Sharp Thresholds of Graph properties, and the k-sat Problem

Ehud Friedgut *
(With an appendix by Jean Bourgain)

July 14, 1998

Abstract

Given a monotone graph property P, consider $\mu_p(P)$, the probability that a random graph with edge probability p will have P. The function $d\mu_p(P)/dp$ is the key to understanding the *threshold behavior* of the property P. We show that if $d\mu_p(P)/dp$ is small (corresponding to a non-sharp threshold), then there is a list of graphs of bounded size such that P can be approximated by the property of having one of the graphs as a subgraph. One striking consequences of this result is that a coarse threshold for a random graph property can only happen when the value of the critical edge probability is a rational power of n.

As an application of the main theorem we settle the question of the existence of a sharp threshold for the satisfiability of a random k-CNF formula.

An appendix by Jean Bourgain was added after the first version of this paper was written. In this appendix some of the conjectures raised in this paper are proven, along with more general results.

* Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel. Email: ehudf@math.huji.ac.il. This Paper is part of a Ph.D. thesis prepared under the supervision of Prof. Gil Kalai. Mathematics Subject Classification (1991): 05C80, 28A35 .

1 Introduction & Definitions

Consider $G(n,p)$ the probability space of random graphs on n vertices with edge probability p. We will be considering subsets of this space defined by *monotone graph properties*. A monotone graph property P is a property of graphs such that
a) P is invariant under graph automorphisims.
b) If graph H has property P than so does any graph G having H as a subgraph.
A monotone symmetric family of graphs is a family defined by such a property.
One of the first observations made about random graphs by Erdös and Rényi in their seminal work on random graph theory [12] was the existence of *threshold phenomena*, the fact that for many interesting properties P, the probability of P appearing in $G(n,p)$ exhibits a sharp increase at a certain critical value of the parameter p. Bollobás and Thomason proved the existence of threshold functions for all monotone set properties ([6]), and in [14] it is shown that this behavior is quite general, and that all monotone graph properties exhibit threshold behavior, i.e. the probability of their appearance increases from values very close to 0 to values close to 1 in a very small interval. More precise analysis of the size of the threshold interval is done in [7].
This threshold behavior which occurs in various settings which arise in combinatorics and computer science, is an instance of the phenomenon of *phase transitions* which is the subject of much interest in statistical physics. One of the main questions that arise in studying phase transitions is: how "sharp" is the transition? For example, one of the motivations for this paper arose from the question of the sharpness of the phase transition for the property of satisfiability of a random $k-$CNF Boolean formula. Nati Linial, who introduced me to this problem, suggested that although much concrete analysis was being performed on this problem the best approach would be to find general conditions for sharpness of the phase transition, answering the question posed in [14] as to the relation between the length of the threshold interval and the value of the critical probability.
In this paper we indeed introduce a simple condition and prove it is sufficient. Stated roughly, in the setting of random graphs, the main theorem states that if a property has a coarse threshold then it can be approximated by the property of having certain given graphs as a subgraph. This condition

can be applied in a more general setting such as that of the k-sat problem, where, indeed, it can be used to demonstrate the sharpness of the threshold.

Let us now define precisely the question with which we wish to deal. Consider A_n, a family of graphs on n vertices, defined by a monotone graph property P_n. Let us define what we mean by a sharp threshold vs. a coarse one, for a series of such properties:
Recall that $G(n,p)$ is actually a product space of $\binom{n}{2}$ copies of the 2 point space endowed with the product measure, and $\mu_p(A)$, the measure of A, is the probability that a random graph with edge probability p will belong to A, and is a monotone function of p. Fix $\epsilon > 0$ and for a property P, and the family A defined by it let p_0 be such that $\mu_{p_0}(A) = \epsilon$, and p_1 be defined by $\mu_{p_1}(A) = 1 - \epsilon$. Define the threshold length δ to be $p_1 - p_0$. There exists $p_c \in [p_0, p_1]$ *the critical* p such that $\mu_{p_c}(A) = 1/2$. Now for a series of properties $P(n)$ we will say that the properties have a sharp threshold if $\lim \delta(n)/p_c(n) = 0$ where $p_c(n)$ is the critical p for $P(n)$. If the ratio δ/p_c is bounded away from zero we will say that properties have a coarse threshold. (Bollobás and Thomason [6] showed that this ratio is bounded from above.) From [14] a coarse threshold for a graph property can only happen for small enough p, i.e. p bounded from above by a negative power of n. The question of understanding coarse thresholds for non-symmetric properties at values of p that are bounded from 0 is also interesting, see [13].
Example: Connectivity has a sharp threshold since the critical p is approximately $\log(n)/n$ where as $\delta \sim 1/n$. On the other hand the property of having a triangle in the graph has a coarse threshold since both the critical p and the length of the threshold interval are of magnitude $1/n$.
The first naive conjecture that one might raise is that a coarse threshold happens only for such properties, i.e. having a certain graph as a subgraph. The following example shows, however, that this conjecture must be slightly modified:
Consider the property "G is a graph on n vertices with a triangle as a subgraph, and at least $\log(n)$ edges". A moment's reflection shows that this property is probabilisticly equivalent to the previous one, and differs from it by a set of graphs with total probability which is negligible. What we suggest in this paper is that the naive conjecture is correct except for such artificial examples.

Before presenting the main theorems here are a few definitions and notations:
A *balanced graph* is a graph with average degree no smaller than that of any

of its subgraphs. A *strictly balanced graph* is one where the average degree is strictly larger than that of any proper subgraph. For example any cycle is a strictly balanced graph, where as two disjoint copies of a cycle make up a balanced but not strictly balanced graph.
For a family of graphs A we will call a graph H minimal if H belongs to A but no subgraph of H does. Let $\|A\|$ denote the number of edges of the largest minimal graph in A, when A is non empty, and define $\|A\| = 0$ when A is the empty family. Throughout this paper c will denote a constant, not necessarily the same one each time it appears. When dealing with graphs, n will denote the number of vertices and $N = \binom{n}{2}$, the number of edges in the complete graph. We will be interested in $p = p(n)$ such that p tends to zero as n tends to infinity. Let $q = 1 - p$.
For a graph H, $|H|$ will denote the number of edges in H , and $v(H)$ the number of vertices. $E(H)$ will denote the expected number of copies of H in $G(n,p)$, $E(H) = p^{|H|}\binom{n}{v(H)}\frac{v(H)!}{|Aut(H)|}$. For graphs H, S we will denote the fact that they are isomorphic by $H \sim S$. For a graph H, let $\Theta(H)$, the orbit of H, be the set of all subgraphs of the complete graph on n vertices which are isomorphic to H. So $E(H) = |\Theta(H)|p^{|H|}$. We define also another function of H which is more convenient to work with: $D(H) = n^{v(H)}p^{|H|}$. Note that for H of bounded size $D(H) \geq E(H) \geq cD(H)$.

Obviously for a property to have a coarse threshold there must be points within the critical interval for which the derivative of the function $\mu_p(A)$ with respect to p is small. More precisely:
Remark: *if $\{A_i\}$ is a series of properties with a coarse threshold, i.e. $\delta(A_n)/p_c(A_n) > C$ for all n then for each n there exists $p^* = p^*(n)$ such that p^* is in the critical interval for A_n and $p^* \cdot \frac{d\mu}{dp}|_{p=p^*} < 1/C$.*
We will attack this aspect of the problem: denoting the slope at a point p by I (for reasons to be explained) give a condition on the family A such that $p \cdot I$ is bounded from above.
We now come to our main theorem:

Theorem 1.1 *There exists a function $k(\epsilon, c)$, such that for all $c > 0$, any n and any monotone symmetric family of graphs A on n vertices, such that $p \cdot I \leq c$, for every $\epsilon > 0$ there exists a monotone symmetric family B such that $\|B\| \leq k(\epsilon, c)$ and $\mu_p(A \triangle B) \leq \epsilon$. Furthermore the minimal graphs in B are all balanced.*

What the theorem essentially means is that a family with a coarse threshold

can be approximated by a family whose minimal graphs are all small. (Notice that any monotone family is characterized by its minimal graphs.)

The following theorem seems at first sight to be slightly less informative than the previous one, it is, however, more suitable for applications, i.e. proving certain properties have a coarse threshold.

Theorem 1.2 *Let $0 < \alpha < 1$. There exist functions $B(\epsilon, c),\ b_1(\epsilon, c),\ b_2(\epsilon, c)$ such that for all $c > 0$, any n and any monotone symmetric family of graphs A on n vertices such that $p \cdot I \leq c$ and $\alpha < \mu_p(A) < 1 - \alpha$, for every $\epsilon > 0$ there exists a graph G with the following properties:*

- *G is balanced*
- *$b_1 < E(G) < b_2$*
- *$|G| \leq B$*
- *Let $Pr(A|G)$ denote the probability that a random graph belongs to A conditioned on the appearance of $\bar{G}$,* a specific copy *of G. Then*

$$Pr(A|G) \geq 1 - \epsilon$$

Note that conditioning on the appearance of, say, a triangle in $G(n,p)$ is not the same as conditioning on the appearance of three *specific* edges $(i,j),\ (j,k),\ (k,i)$ that are the edges of a specific triangle.

These two theorems can also be stated analogously for hypergraphs, and also in a slightly more general setting which is relevant in the case of the k-sat problem:

Consider a k-CNF formula on n boolean variables, i.e. a conjunction of clauses each of which is a disjunction of k of the variables and their negations. A random formula with parameter c consists of cn such clauses chosen uniformly from all $2^k\binom{n}{k}$ clauses. Let

$$P_k(c) = \Pr(\text{a random formula with } cn \text{ clauses is satisfiable.})$$

In section 5 we will prove the following, which was not known for $k > 2$:

Theorem 1.3 *For every fixed $k \geq 2$ there exists a function $c(n)$ such that for every $\epsilon > 0$*

$$P_k(c - \epsilon) \to 1.$$
$$P_k(c + \epsilon) \to 0.$$

The next theorem gives a characterization of the possible values of the critical edge probability for graph properties with a coarse threshold:

Theorem 1.4 *For any $c > 0$ and any $0 < \tau < 1/2$ there exist positive real numbers b_1, b_2, L such that for any monotone graph property A, if $pI < c$ and $\tau \leq \mu_p(A) \leq 1 - \tau$ then $b_1 n^\alpha \leq p \leq b_2 n^\alpha$ with α rational, $\alpha = -k/l$, k and l positive integers and $l \leq L$.*

In other words, coarse thresholds only happen near rational powers of n. Theorems 1.1 and 1.4 each separately imply, for example, the well known fact that connectivity has a sharp threshold. Theorem 1.4 shows this since the critical probability for connectivity is $p = \log(n)/n$. Theorem 1.1 implies this since it is possible to show that at the critical probability it is not possible to approximate connectivity by the property of having a subgraph from a list of graphs of bounded size.

We conjecture that our characterization of coarse thresholds holds in a more general setting, where symmetry plays no role: for any monotone set $A \subset \{0,1\}^n$ define

$$\|A\| = \max\left\{\sum \epsilon_i \mid \epsilon \text{ is a minimal element in } A\right\}.$$

Conjecture 1.5 *There exists a function $k(\epsilon, c)$ such that for all $c > 0$, for any A that is a monotone subset of the probability space $\{0,1\}^n$ endowed with the product measure μ_p, if $p \cdot I \leq c$, then for every $\epsilon > 0$ there exists a monotone set $B \subset \{0,1\}^n$ such that $\|B\| \leq k$ and $\mu_p(A \triangle B) \leq \epsilon$.*

This conjecture seems to be related to conjecture 2.4, that will be presented in the following section, although we are not able to show that one of them implies the other.

