

AD-A094 980

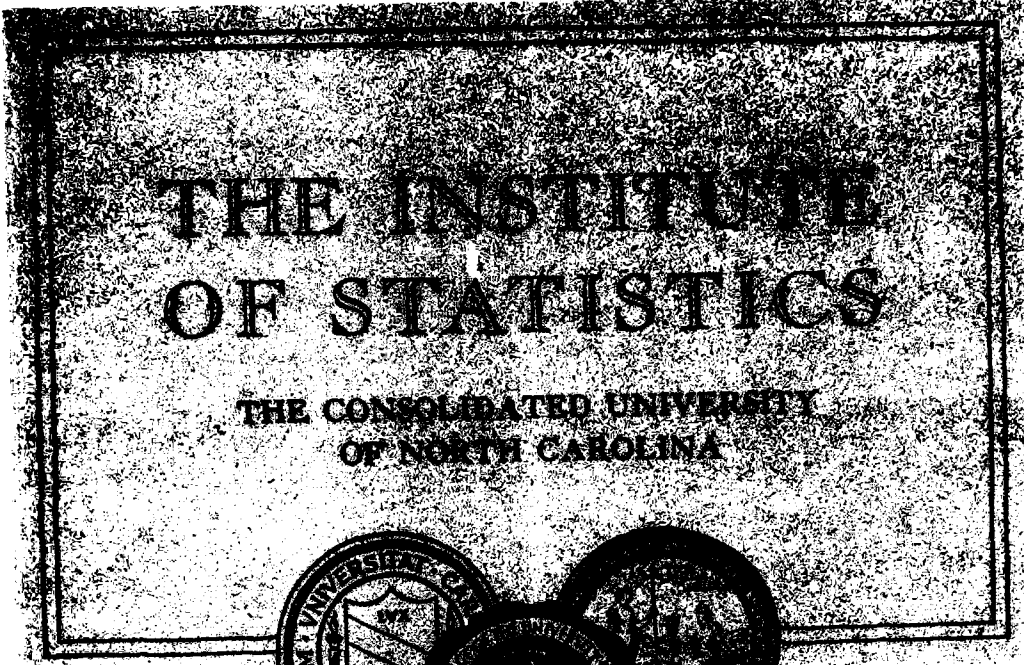
NORTH CAROLINA UNIV AT CHAPEL HILL INST OF STATISTICS F/6 12/1  
ASYMPTOTIC PROPERTIES OF RANDOM SUBSETS OF PROJECTIVE SPACES. (U)  
DEC 80 D G KELLY, J G OXLEY N00014-76-C-0550  
UNCLASSIFIED MIMED SER-1318 NL

1 of 1  
AD  
A094980

END  
DATE  
FILMED  
DTIC

12 LEVEL II

AD A094980



Asymptotic Properties of Random Subsets  
of Projective Spaces

by

Douglas G. Kelly  
and  
James G. Oxley

University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series #1318

December 1980

DEPARTMENT OF STATISTICS  
Chapel Hill, North Carolina

DTIC  
ELECTE  
FEB 12 1981  
S B D

DISTRIBUTION STATEMENT A  
Approved for public release;  
Distribution Unlimited

81 2 12 018

Asymptotic properties of random subsets of  
projective spaces

Douglas G. Kelly<sup>\*</sup> and James G. Oxley<sup>+</sup>  
Departments of Statistics and Mathematics  
University of North Carolina  
Chapel Hill, NC 27514

Abstract

A random graph on  $n$  vertices is a random subgraph of the complete graph on  $n$  vertices. By analogy with this, the present paper studies the asymptotic properties of a random submatroid  $\omega_r$  of the projective geometry  $PG(r-1, q)$ . The main result concerns  $K_r$ , the rank of the largest projective geometry occurring as a submatroid of  $\omega_r$ . We show that with probability one, for sufficiently large  $r$ ,  $K_r$  takes one of at most two values depending on  $r$ . This theorem is analogous to a result of Bollobás and Erdős on the clique number of a random graph. However, whereas from the matroid theorem one can essentially determine the critical exponent of  $\omega_r$ , the graph theorem gives only a lower bound on the chromatic number of a random graph.

Keywords: random graph, random submatroid, critical exponent.

\*Partially supported by ONR Grant No. N00014-76-C-0550.

<sup>+</sup>Partially supported by a Fulbright Postdoctoral Fellowship. Current address: Mathematics Department, IAS, Australian National University, Canberra.

DTIC  
ELECTE  
S FEB 12 1981 D  
B

DISTRIBUTION STATEMENT A

Approved for public release;  
Distribution Unlimited

### A. Introduction

A random submatroid of a matroid  $M$  is obtained from  $M$  by performing a set of independent trials, one for each element of  $M$ , at which the element is deleted with probability  $1-p$  and retained with probability  $p$ . In the study of random graphs such a process is used starting with the complete graph on  $n$  vertices: every simple graph on  $n$  vertices is a possible outcome of the experiment. There are no matroids which are analogous to complete graphs in this sense and so we choose to begin with projective geometries, the random submatroids of which can be thought of as random simple matroids representable over a given finite field. A more complicated model for generating random matroids was proposed by Knuth [7] and implemented by Cravetz [4]. However, this approach does not seem easily amenable to probabilistic analysis.

