Combinatorics, Probability and Computing (2002) 11, 465–474.
DOI: 10.1017/S0963548302005199 Printed in the United Kingdom

A Sharp Threshold for Network Reliability

MICHAEL KRIVELEVICH,[1]†
BENNY SUDAKOV[2]‡ and VAN H. VU[3]§

[1] Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel Aviv University, Tel Aviv 69978, Israel
(e-mail: krivelev@math.tau.ac.il)

[2] Department of Mathematics, Princeton University, Princeton, NJ 08540, USA
and
Institute for Advanced Study, Princeton, NJ 08540, USA
(e-mail: bsudakov@math.princeton.edu)

[3] Theory Group, Microsoft Research, Redmond, WA 98052, USA
(e-mail: vanvu@ucsd.edu;
Web: http://www.math.ucsd.edu/~vanvu)

Received 20 October 2000; revised 5 November 2001

Given a graph G on n vertices with average degree d, form a random subgraph G_p by choosing each edge of G independently with probability p. Strengthening a classical result of Margulis we prove that, if the edge connectivity $k(G)$ satisfies $k(G) \gg d/\log n$, then the connectivity threshold in G_p is sharp. This result is asymptotically tight.

1. Introduction

Reliability problems become more and more important as our modern systems of telecommunications, information transmission, and transportation become more and more complex (the Internet might be a good example to keep in mind). This motivates the theoretical study of network reliability, a topic which has been extensively studied in the past few decades.

† Part of this research was done while visiting Microsoft Research. Supported in part by a USA–Israeli BSF grant and by a Bergmann Memorial Award.
‡ Part of this research was done while visiting Microsoft Research. Research supported in part by NSF grants DMS-0106589, CCR-9987845 and by the State of New Jersey.
§ Current address: Department of Mathematics, UCSD, La Jolla, CA, USA.

One of the most popular abstract models in network reliability problems is the following. Our network can be thought of as a large connected graph where each edge has a certain probability q of failing. We are interested in the probability that the network is still connected. This problem can be formulated in a form which is perhaps more convenient to a graph theorist, as follows. Given a graph G with n vertices and m edges, and a real p between 0 and 1, where $p = 1 - q$ may depend on G, a random subgraph G_p of G is obtained by keeping each edge of G with probability p, independently. The probability that G_p is connected is obviously then a function of the probability p, which will be denoted by $f(G, p)$, or simply by $f(p)$ when the definition of a graph G is clear from the context. We will sometimes refer to $f(p)$ as to the *reliability function* of the graph G.

Estimating $f(G, p)$ seems to be very hard, and there is a vast literature on this issue. The interested reader may check [2, Chapter 7] or [8] for a partial list of references. Several special cases of this problem have been considered in different areas. For instance, if G is the complete graph on n vertices, then G_p is the classical random graph $G(n, p)$, and the connectivity problem is discussed in great detail in Bollobás's book [2]. Another case is when G is the d-dimensional lattice restricted to a compact domain: in this case the problem has been studied in percolation theory, and we refer the interested reader to [4].

In this paper, we investigate the following aspect of network reliability. For a fixed positive constant $x \leqslant 1$ and a graph G, let p_x denote the (unique) value of p where $f(G, p_x) = x$. We say that a family $(G_i)_{i=1}^{\infty}$ of graphs satisfies the *sharp threshold* property if, for any fixed positive $\epsilon \leqslant 1/2$,

$$\lim_{i\to\infty} \frac{p_\epsilon(G_i)}{p_{1-\epsilon}(G_i)} \to 1.$$

The sharp threshold property is very useful from the practical point of view. It implies that the performance of the network is easy to improve. For instance, the fact that $\frac{p_\epsilon(G_i)}{p_{1-\epsilon}(G_i)} \to 1$ implies that, when i is sufficiently large, to increase the reliability of G_i from 0.01 (a very poor network) to 0.99 (a rather reliable network), we need only increase edge reliability by a tiny fraction, a nominal cost for a remarkable improvement! In percolation theory, a sharp threshold is more commonly known as a phase transition, and there is an extensive literature on this phenomenon, motivated by questions from statistical physics (see [4] and its references).