2 Fourier analysis & sketch of the proof

We will now define an orthonormal basis , with respect to μ, for the space of real functions on $G(n,p)$. The use of these functions and their nice properties in a similar setting is introduced by Talagrand in [28]. These functions will be indexed by all subgraphs of the complete graph on n vertices. Let E denote the set of all edges of the complete graph. Define $U_\emptyset$ to be identically equal to 1. for any edge $e \in E$ let U_e be defined as follows:

$$U_e(H) = \begin{cases} -\sqrt{q/p} & \text{if } e \in H \\ \sqrt{p/q} & \text{if } e \notin H \end{cases}$$

For any other graph R define

$$U_R = \prod_{e \in R} U_e.$$

It is not hard to check that these functions indeed are orthonormal with respect to the inner product defined by μ. For any real function f on $G(n,p)$ define $\hat{f}(H)$ as $\langle f, U_H \rangle$. This gives us the Fourier expansion of f :

$$f = \sum_H \hat{f}(H) U_H .$$

For $p = 1/2$ this is the usual Fourier-Walsh expansion of a real function on Z_2^N. For any value of p we define as usual the L_2 norm of f, and Parseval's identity holds:

$$\|f\|_2^2 = \sum \hat{f}^2.$$

A crucial property of $\hat{f}$ that we will use is that if f is symmetric, so is $\hat{f}$, i.e. if $f(H)$ depends only on the isomorphism type of H, the same is true of $\hat{f}$.
For a given edge e define $I_e(f)$, the influence of e on f to be the measure of the set of graphs H such that $f(H) \neq f(H \oplus e)$ where $H \oplus e$ is the graph obtained from H by deleting e if e is an edge of H, or adding it if it is not.
Put $I = \sum_{e \in E} I_e$. Let A be monotone, and $f = \chi(A)$.
The following three lemmas connect I with $\hat{f}$ and with $d\mu(A)/dp$.

Lemma 2.1 (Russo, Margulis) $d\mu(A)/dp = 1/p \int_A h(a) d\mu_p$
where $h(a)$ *is* $|\{a' | a' \notin A,\ dist(a,a') = 1\}|$. *(Here* $dist(a,a')$ *is the Hamming distance.)*

For proofs of this lemma see [23], [25]. An equivalent statement in different notation is:

Lemma 2.2 $d\mu(A)/dp = I$.

(Notice that I is a function of p.)
Remark: These two lemmas imply an easy converse of theorem 1.1 : if $\|A\|$ is small than so is the quantity $p \cdot I$.

Lemma 2.3 $q \cdot p \cdot I_e = \sum_{H|e \in H} \hat{f}^2(H)$. *Consequently* $q \cdot p \cdot I = \sum_H \hat{f}^2(H)|H|$.

One consequence of this last lemma, that we shall use, is as follows:

$$\sum_{H:|H|>L} \hat{f}^2(H) \leq qpI/L$$

For proof of this lemma see [28].

These lemmas seem to suggest that it may be useful to attack our problem via studying the Fourier transform of the characteristic function of a family of graphs.To further gain faith in this approach let us take a look at the Fourier transform of a given family defined by a property that has, as we have seen, a coarse threshold:
Let f be the characteristic function of A, the family of all graphs having a copy of C_3 (a triangle) as a subgraph. Choose p such that the expected number of triangles in $G(n,p)$, $E = E(C_3) = \log(2)$. Standard computations show that $\mu_p(A) \to 1/2$. It is a nice exercise in basic random graph theory to show that $\hat{f}$ exhibits the following asymptotic behavior as n tends to infinity:

$$\hat{f}^2(\emptyset) = \frac{1}{4}$$

$$\sum_{S \sim C_3} \hat{f}^2(S) \to \frac{1}{4}E$$

$$\sum_{S \sim 2 \cdot C_3} \hat{f}^2(S) \to \frac{1}{4}E^2/2!$$

$$\vdots$$

$$\sum_{S \sim k \cdot C_3} \hat{f}^2(S) \to \frac{1}{4}E^k/k!$$

where $k \cdot C_3$ is the graph consisting of k disjoint triangles. Recalling that Parseval's identity gives

$$\sum \hat{f}^2(S) = 1/2$$

and summing these figures gives that asymptotically *all* the L_2 weight of $\hat{f}$ is concentrated on these graphs. The Fourier transform is "announcing":
"f is a function that deals with triangles!"

We now give a short sketch of the proof of the theorem : Given a family A, and its characteristic function f such that $p \cdot I$ is small we will look at approximations of f: g_1, g_2, g_3. First we will truncate $\hat{f}$ i.e. set $\hat{g_1}(S) = 0$ for $|S|$ large and $\hat{g_1}(S) = \hat{f}(S)$ otherwise. Since pI is small lemma 2.3 implies that this will still leave us with a close L_2 approximation of f. Next we will show, and this will take the most effort, that most of the L_2 norm of f, i.e. most of the weight of $\hat{f}^2$ is concentrated on a small number of nicely behaved graphs: balanced graphs such that the expected number of copies of them in a random graph with edge probability p is bounded. So now we define $\hat{g_2}$ to be the same as $\hat{g_1}$ but leave its support only on such "nice" graphs. Next we show that such a function as g_2 "Counts" appearances of these nice graphs, in the sense that its value on a graph H can, with a high probability, be approximated very closely just by knowing the number of subgraphs of H isomorphic to each of our nice graphs. Finally, g_2 is not necessarily Boolean, but the fact that it is close to the original f in the L_2 norm shows that f can also be approximated by a Boolean function g_3 that "Counts" appearances of the nice graphs, e.g. g_3 might be of the form : $g_3(H) = 1$ iff H has as a subgraph a triangle or at least two copies of C_4.

Recalling conjecture 1.5 perhaps this is the place to raise the following conjecture about the Fourier coefficients of any monotone Boolean function on the discrete cube;
Consider the probability space $\{0,1\}^n$ endowed with the product measure μ_p. It is trivial to generalize the definitions of this section to this setting and to define $\hat{f}$ for any $f : \{0,1\}^n \to \mathbf{R}$.

Conjecture 2.4 *Let $f : \{0,1\}^n \to \{0,1\}$ be the characteristic function of A, a monotone subset of $\{0,1\}^n$. Let $p = p_c(A)$. For $\tau > 0$ let*

$$\Omega_\tau = \{S | \hat{f}^2(S) < \tau p^{|S|}\}.$$

Let $p_c \to 0$ $\tau \to 0$ and $n \to \infty$. Then if $p_c \cdot d\mu(A)/dp|_{p=p_c} < c$ then

$$\sum_{S \in \Omega_\tau} \hat{f}^2(S) = o(1)$$

This conjecture is proven in the appendix.

3 Some Lemmas on Random Graphs

In this section we wish mainly to study certain functions on the probability space $G(n,p)$. These functions play a key role in our proof, since we will expand f, the characteristic function of the family of graphs we consider, as a linear combination of these functions. For any graph S we define

$$V = V_S = \sum_{H \in \Theta(S)} U_H.$$

We wish to give an expression that approximates the value of V_S in simple terms. For a given graph R let X_R be the random variable counting the number of copies of R in a random graph. Let $X_\emptyset$ be defined to be identically 1. Although given a certain R, some copies of R appear as subcopies of other subgraphs of S, the value of V_S on a graph is determined by the value of X_R for all R that are subgraphs of S. The following lemma gives a convenient expression for V in terms of the X_R's.

Lemma 3.1 *For any fixed graph S*

$$V_S = (\sqrt{1/qp})^{|S|} E(S) \left(\sum_{R \subseteq S} (-1)^{|R|} \frac{X_R}{E(R)} \right).$$

Proof: Let H be a fixed copy of S, and for every $H' \subseteq H$ let $Y_{H'}$ be an indicator random variable taking the value 1 iff all edges of H' appear in the random graph. U_H is determined by the maximal H' such that $Y_{H'} = 1$,

$$U_H = (-\sqrt{q/p})^{|H|} (-p/q)^{(|H|-|H'|)}$$

This can be expressed as follows:

$$U_H = (-\sqrt{q/p})^{|H|} \left(\sum_{H' \subseteq H} (-p/q)^{(|H|-|H'|)} \sum_{H' \subseteq H'' \subseteq H} (-1)^{(|H''|-|H'|)} Y_{H''} \right). \quad (1)$$

In calculating V_S we sum (1) on all $H \in \Theta(S)$. Using $X_H = \sum Y_H$, we get:

$$V_S = (\sqrt{q/p})^{|H|} \sum_{H'' \subseteq H} (-1)^{|H''|} \sum_{H' \subseteq H''} (p/q)^{|H|-|H'|} \; |\Theta(H)|/|\Theta(H'')| X_{H''}$$

Using $|H| - |H'| = (|H| - |H''|) + (|H''| - |H'|)$ this gives

$$V_S = (\sqrt{q/p})^{|S|} \left(\sum_{R \subseteq S} X_R (-(1+p/q))^{|R|} \frac{|\Theta(S)|}{|\Theta(R)|} (p/q)^{|S|-|R|} \right).$$

□

Remark: The interested reader may take a look once again at the family of graphs having a triangle as a subgraph. The values of the Fourier coefficients as given in section 2 together with lemma 3.1 give a good understanding of the exact structure of the Fourier transform of the characteristic function of the family.
We now take a closer look at the sum defining V_S: Let e be an edge of the complete graph on n vertices. For any graph H we define

$$V = V_{H,e} = \sum_{R \in \Theta(H), e \in R} U_R.$$

Note that

$$V_S = 1/|S| \sum_e V_{S,e}.$$

We wish to study the behavior of such functions, in particular to give a bound B such that for $V = V_{H,e}$, $Pr(|V| > \lambda B)$ decays like $1/\lambda^4$. Note that the expected value of V is 0, since it is orthogonal to the identity.
Let $\tilde{H}$ be a non-empty subgraph of H such that $D(\tilde{H})$ is minimal, i.e.

$$\forall R \subseteq H,\ D(\tilde{H}) \leq D(R).$$

Note that $\tilde{H}$ must be an induced subgraph.

Lemma 3.2 $Prob\Big(|V| \geq \lambda \sqrt{|\Theta(H)|}/(D(\tilde{H})^{1/4})\Big) \leq c\lambda^{-4}/n^2$.