The theorems of this paper may be informally summarized as follows. Fix a prime power  $q$  and for  $r = 1, 2, \dots$ , let  $M_r$  denote  $PG(r-1, q)$ , the projective geometry of rank  $r$  over  $GF(q)$ . Our analysis is unaffected by whether we assume the matroids  $M_r$  to be nested or disjoint. Let  $\omega_1, \omega_2, \dots$  be the random submatroids of  $M_1, M_2, \dots$  obtained by performing sets of independent trials as described above,  $p$  being the fixed probability of retention of an element. We shall assume that  $0 < p < 1$ . For any sequence  $k_1, k_2, \dots$ , we derive the expected values in  $\omega_r$  of the numbers of circuits of size  $k_r$ , independent sets of size  $k_r$ , flats of rank  $k_r$ , and bases (Proposition 1 and Section D). In the cases of the numbers of circuits and independent sets, we show that with probability one these random variables are asymptotic to their expected values (Theorem 3). A consequence of this is Theorem 4 which implies that with probability one there is  $r_0$  such that each  $\omega_r$  for  $r \geq r_0$  has a circuit of size  $r + 1$ , and therefore has rank  $r$  and is connected. In the last section we consider the random variable  $L_r$ , the rank of the largest

			ies
			or
Dist			cial
<b>A</b>			

subspace of  $M_r$  all of whose elements are deleted. We show that with probability one, for all sufficiently large  $r$ ,  $L_r$  takes its value in a set  $V_r$  which contains either a single integer or a pair of consecutive integers. Since the critical exponent  $c_r$  of  $\omega_r$  is just  $r - 1_r$ , a similar statement can be made about  $c_r$  (Theorem 7). Curiously, the asymptotic value of  $c_r$  is  $r - \log_q r + o(\log_q r)$ , and only lower-order terms in the asymptotic expansion involve the value of  $p$ .

The proofs in the last section parallel those of Grimmett and McDiarmid [6], Matula [8,9], and Bollobás and Erdős [3] for analogous results on random graphs. A summary of many of these graph-theoretic results appears in Bollobás's book [2]. It should be noted that in the area of random graphs the terminology used in limiting results is not uniform. In particular, if  $A_1, A_2, \dots$  is a sequence of events, some authors use the term " $A_n$  occurs almost surely" to mean merely that  $1 - P(A_n)$  approaches zero as  $n$  approaches infinity. We have stated our theorems using the term "with probability one"; such theorems are true strong laws in the probabilistic sense.

In general we shall follow Welsh [11] for all matroid terminology which is otherwise unexplained. Some notation and a few simple inequalities will be useful. Remembering that  $q$  is fixed, we define

$$h_r = |M_r| = \frac{q^r - 1}{q - 1};$$

$$[r]_k = (q^r - 1)(q^{r-1} - 1) \dots (q^{r-k+1} - 1), \quad k = 1, 2, \dots, r;$$

$$[r]_0 = 1; \quad [r]_k = 0 \quad \text{if } k < 0 \text{ or } k > r;$$

$$\binom{r}{k} = \frac{[r]_k}{[k]_k}.$$

Evidently,  $h_r = \binom{r}{1}$ .

We will be concerned with the asymptotic growth of the above quantities as  $r$  increases, for various choices of  $k$  depending on  $r$ . The obvious

inequalities

$$q^{j-1} \leq q^j - 1 \leq q^j \quad \text{for } j = 1, 2, \dots$$

and

$$\frac{q^m - 1}{q^n - 1} \geq q^{m-n} \quad \text{if } m \geq n \quad (1)$$

imply that

$$q^{k(r-k)} \leq \begin{bmatrix} r \\ k \end{bmatrix} \leq q^{k(r-k+1)} \quad (2)$$

We also have

$$q^{kr - \binom{k}{2}} \geq [r]_k \geq q^{kr - \binom{k}{2} - k} \quad (3)$$

To sharpen these bounds we notice that

$$[r]_k = q^{kr - \binom{k}{2}} H_{r,k},$$

where

$$H_{r,k} = (1 - q^{-r})(1 - q^{-r+1}) \dots (1 - q^{-r+k-1}).$$

Obviously  $H_{r,k} \leq 1$ . For lower bounds we observe first that

$$H_{r,k} \geq (1 - q^{-r+k})^k$$

which approaches 1 as  $r$  tends to infinity if  $kq^{-r+k}$  approaches 0.

Regardless of the growth of  $k$ , we can obtain a lower bound by using the

inequality  $\prod (1 - a_n) \geq 1 - \sum a_n$  (for  $0 \leq a_n < 1$ ):

$$H_{r,k} \geq \prod_{n=1}^k (1 - q^{-n}) \geq 1 - \sum_{n=1}^k q^{-n} = \frac{q - 2}{q - 1}.$$

Even though the simpler bound  $\frac{q-2}{q-1}$  is zero for  $q = 2$ , the infinite product is never zero.

Combining the above for later reference:

$$q^{kr - \binom{k}{2}} \geq [r]_k \geq q^{kr - \binom{k}{2}} \prod_{n=1}^{\infty} (1 - q^{-n}) \geq \frac{q-2}{q-1} q^{kr - \binom{k}{2}}, \quad (4)$$

and

$$[r]_k \sim q^{kr - \binom{k}{2}} \text{ as } r \rightarrow \infty \text{ if } kq^{-r+k} \rightarrow 0. \quad (5)$$

We will use two standard theorems from probability:

Chebyshev's Inequality. If  $X$  is a random variable with finite variance  $VX$  and expected value  $EX$ , then for any  $\epsilon > 0$ ,

$$P(|X - EX| \geq \epsilon | EX|) \leq \frac{1}{\epsilon^2} \frac{VX}{(EX)^2} = \frac{1}{\epsilon^2} \left( \frac{EX^2}{(EX)^2} - 1 \right).$$

The First Borel-Cantelli Lemma. If  $\{A_1, A_2, \dots\}$  is a sequence of events and  $\sum_{n=1}^{\infty} P(A_n)$  is a convergent series, then with probability 1 there exists  $n_0$  such that none of the  $A_n$  with  $n \geq n_0$  occurs. (That is,

$$P\left(\bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} A_n^c\right) = 1.)$$

As easy consequence of these theorems we have the following lemmas, which we will use repeatedly.

Lemma A. Let  $(X_1, X_2, \dots)$  be a sequence of random variables, and suppose  $\sum_{n=1}^{\infty} \frac{VX_n}{(EX_n)^2}$  is a convergent series. Then  $\lim_{n \rightarrow \infty} \frac{X_n}{EX_n} = 1$  with probability 1.