We would like to address the following central question:

Which families of graphs possess the sharp threshold property?

A different motivation for our study comes from a paper of Pak and the third author [7]. There the above question was considered from a different aspect in relation to phase transitions of random walks. The problems posed in that paper (see Section 13 of [7]) formed the starting point of our study.

In [7] several partial answers to the above question are given in special cases when the graphs in question are highly symmetric. For general graphs, the earliest, and most well-known, result is probably a result of Margulis. For a graph G, let $k(G)$ denote the minimum number of edges one needs to remove in order to disconnect G; if $k(G) = k$ we say that G is k-edge-connected. In [6], as a corollary of a more general theorem, Margulis derived the following result.

Theorem 1.1. *Consider a family $(G_i)_{i=1}^{\infty}$ of graphs. If $k(G_i) \to \infty$, then, for any fixed positive $\epsilon \leqslant 1/2$, we have*

$$\lim_{i\to\infty}(p_{1-\epsilon}(G_i) - p_{\epsilon}(G_i)) = 0. \qquad \square$$

Margulis's theorem implies that a family $(G_i)_{i=1}^{\infty}$ possesses the sharp threshold property if the connectivity $k(G_i)$ of G_i tends to infinity, and $p_{1-\epsilon}(G_i)$ is bounded below by a positive constant. However, this theorem does not provide any information in the case $p_{1-\epsilon}(G_i) \to 0$.

Our main result in this paper is as follows.

Theorem 1.2. *Let $0 < \epsilon < 1/2$. Then, for every $\gamma > 0$, there exist $K(\gamma)$ and $n_0(\gamma)$ such that the following holds. If G is a graph on $n > n_0(\gamma)$ vertices, with average degree d and edge-connectivity $k(G) \geqslant K(\gamma)\frac{d}{\ln n} + 1$, then*

$$\frac{p_{\epsilon}(G)}{p_{1-\epsilon}(G)} \geqslant 1 - \gamma.$$

The above theorem immediately implies the following corollary.

Corollary 1.3. *Let $(G_i)_{i=1}^{\infty}$ be a family of distinct graphs, where G_i has n_i vertices, maximum degree d_i, and it is k_i-edge-connected. If*

$$\lim_{i\to\infty}\frac{k_i \ln n_i}{d_i} = \infty,$$

then the family $(G_i)_{i=1}^{\infty}$ has a sharp connectivity threshold. $\square$

We believe that this result is of interest for a number of reasons. First, it gives a fairly general sufficient condition for a family of graphs to satisfy the sharp threshold property. Second, it strengthens Margulis's result in the case $p_{1-\epsilon}(G_i) \to 0$. It also answers a question posed in [7]. Next, our proof makes use of new and powerful results of Bourgain and Friedgut [3], and it is very different from Margulis's proof and the approaches in percolation theory. Finally, the statement of our theorem is in some sense asymptotically tight, as shown by the following proposition.

Proposition 1.4. *For any constant $a > 1$, there is a constant $0 < \epsilon(a) < 1/2$ such that the following holds. For all sufficiently large n, there exists a graph G on $2n$ vertices, with maximal degree $d = n$ and with edge-connectivity $k(G) = a\frac{d}{\ln n}$, for which*

$$\frac{p_{\epsilon}(G)}{p_{1-\epsilon}(G)} \leqslant \frac{1}{2}.$$

The rest of the paper is organized as follows. In the next section we prove Theorem 1.2 and Proposition 1.4. In Section 3, we show how our main result can be extended to the case of the random matroid process. Finally, the last section contains some concluding remarks.