Proof: The lemma will follow from a bound on the 4th moment of V:

$$E(V^4) = E\left(\left(\sum_{R \in \Theta(H), e \in R} U_R\right)^4\right) \leq cE\left(\sum_S |\Theta(S)| n^{-2} \sum_L \prod_{R \in L} |U_R|\right). \tag{2}$$

Where on the right hand side the first sum is over all isomorphism types of graphs S such that S is a union of 4 copies of H having an edge e in their intersection. The second sum is on L's that are quadruples of copies of H with an edge e common to all 4 of them, such that their union is S.
Let us compute the contribution of a given S to the sum. Assume S is the

union of 4 copies of H. Let c_e be the number of times an edge e of S is covered by these copies.

$$E(\prod U_R) = E(\prod U_e^{c_e}) = \prod E(U_e^{c_e}) = \prod \left(p(-\sqrt{q/p})^{c_e} + q(\sqrt{p/q})^{c_e} \right).$$

We have used the fact that for different edges $e \neq e'$ the values of U_e and $U_{e'}$ are independent random variables. Now if $c_e = 1$ for some edge e then $E(\prod U_R) = 0$. Otherwise the dominant summand corresponding to e is $p(-\sqrt{q/p})^{c_e}$ and

$$E(\prod U_R) \sim p^{|S|}/\sqrt{p}^{(4|H|)}. \tag{3}$$

Note that the number of summands in the double sum $\sum_S \sum_L$ is no more than some constant depending on H. Substituting $E(S) = |\Theta(S)|p^{|S|}$ and (3) in (2) we have

$$E(V^4) \leq c \cdot p^{-2|H|} \max_S E(S)/n^2 \tag{4}$$

where the maximum is only taken over graphs S that can be covered by 4 copies of H with an edge in common, each edge being covered at least twice. Recall that $E(G) \leq D(G)$. Therefore we can replace E by D in (4) and have

$$E(V^4) \leq c \cdot p^{-2|H|} \max_S D(S)/n^2 \tag{5}$$

Claim: The S for which $D(S)$ is maximal among the graphs in question consists of the union of 2 copies of H, whose intersection is exactly $\tilde{H}$.
Using this S in (5) will give

$$E(V^4) \leq c \cdot n^{2v(H)-2}/D(\tilde{H}).$$

Recall that $|\Theta(H)| \approx cn^{v(H)}$, and the desired result follows from Markov's inequality .
Proof of claim: In searching for the best S we are trying to optimize the function $D = n^{v(S)}p^{|S|}$ on S that is double-covered by 4 copies of H with a non empty intersection. Instead let us optimize a function $\tilde{D}$ which allows edges and vertices to be covered only once and is identical with D when S is double covered. Given a graph R *and* a covering of it by copies of H having an edge in common define for any edge or vertex x, $\Phi(x)$ to be 1 if x is covered by more than one copy of H, and $1/2$ if x is covered only once. Now let $\tilde{e}(R) = \sum \Phi(e)$, $\tilde{v}(R) = \sum \Phi(v)$, and $\tilde{D}(R) = n^{\tilde{v}(r)}p^{\tilde{e}(R)}$. It is obvious

that the maximum of $\tilde{D}$ is at least as large as the maximum of D, and that if this maximum is equal to the value obtained by D on the S defined above we are done. Let us build a graph F which is a union of 4 copies of H with a specific edge e belonging to their intersection. Adding the copies of H one by one and keeping track of how much each additional copy contributes to the value of $\tilde{D}$ we have that the first 2 copies contribute $\sqrt{D(H)}$ and the next 2 no more than $\sqrt{D(H)/D(\tilde{H})}$. The conclusion is $D(S) \geq \tilde{D}(F)$ as desired.

□

Corollary 3.3 *Let*

$$\chi = \chi\left\{|V_{e,H}| > \sqrt{|\Theta(H)|}/D(\tilde{H})^{1/4}\right\}.$$

Then

$$\int |V_{e,H}| \cdot \chi \leq c(\sqrt{|\Theta(H)|}/D(\tilde{H})^{1/4})/n^2.$$

Before proceeding to the proof of the main theorem there is one more simple lemma we will need about the number of appearances of a given graph as a subgraph of a random graph.

Lemma 3.4 *Let R be a fixed graph, and X_R be a random variable equal to the number of copies (not necessarily induced copies) of R that appear in a random graph $G(n,p)$. Assume $p = o(1)$. Then*

$$Var(X_R) \asymp E(R)^2 \sum_{H \subseteq R} 1/E(H)$$

where the sum is over all nonempty subgraphs of R.

Proof: $\mathrm{Var}(X_R)$ is given by the formula

$$\sum E(XY) - E(X)E(Y)$$

where the sum is over all pairs (X,Y) that are indicator random variables indicating whether a specific copy of R appeared in the random graph. If X and Y are independent the corresponding summand is 0, otherwise $E(X)E(Y) = o(E(XY))$. Partitioning the sum $\sum E(XY)$ according to the isomorphism type of the intersection of the two copies of R indicated by X and Y gives the desired result.

□

4 The Proof

4.1 Proof of the Main Theorems

In this subsection we present the proofs of theorems 1.1 and 1.4 using some lemmas whose proof we put off to the following subsection. As usual let A be a monotone symmetric family of graphs, and f be its characteristic function. We now wish to extract information on f by analyzing $\hat{f}$. Let H be a graph. Choosing certain bounds L, c_1, c_2 call a graph H modest if
1)$|H| \leq L$.
2)$c_1 \leq E(H) \leq c_2$.
3)H is balanced.

Remark 4.1 *Note that once given the parameters L, c_1, c_2 that define modesty, for sufficiently large n and small p this determines the average degree of the modest graphs (if any indeed exist.) Let this degree be δ. This enables us to define modesty by new parameters L, c_1' and c_2' such that all balanced graphs with no more than L edges and average degree δ are also modest. furthermore c_1' and c_2' are simple functions of c_1, c_2 and L. It will be convenient to always assume such a choice, so we may assume later that if H is modest all subgraphs of H of the same average degree are also modest.*
Moreover, note that for all subgraphs $R \subset H$, $E(R)$ is bounded from below. (Graphs with small expectation have large average degree.)
Note also that the average degree of a finite union of modest graphs must have average degree larger or equal to that of the modest graphs.

Lemma 4.2 *Let A be a monotone symmetric family of graphs on n vertices, and f be its characteristic function. Assume $pI \leq c$. Then for every $\epsilon > 0$ there exist constants L, c_1, c_2 such that for sufficiently large n*

$$\sum_{S \text{ is not modest}} \hat{f}^2(S) \leq \epsilon.$$

This Lemma immediately implies theorem 1.4:
Proof: Let f be such that $pI < c$, and let Ω be the set of graphs S with $|S| \leq L$ and with $c_1 \leq E(S) \leq c_2$. Lemma 4.2 implies that

$$\sum_{S \notin \Omega} \hat{f}^2(S)$$

is small. If $\sum \hat{f}^2 = \tau$ (i.e. $Pr(f = 1) = \tau$), we may conclude that Ω is non-empty, i.e. there exist graphs S with less than L edges with $c_1 \leq E(S) \leq c_2$ which implies that p must be in the range asserted by the theorem. (Note that since $Pr(f = 1)$ is bounded away from 1 we cannot approximate f by the function that is identically equal to 1, corresponding to the case where Ω has only the empty graph as a member in it.)

□

We now present the proof of theorems 1.1 and 1.2:
Proof: Let $S_1, S_2, \ldots, S_l$ be a list of all the modest graphs. For any graph S let C_S be the set of all graphs T on n vertices such that the union of all copies of the S_i's that appear as a subgraph of T is isomorphic to S. We will subdivide the space of subgraphs of the complete graph on n vertices into these disjoint sets and approximate f separately on each. While doing this we will define certain parts of our space in which rare events occur:
1)$\mathcal{C}_1$ is the union of C_S for S which are large.
2)$\mathcal{C}_2$ is the union of C_S for which $\mu(C_S)$ is small.
3)$\mathcal{C}_3$ is the union of C_S for which $\Pr(f = 1)$ is not close to 0 or 1.
Let $\mathcal{C}_4$ be the union of the remaining sets. We will now define a sequence of approximations of f, according to small constants $\epsilon_1, \epsilon_2, \ldots \epsilon_5$. Our final approximation of f will be equal to 1 on a graph R iff R has a subgraph S such that $C_S \subset \mathcal{C}_4$ and f is equal to 1 on most of C_R.

1) Let $g_1 = \sum \hat{f}(S) V_S$ where the sum is only on modest graphs S, and $V_S = \sum_{G \in \Theta(S)} U_G$. By lemma 4.2 we can choose our bounds so that $\|f - g_1\|_2 \leq \epsilon_1$.

2)Let $\mathcal{C}_1$ be the union of all C_S such that $X_{S_i}(S) \geq lE(S_i)/\epsilon_2$, for some i, i.e. the number of copies of S_i appearing is much more than the expected. The measure of $\mathcal{C}_1$ is no more than ϵ_2. Define g_2 to be equal identically to 1 on $\mathcal{C}_1$, and equal to g_1 otherwise. So $\|f - g_2\|_2 \leq \epsilon_1 + \sqrt{\epsilon_2}$.

3)Recalling that $E(S_i)$ is bounded we have that the number of graphs in

$$\{S | C_S \neq \emptyset, C_S \not\subset \mathcal{C}_1\}$$

is bounded i.e., we have a bound on the number of subsets C_S on which g_2 is not identically 1. Say we have M such sets. Let $\mathcal{C}_2$ be the union of all C_S not in $\mathcal{C}_1$ such that $\mu(C_S) < \epsilon_3/M$. So $\mu(\mathcal{C}_2) \leq \epsilon_3$.

Remark: The reason we treat this part of our space separately is because we need a lower bound on the measure of C_G in the proof of lemma 4.14 below. Let $\mathcal{C}$ be the union of all remaining sets C_S, those not in $\mathcal{C}_1 \cup \mathcal{C}_2$.

Remark 4.3 *Note that if S is such that $C_S \subset \mathcal{C}$ then S must be balanced and with average degree the same as all the modest graphs. Furthermore the size of S is bounded from above, and $E(S)$ is bounded from above and below.*

For any $C_S \subset \mathcal{C}$ and for any graph $H \in C_S$ define

$$g_3(H) = E(g_2|C_S),$$

i.e. we replace g_2 by its conditional expectation in C_S. Define g_3 to be equal to 1 on all graphs not in $\mathcal{C}$. We will show that for any $C_S \subset \mathcal{C}$, g_3 is close to g_2 because g_2 is almost constant on C_S in the sense that for any constant δ,

$$Pr(\{|g_2(T) - E(g_2|C_s)| > \delta\}|T \in C_S) \to 0. \tag{6}$$

This will be proven in lemma 4.14 in the following subsection. Recalling there are only M sets C_S in $\mathcal{C}$ we get that by choosing δ small enough by proper choice of $\epsilon_1, \epsilon_2, \epsilon_3$ we have

$$\|g_3 - f\|_2^2 \leq \epsilon_4/2.$$

We now replace g_3 by g_4 which is defined as follows:

4)For any S let g_4 on the graphs in C_S be identically 0 or 1 according to which approximates f better.
We now wish to compare two approximations of f : g_3 which is constant on each C_S, and g_4 which is the best approximation that is both constant on each C_S and Boolean. Let h be the best possible L^2 approximation of f that is constant on C_S. a simple calculation shows that

$$h|_{C_S} = Pr(f(R) = 1|R \in C_S).$$

The following inequality follows by summing over each C_S separately:

$$\|f - g_4\|_2^2 \leq 2\|f - h\|_2^2 \leq 2\|f - g_3\|_2^2 = \epsilon_4. \tag{7}$$

g_4 is the characteristic function of a family B that is a candidate to be the family guaranteed by the theorem. Yet we do not know that B is monotone

and that $\|B\|$ is small.