If  $\liminf_{n \rightarrow \infty} EX_n$  is positive, then with probability 1 there is  $n_0$  such that  $X_n$  is positive for all  $n \geq n_0$ .

Proof: For  $k = 1, 2, \dots$ , let  $A_k$  be the event that there exists  $n_k$  such that  $\left| \frac{X_n}{EX_n} - 1 \right| < \frac{1}{k}$  for  $n \geq n_k$ . By Chebyshev's Inequality, the Borel-Cantelli Lemma, and the hypothesis,  $PA_k = 1$ . Therefore

$1 = P\left(\bigcap_{k=1}^{\infty} A_k\right) = P\left(\lim_{n \rightarrow \infty} \frac{X_n}{EX_n} = 1\right)$ . To prove the second assertion we note

that if  $EX_n$  is positive, then  $P(X_n \leq 0) \leq P(|X_n - EX_n| \geq \frac{1}{2}|EX_n|)$ .  $\square$

Lemma B. If  $VX$  is finite, then  $P(X = 0) \leq \frac{VX}{(EX)^2} = \frac{EX^2}{(EX)^2} - 1$ .  $\square$

Finally, the definition of expectation obviously implies

Lemma C. If  $X$  is a nonnegative integer-valued random variable, then

$P(X \neq 0) \leq EX$ .  $\square$



### B. Some quantities associated with projective spaces

It is well-known (see, for example, [5]) that  $\binom{r}{k}$  equals the number of rank- $k$  subspaces of  $M_r$ . In this section we shall determine the other numerical invariants of  $M_r$  that will be used in the remainder of the paper. We shall need the following

Lemma D. If  $B$  is a basis of  $M_r$ , then there are precisely  $(q-1)^{r-1}$  elements  $x$  of  $M_r$  such that  $B \cup x$  is a circuit.

Proof: We view the projective space  $M_r$  as the submatroid of the vector space  $V(r,q)$  consisting of those non-zero vectors whose first non-zero coordinate is one. Then, by symmetry, we may assume that  $B$  is the natural basis of  $V(r,q)$ . It is clear that  $B \cup x$  is a circuit of  $M_r$  if and only if the vector  $x$  has no zero coordinates. Hence if  $B \cup x$  is a circuit, the first coordinate of  $x$  is 1, while each of the remaining  $r-1$  coordinates can be chosen in  $q-1$  ways from among the non-zero elements of  $GF(q)$ .  $\square$

We now count the members of  $I_{r,k}$  and  $C_{r,k}$  which are respectively the collections of  $k$ -element independent sets and  $k$ -element circuits of  $M_r$ .

A  $k$ -element independent set  $I$  of  $M_r$  lies in precisely one flat of rank  $k$ , namely its closure,  $\bar{I}$ . Therefore

$$|I_{r,k}| = \binom{r}{k} |I_{k,k}|.$$

But  $I_{k,k}$  is the set of bases of  $M_k$  and it is not difficult to show (see, for example, [11, Exercise 16.1.4]) that

$$|I_{k,k}| = \frac{1}{k!} (h_k - h_0)(h_k - h_1) \dots (h_k - h_{k-1}). \quad (6)$$

It follows that

$$|I_{r,k}| = \frac{1}{(q-1)^k} \frac{1}{k!} q^{\binom{k}{2}} [r]_k. \quad (7)$$

To determine  $|C_{r,k}|$ , we note first that  $|C_{r,k}| = 0$  for  $k < 3$ . Thus suppose  $k \geq 3$ . Then

$$|C_{r,k}| = \begin{bmatrix} r \\ k-1 \end{bmatrix} |C_{k-1,k}|.$$

Now, in  $M_{k-1}$ , consider the set of ordered pairs  $(B,C)$  where  $B$  is a basis and  $C$  is a circuit containing  $B$ . By counting the number of such pairs in two different ways, first over circuits and then over bases, we get, using Lemma D, that

$$k |C_{k-1,k}| = (q-1)^{k-2} |I_{k-1,k-1}|.$$

Thus, by (6),

$$|C_{k-1,k}| = \frac{1}{k!} (h_{k-1} - h_0)(h_{k-1} - h_1) \dots (h_{k-1} - h_{k-2}) (q-1)^{k-2}$$

and so

$$|C_{r,k}| = \frac{1}{q-1} \frac{1}{k!} q^{\binom{k-1}{2}} [r]_{k-1} \quad \text{for } k \geq 3. \quad (8)$$

Now suppose that  $\mathcal{D}$  equals  $C_{r,k}$  or  $I_{r,k}$ . Then for  $i$  in  $\{0,1,2,\dots,k\}$  and  $D$  in  $\mathcal{D}$ , the number of members of  $\mathcal{D}$  which meet  $D$  in exactly  $i$  elements does not depend on the choice of  $D$ . We shall call this number  $\alpha_i$  when  $\mathcal{D} = C_{r,k}$  and  $\beta_i$  when  $\mathcal{D} = I_{r,k}$ . These numbers arise in second moment calculations in the next section and the following result bounds them above.

