EVERY MONOTONE GRAPH PROPERTY HAS A SHARP THRESHOLD

EHUD FRIEDGUT AND GIL KALAI

ABSTRACT. In their seminal work which initiated random graph theory Erdös and Rényi discovered that many graph properties have sharp thresholds as the number of vertices tends to infinity. We prove a conjecture of Linial that every monotone graph property has a sharp threshold. This follows from the following theorem.

Let $V_n(p) = \{0,1\}^n$ denote the Hamming space endowed with the probability measure μ_p defined by $\mu_p(\epsilon_1, \epsilon_2, \ldots, \epsilon_n) = p^k \cdot (1-p)^{n-k}$ where $k = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$. Let A be a monotone subset of V_n. We say that A is symmetric if there is a transitive permutation group Γ on $\{1, 2, \ldots, n\}$ such that A is invariant under Γ.

Theorem: For every symmetric monotone A, if $\mu_p(A) > \epsilon$ then $\mu_q(A) > 1 - \epsilon$ for $q = p + c_1 \log(1/2\epsilon)/\log n$. ($c_1$ is an absolute constant.)

1. GRAPH PROPERTIES

A *graph property* is a property of graphs which depends only on their isomorphism class. Let P be a monotone graph property; that is, if a graph G satisfies P then every graph H on the same set of vertices, which contains G as a subgraph satisfies P as well. Examples of such properties are: G is connected, G is Hamiltonian, G contains a clique (=complete subgraph) of size t, G is not planar, the clique number of G is larger than that of its complement, the diameter of G is at most s, etc.

For a property P of graphs with a fixed set of n vertices we will denote by $\mu_p(P)$ the probability that a random graph on n vertices with edge probability p satisfies P. The theory of random graphs was founded by Erdös and Rényi [8, 4], and one of their significant discoveries was the existence of sharp thresholds for various graph properties, that is; the transition from a property being very unlikely to it being very likely is very swift. Many results on various aspects of this phenomenon have appeared since then. In what follows c_1, c_2, etc. are universal constants.

Theorem 1.1. *Let P be any monotone property of graphs on n vertices. If $\mu_p(P) > \epsilon$ then $\mu_q(P) > 1 - \epsilon$ for $q = p + c_1 \log(1/2\epsilon)/\log n$.*

This result verifies a conjecture of Nati Linial [12] and complements a theorem of Bollobás and Thomason [5] who proved that $\mu_q(P) > 1 - \epsilon$ for $q = c(\epsilon) \cdot p$.

Let P be the property "G contains a clique with $k(n)$ vertices" and let $k(n) \sim \log n$. The threshold interval (namely, the interval of edge-probabilities p where $\epsilon \leq \mu_p \leq 1 - \epsilon$,) in this case is of size proportional to $1/\log^2 n$. Perhaps some words of explanation are in order. Consider $G \in G(n,p)$; that is, G is a random graph with n vertices and edge probability p. The length of the interval of probabilities p for which the clique number (= size of maximal clique) of G is almost surely k (where $k \sim \log n$) is of order $\log^{-1} n$. The transition between clique numbers $k - 1$ and k occurs along an interval of length $\approx \log^{-2} n$ and this is precisely the threshold interval of interest to us. At the beginning of this transition period the probability of having a clique of size k is ϵ. This probability rises to $1 - \epsilon$ at the end of this interval, but the probability of having a $(k + 1)$-clique is still small ($\approx 1/\log n$). The value of p must increase by $c \log^{-1} n$ before the probability for having a $(k + 1)$-clique reaches ϵ and another transition interval begins.

Conjecture 1.2. *Let P be any monotone property of graphs on n vertices. If $\mu_p(P) > \epsilon$ then $\mu_q(P) > 1 - \epsilon$ for $q = p + c\log(1/2\epsilon)/\log^2 n$.*

2. SYMMETRIC PROPERTIES

Consider the Hamming space $V_n(p) = \{0,1\}^n$ endowed with the probability measure μ_p defined by $\mu_p(\epsilon_1, \epsilon_2, \ldots, \epsilon_n) = p^k \cdot (1 - p)^{n-k}$ where $k = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$. Put $\{0,1\}_p = V_1(p)$. Let A be a monotone subset of $\{0,1\}^n$; that is, if $(\alpha_1, \alpha_2, \ldots \alpha_n) \in A$ and $\beta_i \geq \alpha_i$ for every i, $1 \leq i \leq n$ then $(\beta_1, \beta_2, \ldots, \beta_n) \in A$. We say that A is *symmetric* if there is a transitive permutation group Γ on $[n] = \{1, 2, \ldots, n\}$ such that A is invariant under Γ.