5) We now define g_5, a monotone boolean function which is constant on each C_S. We will define it such that if $R \subset S$ then $g_5|_{C_R} \leq g_5|_{C_S}$. Call a graph S *decisive* if

$$\Pr(f(T) = 1 | T \in C_S) \notin (\sqrt{\epsilon_4}, 1 - \sqrt{\epsilon_4}).$$

Let $\mathcal{C}_3$ be the union of all sets C_S for S which is not decisive. From (7) we have that $\mu(\mathcal{C}_3) \leq \sqrt{\epsilon_4}$. Let $\mathcal{C}_4$ be $\mathcal{C} \setminus \mathcal{C}_3$. For any S define g_5 on C_S to be equal to 1 iff S has a subgraph R such that $C_R \subset \mathcal{C}_4$, with g_4 equal to 1 on C_R. Since the union $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$ is of small measure, the alterations on C_S that belong to these parts of our space do not affect our approximation much. We will show in lemma 4.8 below that if $R \subset S$, and C_R and C_S belong to $\mathcal{C}_4$ then

$$E(f|C_R) \geq 1 - \sqrt{\epsilon_4} \Rightarrow E(f|C_S) \geq 1 - 2\sqrt{\epsilon_4}$$

and hence on the sets $C_S \subset \mathcal{C}_4$ we do not alter our approximation at all, therefore choosing sufficiently small ϵ_4 we have

$$\|g_5 - f\|_2^2 \leq \epsilon.$$

g_5 is the characteristic function of a symmetric monotone family B with minimal graphs which are balanced and of bounded size, and

$$\mu(B \triangle A) \leq \epsilon.$$

This completes the proof of the main theorem. Furthermore by repeating this process, possibly with a different choice of ϵ_i, $i = 1, \ldots 4$ we can define $\mathcal{C}_4$ such that for $C_R \subset \mathcal{C}_4$ $E(f|C_R) > 1 - \epsilon$. Since we have a bound on $E(R)$ and $|R|$ for all such graphs R, any one of them is a candidate to be the graph guaranteed by theorem 1.2. Let r be a specific copy of such a graph R , and let B_r be the space of all graphs having r as a subgraph. Let $C_r = B_r \cap C_R$. From symmetry

$$E(f|C_r) = E(f|C_R)$$

and from positive correlation of increasing events

$$E(f|B_r) \geq E(f|C_r).$$

In other words, conditioning on the appearance of r the expectation of f is at least $1 - \epsilon$. This concludes the proof of theorem 1.2.

□

4.2 Proof of the main Lemmas

We will need the following observation during the proof:
Remark: If there exists a graph H of bounded size such that $E(H)$ is a constant, then there exists a constant c such that for any other graph G of bounded size and any fixed m, if n is large enough then

$$E(G) < \log(n)^m \Rightarrow E(G) < c.$$

This is true since the fact that $E(H)$ is constant implies that $p = cn^{-v(H)/|H|}$.

We now proceed to prove lemma 4.2:
Proof: Recalling the remark following lemma 2.3 we know that most of the weight of $\hat{f}^2$ is concentrated on graphs H with no more than a certain constant number of edges. We now wish to further characterize the graphs H such that $\sum_{S\sim H} \hat{f}^2(S)$ is significant. A simple calculation shows that for any Boolean function f and any graph S

$$\hat{f}^2(S) \leq (4p)^{|S|}. \tag{8}$$

Summing this on the orbit of S we get

$$\sum_{S\sim H} \hat{f}^2(S) \leq cE(H). \tag{9}$$

So if H has small expectation its orbit does not contribute much to the weight of $\hat{f}^2$. The following lemma shows this is true even if H has a subgraph of small expectation, i.e. if a graph H is such that one does not expect to see a copy of it in $G(n,p)$ then the weight of $\hat{f}^2$ on its orbit is negligible.

Lemma 4.4 *Let H be a graph. Then for every subgraph H'*

$$\sum_{S\in\Theta(H)} \hat{f}^2(S) \leq c \cdot \max_{R\subset H'}\{E(H')/E(R)\}.$$

(Where the expectation of the empty graph is taken to be 1.)

This lemma is of course interesting to us in the case where $E(H')$ is small. If H has a subgraph with small expectation, it has one that is minimal with respect to inclusion, and we can use the lemma with respect to that subgraph.
Proof: Let R be a specific copy of H'. Consider the probability space $\{0,1\}^{E\setminus R}$, where $E \setminus R$ is the set of edges of the complete graph not in R.

We view this as the space of random graphs on n vertices where one copy of H', is fixed (chosen with probability 1), and all other edges are chosen with usual probability p. We define the set of functions $\{U_S\}$ as before, and have a Fourier expansion for any real function on this space.
Define a function g on this space by $g(G) = f(G \cup R)$. Note that g is symmetric in the sense that it is invariant under automorphisms of the complete graph that keep R fixed. Using induction on $|R|$ We will show that

$$\hat{g}(G) = \sum_{R' \subseteq R} \hat{f}(R' \cup G) U_{R'}(R). \tag{10}$$

For any G in this new space define $\tilde{\Theta}(G)$ to be the new orbit of G , under the action of the automorphisms of the complete graph that keep R fixed. Since g is Boolean

$$|\hat{g}(G)| \leq 1/\sqrt{|\tilde{\Theta}(G)|}.$$

Using (10) we have for any graph G such that its edge set is disjoint from R :

$$\left|\hat{f}(R \cup G)/\sqrt{p}^{|R|}\right| \leq c\left(1/\sqrt{|\tilde{\Theta}(G)|} + \sum_{R' \subset R}\left|\hat{f}(R' \cup G)/\sqrt{p}^{|R'|}\right|\right) \tag{11}$$

$$\leq c\left(1/\sqrt{|\tilde{\Theta}(G)|} + \sum_{R' \subset R}\left(\sqrt{p}^{|R'|}\sqrt{|\Theta(R' \cup G)|}\right)^{-1}\right). \tag{12}$$

Note that for $R' \subseteq R$,

$$|\Theta(R' \cup G)| \geq c|\Theta(R')| \cdot |\tilde{\Theta}(G)|. \tag{13}$$

The value of c in this preceding inequality may depend on the graphs involved, however in our case we will be using this inequality for a finite number of graphs, hence it holds with a fixed c. Observe that when $v(R') \cap v(G) = v(R) \cap v(G)$, the two sides of (13) are comparable, i.e.

$$|\Theta(R' \cup G)| \leq c'|\Theta(R')| \cdot |\tilde{\Theta}(G)|.$$

Let G be such that $G \cup R = H$. Multiplying both sides of (12) by $\sqrt{|\Theta(H)|p^{|R|}}$ and using (13) gives

$$\sqrt{\sum_{S \in \Theta(H)} \hat{f}^2(S)} \leq c\left(\sqrt{E(R)} + \sum_{R' \subset R}\sqrt{E(R)/E(R')}\right)$$

which gives the desired result.
What is left to show is the validity of formula (10):
Assume first that $R = e$, R is a single edge. In this case

$$\hat{g}(G) = \sum_{e \notin T} f(T \cup e) U_G(T) p^{|T|} q^{N-1-|T|}.$$

The right hand side of (10) is

$$\hat{f}(G) - \sqrt{q/p}\hat{f}(G \cup e) = \sum_{M} f(M) \left(U_G(M) - \sqrt{q/p} U_{G \cup e}(M) \right) p^{|M|} q^{N-|M|}.$$

Now, for M such that $e \notin M$ the corresponding summand is zero. Setting $M = T \cup e$ and using the fact that $U_{G \cup e}(M) = -(\sqrt{q/p}) U_G(M \setminus e)$ gives the desired result.
For $|R| > 1$ pick $e \in R$ and define $g(G) = h(G \cup e)$, where h is a function on $\{0,1\}^{E \setminus (R \setminus e)}$. We already know that

$$\hat{g}(G) = \hat{h}(G) - (\sqrt{q/p})\hat{h}(G \cup e)$$

and the result follows by using the induction hypothesis on h.

□

Assume that H is such that $W = \sum_{R \in \Theta(H)} \hat{f}^2(R)$ is bounded away from zero by some constant. We already know that H is of bounded size, and with expectation bounded from 0. We will further characterize H, by examining a few cases, and thus prove the following claim:

Claim 4.5 *Under the above conditions H must be balanced, and of bounded expectation.*

Proof: First we define a set of functions $\{f_e\}$ such as those defined by Talagrand in [28], which are a generalization of similar functions defined in [18]. The idea behind these functions is that they measure I_e, the influence of the edge e on f:
For every edge e let f_e be the function defined by:

$$f_e(H) = \begin{cases} q(f(H) - f(H \oplus e)) & \text{if } f(H) = 1 \\ p(f(H) - f(H \oplus e)) & \text{if } f(H) = 0 \end{cases}$$

It is not hard to verify that

$$\hat{f}_e(H) = \begin{cases} \hat{f}(H) & \text{if } e \in H \\ 0 & \text{if } e \notin H \end{cases}$$

By lemma 2.3 $qpI = \sum \|f_e\|_2^2$. A simple calculation gives

$$\|f_e\|_2^2 = 1/2\|f_e\|_1$$

and hence

$$qpI \geq c \sum_e \int |f_e|. \tag{14}$$

Recall that $\tilde{H}$ was defined as a non-empty subgraph of H on which the function D was minimal, and also $W = \sum_{R \in \Theta(H)} \hat{f}^2(R)$. Let $\Theta = \Theta(H)$. For any $R \in \Theta$, $\hat{f}(R) = \sqrt{\frac{W}{|\Theta|}}$. We will now proceed to analyze $\hat{f}$ by looking at the following expansion of f:

$$f = \sum_H \hat{f}(H) V_H = \sum_H 1/|H| \sum_e \hat{f}(H) V_{e,H}.$$

Using the orthogonality of the functions U_H we have that for any two functions f, g

$$\int f \cdot g = \sum \hat{f} \cdot \hat{g}$$

and in particular using symmetry and the definition of f_e

$$\int f_e \cdot \hat{f}(H) V_{e,H} = \sum_{R \in \Theta, e \in R} \hat{f}^2(H)$$

so,

$$W = (1/|H|) \sum_e \int f_e \hat{f}(H) V_{e,H} = (1/|H|) \sqrt{\frac{W}{|\Theta|}} \sum_e \int f_e V_{e,H}. \tag{15}$$

(This technique of calculating the Fourier coefficients is similar to that used in [7].)
Let

$$\chi = \chi \left\{ V_{e,H} > \sqrt{|\Theta(H)|}/D(\tilde{H})^{1/4} \right\}.$$

Then (15) is equal to

$$(1/|H|)\sqrt{\frac{W}{|\Theta|}}\sum_e \left(\int f_e V_{e,H}\chi + \int f_e V_{e,H}(1-\chi)\right).$$

Using corollary 3.3 and the fact that $|f_e| \leq 1$, we get that this is bounded by

$$c_1(1/|H|)\sqrt{\frac{W}{|\Theta|}}\left((c_2+pI)\sqrt{|\Theta(H)|}/D(\tilde{H})^{1/4}\right) = cpI\sqrt{W}/D(\tilde{H})^{1/4} \quad (16)$$

So

$$D(\tilde{H}) \leq c(pI)^4/W^2. \quad (17)$$

We now assume that for a given H, $W = \sum \hat{f}^2 > \tau > 0$ for some constant τ and wish to prove that H is balanced and $E(H)$ is bounded. We will divide this into 3 cases:
Case 1) $\tilde{H}$ consists of a single edge.
In this case (17) implies that $p = c/n^2$. For such values of p $G(n,p)$ is almost surely a disjoint union of edges, and both the lemma and the main theorems hold. The graph property in question may be approximated by the property of having at least k edges for some $k \approx (pI)^2$, and the only graphs H of bounded size for which $E(H)$ is not $o(1)$ are graphs that are a matching, so that $E(H)$ is also bounded from above.

Case 2) $\tilde{H} = H$.
In this case (17) shows that $E(H)$ is bounded by a constant, and lemma 4.4 shows H must be balanced since it has no subgraphs of small expectation.

Case 3) The case where $\tilde{H}$ is properly contained in H but is not a single edge. As before $\tilde{H}$ must be balanced and of expectation that is approximately constant. We will assume also that $\tilde{H}$ is strictly balanced. If it is only balanced but not strictly balanced, replace $\tilde{H}$ by a strictly balanced subgraph of it (Since this subgraph has the same average degree its expectation is a power of the expectation of $\tilde{H}$, and hence also bounded from above and below.)
Choose a specific copy of $\tilde{H}$, call it S, and define a function g on $Z_2^{(N\setminus|S|)}$:

$$g(R) = f(R \cup S)$$

in other words g is the same as f, but its domain is all graphs with set of edges disjoint from that of S. By analyzing g we will show that $H \setminus \tilde{H}$ is balanced, and hence H is balanced.