Lemma E.

$$\alpha_i \leq \begin{cases} \frac{1}{(k-i)!} \binom{k}{i} q^{\binom{k-1}{2} - \binom{i}{2}} (q-1)^{i-1} [r-i]_{k-i-1}, & \text{if } 0 \leq i \leq k-1, \\ 1, & \text{if } i = k. \end{cases}$$

and

$$\alpha_i \leq \frac{1}{(k-i)!} \binom{k}{i} q^{\binom{k}{2} - \binom{i}{2}} \frac{1}{(q-1)^{k-i}} [r-i]_{k-i} \quad \text{for all } i \text{ in } \{0, 1, 2, \dots, k\}.$$

Proof: Clearly  $\alpha_k = 1$ . We now assume that  $i < k$  and let  $X$  be a fixed  $k$ -element circuit of  $M_r$ . It is clear that  $\alpha_i$  is equal to the product of the number of ways to choose an  $i$ -element subset  $Y$  of  $X$  and the number of ways to add a  $(k-i)$ -element set  $Z$  to  $Y$  so that  $Y \cup Z$  is a  $k$ -element circuit meeting  $X$  in  $Y$ . Now  $Y$  can be chosen in  $\binom{k}{i}$  ways. Moreover, if  $N_1$  is the number of choices for  $Z$ , then

$$N_1 \leq \frac{1}{(k-i)!} N_2$$

where  $N_2$  is the number of  $(k-i)$ -tuples  $(p_1, p_2, \dots, p_{k-i})$  such that

- (i) for all  $j$  in  $\{1, 2, \dots, k-i-1\}$ , the element  $p_j$  is not in  $\overline{Y \cup \{p_1, p_2, \dots, p_j\}}$ ; and
- (ii)  $Y \cup \{p_1, p_2, \dots, p_{k-i-1}\} \cup \{p_{k-i}\}$  is a circuit.

On using Lemma D, we obtain that

$$N_2 = (h_r - h_i)(h_r - h_{i+1}) \dots (h_r - h_{k-2})(q-1)^{k-2}.$$

Therefore

$$N_1 \leq \frac{1}{(k-i)!} (h_r - h_i)(h_r - h_{i+1}) \dots (h_r - h_{k-2})(q-1)^{k-2}$$

and thus

$$\begin{aligned} \alpha_i &\leq \frac{1}{(k-i)!} \binom{k}{i} (h_r - h_i)(h_r - h_{i+1}) \dots (h_r - h_{k-2})(q-1)^{k-2} \\ &= \frac{1}{(k-i)!} \binom{k}{i} q^{\binom{k-1}{2} - \binom{i}{2}} (q-1)^{i-1} [r-i]_{k-i-1}. \end{aligned}$$

The last expression is the stated bound on  $\alpha_i$ .

To obtain the bound on  $\beta_i$  we use an argument similar to the above to get

$$\beta_i \leq \frac{1}{(k-i)!} \binom{k}{i} (h_r - h_i)(h_r - h_{i+1}) \dots (h_r - h_{k-1}),$$

and rewriting the right-hand side of this, we obtain the required bound.  $\square$

The last result of this section specifies one further quantity which will be needed in a second moment calculation. Define  $\gamma_i$  to be the number of rank- $k$  subspaces of  $M_r$  which meet a fixed rank- $k$  subspace in a subspace of rank  $i$ . Then it is not difficult to show (see, for example, [1, p. 225]) that

$$\gamma_i = \binom{k}{i} \binom{r-k}{k-i} q^{(k-i)^2}. \quad (9)$$

### C. Existence of circuits and independent sets

Let  $\{k_r\}$  be an arbitrary sequence of positive integers which we will regard as fixed. For simplicity we denote the families  $C_{r,k_r}$  and  $I_{r,k_r}$  by  $C_r$  and  $I_r$ . We also define the random variables  $C_r$  and  $I_r$  to be the numbers of  $k_r$ -element circuits and  $k_r$ -element independent sets in  $M_r$ .

Notice that a  $k_r$ -set  $J$  is a circuit (resp. independent set) in  $M_r$  if and only if  $J$  is a circuit (resp. independent set) in  $M_r$  and none of the elements of  $J$  is deleted. So if we define, for each  $k_r$ -set  $J$  in  $M_r$ ,

$$X_J = \begin{cases} 1, & \text{if none of the elements of } J \text{ is deleted,} \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

then

$$C_r = \sum_{J \in C_r} X_J \quad \text{and} \quad I_r = \sum_{J \in I_r} X_J.$$

Moreover,  $EX_J = P(X_J = 1) = p^{|J|}$ . Therefore we have by (8) and (7)

Proposition 1.

$$EC_r = p^{k_r} |C_r| = \frac{1}{q-1} \frac{p^{k_r}}{k_r!} q^{\binom{k_r-1}{2}} [r]_{k_r-1} \quad \text{provided } k_r \geq 3 \quad (11)$$

and

$$EI_r = p^{k_r} |I_r| = \frac{1}{(q-1)k_r} \frac{p^{k_r}}{k_r!} q^{\binom{k_r}{2}} [r]_{k_r}. \quad (12)$$

The central result of this section is

Proposition 2.  $\sum_{r=2}^{\infty} \frac{VC_r}{(EC_r)^2}$  and  $\sum_{r=1}^{\infty} \frac{VI_r}{(EI_r)^2}$  are convergent series.

The proof is given below. As a corollary of Proposition 2 we get, using Lemma A,

Theorem 3. For every choice of the sequence  $\{k_r\}$ ,

if  $3 \leq k_r \leq r+1$  for all  $r$ , then, with probability one,  $\lim_{r \rightarrow \infty} \frac{C_r}{EC_r} = 1$ ;  
 if  $0 \leq k_r \leq r$  for all  $r$ , then, with probability one,  $\lim_{r \rightarrow \infty} \frac{I_r}{EI_r} = 1$ .

Proposition 1 together with (4) and (5) provide asymptotic expressions for  $EC_r$  and  $EI_r$ , which are almost-sure asymptotic values of  $C_r$  and  $I_r$ .

Since  $EC_r$  and  $EI_r$  are bounded away from zero, we also have from Lemma A:

Theorem 4. For every choice of the sequence  $\{k_r\}$ ,

if  $3 \leq k_r \leq r+1$  for all  $r$ , then with probability 1 there exists  $r_0$  such that  $C_r$  has a  $k_r$ -circuit for all  $r \geq r_0$ ;

if  $1 \leq k_r \leq r$  for all  $r$ , then with probability 1 there exists  $r_0$  such that  $\omega_r$  has a  $k_r$ -independent set for all  $r \geq r_0$ .