2. Main result

In this section we prove our main result. We may, and shall, assume, whenever this is needed, that the number of vertices in our graphs is sufficiently large. To prove Theorem 1.2 it is enough to show that, for all $\epsilon \leqslant \alpha \leqslant 1-\epsilon$, the derivative of $f(p)$ satisfies $p_\alpha f'(p_\alpha) \geqslant 1/\gamma$. Indeed, in this case

$$\begin{aligned} 1 \geqslant f(p_{1-\epsilon}) - f(p_\epsilon) = f'(p_\alpha)\big(p_{1-\epsilon} - p_\epsilon\big) &\geqslant \frac{1}{\gamma p_\alpha}\big(p_{1-\epsilon} - p_\epsilon\big) \\ &\geqslant \frac{1}{\gamma p_{1-\epsilon}}\big(p_{1-\epsilon} - p_\epsilon\big) = \frac{1}{\gamma}\left(1 - \frac{p_\epsilon}{p_{1-\epsilon}}\right). \end{aligned}$$

This implies that $\gamma \geqslant 1 - p_\epsilon/p_{1-\epsilon}$, which is the assertion of the theorem.

Let us recall some terminology. Consider a discrete cube $\{0,1\}^m$ with the probability measure defined by $\Pr_p(x) = p^{|x|}(1-p)^{m-|x|}$ for all $x \in \{0,1\}^m$, where $|x| = |\{1 \leqslant i \leqslant n : x_i = 1\}|$. We say that a vector $x = (x_1, \dots, x_m) \in \{0,1\}^m$ *contains* a vector $y = (y_1, \dots, y_m) \in \{0,1\}^m$ if $x_i \geqslant y_i$ for all $1 \leqslant i \leqslant m$, and denote this by $y \subset x$. A subset $A \subset \{0,1\}^m$ is monotone if whenever $x \in A$ and $x \subset y$, then also $y \in A$. Our proof relies heavily on the following result of Bourgain [3], which provides a sharp-threshold criterion for general monotone properties.

Theorem 2.1. *Let $A \subset \{0,1\}^m$ be a monotone property, α be a positive constant, and $p = o(1)$ satisfy $\Pr_p(A) = \alpha$. If there exists a constant $c > 0$ with the property $p \cdot \frac{d\Pr_p(A)}{dp} < c$, then there exists a $\delta = \delta(c)$ such that, either*

$$\Pr_p(x \in \{0,1\}^m | x \text{ contains } x' \in A \text{ of size } |x'| \leqslant 10c) > \delta,$$

or there exists an $x' \notin A$ of size $|x'| \leqslant 10c$ so that

$$\Pr_p(x \in A | x' \subset x) > \alpha + \delta. \qquad \square$$

The idea of the proof is as follows. Assuming that a threshold for connectivity is not sharp, we know, from Theorem 2.1, that there exists a fixed set of edges whose addition to the random graph G_p changes the probability of connectivity by some constant. On the other hand, the fact that a threshold is not sharp implies that the addition of a large number of random edges to G_p has almost no effect on the connectivity. We show that these two conclusions contradict each other. To do so, we first need to establish a lower bound on the threshold probability for the graph connectivity property.

Lemma 2.2. *Let $G = (V, E)$ be a connected graph on n vertices, with average degree d, and let $0 < p < 1$ satisfy $(1-p)^d \geqslant 1/\sqrt{n}$, i.e., $pd \leqslant \ln n/2$. Then, for any fixed $0 < \alpha < 1$, and sufficiently large n, the probability that a random subgraph G_p is connected is at most α.*

Proof. Let $V_0 = \{v \in V : d(v) \leqslant \frac{3}{2}d\}$. Then $|V_0| \geqslant n/3$, for otherwise $\sum_{v \in V \setminus V_0} d(v) \geqslant |V \setminus V_0|(3d/2) > (2n/3)(3d/2) = nd = 2|E(G)|$, a contradiction.

For every vertex $v \in V_0$ let X_v be an indicator random variable for the event that v is an isolated vertex in G_p. Let X be the total number of such vertices in the random graph

G_p. Clearly $X = \sum_{v \in V_0} X_v$, and G_p is connected only if $X = 0$. It is easy to see that the expected value of X satisfies

$$E[X] = \sum_{v \in V_0} E[X_v] = \sum_{v \in V_0} (1-p)^{d(v)} \geqslant |V_0|(1-p)^{\frac{3d}{2}} \geqslant \frac{n}{3}\left(\frac{1}{\sqrt{n}}\right)^{\frac{3}{2}} = \frac{n^{\frac{1}{4}}}{3}.$$

Next we need to obtain an upper bound on the variance of X.