Remark 4.6 *In order to prove that $H \setminus \tilde{H}$ is balanced , and conclude that H is balanced we must interpret correctly the meaning of average degree of a graph, orbit of a graph, expectation, etc. for a graph in our new space. For any graph in our new space the edge set is well defined, we define the vertex set to be only those vertices not in $v(S)$, and the new orbit is defined by all automorphisms of the complete graph leaving S fixed.*

As we have seen in (10)

$$\hat{g}(R) = \sum_{S' \subseteq S} \hat{f}(R \cup S') U_{S'}(S). \tag{18}$$

Since the weight of $\hat{f}^2$ on any orbit of size $|\Theta(G)|$ is no more than a constant ($\|f\|_2 \leq 1$) we have that $|\hat{f}(G)| \leq c/\sqrt{|\Theta|}$. (And in the case of H we have equality).
Recalling that $E(S') = |\Theta(S')|p^{|S'|}$, and $U_{S'}(S) = (-\sqrt{q/p})^{|S'|}$ this gives

$$|\hat{f}(R \cup S') U_{S'}(S)| \leq c/\sqrt{E(S')}.$$

The fact that S is strictly balanced and of expectation bounded by a constant means that $E(S') >> E(S)$ for all proper subgraphs S' of S.
So putting $R = H \setminus S$ in (18) we have that the dominant summand is the one corresponding to $S' = S$, thus

$$|\hat{g}(H \setminus S)| \sim |\hat{f}(H)|/(\sqrt{p})^{|S|}. \tag{19}$$

Let $W = \sum_{G \in \tilde{\Theta}} \hat{g}^2(G)$ where $\tilde{\Theta}$ is the orbit of $H \setminus S$ in our new space. Note that

$$|\tilde{\Theta}(H \setminus \tilde{H})||\Theta(\tilde{H})| = c'|\Theta(H)|.$$

(19) now gives

$$W = c'|\tilde{\Theta}(H \setminus \tilde{H})|\hat{f}^2(H)/p^{|S|} = c/(|\Theta(\tilde{H})|(p^{|S|}) = c/E(\tilde{H})$$

i.e. W is approximately constant.
The following computation shows that $qpI(g)$ is no more than polylogarithmic:
We will now use Russo's lemma via the interpretation that pI is the expected number of "pivotal" edges, to lend a term from percolation theory:
For a given monotone family of graphs A and a given graph $G \in A$ (which is a subgraph of the complete graph on n vertices) a *pivotal* edge is an edge $e \in G$ such that $G \setminus e$ does not belong to A. If G is a random graph, the number of pivotal edges is a random variable C, and $E(C) = pI$. Let $X_{\tilde{H}}$ be a random variable that counts the number of copies of $\tilde{H}$ in $G(n,p)$. By abuse of notation let S also be the event "S appears in $G(n,p)$". (Recall S is a *specific* copy of $\tilde{H}$.) We have

$$E(C|S) = \sum E(C|X_{\tilde{H}} = i)Pr(X_{\tilde{H}} = i|S)$$

$$= \sum E(C|X_{\tilde{H}} = i)Pr(X_{\tilde{H}} = i)Pr(S|X_{\tilde{H}} = i)/Pr(S)$$

$$= \sum E(C|X_{\tilde{H}} = i)Pr(X_{\tilde{H}} = i) \cdot i/E(\tilde{H}).$$

This expression can not differ by more than a polylogarithmic factor from the expression for the unconditioned expectation of C because of the following lemma.

Lemma 4.7 *Let H be a strictly balanced graph, and p such that $E(H) < C$. Then for every k the probability of having more than* $\log(n)$ *copies of H in $G(n,p)$ is asymptoticaly less than n^{-k}.*

Proof:
Let B be a large integer to be determined, and divide the event of having $\log(n)$ copies of H into two subevents:
1)The number of copies of H that intersect other copies is at most B.
2)More than B copies intersect other copies.
Denoting $\log(n) - B = R$ the Probability of (1) is bounded by a constant times $C^R/R!$, asymptotically less than $n^{-k}/2$ for any fixed B. This follows from the usual way of computing the moments of the random variable X_H, see for example [5].
The event (2) can be described by the existence of a subgraph from a list of subgraphs whose length is a function of B, $l(B)$. Any graph in this list can be described by a union of at least B but no more than $2B$ copies of H.

It follows from the fact that H is strictly balanced that for sufficiently large B the expected number of copies of any graph in this list is smaller than $n^{-(k+1)}$, for all n large enough. So for sufficiently large B the probability of (2) is asymptotically less than $n^{-k}/2$.

□

We now can proceed to analyze $H \setminus \tilde{H}$ in the same manner we analyzed H: The weight of $\hat{g}^2$ on the orbit of $H \setminus \tilde{H}$ is a constant and $qpI(g)$ is no more than a power of $\log(n)$, so we once again can divide into three cases, as before. Even though pI is now logarithmic rather than constant our analysis is the same because of the remark that opens this section: once we know that $E(\tilde{H})$ is constant any graph with expectation bounded by a power of $\log(n)$ must also have bounded expectation. So we once again divide into 3 cases. In cases 1 or 2 we are done, and in case 3 we iterate the computation again. Since the size of H was bounded a priori, this process will terminate after a finite number of steps. This shows that H can be built by Taking $\tilde{H}$, adding a graph which is the minimal strictly balanced subgraph of $H \setminus \tilde{H}$ and so on. This must result with a balanced graph, since in each step we add a balanced graph from our new spaces. We also saw that $E(H)$ is bounded. This completes the proof of claim 4.5.

□

To summarize, if $\sum_{R \sim H} \hat{f}^2(R)$ is not to small, H must be modest:
1)$|H|$ is bounded by lemma 2.3.
2)$E(H)$ is bounded from below by (9).
3)H is balanced and with expectation bounded from above by claim 4.5.
Hence we have completed the proof of lemma 4.2.

□

The following lemma deals with the approximation of the monotone function f in the proof of the main theorems. It implies that the approximation itself is "approximately monotone".

Lemma 4.8 *Let $R \subset T$ be graphs, and $C_R, C_T, \mathcal{C}_4$ be defined as in the proof of the main theorems. Suppose $C_R \subset \mathcal{C}_4$, and $C_T \subset \mathcal{C}_4$. Then*

$$E(f|C_T) > E(f|C_R) - o(1).$$

Proof: Recall that all the four quantities $E(R)$, $E(T)$, $\mu(C_R)$, $\mu(C_T)$ were, by definition of $\mathcal{C}_4$, and by remark 4.3 bounded from above and from below by some constants (that depended on ϵ) and that R and T must be balanced, with a bounded size and average degree δ equal to that of the modest graphs. Let $r \subset t$ be specific copies of R and T. (Perhaps this is not the most successful notation, but as opposed to, say, $\bar{R}$ and $\bar{T}$ it is one that can be noticed by the optically challenged who might not be able to distinguish $C_{\bar{T}}$ from C_T). Define B_r to be the space of all graphs that have r as a subgraph. We will consider all subspaces with the induced conditional probability. Let

$$C_r = B_r \cap C_R,$$

and define C_t analogously. Note that C_R is the disjoint union of $\Theta(R)$ sets isomorphic to C_r. By symmetry we have:

$$E(f|C_R) = E(f|C_r)$$

and

$$E(f|C_T) = E(f|C_t).$$

We will define a mapping $\sigma : C_t \to C_r$ such that σ is 1:1 and measure preserving with respect to the conditional measure on C_t and its image. More precisely, if μ_1, μ_2 are respectively the conditional measures on C_t and $\sigma(C_t)$ then for any $G \in C_t$ $\mu_1(G) = \mu_2(\sigma(G))$. Furthermore σ will be such that

$$f(\sigma(G)) = 1 \Rightarrow f(G) = 1$$

and hence

$$E(f|C_t) \geq E(f|\sigma(C_t)).$$

Therefore it will suffice to show that

$$E(f|\sigma(C_t)) \geq E(f|C_r) - o(1). \tag{20}$$

The definition of σ is very simple: For $G \in C_t$ Define $\sigma(G)$ as the graph obtained from G by deleting the edges in $t \setminus r$. Obviously $\sigma(G) \in C_r$. The question is which graphs in C_r are not in the image of σ. Let $W_1 = \sigma(C_t)$ and $W_2 = C_r \setminus W_1$. The graphs in W_2 can be classified into two types:
1) Graphs that have an edge in $t \setminus r$. (It can be shown that the conditional measure of the set of such graphs is negligible.)
2)Graphs G such that $G \cup t \notin C_t$.

Roughly, the reason for a graph G to belong to W_2 is that there exist a subgraph s in G such that $s \cup t$ is a union of modest graphs, where as $s \cup r$ is not.

Example 4.9 *Take r to be a complete graph on 4 vertices, $x_1, \ldots, x_4$, and t the graph obtained from r by adding on a path x_3, x_5, x_6, x_4. Both r and t are balanced. Now, if a graph in C_r has a path x_5, x_7, x_8, x_6 it will belong to W_2. Another possibility would be a graph with a path x_6, x_7, x_8, x_4. It may be useful to keep these examples in mind for understanding the notion of* an extension*, to be defined shortly.*

For a graph $G \in C_r$ define

$$\alpha(G) = \Pr(\pi(G) \in W_2) \tag{21}$$

where π is a permutation of the vertices leaving the vertices of r fixed, chosen uniformly at random from such permutations. Let

$$a = \frac{\mu(W_2)}{\mu(C_r)}.$$

Let μ_r be the conditional measure on C_r . So

$$a = \int_{C_r} \alpha \; d\mu_r.$$

On the other hand,

$$1 - a = \mu(\sigma(C_t))/\mu(C_r) = \mu(C_t)/(p^{|T|-|R|}\mu(C_r))$$

$$= \frac{\mu(C_T)}{|\Theta(T)|} \frac{|\Theta(R)|}{\mu(C_R)} \frac{p^R}{p^T} = \frac{\mu(C_T)E(R)}{\mu(C_R)E(T)}.$$

Recalling the properties of R and T this shows a is bounded away from 1. Using this and the fact that f is constant on isomorphism classes we have

$$E(f|W_1) = \frac{\int_{C_r} f \cdot (1-\alpha) \; d\mu_r}{(1-a)}$$

But lemma 4.10 below shows that α is essentially constant ($= a$) on C_r, and (20) follows.