Theorem 2.1. *For every symmetric monotone A, if $\mu_p(A) > \epsilon$ then $\mu_q(A) > 1 - \epsilon$ for $q = p + c_2 \log(1/2\epsilon)/\log n$.*

To deduce Theorem 1.1 from Theorem 2.1 (with $c_1 = 2c_2$,) note that the family of edge-sets of graphs on n vertices which satisfy a monotone graph property P is invariant under the action of S_n (the group of permutations of the vertices) on the edges. Note that Theorem 1.1 remains true for monotone properties of random subgraphs of arbitrary finite edge-transitive graphs. In particular, the theorem applies to random subgraphs of the discrete cube and of C_n^k, the product of k copies of an n-cycle (grids on the k-dimensional torus).

Note that the symmetry assumption is needed. If $A_0 = \{(\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \in \{0,1\}^n : \epsilon_1 = 1\}$ then $\mu_p(A) = p$. Bollobás and Thomason [5] proved, using the Kruskal-Katona theorem, that for every monotone A (symmetric or not) if $\mu_p(A) > \epsilon$ then $\mu_q(A) > 1 - \epsilon$ for $q \leq c(\epsilon)p$.

In the rest of the section we will deduce Theorem 2.1 via a Lemma of Margulis [13] and Russo [14], (see also [9], Ch. 2) from a result of Bourgain, Kahn, Kalai, Katznelson and Linial (briefly, BKKKL) [6]. In the next section we will indicate a few extensions of the BKKKL Theorem which imply several extensions of Theorem 2.1. In Section 4 we will give examples showing that our results are tight. In Section 5 we will discuss the connection of the threshold interval to the symmetry group Γ and in Section 6 we will discuss some connections to results by Margulis-Talagrand, Russo and Kruskal-Katona.

For $v \in A$ let $h(v) = |\{w \notin A : dist(v, w) = 1\}|$. Define $\psi_p(A) = \sum_{v \in A} \mu_p(v) h(v)$.

Lemma 2.2 (Margulis, Russo).

$$\frac{d(\mu_p(A))}{dp} = \psi_p(A)/p.$$

Next, we need the notion of influence. Let X be a probability space and let $f : X^n \to \{0, 1\}$ be a measurable map. Define the *influence of the k-th variable* on f, denoted by $I_f(k)$, as follows: For $u = (u_1, u_2, ..., u_{n-1}) \in X^{n-1}$ set $l_k(u) = \{(u_1, u_2, ..., u_{k-1}, t, u_k, ..., u_{n-1}) : t \in X\}$, and define

$$I_f(k) = Pr(u \in X^{n-1} : f \text{ is not constant on } l_k(u)).$$

Given a monotone set $A \subset V_n(p)$ denote by χ_A the characteristic function of A. We will write $I_A(k)$ for $I_{\chi(A)}(k)$. Let $\phi_p(A)$ be the sum of the influences of the variables on χ_A. Note that $\phi_p(A) = \psi_p(A)/p$.

Theorem 2.3 (Bourgain, Kahn, Kalai, Katznelson and Linial). *For every function $f : X^n \to \{0, 1\}$, with $Pr(f^{-1}(1)) = t$, there is a variable k so that*

$$I_f(k) \geq c_3 t' \frac{\log n}{n},$$

where $t' = min(t, 1 - t)$.