In particular, if we choose  $k_r = r + 1$  for circuits we see that with probability 1 there exists  $r_0$  such that for all  $r \geq r_0$ ,  $\omega_r$  has a circuit of size  $r + 1$  and thus is connected and has rank  $r$ .

Proof of Proposition 2.

$$\begin{aligned} EC_r^2 &= \sum_{J_1 \in C_r} \sum_{J_2 \in C_r} P(X_{J_1} X_{J_2} = 1) = \sum_{J_1 \in C_r} \sum_{J_2 \in C_r} p^{2k_r - |J_1 \cap J_2|} \\ &= |C_r| \sum_{J_2 \in C_r} p^{2k_r - |J_1 \cap J_2|} \quad (\text{for any fixed } J_1 \in C_r) \\ &= |C_r| p^{2k_r} \sum_{i=0}^{k_r} p^{-i} \alpha_i, \end{aligned}$$

where  $\alpha_i$  is the number of  $k_r$ -circuits intersecting a fixed  $k_r$ -circuit in  $i$  points.

Therefore, by Lemma E and (8),

$$\begin{aligned} \frac{EC_r^2}{(EC_r)^2} &= \frac{EC_r^2}{p^{2k_r} |C_r|^2} \leq \frac{1}{|C_r|} \left( \sum_{i=0}^{k_r-1} \frac{p^{-i}}{(k_r-i)!} \binom{k_r}{i}_q \binom{k_r-1}{2}^{(i)} (q-1)^{i-1} [r-i]_{k_r-i-1} p^{-k_r} \right) \\ &= 1 + \sum_{i=1}^{k_r-1} p^{-i} \frac{k_r!}{(k_r-i)!} \binom{k_r}{i}_q^{-\binom{i}{2}} \frac{(q-1)^i}{[r]_i} + \frac{k_r!}{p^{k_r}} q^{-\binom{k_r-1}{2}} \frac{(q-1)}{[r]_{k_r-1}} \\ &\leq 1 + \sum_{i=1}^{k_r-1} p^{-i} \frac{k_r!}{(k_r-i)!} \binom{k_r}{i}_q^{-\binom{i}{2}} \frac{q^i}{ir - \binom{i}{2} - i} + \frac{k_r!}{p^{k_r}} \frac{q}{(k_r-1)(r-1)} \end{aligned}$$

where the last step follows by (3). Therefore

$$\frac{VC_r}{(EC_r)^2} \leq \sum_{i=1}^{k_r-1} t_i + \frac{qk_r}{p} \left( \frac{k_r}{pq^{r-1}} \right)^{k_r-1},$$

where

$$t_i = p^{-i} \frac{k_r!}{(k_r-i)!} \binom{k_r}{i} q^{-i(r-2)}.$$

Now

$$\frac{t_{i+1}}{t_i} = \frac{1}{p} \frac{(k_r-i)^2}{i+1} q^{-(r-2)} \leq \frac{r^2}{pq^{r-2}},$$

and thus  $t_{i+1}/t_i \leq 1$  for sufficiently large  $r$ . So for sufficiently large  $r$ ,

$$\frac{VC_r}{(EC_r)^2} \leq k_r t_1 + \frac{qk_r}{p} \left( \frac{k_r}{pq^{r-1}} \right)^{k_r-1} \leq \frac{(r+1)^3}{pq^{r-2}} + \frac{q(r+1)}{p} \left( \frac{r+1}{pq^{r-1}} \right)^2.$$

This is the  $r^{\text{th}}$  term in a convergent series.

Turning now to independent sets, we proceed almost exactly as for circuits.

$$EI_r^2 = |I_r| p^{2k_r} \sum_{i=0}^{k_r} p^{-i} \beta_i$$

where  $\beta_i$  is the number of  $k_r$ -independent sets intersecting a fixed  $k_r$ -independent set in  $i$  points.

Therefore, by Lemma E and (7),

$$\begin{aligned} \frac{EI_r^2}{(EI_r)^2} &= \frac{EI_r^2}{p^{2k_r} |I_r|^2} \leq \frac{1}{|I_r|} \sum_{i=0}^{k_r} \frac{p^{-i}}{(k_r-i)!} \binom{k_r}{i} q^{\binom{k_r}{2} - \binom{i}{2}} \frac{1}{(q-1)^{k_r-1}} [r-i]_{k_r-1} \\ &= 1 + \sum_{i=1}^{k_r} p^{-i} \frac{k_r!}{(k_r-i)!} q^{-\binom{i}{2}} \frac{(q-1)^i}{[r]_i} \end{aligned}$$

This differs only slightly from the upper bound obtained on  $\frac{EC_r^2}{(EC_r)^2}$  in the argument above. A straightforward modification of that argument shows

that  $\frac{VI_r}{(EI_r)^2}$  is the  $r^{\text{th}}$  term in a convergent series.  $\square$

#### D. Expected numbers of bases and flats.

Again we consider as fixed a given sequence  $\{k_r\}$  of positive integers;

and we define the families  $B_r$  and  $F_r$  of bases and  $k_r$ -flats (flats of rank  $k_r$ ) in  $M_r$ , and the random variables  $B_r$  and  $F_r$ , the numbers of bases and  $k_r$ -flats in  $\omega_r$ .

Notice that the results of the previous section imply the existence with probability 1 of an  $r_0$  such that  $\omega_r$  has full rank for all  $r \geq r_0$ , and therefore  $B_r$  almost surely equals  $|I_{r,r}|$  for large  $r$ . In this section we find the expected values of  $B_r$  and  $F_r$  in terms of the Tutte polynomials (see [11, Chapter 15]) of the underlying projective geometries  $M_i$ . We do not obtain asymptotic results. The expected values are given in (16) and (17).