$$\begin{aligned}\mathrm{Var}[X] &= \sum_{v \in V_0} \mathrm{Var}[X_v] + \sum_{v \neq u \in V_0} \mathrm{Cov}[X_v, X_u] \\ &= \sum_{v \in V_0} \mathrm{Var}[X_v] + \sum_{v \neq u \in V_0} (E[X_v X_u] - E[X_v]E[X_u]).\end{aligned}$$

Since X_v is an indicator random variable we deduce that $\mathrm{Var}[X_v] \leqslant E[X_v]$. Note also that, if the vertices u and v are nonadjacent, then X_u, X_v are independent random variables, and thus $\mathrm{Cov}[X_v, X_u] = 0$. On the other hand, for adjacent vertices, we have

$$E[X_v X_u] - E[X_v]E[X_u] = (1-p)^{d(v)+d(u)-1} - (1-p)^{d(v)+d(u)} = p(1-p)^{d(v)+d(u)-1}.$$

Finally the inequality $(1-p)^d \geqslant 1/\sqrt{n}$ implies that $1 + \frac{3}{2}pd < 2 \ln n < \alpha E[X]$. Therefore we conclude that

$$\begin{aligned}\mathrm{Var}[X] &= \sum_{v \in V_0} E[X_v] + 2 \sum_{\substack{v,u \in V_0 \\ (v,u) \in E}} p(1-p)^{d(v)+d(u)-1} \\ &= E[X] + p \sum_{v \in V_0} \sum_{\substack{u \in V_0 \\ (u,v) \in E}} (1-p)^{d(v)+d(u)-1} \\ &\leqslant E[X] + p \sum_{v \in V_0} d(v)(1-p)^{d(v)} \leqslant E[X] + \frac{3}{2} pd \sum_{v \in V_0} (1-p)^{d(v)} \\ &= E[X] + \frac{3pd}{2} E[X] = \left(1 + \frac{3pd}{2}\right) E[X] < \alpha E^2[X].\end{aligned}$$

Now, by Chebyshev's inequality, the probability that G_p is connected has the upper bound

$$\Pr(X = 0) \leqslant \Pr(|X - E[X]| \geqslant E[X]) \leqslant \Pr\left(|X - E[X]| \geqslant \frac{\sqrt{\mathrm{Var}[X]}}{\sqrt{\alpha}}\right) \leqslant \alpha. \qquad \square$$

We are now in a position to prove Theorem 1.2.

Proof of Theorem 1.2. Let α be a real number satisfying $\epsilon \leqslant \alpha \leqslant 1 - \epsilon$, and let p_α be the probability such that $\Pr(G_{p_\alpha}$ is connected$) = \alpha$. First we consider the case when there exists an α with $0 < p_\alpha < 1$ being a constant. Note that, since connectivity is a monotone property, clearly $f(p) = \Pr(G_p$ is connected$)$ is an increasing function of p. Thus by Lemma 2.2, the threshold probability p_α should satisfy $(1 - p_\alpha)^d < 1/\sqrt{n}$. Since p_α is a constant less than 1, the average degree d is at least $\Omega(\ln n)$. In that case, by choosing an appropriate constant $K(\gamma)$, we can make the edge-connectivity $k(G) = K(\gamma)\frac{d}{\ln n}$ arbitrarily large. Therefore we can apply the above mentioned result of Margulis [6] (see also [9]) to derive the assertion of the theorem.

Next we treat the case when $p_\alpha = o(1)$. Let us assume by contradiction that $p_\alpha f'(p_\alpha) < 1/\gamma$. Since clearly no set of edges of a constant size can contain a connected spanning subgraph of G, we obtain from Theorem 2.1 that there exists a constant $\delta(\gamma) > 0$ and a fixed set of edges $e_1, \dots, e_t$, $t \leqslant 10/\gamma$, satisfying

$$\Pr(G_{p_\alpha} \text{is connected} \mid e_i \in E(G_{p_\alpha}), i = 1, \dots, t) > \alpha + \delta. \tag{2.1}$$