Proof of Theorem 2.1: Since A is symmetric the influence of each variable is the same and therefore by Theorem 2.3 the sum of the influences for A in $V_n(r)$ is at least $c_3 \cdot t' \log n$, where $t' = min(\mu_r(A), 1 - \mu_r(A))$. So by Lemma 2.2 for every r such that $\mu_r(A) \leq 1/2$, we have $d\mu_r(A)/dr \geq c_3 \mu_r(A) \log n$. Therefore, $d(\log \mu_r(A))/dr \geq c_3 \log n$. Now we claim that if $\mu_p(A) \geq \epsilon$ then $\mu_q(A) \geq 1/2$ for $q = p + \frac{1}{c_3} \cdot \frac{\log(1/2\epsilon)}{\log n}$. Indeed, $\log(\mu_q(A)) \geq \log(\mu_p(A)) + \int_p^q c_3 \log n \, dr \geq \log(\epsilon) + \log(1/2\epsilon) = \log(1/2)$. So by increasing p by at most $\log(1/2\epsilon)/c_3 \log n$ we reached q with $\mu_q(A) = 1/2$ and by the same token another increase of at most $\log(1/2\epsilon)/c_3 \log n$ will give us $\mu_q(A) = 1 - \epsilon$ as required. $\square$

The case $X = \{0, 1\}_{1/2}$ of the BKKKL theorem was proved by Kahn, Kalai and Linial [10] in response to a conjecture of Ben-Or and Linial [3]. Some words on the proof of this theorem are in place. The proof uses harmonic analysis on

Z_2^n. For a boolean function f on $\{0,1\}^n$ consider the Walsh-Fourier expansion $f = \sum_{S \subset [n]} \hat{f}(S) u_S$, where, $u_S(T) = (-1)^{|S \cap T|}$. (Here we identify vectors in $\{0,1\}^n$ with subsets of $[n]$ in the standard way.) It can be shown that the sum of influences, $\phi_{1/2}(A)$, is equal to: $\sum_{S \subset [n]} \hat{f}^2(S)|S|$. In order to show that $\phi_{1/2}(A)$ is large one has to show that a large portion of the Walsh-Fourier transform of f is concentrated in "high" frequencies. This is shown if all influences are not too large, by applying a certain hypercontractive estimate of Beckner [2]. (In the next section we will show that if all influences are smaller than δ then $\phi_p(A)$ is at least $c \log(1/\delta)$.)

The proof of the BKKKL theorem [6] is similar and is based on a Fourier-Walsh interpretation for influences in arbitrary product spaces. Talagrand [17] (see remark in Section 6) found another proof for the special case $X = \{0,1\}_p$. He replaced the Walsh functions u_S by an appropriate orthonormal basis for $V_n(p)$. All these proofs give constants which are quite realistic.

3. Some extensions

Consider now the situation when p itself is a function of n. We will describe now sharp forms of the above theorems which shows that if $\mu_p(A) = \epsilon$ and $\mu_q(A) = 1 - \epsilon$ then $q = p(1 + o(1))$ always hold when $\frac{\log p}{\log n} \to 0$. If $p \le n^{-c}$ for some $c > 0$ we cannot expect, in general, that $q = p(1 + o(p))$, and it is an interesting problem to understand for which A's this relation holds. We need an improvement of the BKKKL Theorem for the special case of $X = V_n(p)$:

Theorem 3.1. *For every function* $f : V_n(p) \to \{0,1\}$, *with* $Pr(f^{-1}(1)) = t \le \frac{1}{2}$, *there is a variable* k *so that*

$$I_f(k) \ge c_4 \frac{t \cdot \log n}{n \cdot (p \log(1/p))}.$$

This gives (as above) a sharp (up to constants) form of Theorem 2.1:

Theorem 3.2. *For every symmetric monotone property* A, *if* $\mu_p(A) > \epsilon$ *then* $\mu_q(A) > 1 - \epsilon$ *for* $q = p + c_5 \log(1/\epsilon)p \log(1/p)/\log n$.

Corollary 3.3. *Let* P *be any monotone property of graphs. Then if* $\mu_p(P) > \epsilon$ *then* $\mu_q(P) > 1 - \epsilon$ *for* $q = p + c_6 \log(1/\epsilon)p \log(1/p)/\log n$.