#### Bases.

$$EB_r = \sum_{i=0}^r E(B_r \mid \text{rank}(\omega_r) = i) P(\text{rank}(\omega_r) = i),$$

and

$$\begin{aligned} E(B_r \mid \text{rank}(\omega_r) = i) &= \sum_{J \in M_i} E(B_r \mid \text{rank}(\omega_r) = i \text{ and } \omega_r \subseteq J) P(\omega_r \subseteq J \mid \text{rank}(\omega_r) = i) \\ &\quad (\text{where } M_i \text{ is the family of rank-}i \text{ subspaces of } M_r) \\ &= E(B_r \mid \text{rank}(\omega_r) = i \text{ and } \omega_r \subseteq J_0) \end{aligned}$$

for any fixed rank- $i$  subspace  $J_0$  of  $M_r$ . Now such a  $J_0$  is isomorphic to  $M_i$ , so an argument similar to that used for Proposition 1 shows that this last quantity equals  $p^i$  times the number of  $i$ -independent sets in  $M_i$ ; that is,

$$E(B_r \mid \text{rank}(\omega_r) = i) = \frac{p^i}{i!} q^{\binom{i}{2}} \frac{[i]_i}{(q-1)^i}. \quad (13)$$

To find  $P(\text{rank}(\omega_r) = i)$  we use the following theorem of Oxley and Welsh [10]. If  $M$  is a matroid of rank  $i$  on  $h$  elements and  $\omega$  is



a random submatroid of  $M$ , then

$$P(\text{rank}(a) = i) = p^i (1-p)^{h-i} T(M; 1, (1-p)^{-1}), \quad (14)$$

where  $T(M; x, y)$  is the Tutte polynomial of  $M$ . Using this theorem:

$$\begin{aligned} P(\text{rank}(\omega_r) = i) &= \sum_{J \subseteq M_i} P(\text{all elements of } M_r - J \text{ are deleted and } \omega_r \text{ has full rank in } J) \\ &= |M_i|! P(M_r - J_0 \text{ is deleted}) P(\text{a random submatroid of } M_i \text{ has full rank}). \end{aligned}$$

Here  $J_0$  can be any fixed member of  $M_i$ . It follows that

$$\begin{aligned} P(\text{rank}(\omega_r) = i) &= \binom{r}{i} (1-p)^{r-h+i} p^i (1-p)^{h-i} T(M_i; 1, (1-p)^{-1}) \\ &= \binom{r}{i} p^i (1-p)^{h-i} T(M_i; 1, (1-p)^{-1}). \end{aligned} \quad (15)$$

Combining (13) and (15) gives

$$EB_r = \sum_{i=0}^r \frac{p^{2i}}{i!} (1-p)^{h-i} q^{\binom{i}{2}} \frac{[r]_i}{(q-1)^i} T(M_i; 1, (1-p)^{-1}). \quad (16)$$

Notice that the term corresponding to  $i = r$  dominates this sum because  $\omega_r$  almost surely has rank  $r$  for sufficiently large  $r$ .

Flats.  $EF_r$  equals the number of  $k_r$ -flats in  $M_r$  times the probability that a given such flat has full rank in  $\omega_r$ . By (14),

$$EF_r = \binom{r}{k_r} (1-p)^{h-k_r} p^{k_r} T(M_{k_r}; 1, (1-p)^{-1}). \quad (17)$$

E. Largest full subspace.

For  $r = 1, 2, \dots$ , let  $K_r$  be the rank of the largest full subspace of  $\omega_r$ ; that is, the largest subspace of  $M_r$  with no deleted elements. Our main result in this section is Theorem 6, which implies that with probability 1 there is  $r_0$  such that for all  $r \geq r_0$  the random variable  $K_r$  has at most two possible values. Symmetry gives a similar result (Theorem 7) for the rank of the largest subspace of  $M_r$  with no retained elements, and hence for the critical exponent of  $\omega_r$ . (It is merely for convenience of notation that our results are proved for full rather than empty subspaces.)

For an arbitrary integer  $k$ , let  $F_{r,k}$  be the family of rank- $k$  subspaces of  $M_r$ ; then

$$|F_{r,k}| = \binom{r}{k}.$$

Let  $N_{r,k}$  be the number of full rank- $k$  subspaces of  $\omega_r$ . As with circuits and independent sets,

$$N_{r,k} = \sum_{J \in F_{r,k}} X_J$$

where  $X_J$  is defined by (10). Therefore, for any  $J$  in  $F_{r,k}$ ,

$$EN_{r,k} = |F_{r,k}| P(X_J = 1) = \binom{r}{k} p^{h_k}.$$

Moreover,  $K_r < k$  if and only if  $N_{r,k} = 0$ .

In this section "log" will denote base- $q$  logarithms and "ln" natural logarithms. We also let

$$b = \left(\frac{1}{p}\right)^{\frac{1}{q-1}},$$

so that

$$b > 1 \text{ and } p^{h_k} = b^{-qk+1}.$$

For any  $\epsilon \geq 0$ , define

$$d_{r,\epsilon} = \left\lceil \log \frac{r \log r}{\log b} + \epsilon \right\rceil.$$

Notice that if  $0 < \epsilon < 1$ , then either  $d_{r,0}$  and  $d_{r,\epsilon}$  are equal or they differ by 1. It can also be checked that if  $\epsilon$  is a given positive number and  $j$  and  $k$  denote  $d_{r,0}$  and  $d_{r,\epsilon}$ , then for sufficiently large  $r$ ,  $EN_{r,j} \geq 1 > EN_{r,k+1}$ .

Proposition 5. For any  $\epsilon > 0$ ,  $\sum_{r=1}^{\infty} P(K_r > d_{r,\epsilon})$  and  $\sum_{r=1}^{\infty} P(K_r < d_{r,0})$  are convergent series.

The proof is given below. As a corollary we get from the Borel-Cantelli Lemma.

Theorem 6. Suppose  $0 < \epsilon < 1$ . Then with probability 1 there exists  $r_0$  such that for every  $r \geq r_0$ ,  $K_r$  has its value in the set  $\{d_{r,0}, d_{r,\epsilon}\}$  (which may be a singleton or a pair).