Let ϵ' be a positive constant, to be specified later, and let $p_1 = p_\alpha + \epsilon'(1 - p_\alpha)p_\alpha$. Then, by the Taylor expansion of f, together with the fact that $f'(p_\alpha) < 1/(\gamma p_\alpha)$, we obtain

$$\begin{aligned} f(p_1) &= f(p_\alpha) + f'(p_\alpha)(p_1 - p_\alpha) + o(p_1 - p_\alpha) \\ &\leqslant \alpha + \frac{1}{\gamma p_\alpha}\epsilon'\big(1 - p_\alpha\big)p_\alpha + o(\epsilon') \\ &= \alpha + \frac{\epsilon'}{\gamma}\big(1 - p_\alpha\big) + o(\epsilon'). \end{aligned}$$

By choosing an appropriate value of ϵ' we can ensure that the probability that G_{p_1} is connected satisfies $f(p_1) < \alpha + \delta/2$. Note also that, by the definition of p_1, we can view the edge set of the random graph G_{p_1} as a union of two independent copies of the random graphs G_{p_α} and $G_{\epsilon' p_\alpha}$. Let B denote the set of all subgraphs $G' \subset G$ with the property that the graph $G' \cup \{e_1, \dots, e_t\}$ is connected. It is easy to see that, by inequality (2.1), we have $\Pr(G_{p_\alpha} \in B) > \alpha + \delta$.

Next we show that, for any graph $G' \in B$, the union $G' \cup G_{\epsilon' p_\alpha}$ is connected with probability close to one. Indeed, G' becomes connected when adding the edges $e_1, \dots, e_t$. Therefore G' has at most $t + 1$ connected components, and thus there exist at most 2^t possible edge cuts of G which separate the vertices of G'. Each such cut contains at least $k = k(G)$ edges. Recall that, by Lemma 2.2, we have $p_\alpha d = \Omega(\ln n)$. Therefore the probability that at least one of these cuts also separates the vertices of a random graph $G_{\epsilon' p_\alpha}$ is at most

$$2^t(1 - \epsilon' p_\alpha)^k \leqslant 2^t e^{-\epsilon' p_\alpha k} = 2^t e^{-\Omega\left(\frac{\epsilon' \ln n}{d}\right)k} = 2^t e^{-\Omega(\epsilon' K(\gamma))}.$$

By choosing an appropriate constant $K(\gamma)$, we can ensure that this probability is at most $\delta/4$. Finally we obtain a contradiction, since the probability $f(p_1)$ that the random graph $G_{p_1} = G_{p_\alpha} \cup G_{\epsilon' p_\alpha}$ is connected is at least

$$\begin{aligned} \Pr(G_{p_\alpha} \in B)\Pr(G_{p_\alpha} \cup G_{\epsilon' p_\alpha} \text{ is connected} \mid G_{p_\alpha} \in B) &\geqslant \left(1 - \frac{\delta}{4}\right)\Pr(G_{p_\alpha} \in B) \\ &= \left(1 - \frac{\delta}{4}\right)\big(\alpha + \delta\big) > \alpha + \frac{\delta}{2}. \end{aligned}$$

The last case is when $1 - p_\epsilon = o(1)$. Hence both p_ϵ and $p_{1-\epsilon}$ are equal to $1 - o(1)$, and thus their ratio is equal to $1 - o(1) > 1 - \gamma$. This completes the proof of the theorem. □

A graph G is called *vertex-transitive* if, for every pair of vertices v_1 and v_2, there exists an automorphism $\pi : V(G) \to V(G)$ such that $\pi(v_1) = v_2$. By applying our main theorem we can obtain the following result about the connectivity threshold for vertex-transitive graphs.

Corollary 2.3. *Let G be a connected vertex-transitive graph and let G_p be obtained by selecting edges of G randomly and independently with probability p. Then the property 'G_p is connected' has a sharp threshold.*

Proof. Clearly G is regular. Let d be its degree. Since G is a vertex-transitive graph, it is also d-edge-connected (see, *e.g.*, [5], Problem 12.14). Then the result follows immediately from Theorem 1.2. □

An important family of vertex-transitive graphs arises from finite groups. Given a finite group H and a set of generators $S = S^{-1}$ of H, the *Cayley graph* $G(H, S)$ is a graph with vertex set H, in which there is an edge between a and b if and only if $ab^{-1} \in S$. The Cayley graph $G(H, S)$ is easily seen to be connected, because S generates H. The above corollary then implies that the connectivity property of a random subgraph of any Cayley graph has a sharp threshold.