Proof of Theorem 3.1: We rely on the proof from [6]. Note that the proof of Lemma 3 in [6] gives for a monotone 0-1 function on $\{0,1\}_p$, that $w(f) \le cp \log(1/p)$. To see this note that for $X = \{0,1\}_p$ we have the additional inequalities $I_f(j) \le 2p$ for every j, in addition to the inequalities $I_f(j) \le 2^{-(m-j)}$ which always hold. So $\sum_{j=1}^m I_f(j) \le 2p \log(1/p) + p$. Substituting the improved upper bound for $w(f)$ in relation (14) of [6] we get that $\|W_k\|_2^2 \le c \cdot p \log(1/p) \cdot I_f(k)$, and we reach a contradiction in the same way as in the original proof by assuming that $I_f(k) < c_3 \frac{t \cdot \log n}{p \cdot \log(1/p) \cdot n}$, for c_3 sufficiently small. $\square$

The proof of the BKKKL Theorem can be modified to give the following:

Theorem 3.4. *For every function $f : X^n \to \{0, 1\}$, with $Pr(f^{-1}(1)) = t \leq \frac{1}{2}$, If $I_f(k) \leq \delta$ for every k then*

$$\sum_{k=1}^{n} I_f(k) \geq c_7 t \log(1/\delta).$$

Proof: Again we rely heavily on the proof of BKKKL Theorem, (and recycle the constants $c_1, \ldots, c_5$). Let $\delta_k = I_f(k)$. Then, as in the original proof we have that more than half of the weight of $\|f\|_2^2$ is concentrated where

$$|S_1| + |S_2| + \cdots + |S_n| < c_1 t^{-1}(\sum_{k=1}^{n} \delta_k).$$

Substituting relation (18) (from [6]) in relation (20) we can replace the right hand side of relation (20) by $c_2(\sum_{k=1}^{n}(\delta_k)^{3/2})$. It follows that more than half the weight of $\|f\|_2^2$ is concentrated where

$$(|S_1| + \cdots + |S_n|) \cdot \epsilon^{2|S_1|+2|S_2|+\cdots+2|S_n|} \leq c_2 t^{-1} \sum_{k=1}^{n} (\delta_k)^{3/2}.$$

This implies that

$$\sum_{k=1}^{n} \delta_k \geq c_3 t \log(t(\sum_{k=1}^{n}(\delta_k)^{3/2})^{-1}).$$

Now assume that $\delta_k \leq \delta$ for every k and that $\sum_{k=1}^{n} \delta_k \leq c_4 t \log(1/\delta)$. Using the convexity of the function $x^{3/2}$ we get that the maximum value of $\sum_{k=1}^{n}(\delta_k)^{3/2}$ is attained when $\delta_k = \delta$ for $k = 1, 2, \ldots, r$ and $\delta_k = 0$ for $k > r + 1$ where $r = c_4 t \log(1/\delta)/\delta$. Therefore, $\log(t(\sum_{k=1}^{n}(\delta_k)^{3/2})^{-1}) > c_5 \log(1/\delta)$. $\square$

Corollary 3.5. *Let A be a monotone subset of $\{0, 1\}^n$ such that $I_A(k) \leq \delta$ for every p. Then if $\mu_p(A) > \epsilon$ then $\mu_q(A) > 1 - \epsilon$ for $q = p + c_8 \log(1/\epsilon)/\log(1/\delta)$.*

It follows from Theorem 3.4 that if we have a monotone subset A which is invariant under a permutation group Γ with the property that every transitivity class of Γ has at least a elements then the assertion of Theorem 2.1 holds with $1/\log n$ replaced by $1/\log a$. Therefore, the assertion of Theorem 1.1 holds (with a different absolute constant) for symmetric properties of random subgraphs of arbitrary vertex-transitive graphs (in particular, Cayley graphs) on n vertices.