□

Lemma 4.10 *Let $\alpha = \alpha(G)$ be as defined in (21), where G is chosen at random from C_r by the measure μ_r. Then*

$$Var(\alpha) = o(1).$$

Proof: First, we would like to shift to working in B_r, a space with a convenient product measure, and to this end we will extend the definition of $\alpha(G)$ to all graphs $G \in B_r$. Note that a graph in B_r almost surely has no edges in $t \setminus r$, and we will disregard the exceptions to this rule in our calculations. For any graph G let $V(G)$ denote the set of vertices of G. Let $k = |V(t)| - |V(r)|$. Order the vertices in $V(t) \setminus V(r)$, $v_1, \ldots v_k$. For any ordered set $x = \{x_1, \ldots x_k\}$ outside of $V(r)$, let π_x be a permutation of the vertices of the complete graph leaving $V(r)$ fixed such that $\pi(x_i) = v_i$ for $i = 1, \ldots, k$. For a given graph $G \in C_r$ define a set of vertices x to be *problematic* if

$$\pi_x(G) \in W_2.$$

This does not depend on the choice of π.
Returning to example 4.9 in that case a set of two vertices is problematic if they have a path of length 3 between them, or one of them has a path of length 3 connecting it to r. Note that $\alpha(G)$ is exactly the proportion of problematic k-sets. For any problematic set x we can find a set of edges and vertices y in $G \setminus (r \cup x)$ that "are a reason" for x being problematic. More precisely $\pi_x(y) \cup t$ is a union of modest graphs. (In our example these are the paths of length 3 added between two vertices in t.) Call such a set an extension of x. We now want to define extensions and problematic sets for any graph in B_r. For x, a set of k vertices disjoint from $V(r)$ in any graph in B_r and a set of vertices and edges y which is disjoint from r and x, we say that y is an extension of x if

$$(\pi_x(y \cup x) \cup r) \in C_r,$$

but,

$$(\pi_x(y \cup x) \cup t) \notin C_t.$$

Remark 4.11 *Note that we view y as a set of edges and vertices and not as a graph, indeed some edges in y may be such that their end vertices are not in y. We do require, however, that $y \cup x$ be a graph.*

Now extend the definition of a problematic set of vertices in a graph in B_r to include any set of vertices that has an extension. For any graph $G \in B_r$ define $\alpha(G)$ to be the proportion of the problematic sets among all sets of size k. This definition of α coincides with the previous one on C_r.
Now, before shifting to work on B_r note that

$$\mu(C_r)/\mu(B_r) = \mu(C_R)/(p^{|R|}|\Theta(R)|) = \mu(C_R)/E(R).$$

From the definition of $\mathcal{C}_4$, and the fact that $C_r \subset \mathcal{C}_4$, $\mu(C_R)$ is bounded from below, and $E(R)$ is bounded from above (by bounds that depend on ϵ) and hence the relative measure of C_r in B_r is non negligible, so

$$\mathrm{Var}(\alpha|B_r) = o(1) \Rightarrow \mathrm{Var}(\alpha|C_r) = o(1).$$

Lemma 4.12 asserts that $\mathrm{Var}(\alpha|B_r) = o(1)$, hence our result follows.

□

Lemma 4.12

$$Var(\alpha|B_r) = o(1)$$

Proof: Define $X = n_k\alpha$, where $n_k = n!/(n-k)!$. $X(G)$ is the number of problematic k-sets in G. Let $x_1, \ldots x_{n_k}$ be the indicator random variables of the event of the corresponding sets being problematic. So $X = \sum x_i$.
If $E(X) = o(n^k)$ then $E(\alpha) = o(1)$ and since $0 \leq \alpha \leq 1$ $\mathrm{Var}(\alpha) = o(1)$. Hence we may assume that $E(X) = \Omega(n^k)$ and strive to prove that $\mathrm{Var}(X) = o(n^{2k})$.
We have

$$\mathrm{Var}(X) = \sum_i \sum_j \left(E(x_i x_j) - E(x_i)E(x_j)\right).$$

For $i \neq j$ let $x_i \circ x_j$ be the random variable indicating the event that there exist *edge disjoint* extensions of the corresponding sets. The BK inequality ([4]) implies

$$E(x_i \circ x_j) \leq E(x_i)E(x_j).$$

(See also [24] for a more general inequality.)
Let $x_i \diamond x_j = x_i x_j - x_i \circ x_j$. We have

$$\mathrm{var}(X) \leq \sum \mathrm{Var}(x_i) + \sum_{i \neq j} E(x_i \diamond x_j) = \sum_{i \neq j} E(x_i \diamond x_j) + o(n^{2k}). \qquad (22)$$

We now need some notation regarding graphs in the space B_r. Define two graphs G and G' to be of the same isomorphism type if there exists a permutation of the vertices not in $V(r)$ that takes G to G'. Define $\tilde{E}(G)$ as the expected number of copies isomorphic to G in a random graph in B_r.

Let the average degree of the modest graphs be δ. For any set of edges and vertices define the average degree to be the ratio between the number of edges in the set and the number of vertices. So, for an extension y of a set x the average degree is the ratio between the number of edges in y and the number of vertices of y (those not in x and not in $V(r)$). (recall that y is not necessarily a proper graph in the sense that not all edges in y are between vertices that belong to y.) Recall that $t \cup \pi_x(y)$ is the union of modest graphs, and hence has average degree at least δ. Hence the average degree of y must also be at least δ.

We will say an extension y is *nice* if

1) It is a minimal extension.

2) Its average degree is δ.

3) For any $z \subset y$ such that $z \cup t$ is a graph the average degree of z is no larger than δ.

The reason for defining this notion is that our calculations our much simpler when considering such extensions. Whenever $x_i \diamond x_j = 1$ there are minimal extensions causing this. Furthermore, for a given set x, the probability of having a minimal extension with average degree of it or any subextension larger than δ, is $o(1)$. Hence we may concentrate on the events caused by nice extensions: Let $x_i \star x_j$ be the indicator random variable of the event indicated by $x_i \diamond x_j$, but only in the case where there exist nice extensions causing this event. We have

$$\sum E(x_i \diamond x_j - x_i \star x_j) = o(n^{2k}).$$

and hence it suffices to show

$$E\left(\sum x_i \star x_j\right) = o(n^{2k}). \tag{23}$$

For an extension y let $Cl(y)$ denote the graph whose edges are the edges in y. We will need the following property of nice extensions:

Claim 4.13 *Let x be a set of vertices and y a nice extension. Any subgraph of $Cl(y) \setminus r$ has average degree smaller than δ.*

(Note that the claim deals with actual graphs and not extensions, i.e. all edges come with their end vertices.)
Proof: Let y be a nice extension of x, and assume for simplicity of notation that $x = V(t) \setminus V(r)$. Note that from minimality of y there exists a modest graph S such that $Cl(y) \subseteq S \subseteq y \cup t$. This S is the disjoint union of three sets:
a) $S \cap r$.
b) $S \cap (t \setminus r)$
c) All the rest, namely $y \setminus t$.
Note that parts (b) and (c) are not necessarily proper graphs, i.e. they may have edges with only one vertex belonging to them.
Part (c) does not have average degree larger than δ, because y is nice. Part (b) can not have average degree larger than δ, or else its union with r would also have large average degree, but this union is a subgraph of t which has no such subgraphs. Hence part (a) can not have average degree *smaller* than δ. As a subgraph of r it can not have large average degree either, and hence has average degree exactly δ. Now, if $z \subset (Cl(y) \setminus r)$ has average degree δ then $z \cup (S \cap r)$ is modest (it is a subgraph of S which is modest, and has the correct average degree.) Hence $r \cup z$ is a union of modest graphs, which is a contradiction: since y is an extension, there exists a graph in C_r with z as a subgraph, but in C_r the union of all modest graphs is r.

□

Let y be a nice extension of a set x. The graph in B_r consisting of the vertices of x and the edges and vertices of y can take on a finite number of isomorphism types $G_1, \ldots, G_d$. For all these graphs we have

$$\tilde{E}(G_i) = O(n^k). \tag{24}$$

This follows from the fact that there is a copy of G_i, say g whose union with t is a union of modest graphs. We have $E(g \cup t) < c$. (A finite union of modest graphs has bounded expectation.) But $E(g \cup t) \geq cE(t)\tilde{E}(g)/n^k$ and $E(t)$ is bounded from below.
We now can prove (23):
When summing $E(\sum x_i \star x_j)$ we use the fact that if an extension of type R_i intersects an extension of type R_j and their intersection is of type H they form an event such that the expected number of isomorphic events is $\tilde{E}(R_i)\tilde{E}(R_j)/\tilde{E}(H)$. But from the fact that R_i is nice it follows from claim

4.13 that the average degree of H is smaller than δ , or in other words, $\tilde{E}(H) \to \infty$, and

$$\tilde{E}(R_i)\tilde{E}(R_j)/\tilde{E}(H) = o(n^{2k}).$$

Since we have a finite number of such contributions this gives the desired bound.

□

The last brick missing in the proof of the theorems is the following lemma. In the previous section we defined g_2, an approximation of f, and $\mathcal{C}$ as the union of all sets C_G with the following properties:
1)$|G|$ was bounded from above.
2)$E(G)$ was bounded from above and below.
3)$\mu(C_G)$ was not too small.
We promised to show that g_2 is almost determined by the number of appearances of the graphs S_i, in the sense that if $C_S \subset \mathcal{C}$ then for any constant $\delta > 0$

$$Pr(\{|g_2(T) - E(g_2|C_S)| > \delta\}|T \in C_S) \to 0 \tag{25}$$

Recalling the Fourier expansion of g_2 it is sufficient to show this for the functions $\hat{f}(S)V_S$, where S is modest, or using (8):

Lemma 4.14 *Let G be such that $C_G \subset \mathcal{C}$. Let S be modest, and $V = V_S$, then:*

$$\forall \delta > 0 \ Pr(\{|V(T) - E(V|C_G)| > \sqrt{p}^{(-|S|)}\delta\}|T \in C_G) \to 0.$$

Proof: Note that all the modest graphs have the same average degree: the only one that guarantees a bounded (from above and below) expectation. Furthermore $E(S)/E(R)$ is bounded for any R that is a subgraph of a modest graph S. Recalling lemma 3.1 we wish to show that for $R \subseteq S$ $\frac{X_R}{E(R)}$ is almost constant on each C_G. So let us calculate the conditioned variance of X_R in a given C_G. If R is modest then X_R is constant. So we may concentrate on R which is a subgraph of one of the S_i , but not modest itself, i.e. $E(R)$ is large. Recall that lemma 3.4 gave the following expression for the non-conditioned variance of X_R:

$$\mathrm{Var}(X_R) \asymp E(R)^2(\sum 1/E(H))$$

where the sum is over all non-empty subgraphs of R . If none of the modest graphs are subgraphs of R this is $o(E(R)^2)$, since for every subgraph H,

$E(H)$ is also large. So the standard deviation of X_R in this case is o$(E(R))$. Obviously conditioning on an event whose probability is bounded away from zero (being in C_G) can not change this. Let us consider then, the variance of X_R when some S_i is a subgraph of R. C_G was defined by the fact that the union of the modest graphs appearing was isomorphic to G. Let g be a *specific* copy of G. Let B_g be the space consisting of all graphs that have g as a subgraph with the probability measure induced by the conditional probability in $G(n,p)$. As in remark 4.6 we define graphs, orbits, expected number of copies of a graph, etc. in our new space in the natural way. Let $C_g = C_G \cap B_g$ From symmetry we get that the expectation and variance of X_R conditioned on being in C_G is the same as conditioning on being in C_g. Focusing our attention on C_g every copy of R must have g as a subgraph. Therefore X_R now depends on the appearance of copies of certain graphs $T_1, T_2, \ldots, T_k$, such that $T_i \cup g \sim R$. So we may now define X_T so that $X_R = X_T$ (in the space C_g) , but X_T counts the appearance of copies of the T_i's . Since g is the union of all modest graphs in any graph in C_g we may assume T_i has no modest subgraphs in B_g, and $E(H) \to \infty$ for all $H \subset T_i$.

Now, a simple calculation shows that

$$\mu(C_G \cap B_g)/\mu(B_g) = \mu(C_G)/E(G)$$

and from the definition of $\mathcal{C}$, $E(G)$ is bounded from above and $\mu(C_G)$from below, hence if $\mathrm{Var}(X_T) = o(E(X_T)^2)$ conditioned on being in B_g the same must be true on C_g.

