sufficient to prove that

$$\frac{N_h \pi_h}{N_* \pi_*} = o(1/n), \quad \text{where } * = \max\{0, 2k-n\}.$$

For simplicity we assume that $k \leq n/2$ and the case that $k \geq n/2$ is symmetric.

We have (9) and want to prove (10) for every $1 \leq h \leq 2\Lambda$. The left hand side is at least 1, and the right hand side is $o(1)$. So the inequality holds. ■

References

- [1] J. Altschuler and E. Yang. Inclusion of forbidden minors in random representable matroids. *Discrete Mathematics*, 340(7):1553–1563, 2017.
- [2] G. E. Andrews. *The theory of partitions*. Number 2. Cambridge university press, 1998.
- [3] N. Bansal, R. A. Pendavingh, and J. G. van der Pol. On the number of matroids. *Combinatorica*, 35:253–277, 2015.
- [4] C. Cooper, A. Frieze, and W. Pegden. Minors of a random binary matroid. *Random Structures & Algorithms*, 55(4):865–880, 2019.
- [5] W. H. Cunningham. On matroid connectivity. *Journal of Combinatorial Theory Series B*, 30(1):94–99, 1981.
- [6] P. Erdős and A. Rényi. On random graphs i. *Publ. math. debrecen*, 6(290-297):18, 1959.
- [7] P. Gao and P. Nelson. Minors of matroids represented by sparse random matrices over finite fields. *arXiv preprint arXiv:2307.15685*, 2023.
- [8] D. G. Kelly and J. matroid. On random representable matroids. *Studies in Applied Mathematics*, 71(3):181–205, 1984.
- [9] D. G. Kelly and J. G. matroid. Asymptotic properties of random subsets of projective spaces. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 91, pages 119–130. Cambridge University Press, 1982.
- [10] D. G. Kelly and J. G. matroid. Threshold functions for some properties of random subsets of projective spaces. *The Quarterly Journal of Mathematics*, 33(4):463–469, 1982.
- [11] D. E. Knuth. The asymptotic number of geometries. *Journal of Combinatorial Theory, Series A*, 16(3):398–400, 1974.
- [12] W. Kordecki. Strictly balanced submatroids in random subsets of projective geometries. In *Colloquium Mathematicum*, volume 2, pages 371–375, 1988.
- [13] W. Kordecki. Small submatroids in random matroids. *Combinatorics, Probability and Computing*, 5(3):257–266, 1996.
- [14] W. Kordecki and T. Łuczak. On random subsets of projective spaces. In *Colloquium Mathematicae*, volume 62, pages 353–356, 1991.
- [15] W. Kordecki and T. Łuczak. On the connectivity of random subsets of projective spaces. *Discrete mathematics*, 196(1-3):207–217, 1999.
- [16] L. Lowrance, J. matroid, C. Semple, and D. Welsh. On properties of almost all matroids. *Advances in Applied Mathematics*, 50(1):115–124, 2013.
- [17] D. Mayhew, M. Newman, D. Welsh, and G. Whittle. On the asymptotic proportion of connected matroids. *European Journal of Combinatorics*, 32(6):882–890, 2011.
- [18] M. Mitzenmacher and E. Upfal. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.

- [19] P. Nelson. Almost all matroids are non-representable. *arXiv preprint arXiv:1605.04288*, 2016.
- [20] J. G. Oxley. Threshold distribution functions for some random representable matroids. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 95, pages 335–347. Cambridge University Press, 1984.
- [21] J. G. Oxley. *Matroid theory, Second Edition*. Oxford University Press, USA, 2011.
- [22] R. Pendavingh and J. Van Der Pol. On the number of bases of almost all matroids. *Combinatorica*, 38:955–985, 2018.

Appendix

Proof of Remark 12. Since $g_a(b(a)) = 0$, we have

$$b(a) \log b(a) + a \log(q - 1) = a \log a + (b(a) - a) \log(b(a) - a) + \log q \quad \text{for every } 0 < a \leq 1. \quad (13)$$

Differentiating with respect to a gives

$$b'(\log b - \log(b - a)) = \log a - \log(q - 1) - \log(b - a). \quad (14)$$

Implicitly differentiating again gives

$$b''(\log b - \log(b - a))ab(b - a) = a^2(b')^2 - 2abb' + b^2 = (ab' - b)^2 \geq 0.$$

Thus, $b''(a) \geq 0$ for every $0 < a \leq 1$, and so b is convex. Moreover, we found that $(a, b) = (a^*, 1)$ where $a^* = 1 - 1/q$ satisfies (13), implying that $b(a^*) = 1$. Moreover, plugging $(a, b) = (a^*, 1)$ into (14) gives $b'(a^*) = 0$, which means that a^* minimizes $b(a)$. Lastly, observe that $g_1(2) > 0$ and $g_a(1/a) \rightarrow -\log q$ as $a \rightarrow 0$. Since for every fixed $a > 0$, $g_a(y)$ is an increasing function of y on $y \geq a$, it follows that $b(1) < 2$ and that for a sufficiently small, $b(a) > 1/a$. Hence $b(a) \rightarrow \infty$ as $a \rightarrow 0$. ■