The proof of Theorem 3.4 extends without change to:

Theorem 3.6. *For every function $f : X_1 \times X_2 \times ... \times X_n \to \{0, 1\}$, with $Pr(f^{-1}(1)) = t \leq \frac{1}{2}$, if $I_f(k) \leq \delta$ for every k then*

$$\sum_{k=1}^{n} I_f(k) \geq c_7 t \log(1/\delta).$$

Assume now that $X_k = \{0,1\}_{p_k}$ and put $X = \prod_{i=1}^{n} X_i$. Note that for a monotone subset A of X, $\partial \mu(A)/\partial p_k = I_A(k)$. ($\mu(A)$ is a linear function of p_k with slope $I_A(k)$; this is perhaps the shortest derivation of the Margulis-Russo Lemma.) It follows by a similar argument to the proof of Theorem 2.1 that for every subset A of V_n and every n-tuple of probabilities $(p_1, p_2, \ldots, p_n)$, if $\mu_{p_1,p_2\ldots,p_n}(A) > \epsilon$, then the following statement holds: There is a vector of probabilities $(q_1, q_2, \ldots, q_n)$ such that $\mu_{q_1,q_2,\ldots,q_n}(A) > 1 - \epsilon$ and for every δ,

$$|\{i : q_i - p_i \geq c_{10} \log(1/\epsilon)/\log(1/\delta)\}| \leq \frac{1}{\delta}.$$

4. Tightness of the results

Ben-Or and Linial [3] constructed a "tribes" example to show that the $O(\log n/n)$ lower bound on influences is sharp, see also [10]. The following more general examples show that up to multiplicative constants, Theorem 2.1 is sharp when p does not depend on n, and Theorems 3.1 and 3.2 are sharp even when p does depend on n. Let $n = m \cdot r$. Partition $[n]$ into sets $T_1, T_2, \ldots, T_r$ of size m and let A be the set of subsets of $[n]$ which contain T_i for some i. For a given probability q take

$$m = \frac{\log n - \log \log n + \log \log(1/q)}{\log(1/q)}.$$

Then $\mu_q(A)$ is close to $1 - 1/e$. For $p = q(1 + o(1))$, let $\mu_p(A) = t$ and then

$$\frac{d\mu_p(A)}{dp} \approx (1 - t) \cdot \log(1/(1 - t)) \cdot \log n/(p \log(p^{-1})).$$

Therefore the interval $[q, p]$ where $\mu_p(A) = 1 - \epsilon$, is of length $\sim \log(1/\epsilon)q \log(1/q)/\log n$.

There are many other examples for which Theorem 2.1 is sharp. For example, consider the property of subsets S of $[n]$ - "S contains an interval of length $k = k(n)$ of consecutive integers modulo n". Other examples are the properties of subgraphs H of the graph of the n-dimensional cube - "H contains the graph of a k-dimensional face" and, as mentioned below, "H is connected".

Graph properties. Let P be the property "G contains a clique with $k(n)$ vertices". For fixed n and k let $X_{n,p}$ be the number of k-cliques in $G(n,p)$. The expected value of $X_{n,p}$ is

$$\lambda = \lambda(n, k, p) = \binom{n}{k} p^{\binom{k}{2}}.$$

Consider now an arbitrary function $k(n) = O(\log n)$ and let $q = q(n)$ be such that $\lambda(n, k(n), q(n)) \sim 1$. Note that $\log(1/q)/\log(n) \to 0$ iff $k \to \infty$. In this case Corollary 3.3 asserts that the length of the threshold interval is $o(q)$. For $p = q(1 + o(1))$, $X_{n,p}$ can be approximated by a Poisson random variable with mean $\lambda(n, k, p)$, and $Prob(X_{n,p} > 0) = (1 - e^{-\lambda(n,k,p)})(1 + o(1))$. (This can be justified along the lines of [4],

pp. 260-262.) This shows that the actual threshold interval is of length proportional to $q(n)(\frac{\log(q(n)^{-1})}{\log n})^2$ rather than $q(n)\frac{\log(q(n)^{-1})}{\log n}$.

Bounded depth circuits. There is an interesting connection between the complexity of Boolean functions and their threshold behavior. It follows, e.g., from the Hastad Switching Lemma, see [11, 1], that Boolean functions f that can be expressed by bounded-depth, polynomial size circuits have large threshold intervals. Put $a(p) = \mu_p(f^{-1}(1))$. If f is expressed by a depth-two circuit of size N then

$$\lim_{s \to 0}(a(p) - a(p(1 + s/\log N))) = 0.$$

In other words, the length of the threshold interval is at least $c \cdot p_0/\log N$ where $a(p_0) = 1/2$. This result is tight. Most of the examples described in this section and in the next section can be expressed by depth-two circuits of small size. For circuits of depth d, the term $\log N$ should be replaced by $\log^{d-1} N$. (This connection was pointed out to us by Noga Alon and by Joel Spencer.)