Next we show that the result of Theorem 1.2 is nearly tight.

Proof of Proposition 1.4. Set $\epsilon = e^{-4a}$. Let G be a graph which consists of two disjoint copies of a complete graph K_n on n vertices, connected by a matching of size $a\frac{n}{\ln n}$. The maximal degree of G is $d = n$, and its edge-connectivity is $k(G) = a\frac{n}{\ln n}$. Let G_p be obtained by selecting edges of G randomly and independently with probability p. It is easy to see that the probability that G_p is connected is at most $1-(1-p)^{a\frac{n}{\ln n}}$, since we need to choose at least one edge connecting two copies of K_n. One can easily check that $1-t \geqslant e^{-t-t^2}$ for sufficiently small $t > 0$. Therefore, for $p = 3\ln n/n$, the probability that G_p is connected is at most

$$\begin{aligned} 1-(1-p)^{\frac{a\ln n}{n}} &\leqslant 1-e^{(-p-p^2)\frac{a\ln n}{n}} \\ &= 1-e^{\left(-\frac{3\ln n}{n}-\frac{9\ln^2 n}{n^2}\right)\frac{a\ln n}{n}} = 1-e^{-3a-\frac{9a\ln n}{n}} = 1-\epsilon^{\frac{3}{4}+\frac{9\ln n}{4n}} \\ &< 1-\epsilon \end{aligned}$$

for sufficiently large n. This implies $p_{1-\epsilon} \geqslant \frac{3\ln n}{n}$.

On the other hand, if $p = \frac{3\ln n}{2n}$, it is well known (see, *e.g.*, [2]) that the random subgraph of K_n, where each edge is chosen independently and with probability p, is connected with probability tending to one. Therefore, in this case, the probability that G_p is connected equals

$$\begin{aligned} (1-o(1))\left(1-(1-p)^{a\frac{n}{\ln n}}\right) &\geqslant (1-o(1))\left(1-e^{-\frac{pa\ln n}{n}}\right) \\ &= (1-o(1))\left(1-e^{-\frac{3a}{2}}\right) = (1-o(1))\left(1-\epsilon^{\frac{3}{8}}\right) \geqslant \epsilon \end{aligned}$$

as $\epsilon \leqslant e^{-4}$. Hence $p_\epsilon \leqslant \frac{3\ln n}{2n}$, and $\frac{p_\epsilon}{p_{1-\epsilon}} \leqslant \frac{1}{2}$. □

3. Random matroid process

In this section we sketch how our results can be extended to the case of random matroid processes. Let us first introduce some terminology. We define a *matroid* M to be a finite set X and a collection $\mathscr{F}$ of subsets of X, called *independent* sets, satisfying the following properties.

(1) $\emptyset \in \mathscr{F}$, and if $A \in \mathscr{F}$ and $B \subseteq A$ then $B \in \mathscr{F}$.
(2) If U, V are members of $\mathscr{F}$, with $|U| = |V| + 1$ then there exist an $u \in U - V$ such that $V \cup u \in \mathscr{F}$.

(For the theory of matroids see, *e.g.*, [10]). A *base* of a matroid M is an independent set of maximal size, and a subset of X is called *spanning* if and only if it contains a base. One of the main examples of matroids, which we have already discussed in the previous section, is the *cycle matroid of a graph* $M(G)$. Given a graph G, let $X = E(G)$ and let $A \in \mathscr{F}$ if and only if A is an edge set of an acyclic subgraph of G. This defines a matroid $M(G)$. Clearly, if G is connected, then the bases of $M(G)$ are the spanning trees of G, and the spanning sets of this matroid are all connected subgraphs of G. The *rank function* of a matroid is a function $r : 2^X \to \mathbb{Z}$, where $r(A)$ is a size of maximal independent subset of A. The rank of the matroid $r(M)$ is just the rank of the set X. Finally, let $\eta(M)$ be the size of the smallest subset $Y \subset X$ such that $r(X - Y) < r(M)$. This parameter is an extension of the notion of the edge-connectivity number of a graph, since for the case of the cycle matroid of a graph G it is equal to its edge-connectivity.