This theorem translates immediately by symmetry to a result on the rank  $L_r$  of the largest subspace of  $M_r$  with no retained elements and on the critical exponent  $c_r$  of  $\omega_r$ , where  $c_r = r - L_r$ . For  $\epsilon \geq 0$  let

$$d'_{r,\epsilon} = \left\lceil \log \frac{r \log r}{\log b'} + \epsilon \right\rceil$$

where

$$b' = \left(\frac{1}{1-p}\right)^{\frac{1}{q-1}}.$$

Theorem 7. Suppose  $0 < \epsilon < 1$ . Then with probability 1 there exists  $r_0$  such that for every  $r \geq r_0$ ,  $L_r$  and  $c_r$  have their values in the sets  $\{d'_{r,0}, d'_{r,\epsilon}\}$  and  $\{r - d'_{r,\epsilon}, r - d'_{r,0}\}$ , respectively.

We note two more consequences of the above before proving Proposition 5. Firstly, the asymptotic expressions for  $K_r$ ,  $L_r$ , and  $c_r$  have high-order terms that are independent of  $p$ :

$$K_r \sim L_r \sim d_{r,0} \sim \log r + o(\log r), \text{ and } c_r \sim r - \log r + o(\log r).$$

This is in contrast to the growth of the size of the largest clique in a random graph as found in [6,9,2]. Secondly, with probability one,  $K_r$  is eventually greater than two and hence for sufficiently large  $r$ ,  $\omega_r$  is representable only over fields containing  $GF(q)$ .

Proof of Proposition 5. We prove that

$$r^2 P(K_r \geq d_{r,c} + 1) \rightarrow 0 \text{ as } r \rightarrow \infty \quad (18)$$

and

$$r^2 P(K_r < d_{r,0}) \rightarrow 0 \text{ as } r \rightarrow \infty, \quad (19)$$

and the proposition follows.

To prove (18) we notice that for any  $k$ , by Lemma C,

$$P(K_r \geq k) = P(N_{r,k} \neq 0) \leq EN_{r,k} = \binom{r}{k} b^{-q^{k+1}},$$

and so, by (2),

$$P(K_r \geq k) \leq q^{k(r-k+1)} b^{-q^{k+1}}.$$

Now if  $k = d_{r,c} + 1$ , then

$$\frac{r \log r}{\log b} q^c \leq q^k \leq \frac{r \log r}{\log b} q^{1+c}$$

and

$$b^{q^k} \geq q^{r(\log r)q^c} = c^r q^c.$$

So

$$\begin{aligned}
 r^2 P(K_r > d_{r,s} + 1) &\leq r^2 \left( \frac{r \log r}{\log b} q^{1+s} \right)^{r-d_{r,s}} r^{-rq^s} b \\
 &= \left[ \frac{\log r}{r^{q^s-1} \log b} q^{1+s} \right]^{r-d_{r,s}} \frac{b r^2}{r^{q^s d_{r,s}}}
 \end{aligned}$$

which tends to 0 as  $r \rightarrow \infty$ . Thus (18) is proved.

Next we prove (19). For any  $k$ , by Lemma B,

$$P(K_r \leq k) = P(N_{r,k} = 0) \leq -1 + \frac{EN_{r,k}^2}{(EN_{r,k})^2}.$$

Now

$$\begin{aligned}
 EN_{r,k}^2 &= \sum_{J_1 \in F_{r,k}} \sum_{J_2 \in F_{r,k}} P(X_{J_1} X_{J_2} = 1) = \sum_{J_1 \in F_{r,k}} \sum_{J_2 \in F_{r,k}} p^{2h_k - |J_1 \cap J_2|} \\
 &= |F_{r,k}| \sum_{J_2 \in F_{r,k}} p^{2h_k - |J_1 \cap J_2|} \quad (\text{for any fixed } J_1 \in F_{r,k}) \\
 &= \binom{r}{k} p^{2h_k} \sum_{i=0}^k \gamma_i p^{-h_i}
 \end{aligned}$$

where  $\gamma_i$  is the number of rank- $k$  subspaces intersecting a fixed rank- $k$  subspace in a rank- $i$  subspace. Now, because of (9),

$$-1 + \frac{EN_{r,k}^2}{(EN_{r,k})^2} \leq -1 + \sum_{i=0}^k T_i$$

where

$$T_i = \frac{\binom{k}{i} \binom{r-k}{k-i}}{\binom{r}{k}} q^{(k-i)^2} b^{q^i-1}, \quad i = 0, 1, \dots, k.$$

Now, by (1),  $T_0 \leq 1$ ; and (2) implies that

$$T_i \leq b^{q^i-1} q^{k-i(r-2k+i)} \quad (i = 1, 2, \dots, k).$$

Therefore

$$P(K_r < k) \leq \sum_{i=1}^k s_i$$

where

$$s_i = b^{q^i - 1} q^{k-i(r-2k+i)} .$$

Now we show that if  $k = d_{r,0}$ , then for sufficiently large  $r$  the function

$$f(x) = b^{q^x - 1} q^{k-x(r-2k+x)}$$

first decreases and then increases and has exactly one critical point in the interval  $1 \leq x \leq k$ . It will follow that

$$P(K_r < k) \leq \frac{k}{2}(s_1 + s_k) \quad \text{for } k = d_{r,0} \text{ and sufficiently large } r. \quad (20)$$

We use the fact that if  $k = d_{r,0}$ , then

$$\frac{r \log r}{q \log b} \leq q^k \leq \frac{r \log r}{\log b} \quad \text{and} \quad b^{q^k} \leq r^r . \quad (21)$$