5. THE DEPENDENCE OF THE THRESHOLD INTERVAL ON THE PERMUTATION GROUP

The content of this section is the fruit of collaboration with Aner Shalev.

Problem 5.1. *For Γ a permutation group on $[n]$, how large can the threshold interval be for monotone families A which are invariant under Γ?*

Here by the threshold interval we mean the length of the interval $[p, q]$ where $\mu_p(A) = \epsilon$ and $\mu_q(A) = 1 - \epsilon$, for a fixed ϵ, $0 < \epsilon < 1/2$. Given a permutation group $\Gamma \subset S_n$ consider the following class of examples: For every s find a set S, $|S| = s$, such that the orbit of S under Γ is minimal. Consider the family A_S to be those subsets of $[n]$ which contain a set in the orbit of S. We conjecture that (up to multiplicative constants) such an example will give the largest threshold intervals among monotone families invariant under Γ.

If $\Gamma = S_n$ then the length of the threshold interval is proportional to $1/\sqrt{n}$. For a set X let $\binom{X}{r}$ be the set of r-subsets of X. As we already mentioned, for graph properties (i.e., when $\Gamma = S_m$ acting on $\binom{[m]}{2}$,) the threshold interval can be as large as $c/\log^2 n$, $(n = \binom{m}{2},)$ and we conjecture that it cannot be larger. For symmetric properties of r-uniform hypergraphs (i.e., when $\Gamma = S_m$ acting on $\binom{[m]}{r}$,) the threshold intervals can be as large as $\log^{-\frac{r}{r-1}} n$, $(n = \binom{m}{r},)$ and again we conjecture that it cannot be larger. Another example of interest is the group $\Gamma = PSL(d, q)$ acting as a permutation group on d-dimensional projective space over a field of q elements. If q is fixed (say $q = 2$) then considering the monotone families of all sets containing a subspace of dimension $t = \log d + \log\log d (= \log\log n + \log\log\log n)$, we get an example

with threshold interval of length $\frac{1}{\log n \log \log n}$ and we conjecture that no monotone Γ-symmetric property with larger threshold interval exists.

By forming wreath products one can construct groups with intermediate threshold behavior (between $\log^{-2} n$ and $n^{-1/2}$). However, we conjecture that for *primitive* permutation groups Γ there are some gaps in the possible behavior of the largest threshold intervals for properties which are invariant under Γ:

- The length of this interval is proportional to $n^{-1/2}$ for S_n and A_n but at least $\log^{-2} n$ for any other primitive permutation group.
- The length T of the largest threshold interval satisfies $c_1 \log^{-\alpha} n \leq T \leq c_2 \log^{-\alpha} n$ for $\alpha \in \{2, 1\frac{1}{2}, 1\frac{1}{3}, 1\frac{1}{4}, 1\frac{1}{5}, \dots\}$ or for α which tends to one as a function of n in an arbitrary way.
- If Γ does not involve (as factors) large alternating groups then the length of the largest threshold interval is proportional to $1/(\log n \cdot w(n))$ where $w(n) \leq \log \log n$.

This description follows from the following: We conjecture that the value of the largest threshold interval for Γ-invariant properties where Γ is a primitive permutation subgroup of S_n other than A_n or S_n is proportional to $1/\delta(\Gamma)$, where

$$\delta(\Gamma) = \min\{|T| : |\Gamma(T)| \leq 2^{|T|}\}.$$

The family A_T for a set T which realize the minimum in the definition of $\delta(\Gamma)$, seems to have threshold interval of length proportional to $1/|T|$ and this can be proved when $|\Gamma| \geq n^{\log \log n}$ using the detailed knowledge of such permutation groups [7].