Remark: The reason for calculating in the space B_g and not directly in C_g is that the conditional measure in the first space is much simpler than that of the later.

We now repeat the calculation done in lemma 3.4, in the same manner as done in the proof of lemma 4.12 with the sum

$$\sum E(XY) - E(X)E(Y)$$

and the expectations as defined in the space B_g. We get that

$$\mathrm{Var}(X_T) \leq cE(X_T)^2/\mathrm{Min}_{H \subset T_i} E(H).$$

From our remark concerning $E(H)$ we conclude that this is $o(E(T)^2)$, so the standard deviation of X_R is indeed $o(E(R))$.

The above considerations show that $X_R/E(R)$ is almost a constant on any

set C_G , (its standard deviation is o(1)), and hence by lemma (3.1) $\hat{f}(S)V_S = cV_s/\sqrt{|\Theta(S)|}$ is indeed almost constant on our subsets. Moreover for S's such that $\sum_{H\in\Theta(S)} \hat{f}^2(H) = c$

$$\hat{f}(S)V_S = V_s\sqrt{c/|\Theta(S)|} \sim \sqrt{c/E(S)}(\sum(-1)^{(|S|-|R|)}\frac{\tilde{E}(R)}{E(R)})$$

where $\tilde{E}$ is the conditioned expectation of X_R. This completes the proof of the lemma and the theorem.

□

5 The k-sat problem.

The following problem has attracted much attention from physicists and computer scientists, see [20] for a survey on this topic: Let $x_1, x_2, \ldots, x_n$ be Boolean variables and consider a CNF formula, made of clauses of size k of the variables and their negations, i.e. a conjunction of clauses each of which is a disjunction of k of the variables and their negations. A random formula, with parameter M is generated in the following way: Pick M of the possible $2^k\binom{n}{k}$ clauses with uniform probability, and let the formula be the disjunction of the chosen clauses. A property of interest of the formula such obtained is whether it is *satisfiable*, i.e. whether there is an assignment of values to $x_1, \ldots, x_n$ such that the formula takes on the value "true". Denote the probability of such event by $f(M)$. It is obvious that f is a monotone decreasing function of M. It is known that for any given k there are constants c_1, c_2 such that $f(c_1 n) \to 1, f(c_2 n) \to 0$. (see [10].)
Computer simulations suggest that f exhibits a threshold behavior, i.e. that the following is true: There exists a constant c such that for any $\epsilon > 0$, $f((c-\epsilon)n) \to 1, f((c+\epsilon)n) \to 0$.
This was shown to be true for $k = 2$ with the constant $c = 1$, see [10], [17], but for $k \geq 3$ was not known. For $k = 3$ a series of upper and lower bounds have seemed to be slowly converging to the value suggested by simulations ($c = 4.2...$) see [19], [11], [9], [8], [16], [22], [21].
We now show that the existence of a threshold for any given k can be demonstrated by the proof of theorem 1.1. I would like to thank Svante Janson for pointing out the following subtlety to me: What I actually show is not the existence of a constant c but of a function $c(n)$ such that the phase transition

happens within an ϵ neighborhood of $c(n)$, i.e. it is still feasible that though there is a swift transition of f the critical value does not converge to any given value.

First let us consider the dual problem, of the formula being a DNF formula: i.e. a disjunction of k-conjunctions, and the property we shall study is whether or not the formula is a tautology, i.e. does every assignment of values to the Boolean variables yield the value "true" for the formula. This is a monotone increasing property.

Secondly let us consider a model for producing a random formula which relates to the previous model in the same way $G(n,p)$ relates to $G(n,M)$: choose each of the possible clauses independently with probability p, and let the formula be the disjunction of the chosen clauses. For $p \approx M/N$, where $N = 2^k\binom{n}{k}$ this is equivalent to the previous model in the following sense: the question of existence of a "critical" constant c as described above is equivalent to the following question: By abuse of notation define $f(p)$ as the analog of $f(M)$, does there exist a constant c such that for every $\epsilon > 0$, $f((c-\epsilon)n/N) \to 1, f((c+\epsilon)n/N) \to 0$?

Returning to the definitions in the introduction what we are asking is: "Does the property of satisfiability have a sharp threshold?" We claim that the answer is affirmative.

To show this we must first point out the analogy between the case of graphs and the case of DNF formulas. We viewed graphs as a collection of pairs (i,j) with i,j taken from a set of vertices. Our DNF formulas are a slight generalization of hypergraphs: they can be thought of as a collection of k-tuples chosen from a set of variables, with one of 2^k possible labels on each edge, specifying which variables appear with a negation.

The group of graph automorphisims acting on the subgraphs of K_n can be viewed as S_n acting on $\binom{[n]}{2}$, and we only considered properties invariant under the action of this group. In the case of formulas we will consider properties (i.e. families of formulas) invariant under the action of the wreath product of S_n with k copies of Z_2. The property of being a tautology, (or satisfiability) is such a property.

A crucial aspect of the analogy is the following: given a bound on the number of edges (clauses) of a graph (formula), there are only a finite number of isomorphism types.

Following the proof of theorems 1.1 and 1.2 shows that the analogy holds all the way through, and gives for the probability space of all DNF $k-$formulas:

Theorem 5.1 *Consider DNF formulas with clauses of given size k. There exists a function $M(k, \epsilon, c)$ such that for every $c > 0$ and every monotone symmetric family of such formulas, A, such that $p \cdot I \leq c$, for every $\epsilon > 0$ there exists a symmetric monotone family B such that $\|B\| \leq M$ and $\mu(A \triangle B) \leq \epsilon$.*

Here $\|B\|$ is the number of clauses in the largest minimal formula in B. Let $|G|$ be the number of clauses in G. Now define in the obvious manner for a formula G, $E(G)$ to be the expected number of sub formulas isomorphic to G in a random formula. The average degree of a formula G is the ratio between the number of variables and the number of clauses in G. Define a balanced formula to be one with average degree no less than that of any sub formula. The proof of theorem 1.2 gives:

Theorem 5.2 *There exist functions $B(\epsilon, c)$, $b_1(\epsilon, c)$, $b_2(\epsilon, c)$ such that for all $c > 0$, any n and any monotone symmetric family T of k-DNF formulas with n variables such that $p \cdot I \leq c$, for every $\epsilon > 0$ there exists a formula F with the following properties:*

- *F is balanced*
- *$b_1 < E(F) < b_2$*
- *$|F| \leq B$*
- *Let $Pr(T|F)$ denote the probability that a random formula belongs to T conditioned on the appearance of $\bar{F}$,* a specific copy *of F. Then*

$$Pr(T|F) \geq 1 - \epsilon$$

So if a property A of formulas has a coarse threshold, then for a certain value of p in the critical interval, for every $\epsilon > 0$ there exists a "nice" formula F such that the probability of having A conditioned on the appearance of $\bar{F}$, a specific copy of F is at least $1 - \epsilon$.
We will show shortly that for the property of being a tautology one can not produce such a "magic" formula. As a corollary we get theorem 1.3:

Corollary 5.3 *In the space of all DNF $k-$formulas the property of being a tautology has a sharp threshold.*

Proof:
Let p in the critical interval be such that $p \cdot I < c$, and assume w.l.g $\mu_p(T) = 1/2$. ($\mu(T)$ is bounded from 0 and 1 by the definition of the critical interval). Let T be the property of being a tautology. What we will show is that there does not exist a short formula $\bar{F}$ as described in theorem 5.2.
Assume $\bar{F}$ is such a formula. Obviously if $\bar{F}$ has a sub formula $\bar{R}$ that itself is a tautology then $Pr(T|F) = 1 \geq 1 - \epsilon$, however, an unpublished result of M. Tarsi (see [2]) states that if such a formula R, uses r variables it must have at least $r+1$ clauses. The expected number of formulas of such an isomorphism type in a random formula is therefore at most $n^r p^{r+1}$. Since $p = n/N < 1/n$ this tends to zero. But if F is balanced and $E(F)$ is bounded from below so is $E(R)$, hence this is a contradiction.
Let r be the number of variables in the formula $\bar{F}$. Define a *quasi tautology* on r variables to be a formula which is a disjunction of k-conjunctions of variables $x_1, \ldots, x_r$ such that it is satisfied by all but one of the 2^r possible assignments to the variables. Let $\bar{M}$ be a maximal quasi tautology on the r variables (adding any additional clause to it would make it a tautology), such that $\bar{F}$ is a sub formula of $\bar{M}$. From positive correlation of increasing events it would follow from our assumptions that $\Pr(T|\bar{M}) > 1 - \epsilon$. So it is sufficient to show that for any $\tau > 0$ if n is sufficiently large,

$$\Pr(T|\bar{M}) < 1/2 + \tau. \tag{26}$$

Define $p(n)$ to be the critical p such that $\mu_p(T(n)) = 1/2$, where $T(n)$ is the family of tautologies on n variables. Note that $p(n)$ is monotone decreasing as a function of n, so that $\mu_{p(n)}(T(n-r)) \leq 1/2$.
Let $1/2 - \epsilon > \tau > 0$ be some constant. The following claim implies (26):

Claim 5.4 *consider $n-r$ variables and build a random k-DNF formula with $p = p(n)$. Now perform a second stage and add with probability $r^k p$ each of the clauses with less than k variables (corresponding to the clauses in which some of the r variables of the quasi tautology appeared.) The resulting formula is a tautology with probability no more than $1/2 + \tau$.*

To simplify matters we will prove a claim that is even stronger.
After the first stage the probability of having a tautology was less than $1/2$. In the second stage with probability tending to 1 no clauses of size smaller than $k-1$ were chosen (recall that $p \asymp n^{1-k}$.) The expected number of clauses of size $k-1$ that were added can be bounded by a constant c. Define

$d = 2c/\tau$. The probability that more than d clauses were added in the second stage is less than $\tau/2$. Therefore claim 5.4 is implied by the following:

Claim 5.5 *As before start with a random DNF formula on $n - r$ variables with k-clauses and $p = p(n)$, and in the second stage pick at random d different $(k-1)$- clauses, and add them to the formula. The resulting formula is a tautology with probability $< 1/2 + \tau/2$.*

We will prove something even stronger: Assume that in the second stage the clauses added are not of size $k-1$ but of size 1. Still, this does not increase the probability of a tautology to $1/2 + \tau/2$. First we need the following:

Lemma 5.6 *: Let $f(n) = o(\sqrt{n})$. Assume the second stage of building the formula consists of adding $f(n)$ clauses of size k. Then $Pr(T)$ after the second stage is less than $1/2 + \tau/2$.*

Proof: Consider $\mu_p(T)$ as a function of p. We are interested in the slope of this function in a neighborhood of p_c ($p_c = cn/N$). The lemma will follow if we show the slope is $O(N/\sqrt{n})$, since enlarging p by δ results with an expected addition of δN clauses.
Let M be a Hamming ball, the family of all formulas of size larger than Np_c. The following two facts are easy exercises, and the lemma follows from them:
1)$d\mu_p(M)/dp|_{p=p_c} \approx \sqrt{N/p_c}$.
2)This is the maximum possible slope at p_c for all monotone families of formulas.