Given a matroid $M = (X, \mathscr{F})$, let X_p be obtained by choosing elements of X randomly and independently with probability p. Consider the property 'X_p is spanning'. Clearly this property is monotone, and we let p_α denote the value of p such that $\Pr(X_{p_\alpha}$ is spanning$) = \alpha$. Note that, in the case when M is the cycle matroid of a connected graph, the property of being spanning corresponds to the property that a random subgraph G_p is connected. Therefore a natural extension of the result of the previous section is to determine when the property 'X_p is spanning' has a sharp threshold. This is done in the following theorem, whose proof we merely sketch, since it is rather similar to the proof of Theorem 1.2.

Theorem 3.1. *Let $M = (X, \mathscr{F})$ be a matroid and let X_p be obtained by choosing the elements of X randomly and independently with probability p. If $r(M)$ tends to infinity and $(1 - p_\alpha)^{\eta(M)} = o(1)$ for any constant α, then the property 'X_p is spanning' has a sharp threshold.*

Sketch of proof. First consider the case when $0 < p_\alpha < 1$ is a constant. Then $\eta(M) \to \infty$, and therefore we can apply the result of Margulis [6] to derive the sharpness of the threshold.

Now, suppose that $p_\alpha = o(1)$ and that the property does not have a sharp threshold. This implies that, for $p = p_\alpha$, the value of the derivative of $f(p) = \Pr(X_p$ is spanning$)$ is bounded by c/p_α for some constant c. Since the size of a base of M tends to infinity, we obtain, by Theorem 2.1, that there exist a constant $\delta(c) > 0$ and a fixed set of elements $Y \subset X, |Y| = y$ such that $\Pr(X_{p_\alpha}$ is spanning $|Y \subset X_{p_\alpha}) > \alpha + \delta$. On the other hand, the fact that the derivative is bounded by c/p_α implies that there exists a constant $\beta > 0$ with the property that, for $p_1 = (1 + \beta)p_\alpha$, we have $\Pr(X_{p_1}$ is spanning$) < \alpha + \delta/2$. Let $\mathscr{S}$ be the family of all subsets $X' \subset X$ with the property that $X' \cup Y$ is spanning. Then $\Pr(X_{p_\alpha} \in \mathscr{S}) > \alpha + \delta$. In addition, we can view X_{p_1} as a union of X_{p_α} with y independent copies of $X_{\epsilon' p_\alpha}$, for some appropriate constant ϵ' which depends on β. Denote these copies by $X^{(1)}_{\epsilon' p_\alpha}, \ldots, X^{(y)}_{\epsilon' p_\alpha}$.

Now we prove that for any non-spanning subset $T \subset X$, $r(T \cup X_{\epsilon' p_\alpha}) > r(T)$ with probability $1 - o(1)$. First we show that there exist at least $\eta(M)$ elements in X whose addition to T will increase its rank. Indeed, let $T_0 \subseteq T$ be an independent set satisfying $r(T_0) = r(T)$, and let U be the set of all elements of M such that $r(T_0 \cup \{u\}) > r(T_0)$, for every $u \in U$. Obviously $U \cap T = \emptyset$, and, for every element of U, its addition to T will increase its rank. As $r(T_0) < r(M)$, for every base B_i of M there is an element $b_i \in B_i \setminus T_0$ such that $r(T_0 \cup \{b_i\}) > r(T_0)$. Hence $b_i \in U$. This shows that the set U meets every base of M, and thus has cardinality at least $\eta(M)$. The probability that $X_{\epsilon' p_\alpha}$ misses all the elements of U is at most $(1 - \epsilon' p_\alpha)^{\eta(M)} = (1 - p_\alpha)^{\Theta(\epsilon')\eta(M)} = o(1)$. Also note that, when $X_{p_\alpha} \cup Y$ is spanning, the rank of X_{p_α} is at least $r(M) - y$. Finally we have obtained a contradiction, since