As for upper bounds it seems that the issue is to show that if a Boolean function f has Fourier-Walsh coefficients which are "smeared" then the sum of influences must be large. The following (or similar) inequality is needed:

Conjecture: For every Boolean function $f : \{0,1\}^n \to \{0,1\}$

$$\sum \hat{f}^2(S) \log(1/|\hat{f}(S)|) \leq c_3 \sum \hat{f}^2(S)|S|.$$

6. Some connections to earlier work

Connectivity and the Margulis-Talagrand isoperimetric formula. Perhaps the most extensively studied problems (in random graph theory as well as in percolation theory) on critical probabilities and threshold behavior are on connectivity. Erdös and Rényi found the critical probability $p = p(n) = \log n/n$ for a random graph in $G(n,p)$ to be connected and proved that the threshold interval is of length $o(p)$. For connectivity of random subgraphs of the graph of the n-dimensional cube, Burtin found the critical probability $p = 1/2$ and results of Erdös, Spencer and Bollobás show that the threshold interval is of length $O(1/n)$. See [4], pp. 337-346.

Margulis [13] (see also [4] pp. 169-174,) proved a sharp threshold property for connectivity of random subgraphs of k-connected graphs, as $k \to \infty$, and Talagrand [16] proved sharper forms of Margulis' result. Neither the results of this paper nor

the Margulis-Talagrand theorem include the Erdös-Rényi result as a special case. For random subgraphs of the the graph of the n-dimensional cube, the Margulis-Talagrand theorem gives that the length of the threshold interval is $O(n^{-1/2})$, and Theorem 2.1 gives $O(n^{-1})$ which is sharp.

The main tool used by Margulis and sharpened by Talagrand is the following: Let $\partial_v(A) = \{v \in A : h(v) > 0\}$ be the vertex-boundary of A. Margulis [13] proved that

$$\mu_p(\partial_v(A)) \cdot \phi_p(A) \geq g(p, \mu_p(A)),$$

for some positive function $g(p, \mu_p(A))$ in $(0,1)^2$. It would be interesting to understand the connection between the influences of the variables and the quantity $\mu_p(\partial_v(A)) \cdot \phi_p(A)$.

Russo's approximate 0-1 law. Corollary 3.5 is a sharp version of a theorem of Russo [15]. A weaker version of Theorem 2.1 can be derived also from Russo's theorem itself. The derivation is not immediate since in order to apply Russo's theorem it is necessary to show that $I_f(k) = o(1)$ for every k and every p which is not very close to 0 or 1. This follows from

Lemma 6.1. *Let $p, 0 < p < 1$ be fixed. Let B be the Hamming ball in $\{0,1\}^n$ containing all the sets S such that $|S| > np$. Then for every monotone A, $\sum_{k=1}^n I_A(k) \leq \sum_{k=1}^n I_B(k)$.*

Proof: Write $S \succ R$ for $R \subset S$ and $|S| = |R| + 1$.

$$\sum I_A(k) = \sum_{S \succ R} p^{|X|-1} q^{n-|X|} (\chi_A(S) - \chi_A(R)) = \sum_{k=1}^n \sum_{S:|S|=k} p^{k-1} q^{n-k-1} \chi_A(S)(k - np).$$

To maximize this sum A should include precisely those sets S with $|S| > np$. (It can be proved by a similar argument that if B is any Hamming ball (around $(1, 1, \ldots, 1)$) and A is any monotone set such that $\mu_p(A) = \mu_p(B)$ then $\sum I_A(k) \leq \sum I_B(k)$.)

It follows from Lemma 6.1 that $\sum I_A(k) \leq c \cdot \sqrt{\frac{n}{p(1-p)}}$ and therefore if A is symmetric and p is not close to 0 or to 1 then all influences are $o(1)$.

Remark: After this paper was submitted we learned about Talagrand's paper [17] which has some overlaps with the present paper as well as with [6]. Talagrand's main result is a sharp form of BKKKL's theorem for the special case of $V_n(p)$. He also related this result via the Margulis-Russo lemma to threshold phenomena and to Russo's approximate 0-1 law. His results include Theorem 3.1 and the special case of $V_n(p)$ of Theorem 3.4.