□

So we know that if in the second stage we add, say, $n^{1/4}$ clauses of size k we can not increase the probability of a tautology to $1/2 + \tau$. We wish to show that this implies that a constant number of clauses of size 1 will not suffice either. Note that if after the first stage we do not yet have a tautology, the probability of success in the second stage is no more than $1 - (1/2)^d$. In any such case the following lemma will show that a large number of clauses of size k will yield a tautology with probability higher than that of d clauses of size 1:

Lemma 5.7 *For $A \subseteq \{0,1\}^n$ define A to be (d, m, ϵ)-coverable if the probability for the union of a random choice of d sub cubes of co-dimension m to cover A is at least ϵ. Let $f(n)$ be any function that tends to infinity as n tends to infinity. For fixed k, d and ϵ and sufficiently large n any $A \subseteq \{0,1\}^n$ that is $(d, 1, \epsilon)$-coverable is $(f(n), k, \epsilon)$-coverable.*

Proof: Let $A \subseteq \{0,1\}^n$ be $(d,1,\epsilon)$-coverable. This means that sequentially choosing at random d half cubes and building their union covers A with probability not less than ϵ. Now, instead of picking the last half cube, pick at random $\sqrt{f}/d$ cubes of co-dimension k. We will prove below that this decreases the probability of ending with a cover of A by no more than $\epsilon/2d$. A trivial but helpful observation is that first choosing the sub cubes and then the half cubes yields the same result. This enables us to repeat this consideration d times and conclude that A is $(\sqrt{f},k,\epsilon/2)$-coverable. Since ϵ is fixed and f is large this implies that A is (f,k,ϵ)-coverable.
We now prove the above claim, that picking at random $\sqrt{f}/d$ cubes of co-dimension k instead of the last half cube decreases the probability of ending with a cover of A by no more than $\epsilon/2d$.

Our claim will follow if we show that for any $\alpha \geq \epsilon/2d$ a set which is $(1,1,\alpha)$-coverable is $(\sqrt{f}/d,k,\alpha)$-coverable.
For a set A to be $(1,1,\alpha)$-coverable means that it is a subset of the intersection of s half-cubes, where $s \geq 2n\alpha$. We may assume without loss of generality that it is *exactly* the intersection of $s = 2n\alpha$ half cubes. for a given g we will bound from below the probability of g sub cubes of co-dimension k covering A by the probability that at least one of them has A as a subset. The probability of this is approximately $1-(1-(s/2n)^k)^g$. So choosing $g \approx \alpha^{-k}$ gives a cover with probability that is a constant, and hence $g = \sqrt{f}/d$ yields a cover with probability close to 1. This completes the proof of the lemma and with it the proof of the theorem.

□

6 Other Applications

The approach used to solve the k-sat problem can be used to prove sharpness of thresholds in other cases in a similar manner. Here are a few examples:

- **The existence of a perfect matching in a 3-uniform (or r-uniform) hypergraph:** consider a random 3-uniform-hypergraph on $n = 3k$ vertices with edge probability p. The property of interest is that of the existence of a disjoint covering of the vertices by k edges. What is currently known about the value of the critical p for this property is

$$\log(n)/n^2 \leq p_c \leq n^{-4/3}.$$

(See [27] and [15]). The question of showing that $p_c \leq n^{-(2-o(1))}$ is considered to be one of the challenging problems in random (hyper)graph theory. However we may now deduce the sharpness of the threshold: By theorem 1.2 this property has a sharp threshold since it can not be approximated by the appearance of a fixed sub hypergraph . The proof of this is straightforward:
Proof: Assume by contradiction that there exists such a hypergraph $\bar{H}$. Let m be the number of edges in $\bar{H}$, and assume the probability of having a matching conditioned on the appearance of $\bar{H}$ is substantially larger than the unconditioned probability, which is 1/2. The only contribution $\bar{H}$ gives is by using some of its edges for creating a matching. It is not hard to see that adding, say, $m2^m$ edges at random must "help" to achieve a matching even more. But as in the case of the k-sat we know that if $X = o(\sqrt{p_c\binom{n}{3}})$, then adding X edges can not make such a difference.

□

Remark: A similar proof works for the case of "H-factors", the property of having a covering of the vertices of $G(n,p)$ by disjoint copies of some fixed graph H. See [3] for this problem. However in this case, as pointed out to me by Noga Alon, it is not enough to use the fact that $o(\sqrt{E})$ edges (where E is the expected number of edges) do not make a difference. Here one should use the fact that even $o(E)$ edges should not make a difference, or else the threshold would be sharp. This type of proof seems to be easy for some "non-local" properties such as connectivity or having a perfect matching.

- **k-colorability for $k > 2$.** In a paper in preparation [1] it is shown by similar techniques that the property of being non-k-colorable for a fixed k larger than 2 has a sharp threshold. The crux of the proof there is to show that if $G(n,p)$ is non-k-colorable with probability 1/2, this does not change substantially if the color of a fixed number of vertices is predictated.

- **Properties for which the critical probability is** $\log(n)/n$. Such properties have a sharp threshold by theorem 1.4. This reproves the well known facts that connectivity, having a Hamilton cycle and other such properties have a sharp threshold.

7 Consequences of the Appendix

Before dealing with the consequences of the appendix I would like to describe the chronological development of the results in this paper and the appendix. After the first draft of this paper was written Jean Bourgain came up with the results described in the appendix. At that stage theorem 1.1 was stated in a weaker form, with the restriction that p must be close to a rational power of n, and theorem 1.4 was a conjecture. The appendix contains two main results: one of them, proposition 1 is analogous to theorem 1.1 but is placed in a more general setting where symmetry plays no role. The second is proposition 2 which states that conjecture 2.4 is true. Conjecture 2.4 itself was strong enough to imply together with the rest of the paper at that stage that theorems 1.1 and 1.4 were true.
After this Joel Spencer suggested extensions of the arguments in the first version of the paper to get the present strengthened versions of the theorems. This in turn led to a simpler approach which consisted of slight alteration of the original version of the paper, yielding the present version.

Here are some reflections as to the consequences of the results described in the appendix:
∗What is proven in proposition 1 is of course more general than theorem 1.1 since it holds with no assumptions on symmetry. On the other hand in the setting of graphs it does not imply theorem 1.1. However in every application mentioned in this article (k-sat, k-colorability, etc.) it seems that both theorems can be used equaly well to prove sharpness, since they both deal with the possibility of approximating "global" properties by "local" ones. It seems that this will happen for essentially all applications.
∗Proposition 2 gives an immediate proof of theorem 1.4. This proof is presented in the appendix. It also can be used to substantialy simplify the proof of lemma 4.2 which is a key lemma in this paper.

Results similar to those of this paper may be deduced from the appendix in certain cases where there is a group action under which the families considered are invariant, and the number of different isomorphism types of sets with a bounded size is bounded. An intriguing question is what can be said about the possible values of p_c for properties with a coarse thresholds in the case of a family of subsets of $\{1, \ldots, n\}$ that is invariant, say, under the action of the cyclic group, C_n.
∗Finally, it would be interesting to try to prove conjecture 1.5 using the techniques of the appendix.

Acknowledgments

I would like to thank Gil Kalai for his patience and assistance in the preparation of this paper, and guidance throughout my dealing with these topics. I would like to thank Nati Linial for introducing me to the k-sat problem and encouraging me to work on this topic, and both of the above for many long, useful and instructive discussions.
I would also like to thank Dorit Aharonov who is responsible for a quantum leap in the level of readability and correctness of this paper.
I wish to thank Joel Spencer for useful discussions leading to substantial improvements in this paper, as mentioned in section 7.
Thanks to Christian Borgs and Van Ha Vu for pointing out an error in an earlier version of this paper, and to Jeong Han Kim for useful discussions.
Thanks to Svante Janson for useful suggestions, and for pointing out various inaccuracies in earlier versions. His extreme care in refereeing this paper was admirable.
Last but not least I would like to thank Jean Bourgain for writing the appendix to this paper in which he presents a more general and compact approach to the problems treated here.

References

[1] D. Achlioptas, E. Friedgut, A Threshold for k-Colorability, To appear, Random Structures and Algorithms.

[2] R. Aharoni, N. Linial, Minimal Non-2-colorable Hypergraphs and Minimal Unsatisfiable Formulas, Journal of Combinatorial Theory, Series A Vol. 43 no.2 November 1986.

[3] N. Alon and R. Yuster, Threshold functions for H-factors, Combinatorics, Probability and Computing 2 (1993), 137-144.

[4] J. van den Berg and H. Kesten Inequaleties with application to percolation and reliability. Journal of Applied Probability 22 (1985) 556-569.

[5] B. Bollobás, Random Graphs, Academic Press, London, 1985.

[6] B. Bollobás and A. Thomason, Threshold functions, Combinatorica 7 (1986) 35-38

[7] J. Bourgain, G. Kalai, Influence of variables in product spaces under group symmetries, to appear in GAFA.

[8] A.Z. Broder, A. M. Frieze, E. Upfal (1993) On the Satisfiability and Maximum Satisfiability of Random 3-CNF Formulas. In Proc. 4th. Ann. ACM-SIAM Symposium on Discrete Algorithms, pp322-330.

[9] M. T. Chao, J. Franco (1986) Probabilistic Analysis of Two Heuristics for the 3-sat. Problem. Siam J. Comput. 15(4).

[10] V. Chvatal, B. Reed, Mick gets some. Proc. 33rd Ann. FOCS Symp.(1992) 620-627.

[11] A. El Maftouhi, W. Fernandez de la Vega, On Random 3-sat, Combinatorics, Probability and Computing (1995) 4, 189-195.

[12] P. Erdös, A. Rényi, On the evolution of random graphs, Mat Kutató Int. Közl. (1960)5, 17-60.

[13] E. Friedgut, Boolean functions with low average sensitivity depend on few coordinates, to appear, Combinatorica.

[14] E. Friedgut, G. Kalai, Every monotone graph property has a sharp threshold, Proc. Amer. Math. Soc. 124 (1996), pp. 2993-3002 .

[15] A. Frieze, S. Janson, Perfect Matchings in Random s-uniform Hypergraphs. Random Structures and Algorithms, 7 (1995), no. 1, 41-57.

[16] A. Frieze, S. Suen, (1992) Analysis of Two Simple Heuristics on a Random Instance of k-sat. Journal of Algorithms 20, (1996) 312-355.

[17] A. Goerdt (1992) A threshold for satisfiability. In Math. Foundations of Computer Science (I.M.Havel and V.Koubek eds.) Prague, Poland.

[18] 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.

[19] A. Kamath, R. Motwani, K.Palem, P.Spirakis Tail Bounds for Occupancy and the Satisfiability Threshold Conjecture. Proc. 35th FOCS pp. 592-603.

[20] S. Kirkpatrick, B. Selman, Critical Behaviour in the Satisfiability of Random Boolean Expressions, Science, Vol.264 (1994) 1297-1301.

[21] M. Kirouris, E. Kranakis and D. Krizanc, Approximating the unsatisfiability threshold of random formulas, in: Algorithms - ESA' 96, Lecture Notes in Computer Science 1136

[22] T. Larrabee, Y. Tsuji (1992), Evidence for a Satisfiability Threshold for Random 3CNF Formulas. Technical report UCSC-CRL-92-42, University of California, Santa Cruz.

[23] G. Margulis, Probabilistic characteristics of graphs with large connectivity, Prob. Peredachi Inform. 10(1974), 101-108.

[24] D. Reimer, Butterflies, 1995 (to appear).

[25] L. Russo, On the critical percolation probabilities, Z. Wahrsch. werw. Gebiete, 43(1978), 39-48.

[26] L. Russo, An approximate zero-one law, Z. Wahrsch. werw. Gebiete, 61 (1982), 129-139.

[27] J. Schmidt-Pruzan, E. Shamir, A threshold for perfect matchings in random $d-$pure hypergraphs, Combinatorica 5 (1985) 81-94.

[28] M. Talagrand, On Russo's approximate 0-1 law, The Annals of Probability, 1994, Vol.22 No.3 1576-1587.