$$\Pr(X_{p_1} \text{ is spanning}) \geqslant \Pr(X_{p_\alpha} \in \mathscr{S}) \Pr\left(X_{p_\alpha} \cup \bigcup_{i=1}^{y} X^{(i)}_{\epsilon' p_\alpha} \text{ is spanning} \mid X_{p_\alpha} \in \mathscr{S} \right)$$
$$\geqslant (1 - o(1))(\alpha + \delta) > \alpha + \delta/2. \qquad \square$$

Remark. This theorem is less powerful than Theorem 1.2 since its application needs a lower bound on the threshold probability p_α. In the case of the cycle matroid of a graph, this bound can be derived from Lemma 2.2.

4. Concluding remarks

We have provided a fairly general condition for the sharpness of the threshold connectivity in random subgraphs of arbitrary graphs. This condition can be applied to many families of graphs. Combined with known results of the value of the connectivity threshold, our result can be used to estimate from above the width of the threshold interval for connectivity. Putting it somewhat informally, we say that the *width of the connectivity threshold* interval of a random subgraph G_p is the difference $p_{0.99} - p_{0.01}$. Alon proved in [1] that, if G is a k-connected graph of n vertices, and the edge probability $p(n)$ satisfies $p(n) \geqslant c \log n/k$ for a sufficiently large absolute constant $c > 0$, then a.s. the random subgraph G_p is connected. It follows therefore, from Theorem 1.2, that the width of the connectivity threshold interval is $o(\log n/k)$. In many instances this conclusion compares favourably with that of a more general result of Talagrand [9], asserting that the width of the connectivity threshold interval of a k-connected graph G is at most $O(1/\sqrt{k})$.

It is intuitively clear that Bourgain's general threshold sharpness criterion can, and should, be used to establish the sharpness of the threshold of other graph theoretic functions in random subgraphs of arbitrary graphs. Potential applications include the appearance of a cycle in G_p, of a perfect matching, of a Hamiltonian cycle, to mention just a few. While those questions have been studied very extensively for classical random graphs $G(n, p)$ (see, *e.g.*, [2] for a detailed account), nothing, or almost nothing, appears to be known for the case when the ground graph G is different from the complete graph K_n.

The task of obtaining such results for various graphs G seems quite appealing. Also, it would be interesting to get further threshold sharpness results for the random matroid process.

Acknowledgement

We would like to thank the referee for their careful reading, and also for pointing out that the assumption on the maximal degree in Theorem 1.2 (in the original version of the paper) can be replaced by a weaker assumption on the average degree.

References

[1] Alon, N. (1995) A note on network reliability. In *Discrete Probability and Algorithms* (D. Aldous *et al.*, eds), Vol. 72 of *IMA Volumes in Mathematics and its Applications*, Springer, pp. 11–14.
[2] Bollobás, B. (1985) *Random Graphs*, Academic Press, London.
[3] Friedgut, E. (1999) Sharp thresholds of graph properties, and the k-sat problem; with an appendix by Jean Bourgain. *J. Amer. Math. Soc.* **12** 1017–1054.
[4] Grimmett, G. (1989) *Percolation*, Springer.
[5] Lovász, L. (1993) *Combinatorial Problems and Exercises*, North-Holland, Amsterdam.
[6] Margulis, G. (1974) Probabilistic characteristics of graphs with large connectivity. *Problems Info. Transmission* **10** 101–108.
[7] Pak, I. and Vu, V. (2001) On mixing of certain random walks, cutoff phenomenon and sharp threshold of random matroid processes. *Discrete Appl. Math.* **110** 251–272.
[8] Roberts, F., Hwang, F. and Monma, C., eds (1989) *Reliability of Computer and Communication Networks*, Vol. 5 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*.
[9] Talagrand, M. (1993) Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem. *Geometric and Functional Analysis* **3** 295–314.
[10] Welsh, D. J. A. (1976) *Matroid Theory*, Vol. 8 of *Lond. Math. Soc. Monographs*, Academic Press, London.