Talagrand shows that for the influences of a Boolean function f (with $prob(f = 1) \approx 1/2$, say)

$$\sum -I_f(k)/(\log I_f(k)) \geq C,$$

for some universal constant C. This is remarkable because for some other universal constant D, if $\sum -a_i/(\log(a_i)) \geq D$ then there exists a Boolean function f, with

$prob(f = 1) \approx 1/2$ such that $I_f(k) \leq a_i$ for every i, $1 \leq i \leq n$. To construct f consider a "tribes" example with tribes of different sizes.

Kruskal-Katona formulation. A simplicial complex K is a collection of subsets of $[n]$ such that $S \in K$ and $R \subset S$ implies that $R \in K$. Let $f_k(K)$ denotes the number of sets in K of cardinality $k+1$. Switching from the model where each vertex is chosen with probability p to the model where each set of cardinality $[pn]$ is chosen with equal probability we obtain the following strong form of the Kruskal-Katona theorem for simplicial complexes with transitive group action on the vertices. A similar extension of Macaulay's theorem on complexes of monomials (multisets) also follows.

Theorem 6.2. *Let K be a simplicial complex with n vertices which is invariant under a transitive permutation group on its vertices. If $f_k(K) \leq (1 - \epsilon)\binom{n}{k+1}$ then $f_r(K) \leq \epsilon\binom{n}{r+1}$, for $r \geq k + c \log n/n$.*

Acknowledgement. We would like to thank Nati Linial for suggesting the problem and for many other useful suggestions, Noga Alon, Jeff Kahn and Aner Shalev for fruitful discussions and Itai Benyamini for pointing out the paper by Talagrand [17].

REFERENCES

1. N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New-York, 1992.
2. W. Beckner, Inequalities in Fourier analysis, Ann. Math. 102 (1975), 159-182.
3. M. Ben-Or and N. Linial, Collective coin flipping, in *Randomness and Computation* (S. Micali, ed.), Academic Press, New York, 1990, pp. 91-115. Earlier version: Collective coin flipping, robust voting games, and minima of Banzhaf value, Proc. 26th IEEE Symp. on the Foundation of Computer Science, 1985, pp. 408-416.
4. B. Bollobás, *Random Graphs*, Academic Press, London, 1985. Press, Cambridge, 1986.
5. B. Bollobás and A. Thomason, Threshold functions, Combinatorica 7 (1986) 35-38
6. J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson and N. Linial, The influence of variables in product spaces, Israel J. Math. 77(1992), 55-64.
7. P. Cameron, Finite permutation groups and finite simple groups, Bull. London Math. Soc. 13(1981), 1-22.
8. P. Erdös and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960), 17-61.
9. G. Grimmet, Percolation, Springer-Verlag, New York, 1989.
10. J. Kahn, G. Kalai, and N. Linial, The influence of variables on Boolean functions, Proc. 29-th Ann. Symp. on Foundations of Comp. Sci., 68-80, Computer Society Press, 1988.
11. J. Hastad (1988), Almost optimal lower bounds for small depth circuits, in S. Micali, ed., Advances in Computer Research, Vol. 5 :Randomness and Computation, 143-170, JAI Press, Greenwich, CT, 1988.
12. N. Linial, private communication.
13. G. Margulis, Probabilistic characteristics of graphs with large connectivity, Prob. Peredachi Inform. 10(1974), 101-108.
14. L. Russo, On the critical percolation probabilities, Z. Wahrsch. werw. Gebiete, 43(1978), 39-48.
15. L. Russo, An approximate zero-one law, Z. Wahrsch. werw. Gebiete, 61 (1982), 129-139.

16. M. Talagrand, Isoperimetry, logarithmic Sobolev inequalities on the discrete cube and Margulis'
graph connectivity theorem, Geometric and Func. Anal. 3(1993), 296-314.
17. M. Talagrand, On Russo's approximate zero-one law, Ann. of Probab. 22(1994), 1576-1587.

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY OF JERUSALEM, GIVAT RAM, JERUSALEM,
ISRAEL.
  E-mail address: ehudf@math.huji.ac.il

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY OF JERUSALEM, GIVAT RAM, JERUSALEM,
ISRAEL AND INSTITUTE FOR ADVANCED STUDIES, PRINCETON, NJ 08540.
  E-mail address: kalai@math.huji.ac.il