We can rewrite  $f(x)$  as

$$q^k q^{(q^x - 1) \log b - x(r-2k+x)}$$

and it suffices to show that the nonconstant part of the exponent,

$$g(x) = q^x \log b - x^2 - (r - 2k)x ,$$

has the properties claimed above for  $f(x)$ . But

$$g'(x) = q^x \ln b - 2x - r + 2d_{r,0} ;$$

so

$$g'(1) = q \ln b - 2 - r + 2d_{r,0} ,$$

which is obviously negative for large  $r$ . Moreover,

$$\begin{aligned} g'(k) &= q^k \ln b - r \geq \frac{r \log r}{q \log b} \ln b - r \\ &= r(\log r) \ln q - r \end{aligned}$$

which is positive for large  $r$ . Thus  $g(x)$  first decreases and then increases for  $1 \leq x \leq k$ , and so  $g'(x)$ , being continuous, has an odd number of zeros in  $[1, k]$ . But  $g'(x)$  has at most two zeros, since it is the difference between the convex function  $q^x \ln b$  and the linear function  $2x + (r - 2k)$ . So  $g'(x)$  has exactly one zero in  $[1, k]$ , the assertion about  $f(x)$  is proved, and (20) follows. We get

$$\begin{aligned} P(K_r < k) &\leq \frac{k}{2} (b^{q-1} q^{k-r+2k-1} + b^{q^k-1} q^{k-k(r-k)}) \\ &= \frac{b^q}{2bq} kq^{3k-r} + \frac{1}{2b} kb^q q^{k^2+k-kr} \end{aligned}$$

But we can use (21) to show that each of these terms is  $o(r^{-2})$ :

$$r^2 \frac{b^q}{2bq} kq^{3k-r} \leq r^2 \frac{b^q}{2bq} \log\left(\frac{r \log r}{\log b}\right) \frac{1}{q^{r-3k}} \rightarrow 0 \text{ as } r \rightarrow \infty,$$

and

$$\begin{aligned} \log\left(r^2 \frac{1}{2b} kb^q q^{k^2+k-kr}\right) &\leq 2 \log r - \log 2b + \log \log\left(\frac{r \log r}{\log b}\right) + r \log r \\ &\quad + k^2 + k - kr \\ &= r(\log r - k) + o((\log r)^4) \\ &\leq r\left(\log r - \log\left(\frac{r \log r}{\log b}\right) + 1\right) + o((\log r)^4) \\ &\rightarrow -\infty \text{ as } r \rightarrow \infty. \end{aligned}$$

□

References

1. George E. Andrews, The Theory of Partitions, Encyclopedia of Mathematics and its Applications, Volume 2 (Addison-Wesley, Reading, Massachusetts, 1976).
2. Béla Bollobás, Graph Theory. An Introductory Course, Graduate Texts in Mathematics No. 63 (Springer-Verlag, New York, Heidelberg, Berlin, 1979).
3. B. Bollobás and P. Erdős, Cliques in random graphs, Math. Proc. Camb. Phil. Soc. 80 (1976), 419-427.
4. Allan E. Cravetz, Essentials for Matroid Election (M.S. Thesis, Department of Mathematics, University of North Carolina, Chapel Hill, 1978).
5. Jay Goldman and Gian-Carlo Rota, The number of subspaces of a vector space, Recent Progress in Combinatorics, Editor: W. T. Tutte (Academic Press, New York, London, 1969) pp. 75-83.
6. G. R. Grimmett and C. J. H. McDiarmid, On colouring random graphs, Math. Proc. Camb. Phil. Soc. 77 (1975), 313-324.
7. Donald E. Knuth, Random matroids, Discrete Math. 12 (1975), 341-358.
8. David W. Matula, On the complete subgraphs of a random graph, Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications (U.N.C. Press, Chapel Hill, 1970) pp. 356-369.
9. \_\_\_\_\_, Graph-theoretic cluster analysis, Classification and Clustering, Editor: J. Van Ryzin (Academic Press, New York, San Francisco, London, 1977) pp. 95-129.
10. J. G. Oxley and D. J. A. Welsh, The Tutte polynomial and percolation, Graph Theory and Related Topics, Editors: J. A. Bondy and U.S.R. Murty (Academic Press, New York, San Francisco, London, 1979) pp. 329-339.
11. D. J. A. Welsh, Matroid Theory, London Math. Soc. Monographs No. 8 (Academic Press, London, New York, San Francisco, 1976).



REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
	AD-A094980	
4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED	
Asymptotic Properties of Random Subsets of Projective Spaces.	9) TECHNICAL REPORT	
7. AUTHOR(s)	6. PERFORMING ORG REPORT NUMBER	
10 Douglas G. Kelly <del>and</del> James G. Oxley	Mimeo Series No. 1318	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	8. CONTRACT OR GRANT NUMBER(s)	
14) MIMOSA - 2327	15) ONR Grant N00014-76-C-0550	
11. CONTROLLING OFFICE NAME AND ADDRESS	10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS	
Office of Naval Research Statistics and Probability Program Arlington, VA 22217	11) 15	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	12. REPORT DATE	
	December 1980	
	13. NUMBER OF PAGES	
	22	
	15. SECURITY CLASS. (of this report)	
	UNCLASSIFIED	
	15a. DECLASSIFICATION DOWNGRAUING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)		
Approved for Public Release -- Distribution Unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
random graph, random submatroid, critical exponent		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		
<p>A random graph on <math>n</math> vertices is a random subgraph of the complete graph on <math>n</math> vertices. By analogy with this, the present paper studies the asymptotic properties of a random submatroid <math>\omega_r</math> of the projective geometry <math>PG(r-1, q)</math>. The main result concerns <math>K_r</math>, the rank of the largest projective geometry occurring as a submatroid of <math>\omega_r</math>. We show that with probability one, for</p>		

20. sufficiently large  $r$ ,  $K_r$  takes one of at most two values depending on  $r$ . This theorem is analogous to a result of Bollobás and Erdős on the clique number of a random graph. However, whereas from the matroid theorem one can essentially determine the critical exponent of  $\omega_r$ , the graph theorem gives only a lower bound on the chromatic number of a random graph.

